



Theses and Dissertations

2020-04-09

A Longitudinal Investigation of the Effects of a Kindergarten Multi-Tiered Oral Narrative Language Intervention on Later Literacy Outcomes

Tristin Carolyn Hampshire
Brigham Young University

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>



Part of the [Education Commons](#)

BYU ScholarsArchive Citation

Hampshire, Tristin Carolyn, "A Longitudinal Investigation of the Effects of a Kindergarten Multi-Tiered Oral Narrative Language Intervention on Later Literacy Outcomes" (2020). *Theses and Dissertations*. 8974.
<https://scholarsarchive.byu.edu/etd/8974>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.

A Longitudinal Investigation of the Effects of a Kindergarten Multi-Tiered Oral
Narrative Language Intervention on Later Literacy Outcomes

Tristin Carolyn Hampshire

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of
Master of Science

Douglas B. Petersen, Chair
Christopher Dromey
Tyson Harmon

Department of Communication Disorders
Brigham Young University

Copyright © 2020 Tristin Carolyn Hampshire

All Rights Reserved

ABSTRACT

A Longitudinal Investigation of the Effects of a Kindergarten Multi-Tiered Oral Narrative Language Intervention on Later Literacy Outcomes

Tristin Carolyn Hampshire
Department of Communication Disorders, BYU
Master of Science

The purpose of the current study was to examine the longitudinal effects of a multi-tiered narrative language intervention in at-risk students provided in kindergarten on fourth grade reading comprehension. The participants included 686 students from four school districts in the upper Midwest. Twenty-eight kindergarten classrooms were randomly assigned to a treatment or control condition resulting in 14 treatment classrooms and 14 control classrooms. Every student in the study participated in a pretest regarding oral narrative language skills. Students in the control group were considered to be at-risk, average performing, or advanced performing depending on their pretest score. Each student in the treatment group received large group oral narrative language instruction that followed *Story Champs* procedures and was led by the classroom teachers for 14 weeks. The control group engaged in their regular classroom instruction that was established at the commencement of the school year. Students who were unable to meet the narrative retell criterion at pretest and whose oral narrative retell skills did not improve after one month of large group instruction then received additional small group (Tier 2) oral narrative intervention for 10 weeks. Tier 2 intervention followed *Story Champs* small group procedures and was administered by speech-language pathologists. Posttest scores reflecting a significant difference in progress between treatment and control groups in narrative skill in kindergarten are given in Mollie Brough's thesis (Brough, 2019). Reading comprehension was then measured five years later via the state standardized assessment. The results indicated that the at-risk treatment group had similar reading comprehension scores to the average performing, advanced performing and combined average and advanced performing control groups. This study provides preliminary evidence of the effectiveness of a multi-tiered oral narrative language intervention on later reading comprehension skills in at-risk students.

Keywords: oral language, narrative, reading comprehension, multi-tiered intervention

ACKNOWLEDGMENTS

I express my appreciation to my committee members: Dr. Petersen, Dr. Dromey, and Dr. Harmon who provided feedback and support throughout the writing process. I want to thank Trina Spencer for her contributions made in the initial stages of the study and Maureen Statkowski for her assistance and dedication in obtaining fourth grade data. I especially want to thank my chair, Dr. Petersen. I could not have completed this thesis project were it not for his enthusiasm, expertise, encouragement, and many late hours devoted to helping execute this project. Dr. Petersen's positivity and passion for child language was very influential throughout my graduate school experience and will continue to inspire me throughout my future career in speech-language pathology. I would also like to express gratitude to my parents for their constant love, support, and sacrifice throughout my undergraduate and graduate coursework. I could not have accomplished what I have without them. Lastly, I would like to thank my partner, Ross, for believing in me and always encouraging me to follow my dreams

TABLE OF CONTENTS

| | |
|---|------|
| TITLE PAGE | i |
| ABSTRACT..... | ii |
| ACKNOWLEDGMENTS | iii |
| TABLE OF CONTENTS..... | iv |
| LIST OF TABLES..... | vii |
| LIST OF FIGURES | viii |
| DESCRIPTION OF THESIS STRUCTURE AND CONTENT | ix |
| Introduction..... | 1 |
| Academic Language and Reading Comprehension | 1 |
| The Overemphasis of Decoding in Reading Comprehension..... | 2 |
| Improving Oral Language Comprehension Skills | 4 |
| Targeting Reading Comprehension Through Narrative Language Intervention | 7 |
| Research Questions..... | 10 |
| Method | 10 |
| Participants..... | 11 |
| Pretest Measures | 13 |
| Pretest narrative retells..... | 14 |
| Pretest personal story generations..... | 15 |
| Pretest expository language | 15 |
| Pretest narrative writing..... | 15 |
| At-risk, average performing, and advanced performing control groups..... | 15 |
| Intervention Procedures | 16 |

| | |
|---|----|
| Large group (Tier 1) narrative intervention procedures | 16 |
| Small group (Tier 2) procedures | 17 |
| Control group | 20 |
| Matching of participants | 21 |
| Posttest Measures | 21 |
| Michigan Student Test of Educational Progress (M-STEP) | 21 |
| Scoring | 22 |
| Fidelity and Scoring Reliability | 23 |
| Interventionists and fidelity of intervention | 23 |
| Pretest administration fidelity and scoring reliability | 24 |
| Posttest administration fidelity and scoring reliability | 25 |
| Results | 26 |
| Research Question 1 | 26 |
| Research Question 2 | 26 |
| Research Question 3 | 27 |
| Results Summary | 27 |
| Discussion | 28 |
| Oral Narrative Language and Reading Comprehension | 28 |
| Multi-Tiered System of Language Support | 32 |
| Practicality | 33 |
| Early Identification Using Large Group Dynamic Assessment | 33 |
| Limitations | 34 |
| Conclusions | 35 |

| | |
|--|----|
| References..... | 36 |
| APPENDIX A: Annotated Bibliography..... | 43 |
| National Reading Problem..... | 43 |
| The Relationship Between Reading Comprehension and Oral Language Skills..... | 43 |
| Narrative Language Intervention..... | 54 |
| MTSLS Can Help All Students Improve Language Skills..... | 70 |
| Other Related Articles..... | 72 |
| APPENDIX B: Posttest M-STEP Reading Comprehension Measure Sample..... | 78 |
| APPENDIX C: IRB Approval Form..... | 81 |

LIST OF TABLES

| | | |
|---------|--|----|
| Table 1 | <i>Descriptive Information for Kindergarten Treatment and Control Group</i> | |
| | <i>Participants</i> | 12 |
| Table 2 | <i>Descriptive Information for Fourth Grade Treatment and Control Group</i> | |
| | <i>Participants</i> | 13 |
| Table 3 | <i>Posttest Means, Standard Deviations, and T-Test Results for the At-Risk</i> | |
| | <i>Treatment Group and the Average/Advanced Control Groups Combined</i> | |
| | <i>in Reading Comprehension</i> | 26 |
| Table 4 | <i>Posttest Means, Standard Deviations, and T-Test Results for the At-Risk</i> | |
| | <i>Treatment Group and the Average Control Group in Reading Comprehension</i> | 27 |
| Table 5 | <i>Posttest Means, Standard Deviations, and T-Test Results for the At-Risk</i> | |
| | <i>Treatment Group and the Advanced Control Group in Reading Comprehension</i> | 28 |

LIST OF FIGURES

Figure 1. Methods flowchart 19

DESCRIPTION OF THESIS STRUCTURE AND CONTENT

This thesis, *A Longitudinal Investigation of the Effects of a Kindergarten Multi-Tiered Oral Narrative Language Intervention on Later Literacy Outcomes*, is part of a larger study with some of the data reported in Brough's (2019) thesis. The structure of this thesis is written in a hybrid format with the introductory pages following university requirements and the report presented as a journal article. Appendix A includes the annotated bibliography. Appendix B provides an example of the Michigan Student Test of Educational Progress reading comprehension measure.

Introduction

Several studies have found that the Simple View of Reading (SVR) has considerable explanatory power of the reading process (Gough & Tunmer, 1986). The SVR proposes that reading is the product of decoding and language comprehension. Decoding is the ability to apply letter-sound relationships of print to accurately identify a word. Language comprehension is the ability to understand language and extract literal and inferred meanings (Hoover & Gough, 1990; Hoover & Tunmer, 2018). In a cross-sectional analysis on first, second, and third graders to determine what impacts reading comprehension, Language and Reading Research Consortium (2015) found that both pieces, word recognition and listening comprehension, influence reading comprehension and that 90% of the variance in reading comprehension is explained by the SVR.

Academic Language and Reading Comprehension

Children need strong academic language skills to understand what they read and to succeed in school. Academic language is the oral and written language used by teachers, textbooks, and general curriculum that is required for learning. Academic language includes a literate language style. Conjunctions, adverbs, elaborated noun phrases, and mental (i.e., think, understand, guess, etc.) and linguistic (i.e., talk, say, yell, etc.) verbs are the four main contributors to a literate language style and the frequency of use of these features can provide an estimate of the child's literate language ability (Greenhaulgh & Strong, 2001). In a review of the literature regarding evidence-based interventions for elementary grade students with dyslexia, learning disabilities, and/or interventions for elementary grade students, Al Otaiba, Rouse, and Baker (2018) defined academic language comprehension as including meaning-focused skills that consist of vocabulary development, listening comprehension, reading comprehension, and oral language development.

Research has shown that there is a relationship between oral academic language and reading comprehension. For example, Catts, Adlof, and Weismer (2006) separated students into three groups: students with specific reading comprehension deficits, students with specific decoding deficits, and a control group of typical readers. They evaluated and compared the groups' language comprehension skills in kindergarten, second, fourth, and eighth grades. Results showed that in language comprehension measures, the group with specific reading comprehension deficits scored significantly lower than the group with specific decoding deficits and the control group of typical readers. Yet, the group with specific comprehension deficits performed at typical levels in phonological processing measures. This indicates that children who have difficulty with reading comprehension will have difficulty with overall language comprehension. Snowling and Stothard (1998) investigated the relationship between oral language and reading comprehension in a longitudinal study. Students diagnosed with speech-language impairment at age four were reassessed at age 5;6 to examine the extent of their speech and language impairment. They later received language and cognitive assessments at age 15 and 16 years. It was found that if the child's language difficulties were mainly resolved by 5;6 then there was a better prognosis for spoken language, but not necessarily for literacy and phonological processing. It was also found that students with normal early oral language skills scored significantly better on standardized reading comprehension tests at age 15 than students that were diagnosed with language disorder by age 4, even if the language disorder resolved by age 5;6, indicating a relationship between early oral language skills and reading comprehension.

The Overemphasis of Decoding in Reading Comprehension

The SVR postulates that reading consists of both decoding and language comprehension and research has clearly demonstrated that language comprehension is strongly predictive of

reading comprehension outcomes (Foorman, Herrera, Petscher, Mitchell, & Truckenmiller, 2015). Yet, decoding is often the main focus of assessment and instruction for early school-age children struggling with reading (Petersen & Stoddard, 2018). The current focus on decoding in early grades has not been successful in elevating reading comprehension for over 25 years. The majority of children do not understand grade-level reading material when assessed on state and national exams. Only about 37% of fourth graders are reading at grade level in the United States. Culturally and linguistically diverse (CLD) children are reported to score even lower. For example, only approximately 20% of Black children and 23% of Hispanic children are reading at grade level in fourth grade. These data have remained mostly stagnant, and for students performing in the lower quartile, reading scores have even decreased over time (National Assessment of Education Progress [NAEP], 1992, 2017).

The majority of children with reading difficulties have near-normal abilities in decoding but deficits in reading comprehension (Nation & Snowling, 1997; Yuill & Oakhill, 1991). Research has shown that most CLD students learn to decode like everyone else. For example, Nakamoto, Lindsey, and Manis (2007) conducted a study in which they examined the decoding and reading comprehension skill development of Spanish-speaking English-language learners from first through sixth grade. The sample showed typical decoding ability throughout first through sixth grades but showed difficulty with reading comprehension starting in third grade. This study, along with others (Catts, Hogan, & Fey, 2003; Landi & Ryherd, 2017; Lesaux, Crosson, Kieffer, & Pierce, 2010), indicates that difficulty with academic language comprehension is a main factor in the current epidemic of poor reading performance across the United States.

Improving Oral Language Comprehension Skills

Rather than addressing reading comprehension through decoding, reading comprehension should be targeted through oral language comprehension. In particular, oral narrative language forms the foundation of reading comprehension (Spencer, Kajian, Petersen, & Bilyk, 2014). Narratives are an important aspect of oral language skills because they help to close the gap between the language styles that are used within the home and the academic language that is used in school (Westby, 1985). The ability to form a clear narrative requires complex linguistic skills. Narratives are a naturalistic context to target specific language goals, are essential to school curriculum, and are replete with academic language that students need to understand and produce to be successful in school (Petersen, 2011; Petersen et al., 2014).

There is emerging evidence of the effectiveness of narrative-based language intervention on language growth (Petersen, 2011). Oral narrative language intervention in a multi-tiered system of language support (MTSLS) context has been researched recently. This language focused MTSLS is designed to identify the children who are struggling with language and give children the language support they need to have academic success regardless of special education status within the school system. There are three tiers of support in MTSLS. Tier 1 consists of language instruction that typically takes place in the classroom and is primarily directed by the classroom teacher. However, students may need more intensive language intervention and will therefore move to Tiers 2 or 3 depending on performance related to the grade level requirements using screening and progress monitoring assessments. Tier 2 interventions usually consist of small groups and more targeted intervention. Tier 3 intervention is the most intensive language intervention and is often delivered by highly trained professionals, including speech-language pathologists (Al Otaiba et al., 2018).

Gillam, Olszewski, Fargo, & Gillam (2014) examined narrative language intervention in an MTSSL context with CLD preschool and early school-age students at the Tier1 classroom level. Two classrooms were randomly assigned to a treatment or control condition. One classroom received narrative and vocabulary intervention for three 30-minute periods per week for 6 weeks that consisted of three phases which focused on improving story grammar elements, elaboration, and telling complex stories independently while the other classroom had their regular reading and listening comprehension lessons. It was found that children who were in the classroom that received intervention made clinically significant improvements on narrative measures and actually had 3x larger effect sizes for narrative scores as compared to the control classroom, while the children in the control classroom did not make clinically significant gains. Notably, children who were in the high-risk subgroup made greater gains than the low-risk subgroup and caught up in story complexity after intervention.

Not only is narrative intervention effective, but it is feasible and has successfully been implemented by regular education teachers when implemented at the Tier 1 level (Language and Reading Research Consortium, 2015; Spencer, Weddle, Petersen, & Adams, 2017; Stetter & Hughes, 2010). A review of the literature regarding story grammar and reading comprehension found multiple experimental studies that involved regular education teachers successfully providing story grammar instruction (Stetter & Hughes, 2010). Recently, Spencer, Weddle, Petersen, & Adams (2017) conducted a study in which preschoolers participated in *Story Champs* narrative intervention in an MTSSL context that was implemented by Head Start teachers and teaching assistants (Spencer & Petersen, 2012). The students who received *Story Champs* intervention had statistically significant improvements in language comprehension with

moderate effect sizes and teachers and teaching assistants had acceptable fidelity and reliability scores when implementing the *Story Champs* intervention and narrative retell probes.

Oral narrative language intervention has also been found to be effective at the Tier 2 level in an MTSL context in CLD students (Spencer, Petersen, & Adams, 2015). Spencer, Petersen, Slocum, & Allen (2014) identified preschool students who would benefit from Tier 2 language instruction through a dynamic assessment across three Head Start programs. They found 54% of the preschoolers assessed to be candidates for Tier 2 language intervention. The students who were considered to benefit from Tier 2 language instruction were then randomly assigned to a treatment or control group. The treatment group received 18 sessions of *Story Champs* oral narrative language intervention in small groups of four. The treatment group showed statistically significant gains as compared to the control group in their narrative retells.

Individualized, or Tier 3, narrative language intervention has also been shown to improve narrative outcomes for CLD students (Spencer, Kajian, Petersen, & Bilyk, 2014). For example, Petersen et al. (2014) conducted a study examining the effectiveness of an individualized systematic narrative language intervention in children with autism on their personal narratives. Three 6- to 8- year-old boys with autism were examined in a multiple-baseline design to assess improvement in story grammar elements and linguistic complexity elements within their personal narratives. Intervention included 12 individualized sessions that focused on 2-3 story grammar elements and 3-4 linguistic complexity elements that were selected according to the participant's baseline performance. Immediate improvements were evident in the elements targeted for story grammar elements only when the participant was in intervention for that target. It was also found that there were immediate improvements during the intervention phase for seven out of the nine linguistic complexity variables indicating a possible treatment effect for

these variables. This indicates that children with autism can benefit from individualized narrative language intervention.

Recently, Brough (2019) reported the results of a fully implemented MTSS for a language study. Brough administered a pretest to 686 kindergarten students who then were randomly assigned to a treatment or control group. The students in the treatment group received Tier 1 oral narrative language instruction from their classroom teacher two times per week for 14 weeks. Students who were not making adequate progress after one month of large group instruction also received Tier 2 intervention by the speech-language pathologist (SLP). At the conclusion of the study, the students were assessed again in oral narrative retell, personal story generation, expository retell, and narrative writing. After the conclusion of the study, students in the control group who were at risk at pretest were then provided Tier 2 intervention by the SLPs. Results indicated that after 14 weeks of MTSS oral narrative language instruction at the Tier 1 level, kindergarten students saw significant improvements in narrative retells, personal story generations, narrative writing, and expository language when compared to controls. It was also reported that kindergarten students in the Tier 2 treatment group performed similarly or outperformed matching at-risk, average, and advanced students in the control group in every measure except expository retell. This study demonstrated that narrative intervention across various tiers, including when provided to at-risk students, is effective in improving oral and written narrative and expository language skills of kindergarten students.

Targeting Reading Comprehension Through Narrative Language Intervention

It is hypothesized that because of the strong evidence of successful narrative intervention on oral language skills and the strong association between oral language and reading

comprehension that narrative oral language intervention will have positive impacts on future reading comprehension.

Research has clearly indicated that narrative ability in early grades is greatly predictive of later reading comprehension (Catts, Fey, Tomblin, & Zhang, 2002; Dickinson & McCabe, 2001). Griffin, Hemphill, Camp, and Wolf (2004) found oral discourse skills in five-year-old children, specifically their use in narrative clauses, textual evaluation, and character states (use of internal state words) to be predictive of literacy outcomes at age eight.

Not only is narrative ability predictive of reading comprehension, there is emerging evidence that narrative intervention has a positive effect on reading comprehension. In fact, Catts, Nielsen, Bridges, & Liu (2014) found evidence that response to kindergarten narrative and vocabulary intervention was predictive of reading comprehension outcomes in third grade over and above kindergarten word reading measures. When examining the effects of three different interventions (text-comprehension training, oral-language training, and combined text-comprehension and oral language training) Clarke, Snowling, Truelove, and Hulme (2010), found that each group improved in reading comprehension. However, the group that focused specifically on oral-language instruction, which focused on narrative language, vocabulary, and figurative language, had the greatest long-term progress in reading comprehension when compared to text comprehension intervention and combined oral language and text comprehension intervention.

Another study to investigate the effects of language intervention on reading comprehension was a study conducted by the Language and Reading Research Consortium, Jiang, and Logan (2019). Researchers investigated the effects of a language based large group instruction on elementary grade students' oral language and reading comprehension. A large

sample of first through third grade students were randomly assigned to a treatment condition focused on language or a control condition that participated in business-as-usual. The intervention targeted grammar, vocabulary, text structure knowledge, inferencing, and comprehension monitoring. The intervention was delivered by the students' classroom teachers in whole-group instruction for 20-30 minutes a week for 25 weeks. Researchers administered curriculum aligned measures focused on three areas: oral language comprehension monitoring, vocabulary, and text listening comprehension. Comprehension monitoring required students to listen to a passage and identify what did not make sense within the passage. For the vocabulary assessment, children were tested on 32 Tier 2 vocabulary words that were taught during the intervention. Children also listened to a narrative or expository passage and answered three comprehension questions. Researchers then administered a battery of reading comprehension assessments. Results indicated that all students in the treatment group significantly outperformed all students in the control group in comprehension monitoring and vocabulary. For narrative text comprehension, only third graders outperformed their peers in the control group. For expository text comprehension, there was no significant difference across all grades. Vocabulary was the only language outcome to consistently predict reading comprehension across all grades ($p < .001$). This means that overall, the language-based intervention significantly impacted reading comprehension through the mediation of vocabulary, yet, there was not strong evidence for language comprehension monitoring or text listening comprehension to consistently predict reading comprehension. This might be due to low implementation by teachers within the study. There was an average of 80% of the treatment implemented as prescribed, with a range of 8% to 100%.

Several studies have shown that narrative intervention can have an impact on oral language; however, few studies have demonstrated the effectiveness of early oral narrative intervention in a fully realized MTSL context implemented by end users (e.g., teachers and speech-language pathologist) on later reading comprehension. Furthermore, few studies have investigated the long-term impact of MTSL on students in need of Tier 2 oral language intervention. This study followed the Tier 2 students included in the Brough (2019) study from kindergarten to the end of fourth grade. We were interested in seeing if students who received Tier 2 oral narrative language intervention in kindergarten had equivalent or higher scores when compared to a matched control group on reading comprehension in fourth grade.

Research Questions

Our research questions were as follows:

Question 1: Did children in the at-risk treatment group have scores that were similar to a matched combined average/advanced performing control group in fourth grade on reading comprehension measures?

Question 2: Did children in the at-risk treatment group have scores that were similar to a matched average performing control group in fourth grade on reading comprehension measures?

Question 3: Did children in the at-risk treatment group have scores that were equivalent to a matched advanced performing control group in fourth grade on reading comprehension measures?

Method

A flowchart has been provided to aid in the interpretation of the methods section and can be found in Figure 1.

Participants

Approval was acquired from the Brigham Young University Institutional Review Board before the collection of data because of the involvement of human participants. A total of 686 kindergarten students initially participated in this study and were followed longitudinally for five years with a total of 413 students remaining in the study. Four school districts from the upper Midwest United States participated in the current study, resulting in twenty-eight kindergarten classrooms that were randomly assigned to a treatment or control condition ending with 14 treatment classrooms and 14 control classrooms. The initial treatment group had 337 kindergarteners of which 41 were randomly assigned to an at-risk treatment group, and the initial control group had 349 kindergarteners. These 41 at-risk kindergarten students were matched to 41 average performing (language scores at pretest between the 50th and 75th percentile) and 41 advanced performing (language scores at or above the 75th percentile) students from the control group. Five years later, 182 students with M-STEP data remained from the treatment group and 231 students with M-STEP data remained from the control group with 16 participants in the at-risk treatment group remaining, 20 participants in the average-performing control group, and 20 participants in the advanced control group, making 40 participants in the average and advanced-performing control group in fourth grade. The demographic information on the participants that was available from the school districts was acquired to help describe participants. Participant characteristics including ethnicity, dominant language, socioeconomic status (SES), and presence of a disability are displayed in Table 1 and the remaining fourth grade student demographics are displayed in Table 2.

No significant differences between kindergarten groups across gender, $p = .83$, $t = .21$, $F(684) = .17$, special education $p = .33$, $t = .97$, $F(684) = 3.25$, or free/reduced lunch = $.12$, $t =$

1.55, $F(684) = 9.47$ were found from independent sample t tests. However, a significant difference between groups for ethnicity was found $p = .04$, $t = .2.04$, $F(684) = 12.11$, where there were significantly fewer Caucasian students in the treatment group than in the control group.

Table 1

Descriptive Information for Kindergarten Treatment and Control Group Participants

| | Treatment Group | Control Group |
|--------------------------|-----------------|---------------|
| Gender | | |
| Female | 163 (48%) | 166 (48%) |
| Male | 174 (52%) | 183 (52%) |
| Ethnicity | | |
| White | 275 (82%) | 310 (89%)* |
| Hispanic | 16 (5%) | 12 (3%) |
| African American | 24 (7%) | 13 (4%) |
| Asian | 4 (1%) | 2 (1%) |
| Native American / | | |
| Pacific | 12 (4%) | 8 (2%) |
| Islander/Hawaiian | | |
| Other | 6 (2%) | 4 (1%) |
| SES (Free/Reduced Lunch) | 224 (67%) | 251 (72%) |
| Language Impairment | 37 (11%) | 31 (9%) |

Note. Language impairment was determined based on an active individualized education program for language. *significant difference $p < .05$.

An ANOVA was conducted to determine whether there were significant differences in demographic variables for the fourth-grade cohort. There were no significant differences

between the fourth grade at-risk treatment, average control, and advanced control groups for gender, $F(2, 53) = 2.16$, $p = .13$, free/reduced lunch $F(2, 53) = .06$, $p = .94$, and ethnicity $F(2, 53) = .05$, $p = .95$.

Table 2

Descriptive Information for Fourth Grade Treatment and Control Group Participants

| | Treatment Group | Average Control Group | Advanced Control Group |
|--------------------------|-----------------|-----------------------|------------------------|
| Gender | | | |
| Female | 5 (31%) | 5 (25%) | 11 (55%) |
| Male | 11 (69%) | 15 (75%) | 9 (45%) |
| Ethnicity | | | |
| White | 14 (86%) | 19 (95%) | 17 (85%) |
| Hispanic | 1 (6%) | 0 | 2 (10%) |
| African American | 0 | 0 | 1 (5%) |
| Other | 1 (6%) | 1 (5%) | 0 |
| SES (Free/Reduced Lunch) | 5 (31%) | 7 (35%) | 6 (30%) |

Note. Language impairment was determined based on an active individualized education program for language. *significant difference $p < .05$.

Pretest Measures

In the original study, pretest and posttest measures were given to examine if there was a significant difference in the progress of narrative skills for treatment and control groups. The *Narrative Language Measures* subtest of the CUBED (Petersen & Spencer, 2012) was administered to each participant as an oral language screening that included narrative retelling and personal story generation assessments, an experimental expository language measure, and a

narrative writing sample. These assessments were used as pretests and posttests during kindergarten. Each elementary school had a SLP assigned to it with selected paraprofessionals administering all of the pretest and posttest assessments. A more comprehensive description of pre and posttest measures taken during the participants' kindergarten school year reporting a significant difference in progress between treatment and control groups in narrative skill is given in Mollie Brough's thesis (Brough, 2019).

Pretest narrative retells. The NLM is a standardized, criterion-referenced general outcome measure used to assess children's narrative growth and includes 25 parallel forms for each grade from pre-k to third grade and includes standardized administration and scoring procedures. Psychometric analyses display good to excellent reliability and validity for the NLM (Petersen & Spencer, 2012). Of the four subtests included in the NLM (personal-themed narrative retells, personal story generations, story grammar comprehension, and inferential vocabulary comprehension), only the narrative retell and personal story generation subtests were used for the purposes of this study. The subtest for narrative retell measures a child's ability to both comprehend and produce limited aspects of complex language and story grammar elements through the retelling of a personal-themed narrative procedures for the NLM narrative retell included having students listen to a model story and then retell the story with no pictures and only neutral prompts from examiners. The NLM scoring rubric was used to score the children's retells in real time with individual administration of the three stories taking approximately 3-5 minutes. Stories were scored for the accuracy and completeness of story grammar elements (character, setting, problem, feeling, action, consequence, and ending) on a 0-2 scale with weighted points for episodic elements (e.g., problem, action, consequence). Retells were also scored for the prevalence of aspects of language complexity like the use of causal subordinating

conjunctions (because) and temporal subordinating conjunctions (after, when). Story grammar, episodic points, and language complexity points were all added to give the total NLM retell scores.

Pretest personal story generations. To elicit personal stories, students were asked if they had a similar experience to the story they heard during the story retell assessment. Each story was audio recorded and scored using the Story Grammar and Language Complexity sections of the NLM Flow Chart at a later time. The Story Grammar and Language Complexity section examines story grammar and language complexity with each story grammar element or aspect of language complexity awarded 0-4 points depending upon complexity and clarity, making a total of 55 points possible on the NLM Flow Chart.

Pretest expository language. An experimental expository measure was used to measure expository language. The expository measure is a criterion-referenced assessment of informational text comprehension and production. Similar to the *NLM* retell measures, students were asked to listen to information and then retell that information. Examiners scored the retells in real-time. The expository information was constructed to have unfamiliar information that students likely would not have heard before.

Pretest narrative writing. During pretest, students were asked to write one narrative story about a time that they had gotten hurt. At the top of each student's page was a space where students could illustrate their story. The Story Grammar and Language Complexity sections of the NLM Flow Chart were used to score the narrative writing. To aid in the interpretation of the students' writing, the classroom teachers made notes on the students' writing samples.

At-risk, average performing, and advanced performing control groups. In the control group, pretest *NLM* Listening scores were used to determine which students were considered to

be at-risk, average performing, or advanced performing. Students who received an *NLM* score of < 10 or had a current Individualized Education Program (IEP) for language were considered to be at-risk. Researchers established this criterion because the CUBED Examiner's Manual indicates winter benchmark expectations to be an *NLM* score of 10 (Petersen & Spencer, 2012). Students who received *NLM* pretest scores at or above the 50th percentile (local norms, a *NLM* score of 13) and lower than the 75th percentile (a score of 17) were considered average performing. Students with scores ranging from 10 to 12 were not assigned to any of the at-risk control conditions. The advanced performing control group was considered to be high achieving typically developing peers who scored at or above the 75th percentile on the pretest *NLM* Listening (a score of 17 or higher). At-risk students in the treatment group were identified as part of the treatment intervention process.

Intervention Procedures

Large group (Tier 1) narrative intervention procedures. Language intervention was given to the treatment group within a multi-tiered system of language support (MTSLS) framework. Each school's MTSLS team was led by the assigned SLP. Narrative language instruction was given twice a week for 15-20 minutes at the whole classroom level to each class in the treatment group by the classroom teacher. Classroom teachers used the *Story Champs* large group procedures (Spencer & Petersen, 2012; Spencer, Petersen, Slocum, & Allen, 2014) to present narrative language education. During the first week of treatment, the assigned SLPs modeled the first large group *Story Champs* intervention session, and then, using a fidelity checklist, observed and mentored the classroom teachers while they led the second large group *Story Champs* session. These intervention sessions given to the whole class and led by the classroom teachers were considered Tier1 instruction in the MTSLS system, and continued for

14 weeks, yielding a total of 28 sessions. Additionally, assigned SLPs observed and gave feedback concerning the teachers' fidelity of implementation three more times throughout the study.

Steps for the Tier 1 *Story Champs* instruction were similar as those described in Spencer Petersen, Slocum, & Allen, (2014) and can be viewed at <http://www.youtube.com/watch?v=0M-IKtJVg7s>. Pictures of different scenes in the story were displayed for the whole class to see as the teacher modeled the story. The teacher pointed to the picture displaying the part of the story they were modeling while attaching brightly colored story grammar icons to the corresponding picture to teach story grammar elements. Students were also asked to name the different parts of the story (e.g., character, problem, feeling, action, ending) and then produce gestures symbolizing each part of the story while the teacher retold the story. Afterward, children took individual turns, in which they answered questions about parts of the story (e.g., "Who was this story about?" and "What did they do to fix their problem?"). Following the students' remarks, the entire class restated the answer using group responding. Subsequently, students were paired up to tell the whole story to their partner (i.e., peer tutoring). Peers helped monitor, with roles finishing once one partner finished retelling the story. This whole class intervention included multiple stories that increased in complexity over time.

Small group (Tier 2) procedures. Participants in the treatment group who received a pretest *NLM* retell score of 10 or better and could generate a complete episode (initiating event, attempt, and consequence) on one or more of the pretest narrative retells during pretest were not assessed again until the end of the school year at posttest ($n = 194$; 58%). After one month of Tier 1 whole class intervention, participants in the treatment group who were not able to meet this narrative retell criterion at pretest or who already had an IEP for language ($n = 143$; 42%)

were given two additional narrative retell assessments. The purpose of the additional assessments was for researchers to assess whether the whole-class instruction was sufficiently intense to improve the language of students who were considered at-risk. The participants with the lowest scores and who continued to not meet the criterion (*NLM* retell score of 10 and a complete episode) after one month of whole-class instruction, were assigned to receive additional, Tier 2 small group narrative intervention in addition to the Tier 1 large group intervention throughout the remainder of the school year. Tier 2 intervention consisted of instruction given by an SLP within a small group of three or four students. If there were more than four students who qualified for small group instruction in a classroom, then Tier 2 intervention was randomly assigned out of the qualifying students. Forty-one (12%) of the students in the treatment group received small group, Tier 2 intervention. To ensure an appropriate small group size, two participants in the treatment group from two different schools who did not have IEPs and scored a 10 and an 11 on the *NLM* were included in the Tier 2 group. Tier 2 small group sessions were approximately 20 minutes per session and occurred two times each week outside of the Tier 1 classroom narrative instruction time. This means that students who participated in the small group intervention received two 20-minute small group sessions each week for 10 weeks in addition to the two 15-20-minute large group narrative instruction that occurred over 14 weeks. This came to a total of approximately 410 minutes of explicit narrative instruction for the 41 students in the Tier 2 group. Each student who received small group narrative intervention also participated in a weekly progress monitoring assessment outside of narrative intervention sessions. The assessment used was the *NLM* progress monitoring assessment which helped SLPs

determine what each student's specific strengths and weaknesses were to inform the Tier 2 intervention.

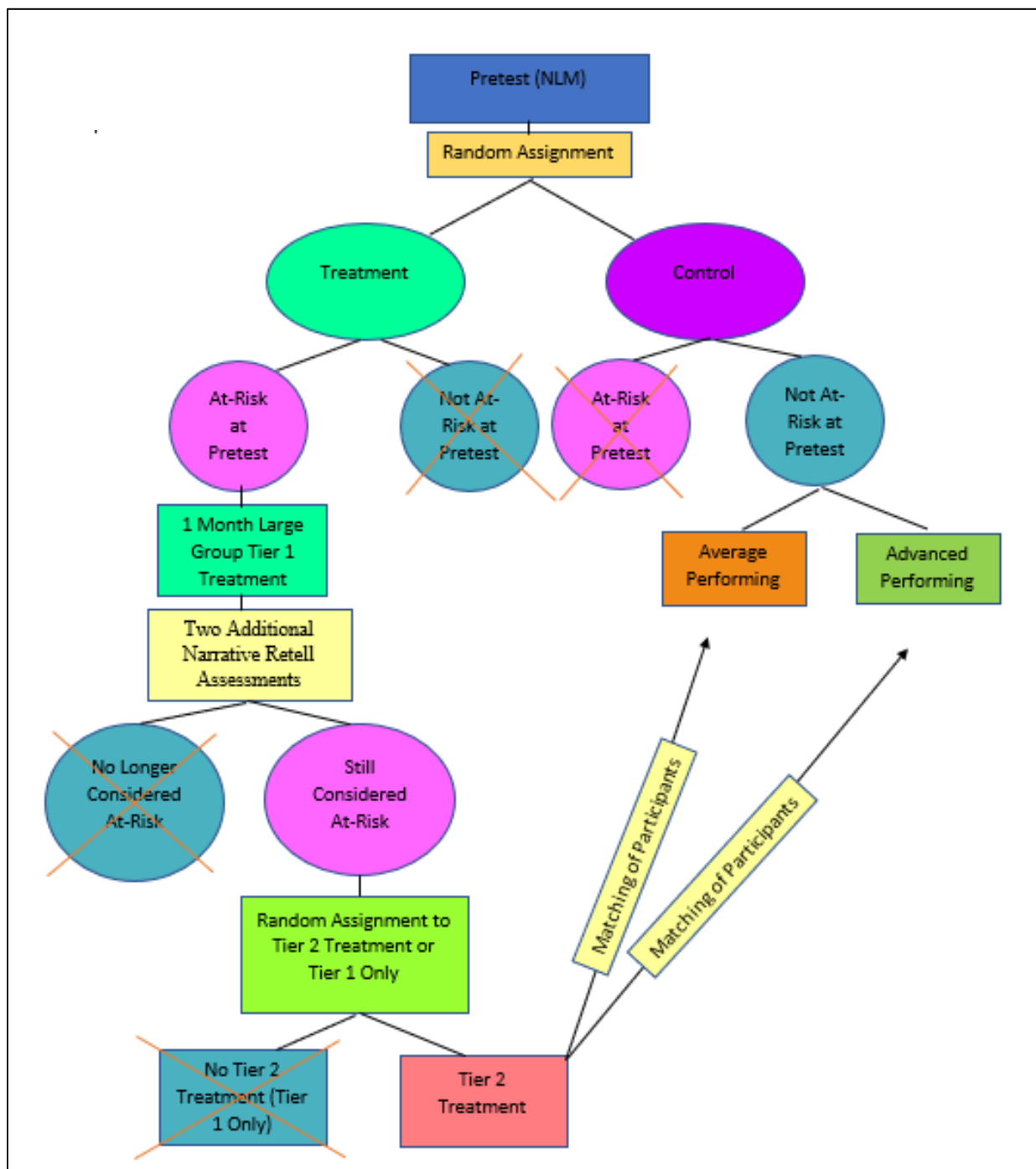


Figure 1. Methods flowchart.

Small group intervention followed the small group procedures of *Story Champs*

(Spencer & Petersen, 2012) with an example of small group procedures shown at <https://www.youtube.com/watch?v=oeQhZbL9vHY&t=302s>. *Story Champs* small group procedures include various personal themed narratives with concomitant pictures that can be dispersed across the table and are large enough for all of the students in the group to see. To help teach parts of a story (setting, problem, consequence, etc.) brightly colored story grammar icons that represented parts of a story were used. To help boost engagement, story games were implemented while students listened to a peer tell a story individually.

Control group. Participants in the control group engaged in their regular classroom instruction that was established at the commencement of the school year. A control group was established to account for various risks to internal validity and to give information on later reading comprehension progress over time under the currently executed curriculum. During the time of intervention, Michigan kindergarten reading standards stated that kindergartners would learn to “ask and answer questions about key details in a text, retell familiar stories with key details, and identify characters, settings, and major events in a story” (Michigan Department of Education, n.d.). Ten key practices were followed by the school districts in early literacy: Intentional and evidence-based efforts that create literacy motivation; read alouds of age-appropriate books; various grouping strategies (e.g. individualized, small group, large group, etc.) depending on the child’s needs; phonological awareness activities; letter-sound relationship instruction; evidence-based and standard-based writing instruction; deliberate aims toward increasing vocabulary; ample reading opportunities and available reading material in the classroom; progress monitoring in students’ language and literacy development; and working with families in advocating literacy (General Education Leadership Network, n.d.). At the conclusion of the kindergarten study, the students who were identified as at risk at pretest in the

control group received Tier 2 intervention by the SLP, thereby contaminating the at-risk control group for longitudinal comparisons.

Matching of participants. Students in the at-risk treatment group were matched to students in the control group in each of the three scoring groups: at-risk, average performing, and advanced performing for comparison of progress in narrative language in the original study. However, because the at-risk control group received Tier 2 intervention in later grades, the current study only examined the at-risk treatment group compared to the average performing and advanced performing control groups. To do this, we matched the at-risk Tier 2 students to 41 students who were considered to be average performing in the control group. These students were matched for school, socio-economic status, gender, and school district to the extent possible. The matching helped us to determine whether at-risk students who received Tier 2 services were able to make sufficient gains so that their performance on the outcome measures were not significantly different from average performing students (not at-risk). The same matching procedures were used to examine whether the students at-risk, who received Tier 2 services, were able to make sufficient gains so that their performance at the end of fourth grade was not significantly different from a matching sample of high achieving typically developing peers (advanced performing control group).

Posttest Measures

At the end of the participants' fourth grade school year, we examined reading outcomes from the state examination, Michigan Student Test of Educational Progress (M-STEP). We obtained these data from the school districts via the lead SLP over those districts.

Michigan Student Test of Educational Progress (M-STEP). The M-STEP is a state-administered online assessment given to all third through eighth grade students across the state of

Michigan. Students are assessed towards the end of the school year usually in the months of April or May. This test measures a student's progress regarding Michigan's academic standards. In fourth grade, the M-STEP assesses students in English language arts (ELA) and mathematics. We analyzed the results from the reading comprehension subtest of the English language arts section.

Scoring. The ELA section includes three types of questions: multiple choice, constructed response, and technology enhanced items with multiple choice options delivered via scantron and technology enhanced items delivered via computer technology. However, each test can be offered in pencil/paper form or through online testing depending on various needs. Multiple choice items were scored by computer as a 1 if correct and as a 0 if incorrect. Technology enhanced items were scored through computer technology. The technology enhanced items could be in various forms: drag and drop, choice interaction, hotspot, matching interaction, keypad input, evidence-based selected response, hot text highlight (line and paragraph), and/or order. Technology enhanced items were scored by computer.

In the reading comprehension subtest, an overall scale score was given. Scale scores are when the raw score is statistically converted to give a standardized score that can be representative of the child's performance level and compared to the normative sample within the reading comprehension section for their particular grade. Scaled scores for the M-STEP can also be described in ranges that are labeled as various performance levels. The performance levels are as follows: not proficient, partially proficient, proficient, and advanced. Each of the performance levels are separated by cut scores determined by a panel of educators and other stakeholders throughout Michigan. The panel determines cut scores based on descriptions explaining what a student should know and be able to do. The descriptions for Performance Levels differ by grade

and subject area; however, comparisons across subjects within a grade can be made to indicate whether students are meeting the Michigan State Standards in their subject and grade which indicates a student's progress towards college and career readiness ("Michigan student test of educational progress: Interpretive guide to M-STEP reports", 2019).

The claims and targets of the M-STEP align with the Michigan English Language Arts (ELA) standards. The claim and target descriptions can be used to understand M-STEP reports and to help teachers design lessons that align with the state curriculum. The claim regarding reading comprehension states that students will be able to closely read increasingly complex literary and informational texts. Some of the targets to reach this claim include using details from a text to support a given conclusion; distinguishing main ideas; inferring meaning of Tier 2 academic words through word structure, word relationships, reference materials, and/or context; independently making inferences from a text; describing relationships between various story grammar elements like character or setting within a text or making comparisons across a text; using knowledge of the text structure to give information about the text; and demonstrating an understanding of word meanings through an understanding of figurative language, word relationships, and nuances of words (Michigan Department of Education, 2015).

Fidelity and Scoring Reliability

Interventionists and fidelity of intervention. Classroom teachers and SLPs participating in the study understood the purpose of the study and behaved as the interventionists. The SLPs engaged in a 4-hour training on the application of MTSLs using the *Story Champs* procedures before serving as interventionists. The SLPs then instructed the classroom teachers on how to apply the Tier 1 instruction. Both the SLPs and classroom teachers studied the *Story Champs* manual, trained with children who were not participants of the study,

and received training and suggestions from the research team. Additionally, classroom teachers were observed by SLPs on at least five different occasions throughout the intervention phase. Each time the SLPs went in to observe teachers during implementation of Tier 1 narrative instruction, a fidelity checklist was filled out. This guided the SLPs when giving feedback to the classroom teachers following the session. During small group sessions, the SLPs used the fidelity checklist to track their own fidelity of carrying out the small group *Story Champs* procedures. The average fidelity of intervention implementation was 97.8% with a range of 91% to 100%.

Pretest administration fidelity and scoring reliability. All of the SLPs who participated in the study attended a 3-hour long training on the *NLM* led by researchers and were taught how to administer and score each pretest prior to the study. Each of the narrative retell and expository retell pretests were administered and scored in real time by the trained SLPs. To ensure reliability, ten percent of the *NLM* retells and expository retells from pretest were randomly selected to be scored by independent scorers. The pretests were then independently listened to and scored in real time by a BYU research team that included undergraduate and graduate students in the communication disorders field. The following formula was used to calculate percent agreement: $\text{Number of agreements} \div \text{agreements plus disagreements}$, multiplied by 100. The mean agreement was 96.4% (range 64% - 100%) for the *NLM* retell and expository retell with inter-raters being blinded to groups. The same BYU research team also scored the personal stories and writing samples. After having read the CUBED manual, the first author of the CUBED assessment taught the research assistants how to use the *NLM* Flow Chart to score personal story generations and writing samples. Research assistants were required to show accurate scoring of the various tests (story retell, personal story generation, writing) and had to have a scoring agreement of 90% or higher with the first author of the CUBED

assessment before they could work on the study. To ensure reliability on the personal narrative and writing sample pretests, fifteen percent of the participants' personal narratives and writing samples were randomly selected to be scored by an independent scorer. Individuals on the research team listened to the participants' audio recordings that had been previously scored by a different research assistant and independently scored the assessments. The following formula was used to calculate percent agreement: $\text{Number of agreements} / (\text{Number of agreements} + \text{Number of disagreements}) \times 100$. The mean agreement was 96.4% (range 64% - 100%) for the personal stories and 84.9% (79% - 100%) for the writing.

Fidelity of test administration was also examined by random selection in 15% of all of the retell narratives, personal narratives, and expository retells from pretest. One research assistant listened to the selected audio recordings and completed a multi-step administration checklist for each test. The percent of accurately completed steps was calculated for each of the selected tests. The overall mean fidelity of test administration was 96.5% (range 88% - 100%) for the *NLM* retells and personal stories and 94.8% (range 76% - 100%) for the expository measure.

Posttest administration fidelity and scoring reliability. All staff that were involved with test administration of the M-STEP were required to have training in the M-STEP procedures. This included each test administrator to have a copy of the Test Administration Manual and to have reviewed the manual prior to conducting the assessment (Gohs, n.d., slide 15).

Results

Research Question 1

Did children in the at-risk treatment group have scores that were similar to a matched combined average/advanced performing control group in fourth grade on reading comprehension measures?

Posttest means and standard deviations in reading comprehension for the at-risk treatment group and the matched average/advanced control groups combined are reported in Table 3. An independent samples t-test was used to compare the at-risk treatment group to the combined average/advanced control group. The at-risk treatment group and the combined average/advanced control group scored similar on reading comprehension with no significant difference ($p = .90$, $t = .13$, $df = 54$).

Table 3

Posttest Means, Standard Deviations, and T-Test Results for the At-Risk Treatment Group and the Average/Advanced Control Groups Combined in Reading Comprehension

| | Groups | | | | | |
|-----------------------|-------------------|-------|--------------------------|-------|----------|----------|
| | At-Risk Treatment | | Average/Advanced Control | | t-test | |
| | Mean | SD | Mean | SD | <i>p</i> | <i>d</i> |
| Reading Comprehension | 1394.88 | 35.90 | 1393.23 | 44.40 | .90 | .04 |

Note. SD = standard deviation.

Research Question 2

Did children in the at-risk treatment group have scores that were similar to a matched average performing control group in fourth grade on reading comprehension measures?

Posttest means and standard deviations in reading comprehension for the at-risk treatment group and the matched average performing control students are reported in Table 4. An

independent samples *t*-test was used to compare the at-risk treatment group to the average performing control group. The at-risk treatment group and the average performing control group scored similar on reading comprehension with no significant difference ($p = .33$, $t = .99$, $df = 34$).

Table 4

Posttest Means, Standard Deviations, and T-Test Results for the At-Risk Treatment Group and the Average Control Group in Reading Comprehension

| | Groups | | | | | |
|-----------------------|-------------------|-----------|-----------------|-----------|----------------|----------|
| | At-Risk Treatment | | Average Control | | <i>t</i> -test | |
| | Mean | <i>SD</i> | Mean | <i>SD</i> | <i>p</i> | <i>d</i> |
| Reading Comprehension | 1394.88 | 35.90 | 1380.95 | 46.04 | .33 | .34 |

Note. *SD* = standard deviation. Equal variances were assumed.

Research Question 3

Did children in the at-risk treatment group have scores that were equivalent to a matched advanced performing control group in fourth grade on reading comprehension measures?

Posttest means and standard deviations in reading comprehension for the at-risk treatment group and the matched advanced control students are reported in Table 5. An independent samples *t*-test was used to compare the at-risk treatment group to the advanced control group. The at-risk treatment group and the advanced control group remained scored on reading comprehension with no significant difference ($p = .241$, $t = -.83$, $df = 34$).

Results Summary

When at-risk kindergarten students were provided more explicit and intense Tier 2 oral narrative language instruction, they were able to not only catch up to their typically developing matching peers in kindergarten on oral language measures (Brough, 2019), but also have

equivalent scores in reading comprehension at the end of fourth grade, even when the at-risk students were compared to students who were identified as having average and advanced oral narrative language in kindergarten. The results of this study indicate that early oral language instruction can have a lasting, meaningful impact on reading comprehension.

Table 5

Posttest Means, Standard Deviations, and T-Test Results for the At-Risk Treatment Group and the Advanced Control Group in Reading Comprehension

| | Groups | | | | | |
|-----------------------|-------------------|-----------|------------------|-----------|----------------|----------|
| | At-Risk Treatment | | Advanced Control | | <i>t</i> -test | |
| | Mean | <i>SD</i> | Mean | <i>SD</i> | <i>p</i> | <i>d</i> |
| Reading Comprehension | 1394.88 | 35.90 | 1405.50 | 40.12 | .41 | .28 |

Note. *SD* = standard deviation.

Discussion

Oral Narrative Language and Reading Comprehension

This is the first randomized controlled longitudinal study that provides evidence for the positive effect of early oral narrative language instruction on future reading comprehension for at-risk students in an MTSL context. The intervention in kindergarten was exclusively focused on oral narrative language but had significant effects on oral and written language outcomes in kindergarten (Brough, 2019), and those effects appeared to extend to fourth grade on reading comprehension. This is demonstrated by the non-statistically significant difference between the at-risk treatment group and the combined control group, average-performing control group, and advanced control group. The average and advanced control group had raw scores that were up to three standard deviations above the mean of the at-risk treatment group at the beginning of kindergarten. After Tier 1 and Tier 2 oral narrative language instruction for 10 weeks in

kindergarten, the at-risk treatment group was no longer significantly different from the advanced control group or the average-performing control group in oral language measures in kindergarten. This study focused on fourth grade outcomes for those students, and results demonstrated that the at-risk students who received oral narrative language treatment in an MTSL context remained similar to the combined, average-performing, and advanced control groups at the end of fourth grade on reading comprehension measures.

These findings provide further evidence that when we are addressing oral language skills with school-aged children, we are building essential skills to aid in reading comprehension (Paul, Norbury, & Gosse, 2018; Westby, 1985). Students are struggling with reading performance on national and state assessments primarily due to deficits in reading comprehension, not decoding (Nakamoto et al., 2007; Nation & Snowling, 1997; Yuill & Oakhill, 1991). This difficulty with reading comprehension is likely due to underdeveloped oral academic language. Oral language has been found to be predictive of reading comprehension (Catts et al., 2002; Dickinson & McCabe, 2001; Griffin et al., 2004). Not only is oral language predictive, but Catts et al. (2014) found that response to kindergarten narrative and vocabulary intervention was causally related to reading comprehension outcomes in third grade over and above kindergarten word reading measures. Clarke et al. (2010) furthered our understanding of the causal relationship between oral language intervention and reading comprehension by finding in a randomized controlled study that oral language instruction alone yielded larger gains in reading comprehension than written language instruction or a combination of oral and written language instruction. Language and Reading Research Consortium et al. (2019) also found that a classroom-based language intervention had a significant effect on students' reading comprehension through the mediation of vocabulary. The current study is the first to provide evidence that kindergarten oral narrative

language intervention may have longitudinal effects on later reading comprehension for at-risk students in an MTSSL context.

The current evidence demonstrating a relationship between oral language and reading comprehension indicates that SLPs are helping with literacy instruction when they are focusing on oral language skills by laying the groundwork to understanding written academic language (Paul et al., 2018). Not only does oral language establish a foundation for written academic language, but Spencer, Kajian, Petersen, & Bilyk, (2014) state that narrative language in particular forms the base for reading comprehension. The current study suggests that significant gains on oral narrative language in kindergarten impacted fourth-grade reading comprehension because the intervention focused on complex academic language, Tier 2 vocabulary, and story grammar elements. The narrative oral language intervention, *Story Champs*, involved students listening to a story replete with complex academic language and then retelling that narrative. Explicit teaching regarding Tier 2 vocabulary words was provided and then those words were included and emphasized within the story. Students were then encouraged to include the Tier 2 vocabulary words taught when retelling their stories. The major focus of the narrative intervention in the current study included explicit teaching of macrostructure story grammar through use of pictures, gestures, visual icons, and peer tutoring.

In alignment with findings from this study, research has shown that vocabulary (Language and Reading Research Consortium, 2015; Language and Reading Research Consortium et al., 2019; Uccelli, Galloway, Barr, Meneses, & Dobbs, 2015), complex academic language (Phillips Galloway, & Uccelli, 2019; Uccelli et al., 2015), and an understanding of story grammar elements (Amer, 1992; Short & Ryan, 1984; Stetter & Hughes, 2010) have a large impact on reading comprehension. Vocabulary appears to have the largest indirect effect on

reading comprehension. For example, when investigating the effects of a language-based classroom intervention on reading comprehension, Language and Reading Research Consortium et al. (2019) found that vocabulary consistently predicted reading comprehension across first through third grade students. Also, Language and Reading Research Consortium (2015) found that although vocabulary did not have a direct, significant effect on reading comprehension, it did have an influence on word recognition and listening comprehension (SVR) which were found to explain 90% of the variance in reading comprehension. Interestingly, vocabulary also had a larger influence on listening comprehension than word recognition which adds further evidence of the relationship between oral language and reading comprehension. Academic language has also been found to be highly predictive of reading comprehension. For example, a study examining a group of fourth through sixth graders by Uccelli et al. (2015) found that core academic language skills made the largest independent contribution to reading comprehension ability over and above word reading fluency, SES, English proficiency designation, vocabulary knowledge, and grade level. Similarly, a recent study examined the relationship between reading comprehension and core academic language skills and found a strong relationship between the growth rate in academic language skills and the rate of growth in reading comprehension skills (Phillips Galloway & Uccelli, 2019). In addition to vocabulary and complex academic language, story grammar is highly predictive of reading ability. A review of the literature regarding story grammar and reading comprehension in students with learning disabilities found various experimental studies demonstrating that story grammar instruction had a positive influence on reading comprehension measures in students both with and without learning disabilities (Stetter & Hugh, 2010).

This study adds to this body of research indicating that a low dose of kindergarten oral narrative language intervention that focuses on vocabulary, complex academic language, and story grammar positively affects later reading comprehension even for students with weak language skills.

Multi-Tiered System of Language Support

It is also likely that the at-risk students were able to make significant gains in reading comprehension because the intervention was provided in an MTSSL context. Many studies have demonstrated the significant effect of narrative language intervention in an MTSSL context on oral narrative language skills (Brough, 2019; Gillam et al., 2014; Petersen et al., 2014; Spencer, Kajian, Petersen, & Bilyk, 2014; Spencer et al., 2015; Spencer, Petersen, Slocum, & Allen, 2014; Spencer et al., 2017). In the current study, at-risk students were given Tier 1 and Tier 2 intervention at kindergarten in which they were able to make significant gains on narrative language skills. In fact, these students were able to make such significant gains that they caught up to their average and advanced peers in the control group after only 14 weeks of intervention. This low dose of early Tier 1 and Tier 2 intervention appeared to have a lasting impact on reading comprehension. This impact on reading comprehension might be because the Tier 2 intervention was individualized to each at-risk student's strengths and weaknesses. During Tier 2 intervention, the SLP administered the *NLM* progress monitoring tool to at-risk students weekly to help the SLP to be aware of what specific areas the child was struggling with and target those areas. It is also possible that there were effects on later reading comprehension because most of the Tier 2 students were typically developing students who may have just needed an extra academic language focus in early grades.

Practicality

Not only was the narrative intervention effective, but a distinctive feature of this study is that all oral narrative language instruction was given reliably by the elementary school teachers and SLPs. All Tier 1 instruction was provided by the classroom teacher and all Tier 2 intervention was provided by the school SLP. Stetter and Hugh (2010) found in a review of experimental studies examining story grammar and its effect on reading comprehension that multiple studies have successfully had regular education teachers provide the story grammar instruction. Language and Reading Research Consortium et al. (2019) also found effects on oral language and reading when general education teachers delivered a language intervention in the classroom. The findings of this study are also potentially applicable to the current education system and replicable in that context.

Early Identification Using Large Group Dynamic Assessment

Under the current “wait to fail” model, students are not identified as at-risk for reading comprehension difficulty until later grades when they are significantly behind their peers. This means that students might go unnoticed for years before it is realized that they were at risk for language and/or reading comprehension deficits. Early identification and intervention can help decrease and sometimes prevent later academic problems. Specifically, it has been found that providing reading intervention in an MTSS context sooner has shown to have significant effects on reading assessments in students who require Tier 2 and Tier 3 intervention (Al Otaiba et al., 2014). Our findings show that children with weaker language skills, the majority of whom do not have a language disorder or are not in special education, can be identified at the beginning of their school career (kindergarten) and in a short amount of time. In this study, these at-risk students were relatively easy to find through use of a large group dynamic assessment. A pretest

was given, and then after one month of large group oral language instruction, the students who were not improving at the Tier 1 level were considered to need Tier 2 instruction. The students who received Tier 2 instruction were able to make gains in oral narrative language along with Tier 1 students.

Limitations

One possible limitation of this study is that oral narrative language skills were not measured in fourth grade. Although evidence has been provided regarding the relationship between oral language skills and reading comprehension skills, without narrative language scores, we can only hypothesize that there were lasting effects on participants' oral narrative language skills from kindergarten intervention, which in turn affected reading comprehension. Future studies may explicitly examine longitudinal oral narrative language outcomes when students are given early oral narrative language intervention in an MTSL context.

Another limitation was a possible confounding variable forcing researchers to take the at-risk control group out of the study. After the study, SLPs and kindergarten teachers were encouraged to implement the oral narrative language instruction, *Story Champs*, to students in the control group who were at risk for future reading difficulty. This means that some speech-language pathologists implemented *Story Champs* with the at-risk matched control students in the years following the study. Therefore, at-risk students in both the treatment and control groups could have received Tier 2 oral narrative language intervention in later grades and that the no-treatment control group could no longer be characterized as having a control condition.

A final limitation is that although this study had a large sample size, only 16 participants in the at-risk treatment group, 20 participants in the average control group, and 20 participants in

the advanced control group remained. This means that the findings of the current study are potentially less generalizable and that future studies should include a larger sample size.

Conclusions

Oral narrative language intervention in an MTSLS context appears to improve future reading comprehension skills in at-risk students. This study showed that language instruction given at Tier 1 and Tier 2 can impact later reading comprehension and can be successfully implemented by regular education teachers. This study also showed that through an MTSLS context, at-risk students can be identified in kindergarten within a relatively short amount of time. These findings are educationally significant and could lead to practices that help ameliorate the current nationwide reading dilemma. Furthermore, we suggest that there should be an increase in research regarding the identification and improvement of students' early oral language weaknesses as they seem to be related to reading comprehension deficits.

References

- Al Otaiba, S., Connor, C., Folsom, J., Wanzek, J., Greulich, L., Schatschneider, C., & Wagner, R. (2014). To wait in Tier 1 or intervene immediately: A randomized experiment examining first grade response to intervention in reading. *Exceptional Children, 81*(1), 11-27.
- Al Otaiba, S., Rouse, A. G., & Baker, K. (2018). Elementary grade intervention approaches to treat specific learning disabilities, including dyslexia. *Language, Speech, and Hearing Services in Schools, 49*, 829-842.
- Amer, A. (1992). The effect of story grammar instruction on EFL students' comprehension of narrative text. *Reading in a Foreign Language, 8*, 711-720.
- Brough, M. (2019). *A large-scale randomized control trial examining the effects of a multi-tiered oral narrative language intervention on kindergarten oral and written narratives and oral expository language* (Unpublished master's thesis). Brigham Young University, Provo, Utah.
- Catts, H., Adlof, S., & Weismer, S. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49*, 278-293.
- Catts, H., Fey, M., Tomblin, J., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research, 45*, 1142-1157.
- Catts, H., Hogan, T., & Fey, M. (2003). Subgrouping poor readers on the basis of individual differences in reading related abilities. *Journal of Learning Disabilities, 36*, 151-164. doi: 10.1177/002221940303600208

- Catts, H., Nielsen, D., Bridges, M., & Liu, Y. (2014). Early identification of reading comprehension difficulties. *Journal of Learning Disabilities, 49*, 451-465. doi: 10.1177/0022219414556121.
- Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science, 21*, 1101-1611.
- Dickinson, D. K., & McCabe, A. (2001). Bringing it all together: The multiple origins, skills, and environmental supports of early literacy. *Learning Disabilities Research and Practice, 16*, 186-202. doi: 10.1111/0938-8982.00019
- Foorman, B. R., Herrera, S., Petscher, Y., Mitchell, A., & Truckenmiller, A. (2015). The structure of oral language and reading and their relation to comprehension in kindergarten through grade 2. *Reading and Writing: An Interdisciplinary Journal, 28*, 655-681. doi: 10.1007/s11145-015-9544-5
- General Education Leadership Network. (n.d.) *Essential instructional practices in early literacy* [PDF file]. Retrieved from https://www.gomaisa.org/downloads/gelndocs/k3_literacy_essentials.pdf
- Gillam, S., Olszewski, A., Fargo, J., & Gillam, R. (2014). Classroom-based narrative and vocabulary instruction: Results of an early-stage, nonrandomized comparison study. *Language, Speech, and Hearing Services in Schools, 45*, 204-219.
- Gohs, P. (n.d.) *M-STEP assessment administration training* [PowerPoint slides]. Retrieved from <http://gomasa.org/wp-content/uploads/M-STEP-Assessment-Administration.pptx>
- Gough, P., & Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*, 6-10. doi: 10.1177/074193258600700104

- Greenhaulgh, K., & Strong, C. (2001). Literate language features in spoken narratives of children with typical language and children with language impairments. *Language, Speech, and Hearing Services in Schools, 32*, 114-125.
- Griffin, T., Hemphill, L., Camp, L., & Wolf, D. (2004). Oral discourse in the preschool years and later literacy skills. *First Language, 24*, 123-147. doi:10.1177/0142723704042369
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2*, 127-160. doi: 10.1007/BF00401799
- Hoover, W., & Tunmer, W. (2018). The simple view of reading: Three assessments of its adequacy. *Remedial and Special Education, 39*, 304-312.
- Landi, N., & Ryherd, K. (2017) Understanding specific reading comprehension deficit: A review. *Language and Linguistics Compass, 11*, 1-24. doi: 10.1111/lnc3.12234
- Language and Reading Research Consortium. (2015). Learning to read: Should we keep things simple? *Reading Research Quarterly, 50*, 151-169.
- Language and Reading Research Consortium, Jiang, H., & Logan, J. (2019). Improving reading comprehension primary grades: Mediated effects of a language-focused classroom intervention. *Journal of Speech, Language, and Hearing Research, 62*, 2812-2828.
- Lesaux, N., Crosson, A., Kieffer, M., & Pierce, M. (2010). Uneven profiles: Language minority learners' word reading, vocabulary, and reading comprehension skills. *Journal of Applied Developmental Psychology, 1*, 475-483. DOI: 10.1016/j.appdev.2010.09.004
- Michigan Department of Education. (2015). *English language arts grade 4 cross walk: Claims-targets-standards* [PDF file]. Retrieved from https://www.michigan.gov/documents/mde/G4_crosswalk_claim_target_standard_500253_7.pdf

Michigan Department of Education. (n.d.). *Michigan K-12 standards English language arts*

[PDF file]. Retrieved from

https://www.michigan.gov/documents/mde/MDE_ELA_Standards_599599_7.pdf

Michigan student test of educational progress: Interpretive guide to M-STEP reports. (2019).

Retrieved from

https://www.michigan.gov/documents/mde/2019_Interpretive_Guide_to_M-STEP_Reports_661956_7.pdf

Nakamoto, J., Lindsey, K., & Manis, F. (2007). A longitudinal analysis of English language learners' word decoding and reading comprehension. *Reading and Writing, 20*, 691-719.

Nation, K., & Snowling, M. (1997). Assessing reading difficulties: The validity and utility of current measures of reading skill. *British Journal of Education and Psychology, 67*, 359-370.

National Assessment of Education Progress. (1992). *1992 Reading Assessment*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

National Assessment of Education Progress. (2017). *2017 Reading Assessment*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

Paul, R., Norbury, C., & Gosse, C. (2018). *Language disorders from infancy through adolescence: Listening, speaking, reading, writing, and communicating*. St. Louis, MO: Elsevier.

- Petersen, D. (2011). A systematic review of narrative-based language intervention with children who have language impairment. *Communication Disorders Quarterly*, *32*, 207-220. doi: 10.1177/1525740109353937
- Petersen, D., Brown, C., Ukrainetz, T., Wise, C., Spencer, T., & Zebre, J. (2014). Systematic individualized narrative language intervention on the personal narratives of children with autism. *Language, Speech, and Hearing Services in Schools*, *45*, 67-86. doi:10.1044/2013_lshss-12-0099
- Petersen, D., & Spencer, T. (2012). The Narrative Language Measures: Tools for language screening, progress monitoring, and intervention planning. *Perspectives on Language and Learning and Education*, *19*, 119-129. doi: 10.1044/lle19.4.119
- Petersen, D., & Stoddard, A. (2018). Psychometric requirements of oral and written language progress monitoring assessments. *Perspectives of the ASHA Special Interests Groups*, *3*, 180-197.
- Phillips Galloway, E., & Uccelli, P. (2019). Examining developmental relations between core academic language skills and reading comprehension for English learners and their peers. *Journal of Educational Psychology*, *111*, 15-31.
- Short, E., & Ryan, E. (1984) Metacognitive differences between skilled and less skilled readers: Remediating deficits through story grammar and attribution training. *Journal of Educational Psychology*, *76*, 226-235.
- Snowling, M., & Stothard, S. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Research*, *41*, 407-418.

- Spencer, T., Kajian, M., Petersen, D., & Bilyk, N. (2014). Effects of an individualized narrative intervention on children's storytelling and comprehension skills. *Journal of Early Intervention, 35*, 243-269.
- Spencer, T., & Petersen, D. (2012). *Story Champs: A multi-tiered language intervention program*. Retrieved from <http://www.languagedynamicsgroup.com>
- Spencer, T., Petersen, D., & Adams, J. (2015). Tier 2 language intervention for diverse preschoolers: An early-stage randomized control group study following an analysis of response to intervention. *American Journal of Speech-Language Pathology, 24*, 619-636.
- Spencer, T., Petersen, D., Slocum, T., & Allen, M. (2014). Large group narrative intervention in Head Start preschools: Implications for response to intervention. *Journal of Early Childhood Research, 13*, 196-217. doi: 10.1177/1476718X13515419
- Spencer, T., Weddle, S., Petersen, D., & Adams, J. (2017). Multi-tiered narrative intervention for preschoolers: A head start implementation study. *NHSA Dialog, 20*, 1-28.
- Stetter, M., & Hughes, M. (2010). Using story grammar to assist students with learning disabilities and reading difficulties improve their comprehension. *Education and Treatment of Children, 33*, 115-151.
- Uccelli, P., Galloway, E., Barr, C., Meneses, A., & Dobbs, C. (2015). Beyond vocabulary: Exploring cross-disciplinary academic-language proficiency and its association with reading comprehension. *Reading Research Quarterly, 50*, 337-356.
- Westby, C. E. (1985). Learning to talk—talking to learn: Oral literate language differences. In C. S. Simon (Ed.), *Communication skills and classroom success: Therapy methodologies for language-learning disabled students* (pp. 181-218). San Diego, CA: College Hill Press.

Yuill, N., & Oakhill, J. (1991). Children's problems in text comprehension: Assessing casual issues. *Journal of Deaf Studies and Deaf Education*, 5, 51-59.

APPENDIX A

Annotated Bibliography**National Reading Problem**

National Assessment of Education Progress. (2017). *NAEP nations report card*. Retrieved from <https://nces.ed.gov/nationsreportcard/>

The Relationship Between Reading Comprehension and Oral Language Skills

Nakamoto, J., Lindsey, K., & Manis, F. (2007). A longitudinal analysis of English language learners' word decoding and reading comprehension. *Reading and Writing, 20*, 691-719.

Objective: The purpose of this study was to describe word decoding and reading comprehension skill development of Spanish-speaking English-language learners (ELLs) from first through sixth grade. The study also examined the degree to which phonological awareness, rapid automatic naming, and oral language measures taken in English in first grade would predict reading skills and reading growth.

Method: The sample included 261 Latino children (120 boys and 141 girls) with a limited knowledge of English as determined by language assessment tests in kindergarten. Analyses began when the children were in first grade. 218 children remained in the study by sixth grade.

Tests administered in kindergarten and first grade were the *Letter-Word Identification* (Woodcock & Munoz-Sandoval, 1995), *Picture Vocabulary* (Woodcock & Munoz-Sandoval, 1995), *Memory for Sentences* (Woodcock & Munoz-Sandoval, 1995), and the *Passage Comprehension* (Woodcock & Munoz-Sandoval, 1995). Tests administered in first grade only were *Sound Matching* (Wagner, Torgesen, & Rashotte, 1998), *Elision* (Wagner et al., 1998), *RAN-Objects and Digits* (Wagner et al., 1998), and the *Memory for Sentences* (Woodcock & Johnson, 1989). Tests administered in all grades were *Letter-Word Identification* (Woodcock & Johnson, 1989; Woodcock et al., 2001), *Word Attack* (Woodcock & Johnson, 1989; Woodcock et

al., 2001), *Passage Comprehension* (Woodcock & Johnson, 1989; Woodcock et al., 2001), and the *Picture Vocabulary* (Woodcock & Johnson, 1989; Woodcock et al., 2001). A growth curve modeling was used to analyze the change in reading scores over time.

Results: When compared to the normative English-speaking sample, the mean National Percentile for Passage Comprehension measures was in the average range for first grade, decreased in third grade, and was below average by sixth grade. The mean National Percentile for Basic Reading Skills (decoding measures) remained average throughout first through sixth grades. Time invariant predictors gave further evidence to past findings by displaying correlations between phonological awareness, RAN, and oral language and reading measures. Students with lower scores on phonological awareness and RAN tasks showed greater initial growth, but a rapid deceleration in word decoding prediction models. Students with lower scores on oral language measures showed greater initial growth, but a rapid deceleration in passage comprehension prediction models.

Relevance to Current Study: The sample showed adequate levels of decoding ability throughout first through sixth grades. However, the sample showed difficulty with reading comprehension in later grades, starting in third grade. Therefore, comprehension strategies and oral language skills should be targeted to see gains in reading comprehension given that evidence has shown a relationship between oral language abilities and reading comprehension.

Catts, H., Adlof, S., & Weismer, S. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49*, 278-293.

Study 1:

Objective: The purpose of this study was to evaluate and compare the language skills of children with specific reading comprehensions deficits to children with specific decoding deficits and to typical readers in eighth grade.

Method: The sample included eighth-grade students divided into three groups: poor reading comprehension but normal word recognition (poor comprehenders), poor word recognition but normal reading comprehensions (poor decoders), and a control group with typical word recognition and reading comprehension (typical readers). There were 57 poor comprehenders, 27 poor decoders, and 98 typical readers. To determine reading comprehension levels, the Passage Comprehension subtest of the Woodcock Reading Mastery Tests-Revised (WRMT-R), the comprehension component of the Gray Oral Reading Test- 3 (GORT-3), and the Qualitative Reading Inventory (QRI-2) was administered. To assess cognitive abilities, the Block Design and Picture Completion subtests of the Wechsler Intelligence Scale for Children- Third Edition (WISC-III). To determine language comprehension, a battery of standardized tests was given to assess receptive vocabulary, grammatical understanding, and discourse comprehension. To assess phonological processing, a phoneme deletion, Pig Latin, and phonological memory measures were given. Subgroups were determined by performance on comprehension and word recognition composite scores. A participant was considered to be a “poor comprehender” if they scored below the 25th percentile in reading comprehension and above the 40th percentile in word recognition. A participant was considered to be a “poor decoder” if they scored below the 25th percentile in word recognition and above the 40th percentile in reading comprehension. A participant was considered a “typical reader” if they scored between the 40th and 84th percentile in both reading comprehension and word recognition.

Results: The subgroups scored significantly different in measures of receptive vocabulary ($F(2, 179) = 31.35, p < .001$) and grammatical understanding ($F(2, 179) = 20.43, p < .001$). Poor comprehenders scored significantly lower than typical readers ($p < .001, d = 1.47$) and poor decoder ($p < .01, d = 0.96$) subgroups in receptive vocabulary measures.

Poor comprehenders also scored lower than typical readers ($p < .001, d = 1.26$) and poor decoders ($p < .001, d = 1.39$) while poor decoders scored comparably to typical readers ($p > .05, d = 1.39$) in discourse comprehension. In the inference tasks, scores were significantly different between subgroups in the distant inference conditions ($F(2, 178) = 5.07, p < .01$), but not the adjacent inference conditions ($F(2, 179) = 2.12, p > .05$). Poor comprehenders scored significantly lower than typical readers ($p < .01, d = 0.61$), but did not differ significantly from poor decoders ($p > .05, d = 0.49$).

In phonological processing measures, subgroups differed significantly in each task. The poor decoders scored significantly lower than the typical and poor comprehender subgroups on the phoneme deletion task, pig Latin task and nonword repetition tasks.

Relevance to Current Study: The results of this study further builds on the simple view of reading and the phonological deficit hypothesis because the poor comprehenders had deficits in language comprehension but showed typical levels of phonological processing while the poor decoders had deficits in phonological processing but typical levels in language comprehension. This indicates that children who specifically struggle with reading comprehension (not decoding or word reading skills) are expected to struggle with language comprehension.

Study 2 (Same reference, but longitudinal data):

Objective: The purpose of this study was to examine word recognition and language abilities of the participants in Study 1 in earlier grades within the same subgroups of poor comprehenders, poor decoders, and typical readers.

Method: The same participants were included in Study 2 as in Study 1. Scores were taken from 2nd and 4th grades. Reading comprehension was assessed using the WRMT-R Passage Comprehension, the GORT-3 comprehension score, and the Diagnostic Achievement Battery, Second Edition (DAB-2). Word recognition was assessed using the Word Identification and Word Attack subtests of the WRMT-R. Language comprehension was assessed using the Concepts and Directions subtest from CELF-3 and the PPVT-R. Discourse comprehension was measured using the Listening to Paragraphs subtest from the CELF-3 and an experimental measure of discourse developed by Cullata, Page, and Ellis. Participants were also assessed in kindergarten with the Test of Language Development- Primary Second Edition (TOLD-P;2). Lastly, phonological processing was measured using a sound deletion task in kindergarten, 2nd, and 4th grades and the nonword repetition task in 2nd grade.

Results: ANOVA analysis showed that the subgroups were significantly different in kindergarten, second, and fourth grades in language composite scores. Poor comprehenders scored significantly less than typical readers and poor decoders in kindergarten, second, and fourth grades in language scores. Poor decoders scored significantly lower than typical readers in kindergarten, but not second and fourth grades. The subgroups also scored significantly different in phonological awareness in kindergarten, second, and fourth grades. Poor decoders scored significantly lower in than typical readers in phonological awareness in kindergarten, second and fourth grades, and significantly lower than poor comprehenders in second and fourth grades.

Typical readers and poor comprehenders did not score significantly different from each other in second and fourth grades. In reading comprehension, poor comprehenders and poor decoders scored significantly lower than typical readers in second and fourth grades.

Relevance to Current Study: Study 2 showed that the double dissociation (poor comprehenders had poor language skills, but typical phonological skills and vice versa for poor decoders) was evident in earlier grades (kindergarten, 2nd, and 4th grades). The studies indicate that children with specific problems in reading comprehension have deficits in overall language comprehension in both early and later grades. However, study 2 also showed that poor decoders and poor comprehenders might be less differentiated in younger grades. (This creates the problem of poor comprehenders being classified as poor decoders so the article goes on to state that the classification system should be based on the simple view by classifying poor readers on basis of strengths and weaknesses in language comprehension and word reading).

Language and Reading Research Consortium. (2015). Learning to read: Should we keep things simple? *Reading Research Quarterly*, 50, 151-169.

Objective: A cross sectional analysis on 1st, 2nd, and 3rd graders to determine which areas have the greatest impact on reading comprehension. Word reading accuracy, word reading fluency, listening comprehension, reading comprehension, and vocabulary were all measured.

Questioning the simplicity of the simple view of reading.

Method: Multiple assessments were used to measure reading comprehension, word recognition accuracy, word recognition fluency, listening comprehension, and vocabulary to students in grades 1-3. Structural equation modeling was then used to determine the relationship between the variables and the theoretical models.

Results:

1. Word recognition and listening comprehension do influence reading comprehension. Both variables were significantly different from zero in the positive direction. Approximately 90% of the variance in reading comprehension in this study was explained by the simple view of reading model.
2. Word recognition had a stronger influence on reading comprehension in 1st grade than listening comprehension. In 2nd and 3rd grade listening comprehension had a stronger influence on reading comprehension than word recognition.
3. Word reading accuracy impact decreases after 1st grade and is not significant by grade 3. Word reading fluency only significant in grade 3. Word reading accuracy, word reading fluency, and listening comprehension all independently influence reading comprehension.
4. Vocabulary influences word recognition and listening comprehension and therefore influences reading comprehension. However, vocabulary did not have a significant effect on reading comprehension directly and thus only affects it indirectly. Vocabulary had a slightly stronger effect on listening comprehension than on word recognition for every grade.

Relevance to current study: Although word recognition had a stronger influence on reading comprehension in 1st grade than listening comprehension, it has less of an impact in later grades as soon as 2nd grade. This implies that we should still target decoding, but we should also target listening comprehension skills to improve long term reading comprehension because that becomes the more dominant force in reading comprehension. The simple view of reading is lacking developmental changes in its subcomponents. Vocabulary influences reading

comprehension indirectly by having an influence on word recognition and listening comprehension.

Al Otaiba, S., Rouse, A. G., & Baker, K. (2018). Elementary grade intervention approaches to treat specific learning disabilities, including dyslexia. *Language, Speech, and Hearing Services in Schools, 49*, 829-842.

Objective: This is a narrative review of the literature to describe current evidence-based practice intensive interventions for elementary grade students with dyslexia, learning disabilities, and or intensive reading and writing needs to inform professional development efforts.

Results: Response to intervention (RTI) or multitiered systems of support (MTSS) have been implemented in the education system since 2004 to provide early literacy intervention. MTSS includes behavior, social/emotional learning supports, and a greater variety of academic skills than RTI models which mainly assess reading and math. There are three tiers of support. Tier 1 consists of general education that follows the core curriculum. However, students may need more intensive intervention and will therefore move to tiers 2 or 3 depending on performance related to the grade level requirements using screening and progress monitoring assessments. Tier 2 interventions consists of small groups and more targeted intervention. Tier 3 intervention is the most intensive intervention. Tier 3 could include speech and language services, special education services, and/or a 504 plan.

This review also takes on the Simple View of Reading and Writing and the Taxonomy Intervention Intensity. The Simple View of Reading and Writing proposes that successful reading relies on both code-focused skills and meaning-focused skills and successful writing is relies on spelling and ideation. The Taxonomy of Intervention Intensity describes how to meet the needs of the individual through intensity of intervention by manipulating dosage, targeting

the students' skills deficits and strengths, generalization, giving behavioral and motivational support, and using a data-based process to ensure individualization.

The narrative review splits intervention for elementary grade students with dyslexia, learning disabilities, and/or reading and writing needs into three categories: code-focused reading skills, meaning-focused reading skills, and specific intervention programs. Code-focused reading skills consists of phonemic awareness, phonics, and fluency. Meaning-focused reading skills consisted of vocabulary development, listening comprehension, and reading comprehension. It was also emphasized that meaning-focused reading skills consists of oral language development, including semantics, morphology, syntax, and pragmatics. Specific intervention programs included various specific programs targeted to help students with dyslexia.

Relevance to Current Study: Review of the literature demonstrates that MTSS and oral language intervention have shown to be beneficial and related to reading comprehension.

Snowling, M., & Stothard, S. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Research, 41*, 407-418.

Objective: The purpose of this study was to conduct a longitudinal study on a group of 4-year-old children diagnosed with speech-language impairment to assess their language and cognitive outcomes at age 15 and 16 years.

Method: The participants of the study included 71 children who were diagnosed as having a language impairment between the ages of 3;9 and 4;2 that was not due to low intelligence, hearing loss, physical defect, bilingual background, or associated with a syndrome. At age 4;0 participants were evaluated for non-verbal abilities. Nineteen of the children had nonverbal IQs below 70 and were placed in the group "general delay". At age 5;6 the children were reassessed to find the range and extent of their speech-language impairments. 32 children met the criteria

for satisfactory speech-language status and were put into the group “resolved SLI”. The remainder of the children were placed in the “persistent SLI group”. The control group for this study was taken from five schools in Tyneside that were aged 15-16 years.

Receptive vocabulary was tested using the Long Form of the British Picture Vocabulary Scale (BPVS), expressive vocabulary was tested using the WISC-III subtest, general comprehension was tested using the WISC-II Verbal Comprehension subtest, grammatical understanding was tested using the Test for Reception of Grammar (TROG), naming was tested using the Graded Naming Test, sentence repetition was tested using The Clinical Evaluation of Language Fundamentals-Revised (CELF-R), nonword repetition was tested using The Children’s Nonword Repetition Test (CNRep), phonological awareness was tested by administering a Spoonerism task, literacy was tested using the Wechsler Objective Reading Dimensions (WORD), nonverbal ability was assessed using the WISC-III Picture Completion Block Design subtests. The children also participated in a psychiatric interview to evaluate psycho-social outcome. Children were tested in one test session that lasted approximately 2 ½ hours. The tests were standardized relative to the control group.

Results: The control group scored consistently with the standardization sample of the tests. There were significant differences in each of the groups on every spoken language test. The resolved SLI group and control groups earned similar scores on each test except for the sentence repetition, nonword repetition, and spoonerisms task, where they received significantly lower scores. The persistent SLI group scored significantly higher than the general delay group in every test except for the TROG. Both the persistent SLI group and the general delay groups scored significantly lower than the control group in all spoken language skills. On the WORD literacy evaluation, the resolved SLI group, persistent SLI group, and general delay group all

scored significantly lower than the control group. This indicates that children who have language difficulties present at age 5;6 are at a higher risk of language, literacy, and academic difficulties through adolescence. If the child's language difficulties are mainly resolved by 5;6 then there is a better prognosis for spoken language, but not necessarily for literacy and phonological processing.

Relevance to current work: This study demonstrates that students with normal early oral language skills score significantly better on standardized reading comprehension tests at age 15 than students that were diagnosed with SLI by age 4, even if the SLI resolves by age 5;6 indicating that there may be a relationship between early oral language skills and reading comprehension.

Stetter, M., & Hughes, M. (2010). Using story grammar to assist students with learning disabilities and reading difficulties improve their comprehension. *Education and Treatment of Children, 33*, 115-151.

Objective: Review of the literature regarding story grammar and its effect on helping students with learning disabilities improve reading comprehension skills.

Method: The Education Research Information Center database was searched using keywords related to story grammar and reading comprehension in students with learning disabilities. Only peer reviewed studies that were considered to be experimental or quasi-experimental were included in the study.

Relevance to current study: Multiple studies successfully had the regular teachers provide the story grammar instruction. Story grammar intervention positively impacted reading comprehension in both students with and without learning disabilities.

Narrative Language Intervention

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science, 21*, 1101-1611.

Objective: The purpose of this study was to examine the efficacy of text comprehension intervention, oral language intervention, and combined text comprehension and oral language intervention relative to a control group in children with specific reading-comprehension difficulties.

Method: A randomized controlled trial was conducted in various schools with children clustered. 3 different types of intervention (oral language-OL, text comprehension-TC, & combined text comprehension and oral language-COM) were given by the same teaching assistant in each school. A control group was included to show the difference in gains between the students with intervention and the students with just standard classroom instruction. Eight participants in each of the 20 schools were randomly assigned to either OL, TC, COM, or control. The students were assessed at pretest (Time 1), 10 weeks after intervention (Time 2), 20 weeks after intervention (Time 3), and a delayed follow-up of approximately 11 months after intervention finished (Time 4).

Participants for the study were required to have a specific reading-comprehension impairment. To identify reading-comprehension deficits, group-administered measures of spelling (Wechsler Objective Reading Dimensions- WORD), nonverbal IQ (Raven's Standard Progressive Matrices-Raven) and listening comprehension (Neale Analysis of Reading Ability: Second Revised British Edition, NARA II) were given to children in 23 different schools with varying socioeconomic backgrounds in Yorkshire, England. Of the identified children with the lowest listening-

comprehension, only the students with age-appropriate nonverbal and spelling ability received further testing in reading comprehension (NARA II) and reading accuracy (Test of Word Reading Efficiency-TOWRE). The children that had at least a one standard deviation discrepancy between reading comprehension (TOWRE) and reading fluency (NARA II) scores were selected for the study, resulting in eighty-four children.

The NARA II and WIAT II were used to measure reading comprehension questions. Vocabulary was also assessed using a vocabulary subtest from the Wechsler Abbreviated Scale of Intelligence (WASI) and a bespoke vocabulary test designed by the researchers of the study that contained Tier 2 level words.

The TC intervention consisted of metacognitive strategies, reciprocal teaching with text, inferencing from text, and written narrative. This program included solely written text. The OL program consisted of vocabulary, reciprocal teaching with spoken language, figurative language, and spoken narrative. This program was focused on spoken language. The COM program combined all eight components of the OL program and TC program with half of the intervention time spent on OL and half of the time spent on TC intervention.

Results: On the WIAT II, each intervention group showed gains after intervention (Times 1-3), while the control group showed a decrease in WIAT II scores. Follow-up (11 months after intervention) data showed that the OL group showed further gains than the other groups. On the NARA II, every group including the control group made improvements. Considering the control group had a decrease in scores on the WIAT II, this indicates that the NARA II may be susceptible to practice effects. In vocabulary, there were significant gains for the OL group on the taught and non-taught words and significant gains for only taught words in the COM group relative to the control group at Time 3. On the WASI, there were significant gains for the OL

group relative to the control group at Time 3, but the gain fell to nonsignificant levels at Time 4. When looking at vocabulary, the OL group and the COM group showed significant gains compared to the control group. Overall, each group had significantly greater gains than the control group in reading comprehension. It should be noted that OL, TC, and COM interventions all produced statically significant gains in reading comprehension and that the long-term progress in reading comprehension was greatest for the children who received OL intervention, not COM. This indicates that the amount of time spent on oral language was vital to success in long-term reading comprehension gains. Gains in reading comprehension in the OL and COM groups were also partly or wholly mediated by gains in vocabulary.

Relevance to current work: This study demonstrates that oral language intervention produced significant gains in reading comprehension and had the greatest long-term progress in reading comprehension when compared to text comprehension intervention, combined oral language and text comprehension intervention, and a control group. The study also demonstrates that gains in reading comprehension are partly or wholly mediated by gains in vocabulary.

Language and Reading Research Consortium, Jiang, H., & Logan, J. (2019). Improving reading comprehension primary grades: Mediated effects of a language-focused classroom intervention. *Journal of Speech, Language, and Hearing Research, 62*, 2812-2828.

Objectives: To investigate the effects of a language intervention (Let's Know!) on elementary grade students' language skills, specifically comprehension monitoring, vocabulary, and language comprehension. This study also examined the effect of the language intervention on reading comprehension skills and how it related to their language skills.

Methods: 997 first through third graders. Students were randomly assigned to a treatment condition focused on language or a business-as-usual control condition. The intervention targeted

grammar, vocabulary, text structure knowledge, inferencing, and comprehension monitoring.

The intervention was given in a large classroom instruction, 20-30 minutes a week for 25 weeks. An average of 80% of the treatment was implemented as prescribed, with a range of 8% to 100%. Researchers utilized curriculum aligned measures focused on three areas: comprehension monitoring, vocabulary, and text comprehension. Comprehension monitoring included students listening to a passage and identifying what did not make sense within the passage. Children were taught 32 Tier 2 vocabulary words. Children listened to a narrative or expository passage and answered three comprehension questions. Researchers then administered a battery of reading comprehension assessments.

Results: Results indicated that the all students in the treatment group significantly outperformed all students in the control group in comprehension monitoring and vocabulary. For narrative text comprehension, only third graders outperformed their peers in the control group. For expository text comprehension, there was no significant difference across all grades. Vocabulary was the only language outcome to consistently predict reading comprehension across all grades ($p < .001$).

Relevance to current study: Demonstrated that classroom-based language intervention significantly impacted reading comprehension through mediation of vocabulary. There was not strong evidence for language comprehension monitoring or text listening comprehension to consistently predict reading comprehension. This might be due to low implementation by teachers within the study. There was an average of 80% of the treatment implemented as prescribed, with a range of 8% to 100%.

Barton-Hulsey, A., Sevcik, R., & Romski, M. (2017). Narrative language and reading comprehension in students with mild intellectual disabilities. *American Journal on Intellectual and Developmental Disabilities, 122*, 392-408.

Objective: The purpose of this study was to describe the narrative ability and to evaluate the relationship between oral narrative language skills and reading comprehension in elementary children with mild intellectual disability by assessing the children's microstructure elements of their oral narratives (MLU, intelligibility, total utterance length, and number of different root words used), macrostructure elements of their oral narratives (defined by Narrative Scoring Scheme NSS), looking at differences in performance of macrostructure elements that reflect strengths and weaknesses in specific components of narrative language ability, and seeing if narrative ability predicts reading comprehension skills.

Method: Prior to the study, participants had 120 hours of reading intervention focused on phonology as part of a larger research study. The participants were from 11 different public elementary schools and were required to have a diagnosis of mild intellectual disability. A total of 102 students were included in the study with 56 being male and 46 being female. Participants were instructed to look at each page of the wordless picture book *Frog Goes to Dinner* and then asked to tell the story to the examiner. The narratives were videotaped and later transcribed on Systematic Analysis of Language Transcript (SALT). Calculations of the microstructural elements (mean length utterance in morphemes, intelligibility, total utterance length, and number of different words) included were completed by the SALT software. Macrostructure elements (introduction, character development, mental states, referencing, conflict/resolution, cohesion, and conclusion) were scored on a rating scale using the Narrative Scoring Scheme (NSS). Additionally, a battery of standardized assessment was also administered to the participants

including the word attack and passage subscales in the Woodcock Reading Mastery Test-Revised, The Peabody Picture Vocabulary Test-III Form A, the Expressive Vocabulary Test, and the CELF-4.

Results: Decoding ability, narrative microstructure, and narrative macrostructure combined were statistically significant ($R^2 = .657$, $p < .001$). Adding just decoding to the equation created a statistically significant variance of $R^2 = .523$, $p < .001$. Adding narrative microstructure created a statistically significant increase of $R^2 = .069$, $p < .01$. Finally, adding narrative macrostructure to the equation created a statistically significant increase of $R^2 = .082$, $p < .001$. This shows that microstructure accounted for 6.9% more of the variance in reading comprehension than decoding alone predicted and narrative macrostructure skills accounted for 8.2% more of the variance in reading comprehension skills further than measures of decoding skills and microstructure skills alone predicted. Another finding was that overall participants had better skills in describing the introduction, conflicts/resolutions, and characters than they were at using mental state verbs, clear references, and cohesive language within macrostructure elements.

Relevance to current study: This study showed that relative strengths and weaknesses in specific narrative macrostructure components contributed to a variance in reading comprehension skills as compared to microstructure elements with children with disabilities. Narrative macrostructure abilities are predictive of reading comprehension skills beyond what measures of decoding skills and components of microstructure would predict.

Catts, H., Nielsen, D., Bridges, M., & Liu, Y. (2014). Early identification of reading comprehension difficulties. *Journal of Learning Disabilities, 49*, 451-465. doi: 10.1177/0022219414556121.

Objective: The purpose of this study was two-fold. First, researchers aimed to find if language ability in kindergarten added to current third grade reading comprehension prediction measures of word reading taken in kindergarten. The second purpose was to examine if response to kindergarten language intervention also added to the third-grade reading comprehension prediction.

Methods: The sample included 266 kindergarten children on one reading outcome measure and 264 kindergarten children on the other. Every participant was administered a battery of screening measures that measured letter knowledge, phonological awareness, rapid automatized naming, nonword repetition, vocabulary, grammar, narration, and word reading. Participants that were selected into the study based on initial risk were then randomly assigned to either an intervention condition or an at-risk control condition. Intervention consisted of half of the time spent on narration and vocabulary with the rest of the time spent on phonological awareness and letter knowledge. The participants' response to intervention was assessed in various ways. Response to narrative intervention was measured by comparison of pretest-posttest performance on the Test of Narrative Language (TNL). Response to vocabulary intervention was measured by comparison of children's knowledge of target words at pretest, midyear, and end-of-year (posttest). At the end of third grade, students' reading comprehension was then assessed using standardized assessment administered by districts and an experimental measure designed by researchers. This study used an approach that divided children into either having a reading disability (RD) or not (non-RD) based solely on reading comprehension outcomes.

Results: Each correlation between each kindergarten screening assessment and third grade reading comprehension measure were significant. To find if each kindergarten language ability predicted reading comprehension over and above other screening measures, a nested approach was used. It was found that language measures administered at the beginning of kindergarten added significantly to the prediction of reading comprehension at the end of third grade. Though the contribution was small, language measures predicted reading comprehension over and above kindergarten measures related to word reading assessed in kindergarten and second grade. It was also found that children's response to narrative and vocabulary language intervention was predictive of third grade reading comprehension.

Relevance to current study: This study provides evidence that early language assessment measures are predictive of later reading comprehension outcomes. This study also gives evidence that response to narrative and vocabulary language intervention in kindergarten are predictive of third grade reading comprehension. Understanding that early narrative oral language abilities and response to early oral narrative instruction leads to the question that we are asking of if early oral narrative intervention can help improve later reading comprehension.

Gillam, S., Olszewski, A., Fargo, J., & Gillam, R. (2014). Classroom-based narrative and vocabulary instruction: Results of an early-stage, nonrandomized comparison study.

Language, Speech, and Hearing Services in Schools, 45, 204-219.

Objective: This is a preliminary study meant to provide an assessment of the impact of a narrative and vocabulary program provided by an SLP in the regular classroom setting.

Methods: The participants were 43 children in 2 different first-grade classrooms. The majority (86%) of students were from underrepresented ethnic groups and 75% qualified for free or reduced lunch. The participants were divided into high risk and low risk subgroups depending on

their performance on the Test of Narrative Language (TNL). Children in both the experimental classroom and the control classroom were assessed before and after the intervention by generating a spontaneous narrative in response to single-scene prompts. These self-generated narratives were transcribed on SALT and then scored using the progress monitoring tool MISL. MISL is designed to measure macrostructure and microstructure elements. Children were also assessed before and after intervention with a criterion-referenced vocabulary probe that measured the students' understanding of words that were related to story grammar, literacy, knowledge, feelings, verbs, and adjectives. In the experimental classroom, an SLP provided narrative and vocabulary instruction for three 30-minute periods per week for 6 weeks. In the control classroom, the regular reading and listening comprehension lessons were continued. The narrative program had three phases. Phase I focused on story grammar elements by hearing and telling stories with simple episodes consisting of an initiating event, attempt, and consequence. Phase II focused on elaboration and included lessons to make stories more complex by including complicating actions in their stories and use of more complex language like coordinated and subordinated conjunctions, adjectives, adverbs, metacognitive verbs, and dialogue. Phase III focused on helping the children to create and tell complex and elaborated stories independently. Vocabulary was also targeted by teaching words specific to story grammar elements (e.g. character, setting), book concepts (e.g. author), internal responses (e.g. frustrated), adverbs (e.g. quickly), verbs (e.g. discover), adjectives (e.g. sneaky) and words specific to the wordless picture books that they were using in intervention (e.g. alley cat).

Results: Children who were in the classroom that received intervention made clinically significant improvements on narrative and vocabulary measures while the children in the control classroom did not. When looking at pre- to post-test scores between the two classrooms, it was

observed that the children in the experimental classroom actually had 3x larger effect sizes for MISL scores as compared to the control classroom. The children who were in the high-risk subgroup actually made greater gains in narration and caught up in story complexity after intervention. However, the high-risk subgroup had fewer gains in vocabulary than the low-risk subgroup indicating that the high-risk children may have needed a more intensive, explicit instruction to learn vocabulary.

Relevance to current study: This study gives evidence that narrative intervention at the classroom level can create gains in for both low-risk and high-risk students in narration for culturally diverse students.

Spencer, T., Kajian, M., Petersen, D., & Bilyk, N. (2014). Effects of an individualized narrative intervention on children's storytelling and comprehension skills. *Journal of Early Intervention, 35*, 243-269.

Objective: Researchers were examining the effects of an individualized narrative language intervention on the language skills of preschoolers with disabilities as measured by narrative retells, comprehension questions, and personal stories. This study also examined parent and teacher perceptions of the social validity of the individualized narrative intervention.

Methods: The participants included five preschoolers with developmental disabilities who were participating in a special education program. A multiple baseline and multiple probe experimental design was used to examine the effects of intervention. Baseline, intervention, and follow-up conditions were staggered across the five children. Intervention included 24 10-15-minute sessions that consisted of visual supports and retell and personal storytelling practice, on retells, personal stories, and story comprehension. *Story Champs* stories were used in intervention. *Story Champs* focuses on story grammar and complex language features through

story telling. The Narrative Language Measures: Preschool (NLM:P) was used to measure narrative retells, story comprehension, and personal story outcomes.

Results: Each child made gains in narrative retells with and without pictures, however, there were not consistent retells without pictures that are more similar to stories produced by typical peers among all participants. Each participant except one improved in story comprehension. Every participant improved in personal story generations.

Relevance to current study: Reasonable evidence is given for improvements in narrative retells, story comprehension, and personal story outcomes from a 12-week intervention that focused on individualized narrative oral language intervention.

Spencer, T., Weddle, S., Petersen, D., & Adams, J. (2017). Multi-tiered narrative intervention for preschoolers: A head start implementation study. *NHSA Dialog*, 20, 1–28.

Objective: This study examined the effects of multi-tiered narrative intervention in Head start students when delivered by Head Start teachers and teaching assistants. Researchers examined the efficiency, fidelity, reliability, and feasibility when implemented by teachers.

Method: The study included 105 preschoolers across six classrooms. Assessments of story retelling and language comprehension were taken in fall, winter, and spring. Teachers and teaching assistants administered the narrative intervention in an MTSS model (large group instruction, small group instruction, individual narrative intervention, and progress monitoring probes using a narrative retell task) after training, modeling, and coaching from the research staff. Story Champs was used to teach story grammar elements. Researchers also monitored fidelity and accuracy in which teachers and teaching assistants administered Story Champs lessons, progress monitoring probes, and scoring of progress monitoring narrative retells. Teachers also completed feasibility questionnaires throughout the year documenting their

comfort level with the assessments, support needed, efficiency of the program, and the students' engagement.

Results: The students who received Story Champs intervention had statistically significant improvements in language comprehension with medium effect sizes. Teachers and teaching assistants had acceptable fidelity and reliability scores when implementing the Story Champs intervention and narrative retell probes. Teacher feasibility reports showed improvements as the school year progressed.

Relevance to current study: This study gives evidence of narrative intervention producing significant improvements in narrative skills. It also shows the feasibility of narrative intervention by having classroom teachers implement the intervention.

Spencer, T., Petersen, D., & Adams, J. (2015). Tier 2 language intervention for diverse preschoolers: An early-stage randomized control group study following an analysis of response to intervention. *American Journal of Speech-Language Pathology, 24*, 619-636.

Objective: This study examined the percentage of Head Start students identified to have needed Tier-2 intervention as determined by use of a dynamic assessment. The study also examined the effects of Tier-2 language intervention on narrative language measured through narrative retells and personal generations as compared to a control group.

Methods: The participants in the study included preschool students from three Head Start classrooms. There were two phases. In the first phase, researchers identified children that would benefit from Tier 2 intervention by using an individualized dynamic assessment. The NLM was used to measure the children's language skills before (pretest) and after (posttest) the teaching phase that consisted of 3 days of whole-class narrative instruction in the dynamic assessment process. Children who scored at or above a total score of 8 were considered levelers. Children

who scored below 8 and then above 8 at posttest were classified as responders. Children who made gains from pretest to posttest but never scored above 8 were classified as gainers. Finally, children who made no gains from pretest to posttest and scored below 8 were classified as minimal responders. Children who were classified as gainers and minimal responders were considered as candidates for Tier 2 language intervention.

22 children were considered to benefit from Tier 2 language intervention and progressed to the second phase of the study. In the second phase, children that had been identified to benefit from Tier 2 intervention were randomly assigned to treatment or control groups. There were four students from each classroom (total of 12 students) that were assigned to treatment. The children that were in the treatment group received 18 sessions of Story Champs oral narrative language intervention in small groups of 4 while the children in the control group did not receive language intervention other than what was provided in their Head Start classroom. The participants were assessed on a story retell before and after the intervention as well as after a 4-week maintenance period using the NLM. The NLM was considered a proximal measure because of how well it aligns with the intervention Story Champs. and The Renfrew Bus Story was used as a secondary outcome measure and was considered a distal retell measure. Additionally, students also were scored on personal story generations

Results: Phase 1 results- 17% of the students were classified as levelers, 17% were classified as responders, 24% were classified as gainers, and the remaining 42% were classified as minimal responders. Except for five students which were unable to participated in testing were considered to benefit most from Tier 3 language support, the remaining gainers and minimal responders (54%) were considered to benefit from Tier 2 language intervention and continued to Phase II of the study.

Phase II results- ANOVA results showed that there was no statistical difference between the treatment and control groups at pretest. There were statistically significant differences between the treatment and control groups on the NLM. There were statistically significant differences between the groups as measured by the Renfrew Bus Story assessment, however, there were no differences in sentence length analysis. No statistically significant differences were found in personal story generation between the treatment and control groups.

Relevance to current study: Narrative intervention in an MTSS context at the small group level has shown to help children have significant gains in narrative retells as measured by the NLM and the Renfrew Bus Story.

Brough, M. (2019). *A large-scale randomized control trial examining the effects of a multi-tiered oral narrative language intervention on kindergarten oral and written narratives and oral expository language* (Unpublished master's thesis). Brigham Young University, Provo, Utah.

Objective: This study examined the effects of an oral narrative language intervention in an MTSS context on kindergartener's oral and written narrative and oral expository skills.

Method: The participants included 686 kindergarten students from four different school districts. The participants were randomly assigned at the classroom level to a treatment group or a control group. The treatment group received Tier 1 oral narrative language instruction following *Story Champs* procedures that was led by the kindergarten teachers. The students that saw no improvement in their narrative retell skills after one month of instruction then received Tier 2 instruction. Tier 2 instruction included oral narrative intervention in small groups led by an SLP. Narrative retell, personal story generation, narrative writing, and expository retell scores were

analyzed during posttest. Treatment group samples were matched according to at-risk, average, and advanced students in the control group. Scores were then compared across all measures.

Results: The results indicated that the treatment group made significant gains when compared to the control group in oral narrative retells, personal story generations, narrative writing, and expository retell skills. Students who needed Tier 2 intervention performed similarly to or outperformed their at-risk, average, and advanced matches in every measure except expository retell.

Relevance to current study: This study demonstrated that narrative intervention across various tiers is effective in improving narrative and expository language skills of kindergarten students.

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

Objective: This study aimed to see if oral discourse skills in preschool were predictive of later literacy outcomes.

Method: Children were assessed at home with a parent present. At age five, children's discourse abilities were assessed through one narrative (play narrative) and one nonnarrative (picture description) task. The play narrative consisted of the child receiving a set of toy animals, an interviewer introducing the story prompt involving verbal conflict, and the child being asked to tell the rest of the story. The picture description task consisted of the child being asked to describe a picture of a complex scene in a way that another child would be able to draw it. Both assessments were videotaped and transcribed later using the conventions of the Child Language Data Exchange System. Language was also assessed at age five using the Index of Productive Syntax to measure morphosyntactic complexity in a conversational setting. At age eight, literacy was assessed through a reading comprehension assessment and a written narrative. The reading

comprehension assessment consisted of The Gray Oral Reading Test (GORT-4). This includes oral reading and comprehension questions. Written narratives were measured by asking children to write a story about a picture sequence they were given. Two elementary language arts specialists then used holistic scoring procedures to assess the overall quality and to rate the written samples.

Results: Not only did this study add to the body of research that there is a relationship between narrative ability and later literacy, this study had more specific findings that control of text-level macrostructures, use of narrative evaluation and provision of elaborated information might support reading and writing development. It was found that within play narration task, narrative clauses, textual evaluation skills, and character states were significantly predictive of later reading comprehension skills. In the picture description task, descriptive information was significantly predictive of reading comprehension skills.

Relevance to current study: This study demonstrated that particularly skills within narrative clauses, textual evaluation skills, and character states were significantly predictive of later reading comprehension skills.

Petersen, D., Brown, C., Ukrainetz, T., Wise, C., Spencer, T., & Zebre, J. (2014). Systematic individualized narrative language intervention on the personal narratives of children with autism. *Language, Speech, and Hearing Services in Schools, 45*, 67-86.

doi:10.1044/2013_lshss-12-0099

Objective: This study aimed to examine the effectiveness of an individualized, systematic language intervention on personal narratives in children with autism.

Method: The participants included three 6- to 8- year old boys with autism. Information about the participant's language ability was assessed through a play-based conversation sample and

two narrative retell samples elicited using the TNR prior to baseline. Two single-subject designs were used in the study. One multiple-baseline design was used across participants to examine the effect of the intervention on personal narratives. Another multiple-baseline design was used across behaviors. Children were assessed during baseline, intervention, and maintenance times. To do this, a clinician modeled a personal story and then the child was asked to tell a story that had happened to them that was similar to the story they just heard. These stories were then scored using the TPG scoring guide. Intervention included 12 individualized intervention sessions. These sessions focused on 2-3 story grammar elements and 3-4 linguistic complexity elements for each individual that were selected from the participant's baseline performance.

Results: Baseline measure for story grammar elements showed flat baseline performance for all seven elements. Immediate improvement in elements targeted was evident for all participants only when the participant was in intervention for that target. All seven variables showed PNDs ranging from 45% to 100%. When examining effects on linguistic complexity, it was found that there was immediate improvements during the intervention phase for seven out of the nine variables indicating a possible treatment effect for these variables. However, there was mixed evidence regarding the maintenance 2 and 7 weeks after intervention.

Relevance to current study: Children with autism can benefit from individualized narrative language intervention.

MTSLS Can Help All Students Improve Language Skills

Al Otaiba, S., Connor, C., Folsom, J., Wanzek, J., Greulich, L., Schatschneider, C., & Wagner, R. (2014). To wait in Tier 1 or intervene immediately: A randomized experiment examining first grade response to intervention in reading. *Exceptional Children, 81*(1), 11-27.

Objective: To assess the effects of Dynamic RTI and Typical RTI on reading comprehension by the end of first grade. This study also examined the relationship between assignment to specific tiers and standardized reading comprehension scores and if the level of prediction changes when comparing Dynamic and Typical RTI.

Method: 34 first grade classrooms were randomly assigned to either the Typical RTI condition or the Dynamic RTI condition at the classroom level. Five different reading assessments were used to screen and monitor students' progress. Tier 1 students participated in the normal core reading program. Tier 2 and Tier 3 students had additional intervention that was aligned with the core reading program and led by trained project staff. Tier 2 students received small group intervention for 30-minutes two times a week. Tier 3 students received small group intervention for 45-minutes four days a week. The only difference between the Dynamic RTI and the Typical RTI groups was when the students were provided with Tier 2 and Tier 3 intervention. The Typical RTI group was meant to mimic what is being implemented in the districts. Typical RTI group, every student began in Tier 1 regardless of their initial reading scores. It was not until the second screening which occurred eight weeks later were students provided with Tier 2 intervention depending on if they responded to Tier 1 intervention. It was not until the third session that students who were not responding to Tier 2 intervention were then provided with Tier 3 intervention. In the Dynamic RTI group, students were provided with Tier 2 or Tier 3 intervention depending on their initial reading scores and then could move up or down at each following screening.

Results: There was no significant difference at pretest (Fall assessments) between the Dynamic RTI group and Typical RTI group. Students in the Dynamic RTI group scored significantly higher than students in the Typical RTI in reading after intervention. Results also showed that

there were no significant differences in growth between the Tier 1 students in either group. The Tier 2 students in the Dynamic RTI group had significantly higher scores than the Tier 2 students in the Typical RTI group. The Tier 3 students in the Dynamic RTI group scored higher than the Tier 3 in the Typical RTI group.

Relevance to current study: Providing reading intervention at the appropriate Tier level sooner has significant effects on reading assessments in Tier 2 and Tier 3 students.

Fuchs, D., & Deshler, D. (2007) What we need to know about responsiveness to intervention (and shouldn't be afraid to ask). *Learning Disabilities Research & Practice*, 22, 129-136.

Summary: Response to intervention (RTI) can provide early intervention and a valid way for disability identification. Learning at the classroom level is often referred to as Tier 1. Student responsiveness is evaluated. If students are unresponsive at Tier 1 intervention, then Tier 2 instruction is delivered.

Other Related Articles

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

Objective: The purpose of this study was to evaluate the relationship between code-related and oral language precursors and their effect on reading in later grades.

Method: 626 four-year-old age students attending Head Start were assessed in language and literacy skills in the spring of preschool, kindergarten, first, second, third, and fourth grades. The assessments measured code-related skills, oral language skills, and reading. The code-related skill evaluations were given the spring of the participants' Head Start and kindergarten years. The code-related assessment that was administered was the Developing Skills Checklist

including subtests in Memory, Auditory, Print Concepts, and Writing and Drawing Concepts. Oral language was assessed using the Renfrew Bus Story in preschool, One Word Picture Vocabulary Test in preschool and kindergarten, Clinical Evaluation of Language Fundamentals (CELF-P) in kindergarten, and the Peabody Picture Vocabulary Test (PPVT-R) in all grades tested. The reading measures included standardized tests using various subtests from the Stanford Achievement Test- Eight Edition (SAT), Wide Range Achievement Test- Revised (WRAT-R), and the Woodcock Reading Mastery Test- Revised (WRMT-R) in each consecutive grade beginning in first grade and ending in fourth. It should be noted that reading measures were divided into two domains of reading accuracy and reading comprehension.

Results: Findings indicated that reading is comprised of two distinct constructs, decoding and oral language. Print knowledge and phonological awareness predicted decoding ability, as expected. Oral language and code-related skills had a very strong relationship at preschool. In grades 1 and 2, oral language and code-based reading skills were non-significant.

The results showed that oral language skills predicted 48% of the variance in code-related skills in preschool, but less than 10% of the variance in code-related skills in later grade. This indicates that there is a strong relationship between oral language and code-related skills in preschool years but not in later years. The relationship between oral language skills and code-related reading ability was not statistically significant in first and second grades. However, the relationship's influence re-emerged in third and fourth grades by displaying 7% of the variance in reading comprehension.

Another finding was that there was longitudinal continuity in both oral language and code-related skills. 90% of the variance in kindergarten oral language ability was deemed by preschool oral language ability, 96% of the variance in first and second grade oral language ability was

deemed by kindergarten oral language ability, and 88% of the variance in third and fourth grade oral language ability was deemed by first and second grade oral language ability. In the code-related domain, 38% of the code-related skill ability was deemed by preschool code-related oral language ability.

Researchers also found a direct relationship between a child's kindergarten code-related skills and their reading ability in early elementary school. There was a 58% variance in first grade reading ability deemed by kindergarten code-related skills. In second grade there was a 30% variance in second grade reading ability deemed by kindergarten code-related abilities.

Lastly, researchers found that a child's reading comprehension in third and fourth grades had an 18% variance deemed by the child's previous reading achievement, 16% variance deemed by their current reading accuracy, and a 7% variance deemed by their concurrent language skill.

This indicates that reading comprehension in third and fourth grade was significantly affected by the child's prior reading ability, current reading accuracy, and a child's concurrent language skill.

Relevance to current work: Early oral language skills predict variance in a child's code-related abilities, skills necessary for reading. However, this relationship diminishes over time. There is a 7% variance in reading comprehension that is accounted for by concurrent oral language abilities in third and fourth grades. Reading accuracy in later grades is greatly influenced by prior word recognition and decoding abilities, but reading comprehension is influenced by previous reading ability, concurrent reading accuracy, and concurrent language ability. These findings support the idea that oral language is the foundation for reading comprehension.

Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, 56, 72-110.

Objective: This study is a meta-analysis of studies that evaluates vocabulary instruction and its effect on reading comprehension and understanding of word meanings.

Method: Studies were selected from a computer search of the ERIC document service, past reviews and bibliographies, and cross-checking references. Studies were required to have a control group and required to provide statistical information needed to derive an effect size. Teaching methods were classified by setting and method factors. Setting was determined by two factors: if the intervention was in a group or single setting and the amount of time given to instruction. Method descriptions were rated by two authors and a graduate student on (a) whether or not a method gives the student examples of each to-be-learned word in context, (b) the types of activities that are required to learn the word, and (c) the number and type of exposures to information about each word. Raters agreed at least 80% of the time and within one category at least 95% of the time.

Two types of comprehension measures were used in studies that reported the effect of vocabulary instruction on comprehension. They were global comprehension measures and word-specific measures. The studies that were looking at the effects of various teaching methods on word knowledge used three types of measures: global vocabulary measures, definitional word-specific measures, and contextual word-specific measures.

There were two different types of control groups used. When looking at the effects of supplying or not supplying vocabulary content on things like reading comprehension or when comparing different teaching methods, a “no-exposure” control group was used. This is when the control group does not have any type of exposure to the vocabulary words. The other control group utilized was called a “no-instruction” group where the students were given the target words and definitions with the instruction to study them however they would like. Effect sizes were

calculated either using the “no-exposure” control group or the “no-instruction” control group. In the vocabulary/comprehension relationship, only the “no-exposure” set was used. In the comparison of different vocabulary teaching methods, the “no-exposure” set was reported first, with the “no-instructions” set results used to confirm and expand the results.

Results: Vocabulary instruction produced a mean effect size of .97 (SD = .81, N =41). This means that generally the children who were at the 50th percentile of the children receiving vocabulary instruction scored as well as the children at the 83rd percentile of the control groups on passage comprehension measures. Another finding was that vocabulary had a significant general facilitative effect on reading comprehension in standardized test’s reading passages. There was a .30 mean effect size (SD = .22, N = 15). Meaning that generally students who were at the 50th percentile of the children receiving vocabulary instruction scored as well as students at the 62nd percentile of control groups in global reading measures. There was also a significant effect on global measures of vocabulary knowledge for the children that received vocabulary instruction. The mean effect size for this was .26 (SD = .29, N = 17).

In the method comparisons section, the methods that appeared to produce the greatest effects on comprehension and vocabulary measures were the methods that contained both definitional and contextual information about each targeted word. Keyword methods produced strong effects on measures of definitional and contextual vocabulary knowledge. Time allocation and effect size were not significant on vocabulary measures but were strong for passage comprehension measures. This might indicate that words need to be fully learned to assist in comprehension.

Relevance to current study: Vocabulary instruction has a significant effect on the comprehension of passages containing taught words. It also has a significant effect on comprehension of passages that do not necessarily contain the taught words. Vocabulary instruction should contain

both definitional and contextual information to be most effective. Oral language intervention includes vocabulary instruction which in turn leads to improved reading comprehension.

APPENDIX B

Posttest M-STEP Reading Comprehension Measure Sample

Grade 4 ENGLISH LANGUAGE ARTS Sample

Read the passage. Then, answer questions 1 through 7.

Man's First Flight

by Kiera Downie

Orville and Wilbur Wright became famous when they flew their airplane, the Wright Flyer, into the pages of history. But humans had been flying for many years before that famous event. Hot air balloons were the first way humans flew. The idea for these balloons came from China over a thousand years ago. The Chinese made a lantern to use as a signal. It was a balloon made of paper, using a candle to both light it and carry it upward. It wasn't long before people began to think that if they could make a small balloon fly, they could make a big balloon fly, too—one big enough to hold a person.

Today, the hot air balloon design is the same as the balloons in China, although the materials are a little different. We now make the balloons from nylon, a strong and flexible material. They are attached to large baskets that are made of wicker and big enough to carry people. Wicker is woven wood that is strong and lightweight. The strength helps the basket hold the passengers. The light weight makes it easy for the balloon to carry the basket.

The hot air balloon flies by a simple design. The balloon is filled with hot air. Hot air weighs less than cold air. So when the hot air is trapped inside of the balloon, the balloon's response is to rise up in the cooler air surrounding it.

In order to make sure the balloon continues to float, the air is heated by burners. The burners are filled with propane which is the same fuel used in outdoor gas grills. Just like a grill, the propane is lit and burns right beneath the opening at the bottom of the balloon. That flame heats the air inside the balloon and makes it rise into the air. The balloon's pilot must turn the burner on and off to heat the air. In this way, the pilot makes the balloon move up and down. But how does a hot air balloon move from side to side?

Hot air balloons travel on natural air currents. An air current is a flow of air over the earth. We feel air currents as wind on our faces. All around the world, air flows in different directions. These currents flow in layers above the earth. Sometimes one current will flow east, but the current above it will flow west. A hot air balloon pilot uses the burner to lift the balloon into different currents. The balloon moves east, west, north, or south depending on the current it's in.

Of course, a hot air balloon also has to land. To land, the pilot has to slowly cool the air inside. The pilot opens a flap at the top of the balloon. The flap lets in cool air and releases hot air from the balloon. As the air slowly cools, the balloon drops from the sky. It is important the pilot lets the cool air in slowly, or the balloon will fall too quickly. The balloon drifts downward and eventually comes to a stop on the ground.

Once the balloon lands, the pilot releases all of the remaining air. This is called deflation.

When the balloon is deflated, it lays flat as a pancake on the ground, and the passengers can leave the basket.

Hot air balloons are difficult to pilot. They only move as fast as the air currents will carry them. Because of this, we don't fly balloons to work or school. However, hot air balloons are a wonderful way to see the earth from up in the clouds. It's strange to think that a simple idea for a lantern led to the modern use of hot air balloons. It's even more strange when you learn that the way balloons fly isn't much different from the way those lanterns flew.

1. Which sentence from the passage supports the conclusion that the Chinese discovered that hot air is lighter than cold air?

A. "Hot air balloons were the first way humans flew."
 B. "The idea for these balloons came from China over a thousand years ago."
 C. "The Chinese made a lantern to use as a signal."
 D. "It was a balloon made of paper, using a candle to both light it and carry it upward."

2. This question has two parts. First, answer part A. Then, answer part B.

Part A

Which sentence **best** describes the author's main idea in paragraph 1?

A. The hot air balloon was invented before the airplane.
 B. Human flight was the idea of Orville and Wilbur Wright.
 C. Human flight, which is important to history, is over a thousand years old.
 D. The hot air balloon, the first way humans flew, was based on ancient Chinese lanterns.

Part B

Which detail from the passage **best** supports your answer in part A?

A. "...they flew their airplane, the Wright Flyer, into the pages of history."
 B. "...humans had been flying for many years before that famous event."
 C. "Hot air balloons are the first way humans flew."
 D. "The idea for these balloons came from China, over a thousand years ago."

3. This question has two parts. First, answer part A. Then, answer part B.

Part A

Which conclusion about the author's purpose is supported by the passage?

A. to explain how a hot air balloon works
 B. to describe the history of human flight
 C. to explain how humans changed the way people flew
 D. to describe how the modern hot air balloon was created

Part B

Which sentence from the passage **best** supports your answer in part A?

- A.** “Orville and Wilbur Wright became famous when they flew their airplane, the Wright Flyer, into the pages of history.”
- B.** “Today, the hot air balloon design is the same as the balloons in China, although the materials are a little different.”
- C.** “The hot air balloon flies by a simple design.”
- D.** “However, hot air balloons are a wonderful way to see the earth from up in the clouds.”

APPENDIX C

IRB Approval Form



Vice President for Research & Economic Development
1000 E. University Avenue, Department 3355 • Room 305/308, Old Main • Laramie, WY 82071
(307) 766-5353 • (307) 766-5320 • fax (307) 766-2608 • www.uwyo.edu/research

July 15, 2014



Protocol # 20140715DP00480

Re: IRB Proposal “ ”

Dear Doug:

The proposal referenced above (received July 10, 2014) qualifies for exempt review and is approved as one that would not involve more than minimal risk to participants. Our exempt review and approval will be reported to the IRB at their next convened meeting September 25, 2014.

Any significant change(s) in the research/project protocol(s) from what was approved must be submitted to the IRB on the Protocol Update Form for review and approval prior to initiating any change. Further information and the form referenced above may be accessed at the “Human Subjects” link on the Office of Research and Economic Development website: <http://www.uwyo.edu/research/human-subjects/index.html>

You may proceed with the project/research and we wish you luck in the endeavor. Please feel free to call me if you have any questions.

Sincerely,

Ashley Guritza

Ashley Guritza
Associate General Counsel, Compliance
On behalf of the Chairman,
Institutional Review Board