ABSTRACT

Title of Dissertation: INVESTIGATING THE THEORETICAL AND INSTRUCTIONAL RELATIONS BETWEEN THEORY OF MIND AND READING COMPREHENSION

Jerae Hutchison Kelly, Doctor of Philosophy, 2022

Dissertation Directed by: Dr. Kelli D. Cummings
Counseling, Higher Education and Special Education

Research on Theory of Mind (ToM) and reading comprehension is a lively and active field with numerous publications a year. ToM describes a child's ability to identify and reason about the mental states of others (e.g., think, believe, intend, want). The burgeoning findings from this research suggests ToM plays an important role in the reading comprehension outcomes of narrative texts, specifically through the linguistic comprehension domain of the Simple View of Reading (SVR) by supporting inference making. Despite the converging findings, however, the research field has yet to translate the research into practice. The purpose of the research in the present dissertation is to offer the Theory of Mind and reading comprehension research field a glimpse into how we can translate the research into practice and to recount the effects of my efforts. The overall aim of the present dissertation is to explain how ToM contributes to narrative reading comprehension as a skill of social inference making.

The present dissertation is structured into five parts consisting of an introduction that outlines the dissertation and establishes the theoretical frame (Chapter 1), a synthesis of quantitative studies on the relation between ToM and reading comprehension (Chapter 2), a mixed method descriptive analysis of secondary data of preschool teachers' interactive read alouds (Chapter 3, Study 1), an intervention study focused on using ToM to teach theme identification (Chapter 4, Study 2), and a conclusion (Chapter 5).

The findings from the synthesis reveal a small and mediated relation via listening comprehension between ToM (almost solely measured using false-belief tasks) and reading comprehension and point to the explanatory power of second order and advanced ToM, specifically. Additionally, only one intervention study was included in the literature corpus for the synthesis. The findings from Study 1 suggest that preschool teachers use ToM language, defined as references to mental states, during their interaction read aloud instruction and that such use supports comprehension. The findings from Study 2 indicate that situating ToM, as a skill of social inference making during dialogic reading activities, combined with explicit instruction on theme identification is an effective method of increasing the narrative comprehension of less-skilled comprehenders. The results of the dissertation as a whole help clarify how ToM functions within the linguistic comprehension domain as a skill of social inference, making for narrative comprehension in conjunction with other known variables of reading comprehension.

In sum, the present dissertation offers the ToM and reading comprehension research field an applied investigation of ToM. The present dissertation is the first of its kind in this endeavor. Limitations and future research directions are also discussed.

INVESTIGATING THE THEORETICAL AND INSTRUCTIONAL RELATIONS BETWEEN THEORY OF MIND AND READING COMPREHENSION

By

Jerae Hutchison Kelly

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:

Associate Professor Kelli D. Cummings, Chair
Professor Ana Taboada Barber
Professor Susan De La Paz
Professor Jennifer D. Turner
Professor Kelly B. Cartwright

Dedication

For those that believe in the power of stories to change who we are.

Acknowledgments

I truly have so many people to thank. I must start by thanking my family. My grandfather made many sacrifices in his life for the benefit of his family and the country at large. Without him, my family would not be the same. My grandmother was one of the first special education teachers in the country in 1975 when the Education for All Handicapped Children Act was passed. She paved the way for my mother and me, setting a bright and shining example of fortitude and strength.

To my husband Kevin, thank you for teaching me what unconditional love truly is. I couldn't have done this without you. I love you. Onyx and Ruby were the best pandemic coworkers.

To my mom and dad, there are not enough words to describe all the ways in which you have supported and guided me. Mom, it is without a doubt that your steadfast belief in me and unwavering support of me is the most prominent reason I am here. Dad, your support throughout my life has meant the world to me as well. You are the reason I am the confident risk-taker that I am. The resiliency you have shown throughout your life inspires me every day. I thank you both from the deepest part of my heart for all the personal sacrifices you both have made for me throughout the years and for loving me so fiercely. I share in accomplishing this feat not because of you both, but with you.

To my older brother, Fred, and sister, Chalen, you both have provided me with guidance throughout my life. Fred, you've always modeled true dedication and exemplary brilliance in all you do. Chalen, you've always demonstrated a deep commitment to social justice and commitment to continuous learning. You have also always encouraged me and supported me in

following my own path. Thank you both for your encouragement, support, and for setting such fine examples.

I next must thank my committee members. Thank you for all the guidance and mentorship you have provided me during my time at Maryland. To my advisor, Kelli, I feel that I couldn't have found a more nurturing advisor if I tried. Your firm belief in my abilities as a researcher was the perfect antidote when imposter syndrome tried to take hold. The wisdom you have imparted to me is invaluable and I will forever be indebted to you. I share in this accomplishment with you. Ana, I have immense gratitude for you. I hope our shared passions continue to produce quality research for the field. Susan, your support and feedback on developing my intervention idea really helped shape my thinking. Dr. Turner and Dr. Cartwright, thank you for offering your expertise and for all your support.

To my project Propell cohort, Alex, Jen, Erin, Daniel, Brittany, and Cam, thank you for all the ways in which we supported each other as we forged through the program. Alex especially, you helped me in more ways than I have room to write. Cuba is one of my most beloved memories. To my CLD writing group, thank you for showing me the ropes and becoming a source of great support throughout my PhD journey. Lydia especially, thank you for always being my number 1 fan and cheerleader.

Now, all my friends. Kelsey, Nikki, Caitlin, Chelsey, Maggie, Natasha, Rachel, and Anj, thank you for always being there when I need you to be and bringing me so much joy. Each one of you has played a unique and integral role in my journey. You all inspire me to better than I was the day before.

My Urban Teachers crew: Kyair, Savannah, Glenna, and Leroy, our pandemic Zoom hangouts is something I will cherish. Marcus your messages of encouragement always put a

smile on my face. Corey and Liza, I simply would not have made it through without you. Tina, the conversations we have had over the years have made me a better teacher and a better human. I look forward to continually learning from you and supporting you on you journey as an educator. All of you continue to inspire me in ways I can't quite put into words.

To all the Urban Teachers faculty, thank you for providing me a foundation of knowledge and offering me a chance to start a career in teaching as a TFA reject. A special shout out to Katryna for giving me the idea to go for a PhD in the first place. Thank you for seeing something in me that I couldn't see myself. To all the ACCE Eagles and Ben Franklin Bayhawks, thank you for serving the students and families of Baltimore City alongside me.

Finally, to my past students and their families in Baltimore City. Thank you. You taught me more than you will ever know.

And to those I may have missed thanking directly, if you're reading this, I thank you too.

Table of Contents

Dedication ii

Acknowledgments iii

Table of Contents vi

Chapter 1: Introduction 1

- What is Theory of Mind? 3
- Statement of the Problem 6
- Outline of the Dissertation Manuscript 12
- Theoretical Framework: Reading Comprehension and Theory of Mind 15
  - The Construction-Integration Model of Reading Comprehension and Theme 16
    - Inference Making. 18
  - Prominence of Reader Generated Knowledge-Based Inferences 21
  - Background Knowledge and Narrative Comprehension 23
- Potential Contribution 27

Chapter 2: What do We Know About Theory of Mind and Reading Comprehension and Where Do We Go from Here? 29

- Theory of Mind Defined 30
  - A Focus on False Belief 33
- Theoretical Support for the Importance of Theory of Mind to Reading Comprehension 35
- The Present Synthesis 37
- Method 38
  - Search Procedures 38
  - Coding 39
- Results 40
  - Study Features 41
  - Measures 42
  - Analytic Findings 44
- Discussion 45
- Work on Theory of Mind and Comprehension Since 2018 48
- Limitations 51
  - Limitations of Synthesis Studies 51
  - Limitations of the Present Synthesis 52
- Future Directions 53
- Conclusion 55

Chapter 3: A Reading Comprehension View of the Prevalence and Purpose of Mental-State Language in Preschool Interactive Read-Aloud Instruction 56
Guiding Theory 57
Shared Reading 58
Importance of Dialogic Interaction 60
A Focus on Narrative 62
The Present Study 63
Study Overview 64
Participants and Setting 65
Book Details 66
Part 1: Pilot Coding 66
Part 1 Method 66
Measures 66
Procedures 67
Data Analysis 68
Part 1 Results 69
Part 1 Discussion 70
Part 2: Development of Additional Codes 72
Part 2 Method 72
Procedures 73
Coding 74
Trustworthiness 76
Part 2 Results 76
Character Traits as Mental State Types 76
Behavior Descriptions as Mental State Types. 77
Non-Examples 79
Part 2 Discussion 80
Part 3: Final Coding and Data Analysis 81
Part 3 Method 81
Coding Scheme. 81
Quantitative Coding Procedure and Analysis. 81
Coding Reliability. 82
Qualitative Coding Procedure and Analysis. 83
Researcher Positionality 85
Part 3 Results 86
Research Questions 1 and 2. 86
Research Question 3 87
Ms. Taylor 87
Ms. Hanna 89
Ms. Martin 92
Part 3 Discussion 94
Teacher Variation in Mental State Language Use 96
General Discussion 99

Implications for Practice 100
Limitations and Future Research Areas 102
Conclusion 103
Chapter 4: Teaching Theme Using Theory of Mind and Inference Making During Interactive Read Alouds of Authentic Children's Literature 105
Comprehension, ToM, and Inference Making 108
Theory of Mind Training: A Blueprint for Knowledge Activation 112
Inference Training: How to Integrate Activated Knowledge 114
Theory of Change 116
Method 117
Participants and Setting 117
Selection Process and Measures 119
Intervention Procedures 121
Dependent Variable and Measure 122
Book Selection 123
Experimental Design 124
Baseline 124
Explicit Instruction Session 125
Treatment + assessment probe sessions 125
Post-Intervention 126
Procedural Fidelity, Reliability, and Interobserver Agreement 126
Data Analysis 127
Results 127
Kai 127
Kelani 128
Kelsey 128
Social Validity 128
Discussion 129
Explicit Instruction 129
Theme and Background Knowledge 131
Limitations and Future Research Directions 133
Implications for Practice 135
Conclusion 136
Chapter 5: Conclusion 138
Studies Review 139
Summary 143
Implications for Classroom Practice 146
Future Research Directions 147
Conclusion 149

Tables 151
Chapter 1 Table 1 151
Chapter 1 Table 2 152
Chapter 2 Table 1 153
Chapter 2 Table 2 155
Chapter 2 Table 3 157
Chapter 2 Table 4 159
Chapter 3 Table 1 163
Chapter 3 Table 2 166
Chapter 3 Table 3 167
Chapter 3 Table 4 168
Chapter 3 Table 5 169
Chapter 3 Table 6 170
Chapter 3 Table 7 171
Chapter 3 Table 8 172
Chapter 4 Table 1 173
Chapter 4 Table 2 175
Chapter 4 Table 3 176
Figures 177
Chapter 2 Figure 1 177
Chapter 3 Figure 1 178
Chapter 3 Figure 3 180
Chapter 4 Figure 4 181
Chapter 3 Figure 5 182
Chapter 3 Figure 6 183
Chapter 4 Figure 1 184
Chapter 4 Figure 2 185
Appendix A: Intervention Materials 186
Appendix A Figure 1 186
Appendix A Figure 2 187
Appendix A Figure 3 188
Appendix A Figure 4 189
Appendix A Figure 5 190

Appendix A Figure 6 192
Appendix A Figure 7 193
Appendix B: Sample Lesson 194
References 198

Chapter 1: Introduction

In the 21st century, reading assessment tasks present increased complexity reflecting the need for deep comprehension (Alexander & Disciplined Reading and Learning Research Laboratory [DRLRL], 2012; Graesser, 2015). Deep comprehension, most succinctly, represents learning from text (Goldmanm & Pellegrino, 2015). Readers who demonstrate deep comprehension of texts display the use of strategies to comprehend, such as understanding the author's purpose, the text's structure, and the main idea or theme (Graesser, 2007). The employment of deep comprehension reading strategies allows students to synthesize information across texts and use their background knowledge from outside the text (as well as new knowledge gained from the text) to complete tasks of analysis, evaluation, and critique (Graesser, 2015). In other words, it is not enough for a reader to simply comprehend a text; they must comprehend a text *and* use their comprehension to complete a task of application, often across texts.

Deep comprehension strategies help readers complete such complex tasks. For example, it is easier to synthesize a set of texts if one can identify the thematic whole of the set as well as the theme of each one individually. *Strategies* differ from *skills* in that readers flexibly apply strategies across text types and for different reading purposes. Skills, on the other hand, represent the abilities readers need to master in order to proficiently perform the strategies, regardless of text type and reading purpose (Afferbach et al., 2008). Strategies of deep comprehension require higher-order cognitive and linguistic skills of comprehension, such as inference making, vocabulary, and background knowledge.

To showcase these three skills in action, let me present the first paragraph of Harper Lee's American classic, *To Kill a Mockingbird*:

> When he was nearly thirteen, my brother Jem got his arm badly broken at the elbow. When it healed, and Jem's fears of never being able to play football were assuaged, he was seldom self-conscious about his injury. His left arm was somewhat shorter than his right; when he stood or walked, the back of his hand was at right angles to his body, his thumb parallel to his thigh. He couldn't have cared less, so long as he could pass and punt. (Lee, 1960, pg. 1).

You might notice the general contribution of vocabulary by the word *assuaged*, as well as the interaction between background knowledge and vocabulary by the specific use of the words *pass* and *punt* in reference to football. You also might note inference making between the pronouns *he*, *his*, and *my brother Jem* across the paragraph, along with other propositional relations between *arm badly broken*, *it*, and *injury*.

Also apparent is inference making in reference to Jem's love of football and self-confidence; it is an inference because neither are explicitly stated within the text. Textual information about Jem's injury suggests to the reader that his injury makes his arm look misshapen, leading the reader to infer that his injury may cause people to look at him funny, resulting in his embarrassment. Further knowledge from the text, however, suggests Jem is "seldom self-conscious" and "couldn't have cared less." Given this knowledge, even if Jem could have felt embarrassed by his injury, a reader's understanding of such references to Jem's mental state (i.e., his thoughts, beliefs, and feelings) about his injury clarifies something about his character—that he is self-confident even if he has a misshapen arm. You also notice here the role of another higher-order skill of comprehension: comprehension monitoring. Indeed, instead of caring about what his arm looks like, what Jem does care about is playing football (i.e., *so long as he could pass and punt*).

Inferences such as those made about Jem's character traits are the result of a two-stage process representing an initial stage of knowledge activation and a subsequent stage of knowledge integration (Kendeou, 2015). Activated knowledge from the text or a reader's background is integrated to a reader's overall understanding of the text to complete an inference. In the example above, knowledge about Jem is integrated into our understanding of the overall story by providing information about Jem's character traits. What the example demonstrates is that for narrative texts, such as *To Kill a Mockingbird*, the knowledge needed for inference making is inherently *social* (i.e., understanding the likelihood someone might feel embarrassed by a misshapen arm and why). Therefore, in addition to standard prior knowledge about the topic of football, having greater social knowledge strengthens inference making while reading this passage. Mar and Oatley (2008) emphasize that such social inference making processes result from the fact that stories embody abstractions of social worlds, which readers engage in to practice and learn social understanding processes. In accordance with this view, past research does suggest the social inferential processes that constitute one's social knowledge in the real-world liken to inference making for narrative comprehension (Graesser et al., 1994; Heyes & Frith, 2014). In Kim's Direct and Inferential Model of Reading (2017, 2019, 2020), Theory of Mind (ToM) is a socio-cognitive variable of comprehension that is a measurable way to capture such social inference making.

What is Theory of Mind?

ToM refers to the ability to accurately attribute mental states—(i.e., think, believe, want, intend, feel)—to oneself and others (Wellman, 2002) and to use such attributions to reason on one's own and others' social behaviors. The process of mental attribution and reasoning defined as ToM helps a child act pro-socially by, for example, attending to the preferences of others. If a

child's friend prefers regular milk over chocolate milk, a child might get their friend regular milk instead of chocolate, even if the child's own preference is for chocolate milk. In this example, the child first recognizes their friend's preference (a mental state) and then performs a social behavior in response (the act of getting their preferred milk). In this way, the emergence of ToM is foundational to social cognition because it supports the child's understanding of the contingent relationship between mental states and subsequent displays of behavior (Fiske & Taylor, 2013; Perner & Wimmer, 1985; Wimmer & Perner, 1983), which sets the stage for the development of more complex socio-cognitive skills, such as empathy (Dvash & Shamay-Tsoory et al., 2014).

ToM's development is generally divided into three stages: first order, second order, and advanced. First order refers to the ability to differentiate one's mental state from another's (Wimmer & Perner, 1983). Second order refers to the recognition that others can think something about another person's thoughts (e.g., Jane *thinks* that John *thinks* the cake is in the break room; Perner & Wimmer, 1985). Advanced ToM refers to and develops in order of understanding social reasoning, recognizing transgressions of social norms, and reasoning about ambiguity (Osterhaus et al., 2016). The categories' developmental periods are typically between the ages of three and five, six and eight, and adolescence to adulthood, respectively.

First order ToM development is further divided into stages by type of mental state understanding. The ToM scale (Wellman & Liu, 2004) establishes the development of ToM in reference to increasingly complex mental state categories. First, children begin with developing an understanding of desire (e.g., the mental state of *to want*). Next, children come to understand beliefs (e.g., *to think* or *believe*). Beliefs are unique mental states in that they have the propensity to be false. The understanding of belief states develops in two chronological developmental components respective of the propensity for falsehood: an understanding of diverse beliefs (when

both the child and actor are ignorant of what is true) and an understanding of false beliefs (when the actor does not know what belief is true, but the child does).

A false belief is one held by an agent that is incongruent with reality (Wellman et al., 2001). Identifying false beliefs, when a child knows what is true, requires the child to reason theoretically when considering another's false belief. As a result, such false-belief tasks are the most commonly used measurement of ToM because of the requisite for accurate mental state attribution for task proficiency. Consequently, false belief understanding is considered a proxy for ToM across the literature (Astington, 2003). The last type of mental state understanding to develop based on the ToM scale is the understanding of emotion (Wellmam & Liu, 2004). The understanding of emotional mental states (e.g., happy, sad, excited) parallels that of belief in that children develop an understanding of real emotions (i.e., actual felt emotion) before they understand apparent emotion (i.e., what is displayed). In other words, children develop an understanding of emotion by first recognizing appropriate emotional responses and only later do they develop an understanding that sometimes one's expression of emotion may not match the actual felt emotion. The ToM scale's development pattern presents as one of mediation, meaning that the subsequent acquisition of more advanced mental state knowledge occurs after the acquisition of the previous (Wellman & Liu, 2004).

Cultural variation within the ToM scale sequence appears to exist (Wellman et al., 2006, 2011). In a comparison study of Chinese and Western (American & Australian) children, the sequence of mental state understanding differed across cultures. All children, however, did develop ToM suggesting that its development is universal. Chinese children developed an earlier understanding of knowledge and ignorance in contrast to Western children who developed an earlier understanding of desire and belief (Wellman, 2006; see also Doan & Wang, 2010).

Wellman and colleagues attribute the sequential difference in mental state understanding to the socio-cultural nature of ToM development; although innate biological processes contribute to ToM development, environmental experiences contribute as well (Wellman, 2017).

Such interaction effects results in the presence of differential rates and capacities for ToM across individuals (Hughes et al., 2015; Wellman, 2011). Language, for example, functions as a significant contributor to ToM development (Astington & Baird, 2005; Miller, 2012; Milligan et al., 2007; Tompkins et al., 2019). Children in linguistically rich environments develop ToM earlier than children in environments that are less so (Tompkins, 2015). Encouragingly, trainings for those who display ToM delays or overall deficits appear effective, suggesting ToM is a malleable skill and responsive to intervention (Hofmann et al., 2016).

Statement of the Problem

National Assessment of Educational Progress (NAEP) scores indicate that students consistently struggle in reading in Grade 4, often referred to as the Grade 4 slump (Chall et al., 1990). The average Grade 4 NAEP score in 2017 was 222, and the score to achieve proficiency was 238 (United States Department of Education, NAEP, 2019a). In other words, less than half (44%) of the students assessed scored proficient or above on the NAEP. Approximately 64% of the skills assessed within the NAEP in Grade 4 require ToM skill, and 100% of the skills listed under the proficient threshold require ToM skill, indicating that students will likely not achieve an advanced score without ToM proficiency (Chapter 1, Table 1).

The Common Core State Standards also reflect the need for ToM beginning before Grade 4 (National Governors Association Center for Best Practices, 2010). For example, the Reading Literature (RL) CCSS Grade 2 standard RL.2.3 reads, “Describe how characters in a story respond to major events and challenges.” Within a year, the standard RL.3.3 stretches to,

"Describe characters in a story (e.g., their traits, motivations, or feelings) and explain how their actions contribute to the sequence of events." The contribution of ToM is not limited to reading literature. The Reading Informational Text CCSS Grade 2 standard RI.2.6 reads, "Identify the main purpose of a text, including what the author wants to answer, explain, or describe," and by the end of Grade 3 RI.3.6, students must "distinguish their own point of view from that of the author of a text." Across domains, ToM proficiency is integral to mastery.

Many argue that the Grade 4 slump represents the challenges students face in transitioning from learning to read to reading to learn (Chall, 1983), but it is actually more nuanced (Connor et al., 2014; McNamara et al., 2017). Most modern interpretations of the slump assert the development of vocabulary, inferencing, and other background skills during pre-, emergent, and beginning reading support reading comprehension at later stages and become increasingly crucial once students transition to learning from texts (Catts et al., 2006; Cho et al., 2019; McArthur et al., 2000; Nation et al., 2010; Reese et al., 2009; Spencer et al., 2014; Tompkins et al., 2013; van den Broek et al., 2011) and support reading strategies conducive to deep comprehension (LaRusso et al., 2016). Such skills often do not exert an effect on reading comprehension assessment until later in development (e.g., around Grade 4; Oakhill & Cain, 2012) because so much of a student's comprehension score depends on word reading until they reach automaticity. The Grade 4 slump thus may be the result of poorly developed higher-level comprehension skills during pre-, early, and beginning stages of literacy development, the corresponding developmental time period in which ToM, a higher-order reading skill, is also developing.

Perhaps, or because, comprehension assessments in the early grades are constrained by word reading development, early grades' instruction tends to focus less on higher-level

comprehension skills (Paris & Paris, 2003). The reasons for this lack of instructional emphasis are not well defined. One possibility may be that teachers emphasize decoding and related skills because beginning reading measures used in schools place a greater emphasis on such skills (Catts et al., 2015; Paris & Paris, 2006). Another possibility is the result of a more general overcorrection by the education system in response to the perils of the whole language era on word reading skills (Castles et al., 2018; Semingson & Kerns, 2021). Either way, a lack of skill development during these stages contributes to students' struggles with reading comprehension, sometimes even despite the development of sufficiency in word recognition and decoding (Cain & Oakhill, 2006; Geva & Massey-Garrison, 2013; Landi & Ryheard, 2017; Petscher et al., 2018).

Another important piece of the Grade 4 slump puzzle is the influence of reading comprehension assessments. In general, the use of different reading comprehension assessments results in varying importance of different skills. For instance, inferential and literal questions and different text passage types (narrative vs. expository) draw on different cognitive and linguistic skills (Eason et al., 2012; Spencer et al., 2019). Further, these differences also exist across development. In early grades for example, decoding and related skills fully explain reading comprehension outcomes (Cutting et al., 2006; Lonigan et al., 2018). Once students can read with fluency, only then is variance explained by other components (e.g., fluency; Silverman et al., 2013). The importance of higher-order reading skills thus becomes negligible at best if the student's decoding skills are still in the process of developing. This conclusion suggests that, to a degree, assessments of *reading* comprehension may obfuscate the early developmental importance of higher-order comprehension skills because reading comprehension assessments, such as the NAEP, have a built-in assumption that students are reading fluently (Carpenter &

Paris, 2005). As a result, for students with non-fluent decoding skills, assessments of reading comprehension fail to measure true comprehension.

Nonetheless, measures of reading comprehension in the early grades draw differently on a variety of higher-order comprehension processes (Kendeou et al., 2012). In addition, literacy instruction focused on the development of higher-order reading skills in the early grades supports students' reading comprehension success in upper elementary, middle, and high school (Cain et al., 2001, 2004; Lynch et al., 2008; Oakhill & Cain, 2012; Oakhill et al., 2019; Tompkins et al., 2013; van den Broek et al., 2017), even if such skills are not detectable in early reading comprehension assessments. Early interventions on higher-order variables may be of particular benefit because of their developmental importance over time (Catts et al., 2016; Foorman et al., 2018). ToM, the variable examined throughout this dissertation, presents as one such higher-order skill that is malleable in response to intervention and influential to comprehension (Dore et al., 2018; Hofmann et al., 2016; Kim, 2020).

Dissertation Purpose

Extensions and critical examinations of current reading comprehension theories can help us better understand how higher-order reading skills, such as ToM, contribute to deeper reading comprehension (Graesser, 2015; Kendeou & O'Brien, 2018; Snow, 2018). Perhaps the most well-known and documented theory is the Simple View of Reading (SVR), which names decoding and listening comprehension as the only explanatory variables of reading comprehension (Hoover & Gough, 1990). The SVR has proven robust in its vast application over the last three decades, but its simplicity has limitations. Mainly, listening comprehension and decoding are not independent processes as the original SVR proposed (Hoover & Gough, 1986). Additionally, many skills influence both domains, which accounts for the substantial shared

variance between them (Lonigan et al. 2018), and both domains contain numerous sub-skills (Tunmer & Chapman, 2012).

The outward simplicity of the SVR framework further veils the interconnectedness between the two domains. Tunmer & Chapman (2012) and Cartwright et al. (2019) have shown the listening comprehension domain directly influences the decoding domain and also influences reading comprehension through decoding, making the relationship between the two domains more complex than originally proposed (see Hoover & Gough, 1986). Further, ample evidence supports the idea that there is a gradual shift in the proportion of variance explained by the two main components of the SVR as students move from novice to proficient readers (Adlof et al., 2010; Catts et al., 2003; Foorman, et al., 2018; Hjetland et al., 2019; Kendeou, et al., 2009; Language and Reading Research Consortium [LRRC] & Chiu, 2018; Peng et al., 2019; Stanley et al., 2018; Tilstra et al., 2009). The listening comprehension domain grows in importance as students continue to develop higher-order listening comprehension subskills, while the subskills of the decoding domain reach ceiling effects as students reach mastery (Paris, 2005),

In part to address the complexity of the linguistic comprehension domain, one avenue of research focuses on examining reading comprehension models inclusive of SVR components, but with a greater emphasis on delineating the higher-order cognitive and linguistic skills subskills of the listening comprehension domain (Oakhill & Cain, 2007; Cromely & Azevedo, 2007; Francis et al., 2018; Joshi & Aaron, 2000). The specific role of socio-cognitive skills, such as ToM, in reading comprehension is an understudied but burgeoning area of research within this larger corpus of literature (Atkinson et al., 2017; Kim, 2017, 2019, 2020; McIntyre et al., 2018; Ricketts et al., 2013; Strasser & Del Rio, 2013; Tong et al., 2019). A synthesis of studies that examined the relation between ToM and reading comprehension

indicates the presence of a small relation (β = -0.17 to β = 0.45 with a median of 0.20), which is mediated by linguistic comprehension (Kelly et al., in progress). This finding provides some evidence that ToM may be a subskill of the linguistic comprehension domain in the SVR that deserves consideration as a potential novel driver of reading comprehension.

Furthermore, efforts to provide clarity on the subcomponents subsumed within SVR's linguistic comprehension domain are both theoretically and practically driven. In addition to informing theories and models of comprehension, such efforts may also reveal underlying causes of reading comprehension-specific deficits purported to arise from underlying deficits in linguistic comprehension later in development (Longian et al., 2018; Petscher et al., 2018). To these ends, the overall purpose of my dissertation is to investigate ToM's role in reading comprehension theory and instruction in order to translate this burgeoning research in practice.

I adopt the view that ToM is a skill subsumed within the linguistic comprehension domain of the SVR (Kim, 2017) and a skill of inference making (Dore et al., 2018). I aim to specify *how* a socio-cognitive inference-making process, such as ToM, supports the *process* of reading comprehension. That is, how ToM skill functions as a skill of inference making during reading. To achieve my aim, I 1) evaluated how preschool teachers use ToM language (defined as references to characters' mental states) to provide comprehension instruction during classroom read aloud activities (Study 1, Chapter 3) and 2) designed, implemented, and evaluated a read aloud listening comprehension intervention for students aged 8–9 informed by findings from ToM and inference training literature (Study 2, Chapter 4). The present dissertation makes unique contributions to the literature. First, it informs the research field by more precisely identifying the role ToM plays in reading comprehension and early grade literacy

instruction. Second, it examines the utility of ToM as a way to implement a comprehension-focused read aloud intervention. Finally, it identifies areas for future research.

Outline of the Dissertation Manuscript

Although the reading comprehension domain is bound to printed text, comprehension occurs across mediums, including in the absence or presence of a visual stimulus, like television, or in spoken discourse, such as oral storytelling (i.e., Discourse Comprehension; Graesser et al., 1997). Therefore, throughout the present dissertation, I use the term comprehension when referring to contexts that can apply across reading and listening domains. For example, ToM is a general comprehension skill because it supports both reading and listening comprehension. I use reading comprehension, listening comprehension, and story or narrative comprehension when such differences are needed.

In Chapter 1 (the current chapter), I provide a statement of the problem that my dissertation addresses and present the theoretical framework for the entire work. Chapter 2 is a synthesis of literature where I examine the relation between ToM and reading comprehension. The research questions addressed include: (1) To what degree does ToM correlate with reading comprehension, and is there a direct effect between ToM and reading comprehension? and (2) Under what conditions, and for whom, is the relation between ToM and reading comprehension strongest? Chapter 2 represents the first synthesis to examine the relation between ToM and reading comprehension, and the findings provide clarity on the relation between ToM and reading comprehension and help inform further research efforts.

Chapter 3 contains a descriptive mixed method study that addresses gaps in the literature by examining to what extent teachers use mental state language during the interactive read aloud instruction (IRA), to what degree their use varies by mental state type, and how their use

supports comprehension instruction. IRA is a common instructional practice of shared reading between teacher and student(s) in early childhood and elementary settings, which consists of the classroom teacher reading aloud a text with students in either a whole or small group setting (Lennox, 2013). The interactive modifier specifies that an IRA purposely incorporates dialogic interaction between the teacher and student (Wiseman, 2011).

There is a critical period for ToM development between the ages of three- and five-years-old as this is when first order ToM develops (Wimmer & Perner, 1983). These ages correspond with early childhood classroom settings. Literacy practices that support higher order skills of comprehension, such as IRAs, are prevalent during these grades (Pentimonti et al., 2011) and also support ToM development (Martucci, 2016). Teachers already utilize mental-state language during IRA instruction (Ziv et al., 2014; 2015), but there is scant, if any, research examining *how* teachers use such language in the context of comprehension instruction during an IRA. The study described in Chapter 3 aims to answer the following research questions: (1) With what frequency do preschool teachers use mental state language during shared reading instruction? and (2) To what degree do different types of mental state language occur during shared reading instruction? These first two research questions provided the data needed to address my third primary research question, (3) How, if at all, does the mental state language use support ToM and comprehension development? The findings from the study more precisely identify ToM's role in comprehension instruction by identifying ToM in teacher's current practice and better inform future reading comprehension intervention efforts using ToM.

Chapter 4 includes an intervention study where I examined the extent to which a listening comprehension intervention, Theme and Inference Making for Early Readers (TIER), affects students' listening comprehension, as measured by the deep comprehension strategy of theme

identification. TIER uses the inherent mental state language of stories to guide ToM structured inference making during a dialogic read aloud to improve the listening comprehension outcomes of participants aged eight- to nine-years-old who are less-skilled comprehenders. The research question is 1) To what extent does read aloud instruction inclusive of ToM structured inference making improve the listening comprehension of students aged eight- to nine-years-old who are less skilled in comprehension, as measured by their theme identification? The findings identify the utility of using ToM in future early reading comprehension intervention efforts.

During the time ToM is developing, students' basic decoding skills are also developing, which presents a challenge when designing reading comprehension interventions (see LaBerge and Samuels' [1974] theory of information processing). In order to best use ToM for reading comprehension intervention, students must have a foundational set of decoding skills. I address this constraint by including participants in Study 2 who are in Grades 2 and 3, where most students' basic decoding skills and fluency are well-enough intact to support comprehension processes during reading (see Good & Kaminski, 2002).

Second, Kelly et al.'s (in progress) synthesis suggests the relation between ToM and reading comprehension is mediated by listening comprehension. Therefore, my intervention specifically targets listening comprehension. I conducted the intervention by reading the texts to the students and eliciting all student responses orally. However, I still included a standardized measure of reading comprehension as a pre-assessment to identify students less-skilled in comprehension (< 50th percentile). Kelly et al.'s synthesis also suggests that second order ToM may more strongly predict reading comprehension outcomes than a measure of first order or unitary ToM. The development of second order ToM at this stage is maturing in most students,

and variance in development is likely detectable, providing a favorable time for intervention (Perner & Wimmer, 1985).

Theoretical Framework: Reading Comprehension and Theory of Mind

There are numerous theories and models of comprehension that shed light on the skills, processes, strategies, and variables that impact reading comprehension. The SVR represents a product approach to reading comprehension, establishing reading comprehension as the result, or product, of decoding and listening comprehension (shown as $RC = D \; x \; LC$; Hoover & Gough, 1990). Struggling readers represent those with deficits in the decoding domain, the linguistic comprehension domain, or in both domains (Gough & Tunmer, 1986). As such, reading comprehension instruction and intervention should aim to address skills of the decoding or linguistic comprehension domains in order to have a direct or mediated effect on reading comprehension. The skill of interest in the present dissertation , ToM, is subsumed under linguistic comprehension domain (see Kim, 2020). The study described in Chapter 3 builds upon previous research that documents how teachers' comprehension instruction supports ToM (Lecce et al., 2021; Martucci, 2016; Ziv et al., 2014; 2015) to suggest that the use of ToM language supports, not only the development of ToM, but also students' comprehension.

To investigate the application of the empirical findings from Chapter 3, I designed TIER, the intervention reported in Chapter 4. TIER uses theme identification. a deep comprehension strategy, as the dependent variable outcome measure. Theme identification is a skill of deep comprehension because it requires comprehension beyond basic recall. Theme identification refers to a reader's ability to connect what the story is about to a larger universal concept beyond the text (i.e., the lesson or moral of a story; van Dijk & Kintsch, 1983). Typically, readers identify a story's theme after reading by generating an inference that links their background

knowledge of universal world concepts with the story's overall gist. Theme identification, therefore, represents the product of reading comprehension (i.e., how a student shows their comprehension a text or story), and inference making represents the process by which reading comprehension occurs (i.e., how a student constructs and extracts meaning from the text as they read). In the case of narrative comprehension, socio-cognitive inference making supports comprehension. In the present dissertation, I situate ToM as a proxy measure for such socio-cognitive inference making.

Models of reading comprehension that include a more detailed process orientation are more instructionally informative than the simplicity of the SVR. For example, The National Reading Panel and the RAND report presents a wider product view of reading comprehension than the SVR. The RAND report defines reading comprehension as the process by which one makes meaning from text and describes reading comprehension as a complex, multi-component process comprising an array of cognitive and phonological variable interactions influenced by reader, text, and activity characteristics (Snow, 2002). The consideration of the interaction effects across these three variables helps inform the more precise identification of under what conditions and for whom comprehension skills affect reading comprehension outcomes (Francis et al., 2018, Joshi & Aaron, 2000). Therefore, the intervention I propose in Chapter 4 focuses on less-skilled comprehenders, comprehending narrative texts, and completing the comprehension strategy of theme identification.

The Construction-Integration Model of Reading Comprehension and Theme

In contrast to both the SVR and RAND report, models are fully processed orientated models. When considering higher-order reading comprehension skills, such as ToM, from an intervention perspective, a process view of reading comprehension is more informative than a

product view (Rapp et al., 2007) because it is essential to understand how reading comprehension processes aid the reader in constructing and extracting meaning as they read the text. In other words, we have to ask, "How do good readers identify theme?' in order to understand why and where the process breaks down (McNamara & Kendeou, 2017). For this, I turn to online models of reading comprehension.

Online models of reading comprehension describe what occurs in a reader's mind while engaged in reading. The most influential model of online reading comprehension is Kintsch's Construction-Integration (C-I) Model (Kintsch, 1988). From the C-I Model view, as a reader engages with and moves through the text, a cohesive mental representation of what the text is about gradually emerges. Cohesiveness is built from the reader's ability to identify the saliency of the novel text input with the ongoing mental representation and appropriately integrate or update the content by amending or adding to the evolving mental representation (Kendeou et al., 2013; Pefetti & Stafura, 2014; van den Broek et al., 2005). The integration or addition of novel text input is achieved through inference making due to texts' implicit nature (Kendeou, 2015). By supporting the mental representation formation process, inference making builds cohesion by creating connections between distinct textual elements to form an overall understanding of the text's structure (Kintsch & van Dijk, 1978; van den Broek et al., 2015). More cohesive mental representations illustrate stronger comprehension than less cohesive mental representations.

The understanding of the overall narrative structure refers to the comprehension of the text's macrostructure (Kintsch & van Dijk, 1978). Macrostructure comprehension signifies a reader's ability to break down the text into meaningful propositional parts and to its general gist. In the context of narrative comprehension, the general gist is what is most important for theme

identification because it provides what the story is generally about (Kintsch, 1988). Still, although general gist identification supports theme identification, it is ultimately insufficient if the reader lacks background knowledge because of the requirement that sufficient theme identification includes a connection to a broader universal concept (McCarthy et al., 2019; Lehr, 1988; Williams et al., 2002). In the present dissertation, I situate ToM as a specific socio-cognitive inference-making skill that assists theme identification in two ways. First, ToM supports the comprehender in understanding the overall structure of the text through online inference making that coheres the causal relations between a goal-centered protagonist, the central problem, and the central problem's resolution (macrostructure comprehension; Graesser et al., 1994; Trabasso & van den Broek, 1985; Trabasso & Wiley, 2005; Trabasso et al., 1989; van den Broek, 1990; 1994; Zwaan, 2005). Second, ToM also provides the background knowledge needed to connect macrostructure comprehension to broader universal concepts, such as themes of love, courage, and redemption (theme identification; Lehr, 1988; McCarthy et al., 2019; Williams et al., 2002).

Inference Making. An inference is the end result of knowledge activation and integration (Kendeou, 2015). Knowledge, either from the text or from a reader's background, is activated during reading and then subsequently integrated into the readers ongoing mental representation of what the text is about. In general, the types of inferences readers make to form a cohesive mental representation describe local and global cohesion process which are either text- or knowledge-based (Cain & Oakhill, 1999; Graesser et al., 1994; Gygax et al., 2007; Kispal, 2008; van den Broek et al., 2015). Local inferences cohere text elements at the sentence level. Specifically, the ubiquity of propositions and related referents in texts necessitate the generation of local inferences to comprehend anaphoric relations (i.e., the relation between a

subject and a pronoun) between and across sentences (Kendeou, 2015; van den Borek, 2015; Cain & Oakhill, 1999). For example, if a text reads, “Sally wasn’t feeling well. She put the kettle on to make some tea.” A reader must infer the pronoun “She” in the second sentence relates to the subject Sally in the previous sentence. This shows a local text-based inference because the text supplies the information needed to relate the subject with the appropriate pronoun.

As a reader moves through a text, the information garnered from the text is stored in working memory. The information available to the reader to generate a local text-based inference is supplied from such textual information stored in working memory (Cook & O’Brien, 2014). In the above example, the name Sally is stored in working memory as a character within the story. The stored information within working memory fluctuates in activation in response to novel text input. As soon as the reader reads the word “She” in the subsequent sentence, the stored information about the character Sally activates through its resonance with the name Sally. The generation of the inference needed to understand the word is automatically made by the reader, and connecting “She” to Sally is effortless (Cook & O’Brien, 2017).

If the local text input does not automatically activate information within working memory, the inference becomes a global text-based inference. Global text-based inferences cohere text input more broadly across temporal text events. The reactivation of prior information from the text stored in working memory supplies the needed information to generate the appropriate inference (Cook & O’Brien, 2014). For example, prior text input may have indicated that Sally’s mom always made her tea when she felt ill as a child. When confronted with the text input, “Sally wasn’t feeling well. She put the kettle on to make some tea.” a reader would generate an inference that coheres the stored information about the action of Sally’s mom with Sally’s act of making tea.

Knowledge-based inferences, on the other hand, assist readers in understanding more implicit meanings within the text. Readers make knowledge-based inferences when information from the text stored in working memory is unavailable or insufficient (Cain et al., 2004, Linderholm, 2002). The information needed for the reader to generate the inference is instead supplied by the reader's recruitment of relevant background knowledge (Bohn-Gettler & Kendeou, 2014). In the previous example with Sally, "Sally wasn't feeling well. She put the kettle on to make some tea." a reader would need the knowledge that hot tea often helps someone feel better when sick, that to "put a kettle on" means to boil water for tea, and that tea is made from steeping herbs in hot water. This set of inferences refers to local knowledge-based inferences.

Knowledge-based inferences are where we see ToM come into play. In addition to this set of knowledge-based inferences, another local knowledge-based inference is made that connects Sally's want (mental state) to her subsequent behavior of making tea to achieve her intention of feeling better. This inference is made to establish causality between Sally's mental state and subsequent behavior, which is of central importance to narrative texts that reflect strong causal structure (Kintsch, 1995; Lynch et al., 2008; Trabasso & van den Broek, 1985; Trabasso et al., 1989; van den Broek, 1990; van den Broek et al., 2013).

Global knowledge-based inference making refers to the deep comprehension reading strategy of identifying the theme of a story (Kispal, 2008; Shanahan et al., 2010). For example, common narrative themes represent universal concepts such as happiness, hope, grief, or friendship. To identify a theme of happiness requires ToM; a reader must first attribute the mental state of "happy" to character(s) within the story, and perhaps other mental state feelings as they read, and then extract the general gist of happiness from the story's context to create a

universal statement about what that the story teaches about happiness. See Chapter 1 Table 2 for further inference type examples based on the classic fable of "Little Red Riding Hood."

Prominence of Reader Generated Knowledge-Based Inferences

The demarcation between text-based and knowledge-based inferences is somewhat artificial. Effectively, all inferences are knowledge-based; what differs is the derivation of the knowledge used to generate the inference. Is the knowledge derived from the text, a reader's background knowledge, or both? The answer to this question is likely both given that some transference of text-based knowledge occurs while reading (Kintsch, 1994). Therefore, the amount of text-generated and reader-generated knowledge activated and integrated is a scalable continuum rather than a binary (Graesser & Kruez, 1993). The importance of reader-generated knowledge in particular, however, is important to consider for four main reasons. First, texts vary in cohesiveness; authors routinely do not provide all the information necessary to fully comprehend the text, rendering texts inherently implicit and inference making using background knowledge necessary for reading comprehension (Kendeou, 2015).

Second, a reader's working-memory capacity plays a role in determining the degree to which a reader must use their background knowledge to successfully make a knowledge-based inference in order to build a cohesive mental representation of the text (Bohn-Gettler & Kendeou, 2014; Currie & Cain, 2015). Readers with low working-memory capacities become more dependent on the generation of background knowledge-based inferences due to the limited amount of information stored in their working memory (see Hoover & Gough, 1990; Stanovich, 1980). In effect, text-based global inferences (the upper right quadrant of Chapter 1, Table 2) become background knowledge-based inferences when readers cannot access sufficient textual knowledge from working memory in order to effectively integrate the novel text

input (Bohn-Gettler & Kendeou, 2014, Currie & Cain, 2015). In the Sally example from the previous section, a reader may not have access to the text information about Sally's mom always making her tea whenever she felt ill. A reader would then fall back on their prior knowledge that tea is often used to help people feel better when ill, and Sally wants to feel better, which explains why she is making tea. However, although this inference generation serves local cohesion, the overall text cohesion is reduced because of the missing information about Sally's mom.

Third, another consideration is that this strategic use of background knowledge only occurs if the lack of textual information stored in working memory does not meet the standards of text coherence determined by the reader (van den Broek et al., 2011). Readers naturally strive for coherence while reading. Each reader comes to a text with a certain set of standards of coherence that are internally generated or externally placed upon the reader (van den Broek et al., 1995, 2001, 2011). As a result, readers vary in their standards of coherence based on various factors, such as purpose for reading (e.g., skimming vs. learning) and reader's comprehension monitoring skill (van den Broek et al., 1995, 2001). Readers with a purpose of skimming the text have lower standards for coherence because their purpose for reading is less dependent on a thorough understanding of textual details. Readers with weaker comprehension monitoring skills have lower standards for coherence due to weaknesses in identifying areas of the text that are incongruent or in identifying their own misunderstanding. Standards of coherence thus reflect both reader-level and text-level characteristics that impact reading comprehension by influencing to what degree readers make inferences and what kind of inferences readers make.

Fourth and finally, background knowledge-based inferences are necessary for readers to achieve the situation level representative of deep comprehension (Kintsch, 1988). The connections readers make between the text and their background knowledge is defined as the

formation of a situational model of comprehension (van Dijk & Kintsch, 1978, 1983). In order for a reader to achieve a situational model level understanding of what the text is about, a reader must form connections between their prior knowledge and the new information learned from the text (McCarthy, 2015, McCarthy & Goldman, 2015). The formation of a situational model is the end product of the active comprehension processes involved in the online formation of a mental representation of the text (Kintsch, 1988). Therefore, theme identification becomes a strategy of deep comprehension supported by higher-order reading skills because it represents a reader's ability to use their comprehension and background knowledge to make connections beyond the text to universal concepts that apply across texts (Graesser, 2015).

Background Knowledge and Narrative Comprehension

Beyond the C-I Model, background knowledge is consistently identified across the literature as critical to comprehension (McNamara & Magliano, 2009; Cook & O'Brien, 2014). Background knowledge may refer to two distinct types used across the literature: domain and topic-specific background knowledge. Domain background knowledge refers to general knowledge about a particular content, and topic-specific background knowledge refers to knowledge of individual topics within contents (McCarthy & Goldman, 2019). For example, a reader may have a general understanding of the content domain of science, but lack knowledge about the specific scientific topic of physics.

Dore et al. (2018) posit ToM is a "hidden" variable in reading comprehension because it supports reading comprehension through inference making. What is less clear is *how* ToM functions to support inferencing. Dore et al. present numerous alignments with seminal reading comprehension theories that generally support two ways in which ToM serves reading comprehension via inference making: 1) ToM serves as a function of knowledge that supports

meaning-making processes from text by supporting the comprehension of a character's mental states, or 2) ToM represents an inferential ability to engage in abstract thought around a character's mental states.

The parallels between ToM and inference making as both skills of abstract thought are a product of Piagetian child development, where a child develops an ability to understand the potential for falsehood in one's mental representations (Flavell, 1963; Gopnik & Astington, 1988). Prior research suggests that, once readers develop beginning ToM, they adopt the mental perspective of the main character, and this skill increases across development (Fecica & O'Neill, 2010; see also Dore et al., 2018). Further, readers who are better able to adopt characters' perspectives exhibit better comprehension (Emery, 1996; Hodges et al., 2018). Characters' mental states within narratives follows a predictable causal narrative structure (Lynch et al., 2008; Trabasso, 1985, 1989; Trabasso & Wiley, 2005; van den Broek et al., 1999; Zwaan, 1995). ToM is important because it refers to the skill of identifying others as intentional actors (Tomasello & Rakoczy, 2003), meaning the behavior of others serves a causal purpose, exhibiting a contingency between displays of behavior prompted by mental states. In the case with Sally, she wanted (mental state) to feel better, so she made tea (behavior). Readers who infer these causal relations between characters' mental states and succeeding behaviors exhibit better understanding of narrative structure via tests of narrative recall (Astington, 1990; Curenton, 2004, 2011; Pelletier & Astington, 2004; Riggo & Cassidy, 2009; Ronfard & Harris, 2014).

One could argue that this occurrence is a general inferential ability related to general reasoning (Frye et al., 1995; Guajardo & Turley-Ames, 2004; Guajardo et al., 2009). Yet, recent evidence suggests socio-cognitive processing (i.e., a group of processes inclusive of ToM,

perspective-taking, and empathy) continues to advance into adulthood and contributes to narrative recall over and above causal understanding of the texts across childhood, adolescence, and adulthood (Pavias et al., 2016). Such a finding suggests ToM plays a unique socio-cognitive inferential role in reading comprehension distinct from general reasoning.

Conceptualizing ToM in this way is important because, even if a reader has the social knowledge needed to accurately generate an inference, prior research indicates they may not always use it, or may use it inaccurately, making knowledge and integration related but separate constructs (Cain et al., 2001; Oakhill & Cain, 2007). Naming ToM as a measure of social knowledge that encapsulates both a knowledge component (mental state attribution) and an integration component (predicting and explaining behaviors) can help identify errors in comprehension regardless of error type. Therefore, in the present dissertation, ToM serves as a measurable proxy for the socio-cognitive inference making needed for narrative comprehension in line with recent frameworks of reading comprehension put forth by Dore et al. (2018) and Duke and Cartwright (2018) that identify that the contribution of ToM to reading comprehension is in the generation of inferences in response to characters' mental states.

My assertion is constrained, however, first to domain background knowledge because topic specific background knowledge is still important for narrative comprehension. For example, when reading a sports narrative about baseball, students need to know what baseball is and how the game is played. Second, vocabulary is crucial to inference making and shares much in common with background knowledge (Cain & Oakhill, 2014; Cain et al., 2001, 2004; Currie & Cain, 2015; LRRC et al., 2019; Oakhill et al., 2015). It is thus not enough for readers to develop ToM, but they must also have the language of ToM to show their knowledge through the enactment of deep comprehension strategies such as theme identification. If a reader cannot

attribute the appropriate mental state to a character within a story from implicit information in the text or does not understand the explicit mental state reference, their ToM knowledge-based inferences will not aid comprehension performance, even if they understand the character as an intentional actor guided by their mental states. Hence, mental state vocabulary knowledge (think, believe, want, intend, disappoint, hope, etc.) needs to be developed alongside ToM.

Despite these constraints, my assertion fills a literary lacuna because, notwithstanding a wealth of research on background knowledge and expository text comprehension, what knowledge-based inference making means in the context of narrative comprehension remains significantly less explored. There appears to be a convergence of research that asserts narrative texts in general are easier to understand than expository (Best et al., 2008; Diakidoy, et al., 2005; Sáenz & Fuchs, 2002). Narratives represent everyday life, and, therefore, it is generally assumed narratives do not require specialized content-knowledge or academic vocabulary (Graesser et al., 2011). Aside from their relative easiness, narratives do, however, present genre-specific difficulty for some reader populations. Specifically, meta-analytic findings support the presence of narrative comprehension specific deficits in children with autism, a population routinely identified with atypical ToM development (Baron-Cohen et al., 1985) that likely stems from their lack of social knowledge (Brown et al., 2013; Loukusa & Moilanen, 2009; McIntyre et al., 2017). In fact, emerging evidence suggests a direct relation between ToM and reading comprehension exists in children with autism over and above the two main components of the SVR (McIntyre et al., 2018; Tong et al., 2019; Ricketts et al., 2013). Further evidence suggests children with autism even prefer expository texts over narrative (Davidson & Weismer, 2018).

Moreover, the Coh-Metrix tool identifies genre specific differences between narrative and expository texts; narratives are generally less cohesive than expository texts (Graesser et

al., 2004, 2011). The lack of cohesiveness for narrative texts underscores the importance of background knowledge-based inferences for narrative comprehension, especially at deeper levels of comprehension (Eason et al., 2012). Such differences may result in comprehenders with delayed or atypical ToM development who may experience genre specific difficulty due to generating fewer inferences, which culminates in a fractured understanding of the narrative structure and less coherent mental representation overall. Without ToM, then, the achievement of a situational level understanding of what a narrative text is about and the enactment of strategies of deep comprehension such as theme is not likely to occur (Kintsch, 1988).

As a result, narratives may indeed require specialized content-knowledge that we can define as ToM and specialized vocabulary we can define as mental state language (see Cartwright, 2015). Given the contemporaneous emergence of ToM, inference making, and narrative comprehension at a formative time in child development (Dore et al., 2018), children likely develop a schema for narratives that assists them in comprehending stories (Rumelhart, 1980). It stands to reason that in children with delayed or atypical ToM development, narrative schemas may be less robust, if not non-existent, and thus may explain narrative genre specific difficulty and aversion.

Potential Contribution

The identification of ToM as a needed specialized skill of socio-cognitive inference making for narrative comprehension leads to important considerations for reading comprehension theory and instruction. Narratives are the main mode of engagement for pre-, emergent, and beginning readers (Lynch et al., 2008; Pentimonti, 2011). Weak narrative comprehension resulting from delayed or atypical ToM development could result in the prevalence of two main consequences for reading comprehension. First, as described above, such

readers may suffer from decreased inference-making skill due to a general lack of inference generation. Second, if a reader has a sufficient level of knowledge to access a text, the relation between knowledge and reading comprehension becomes bi-directional (O'Riley et al., 2019). Readers with delayed or atypical ToM development may experience a Matthew Effect (see Stanovich, 1986), whereby the gap in overall knowledge between readers with strong ToM and readers with weaker ToM becomes larger due to missing out on crucial learning from texts, further attenuating reading comprehension (Cain & Oakhill, 2011). These occurrences, over time, could result in cascading effects across students' lifespan with delayed or atypical ToM development.

In sum, proficient narrative text comprehension during literacy's pre-, early, and beginning stages lays a foundation for the comprehension of more complex narratives and content area texts in later elementary, secondary, and post-secondary settings. Adept inference-making skills and sufficient background knowledge for narrative comprehension—defined as ToM—can serve as an informative heuristic for both necessary skills in research and practice. Thus, it is important to frame ToM as a proxy for socio-cognitive inference-making for narrative comprehension with beginning readers. In doing so, comprehension instruction during the pre-, early, and beginning stages of reading may better meet all students' comprehension needs.

Chapter 2: What do We Know About Theory of Mind and Reading Comprehension and Where Do We Go from Here?

Reading comprehension is most succinctly described as the product of linguistic comprehension and decoding (the Simple View of Reading [SVR]; Hoover & Gough, 1990). Over the past thirty years, the SVR has proven robust, but its simplicity can be limiting. (Duke et al., 2021). In particular, although the decoding and linguistic comprehension domains contain multiple subcomponents, linguistic comprehension is acutely complex and, unlike the decoding domain, is less constrained by a developmental ceiling (Lerväg et al., 2018; Tunmer & Chapman, 2012). The influences of both domains on reading comprehension vary based on development and change based on interactions between the text, reader, and task (Foorman et al., 2018; Francis et al., 2018). Thus, current research emphasizes the importance of novel, higher-order subcomponents subsumed within the linguistic comprehension domain of the SVR in order to better understand their influence on reading comprehension (Petscher et al., 2020).

Researchers examining such higher-order variables of comprehension sometimes aim to extend the SVR by more precisely presenting the makeup of subcomponents subsumed within the linguistic comprehension domain. Some of the typical variables included in such models are executive functions (EFs), comprehension monitoring, inference making, and background knowledge (see, for example, Cromley & Azevedo, 2007). A lesser-studied but burgeoning subcomponent of linguistic comprehension is the socio-cognitive variable, Theory of Mind (ToM). Given the proliferation of ToM's emergence in education research broadly, and in reading comprehension specifically, my goal in this chapter is to support the field's continuing interest in ToM by synthesizing and critiquing ToM andreading comprehension literature corpus from 1983 to 2018.

Theory of Mind Defined

ToM falls under the purview of social cognition—a broad term encompassing a handful of constructs that attempt to describe one's ability to construct meaning about the social world (Fiske & Taylor, 2013). ToM, also called mentalizing, refers to one's ability to accurately attribute mental states to oneself and others (i.e., what others think, believe, want, intend, feel) and to use such attributions to reason about one's own—and predict others'—social behaviors (Wellman, 2002). Children's acquisition of ToM is a developmental process that details an increased understanding of mental states (Wellman & Liu, 2004; Wellman et al., 2001). Prior research divides ToM development generally into three categories of increasing complexity: first order, second order, and advanced. Children develop first order ToM between the ages of three and five years by recognizing that others can experience different mental states than their own (Wimmer & Perner, 1983). Next, between ages six and eight years, second order ToM develops, which refers to the skill of identifying that others can think something about someone else's thoughts. (e.g., Jane *thinks* that John *thinks* the cake is in the break room; Perner & Wimmer, 1985). Later, in middle childhood and adolescence, more advanced ToM understanding, such as social reasoning, recognizing transgressions of social norms, and reasoning about ambiguity (i.e., irony or white lies) begins to emerge (Osterhaus et al., 2016). Some scholars posit competence with such advanced ToM development continues into adulthood (Apperly, 2010).

ToM development is also broken down into stages by type of mental state understanding. The ToM scale (Wellman & Liu, 2004) establishes the development of ToM understanding in reference to increasingly complex mental state categories. Although the evidence affirms that some degree of ToM develops in all persons, cultural variation within the ToM scale sequence

does appear to exist, mainly between European American and Non-Western cultures (Wellman et al., 2006, 2011). For example, Chinese children appear to develop an earlier understanding of knowledge vs. ignorance in contrast to Western (American and Australian) children who develop an earlier understanding of desire versus belief (Wellman et al., 2006).

Based on the ToM scale (Wellman & Liu, 2004), first, children begin with developing an understanding of desire (e.g., the mental state of *to want*). Next, children come to understand belief mental states (e.g., *to think* or *believe*). Belief mental states are unique in that they are epistemic with the propensity to be false. The understanding of belief states develops in two chronological developmental components respective of the propensity for falsehood, an understanding of diverse beliefs (when both the child and actor are ignorant of what is true) and an understanding of false beliefs (when the actor does not know what belief is true, but the child does).

The last type of mental state understanding to develop based on the ToM scale is an understanding of emotion (Wellman & Liu, 2004). The understanding of emotional mental states (e.g., happy, sad, excited) parallels that of epistemic mental states in that children develop an understanding of real emotions (i.e., actual felt emotion) before they understand apparent emotion (i.e., what is displayed). In other words, children develop an understanding of emotion by first recognizing appropriate emotional responses and only later do they develop an understanding that sometimes the expression of emotion may not match the actual felt emotion.

The emotion understanding component of ToM is often referred to as affective ToM (opposed to cognitive ToM). A less-prevalent measure of ToM, which has been used in reading comprehension research, is the Reading the Minds in the Eyes Test (RMET; Baron-Cohen, 2001). The RMET is a measure of ToM that predominantly assesses affective ToM. The

RMET presents images of individuals' eyes to assess how well participants can discern the mental state of the person depicted in the images. In the child version of the test, participants are shown 28 black-and-white pictures of eyes and given four possible answer choices (e.g., angry, surprised, hate, relaxed). The participant is asked to identify which emotion is shown in the picture. The RMET was originally designed for and is used most often to identify the ToM skill in children and adults with autism.

Measures of Advanced ToM typically use scenario-based, task-like false-belief tasks that assess more advanced epistemic states. The most common advanced ToM assessment is Strange Stories (Happé, 1994). Strange Stories is a set of written narratives either read aloud by the assessor or read by the participant with corresponding questions about the characters' mental states and comprehension control question(s). There is a video version of the Strange Stories task, *The Silent Films Task* (Devine & Hughes, 2013). As with the RMET, advanced measures of ToM were originally designed to assess social cognition competence in clinical samples of students with autism.

There is a multitude of ToM measures (Beaudoin et al., 2020). I highlighted the RMET because it is the most popular affective measure of ToM, and Strange Stories because it is the most popular advanced measure within the research on ToM and reading comprehension. Further, because both the RMET and Strange Stories were originally designed to distinguish between typical and clinical populations and in adult samples, their reliability and validity with typically-developing children and adolescents is debated (Hayward & Homer, 2017; Hutchins et al., 2012). Relatedly, still debated is whether it is appropriate to measure cognitive, affective, first order, second order, and advanced ToM distinctly or if it is more appropriate to measure ToM as a unitary construct (Warnell & Redclay, 2019).

A Focus on False Belief

False-belief tasks are the most common way research assesses ToM. A false belief is a belief held by an agent that is not congruent with reality (Wellman et al., 2001). Identifying false beliefs, when a child knows what is true, requires them to reason theoretically when considering another's false belief. Such false-belief tasks are the most commonly-used measure of ToM because they require accurate mental-state attribution for task proficiency. Wimmer and Perner (1983) first put a measure of false belief, the classic change in location task with Maxi, into action in order to measure ToM in children. This standard false-belief task about an unexpected location typically entails two actors and an object. Unaware to a first actor, the object is moved from one place to another by the second actor (i.e., from the blue cupboard to the red one). The child is then asked where the first actor will look for the object once the actor reenters the scene. It is here that a false belief occurs. The belief about where the object is for the actor is false if the child selects the red cupboard because the actor is ignorant of the switch from the blue cupboard to the red cupboard. Since Wimmer and Perner's pivotal study, false-belief measures are the dominant method with which ToM is measured. Consequently, false-belief understanding is considered a proxy for ToM across the literature in which it is used (Astington, 2003).

False-belief tasks present as a sound measure of ToM because the mental state reasoning for passing such tasks is more difficult (Wellman & Liu, 2004). The use of false-belief tasks to measure ToM is not without critique, however. (Bloom & German, 2000; Lewis et al., 2017; Quesque & Rossetti, 2020; Westra & Carruthers, 2017). Namely, the tacit contribution of linguistic competence in false-belief measures leads to questions regarding their use as a proxy measure for ToM. Language functions as a significant contributor to ToM

development (Astington & Baird, 2005; Milligan et al., 2007; Tompkins et al., 2019). Children in linguistically-rich environments develop ToM earlier than children in less linguistically-rich environments (Tompkins, 2015), and children with specific language impairments often display corresponding ToM deficits (Nilsson & de Lopez, 2016). Milligan et al. (2007) provided a meta-analysis to evaluate the relation between language and false belief understanding. Findings revealed language competency was a better predictor of ToM performance ($r = 0.56$) than ToM performance was of language competency ($r = 0.36$; Cohen et al., 2003). Milligan and colleagues' meta-analysis identified mental-state vocabulary knowledge and syntactic understanding, especially understanding sentential complements (e.g., Jane thinks *that* it will rain today), as unique contributors to performance on measures of false-belief understanding. Additionally, children with deficits in executive functions often display commensurate deficits in ToM when measured using false-belief tasks (Yeh et al., 2017).

Other findings from linguistics studies suggest that children who struggle with interpreting false-belief tasks may instead struggle with interpreting the pragmatic language demands of those tests rather than ToM specifically (Lewis et al., 2017; Westra & Carruthers, 2017). Pragmatic language refers to the ways in which children use and interpret language in different contexts and for specific purposes (Roberts, 2012). For example, a child may interpret the intended meaning of the following sentence, "Dan thinks it's cold outside." in two ways: 1) communicating information about the weather (a direct association), or 2) communicating about Dan's mental state about the weather (an indirect association). The latter is an indirect association because it cannot be seen and, therefore, must be inferred. In false-belief tasks, Dan's mental state is false (i.e., the weather is actually warm outside), making Dan's thought incongruent with reality. Therefore, if a child is asked if this sentence is true or false, and

the child interprets the intended meaning to be about the weather and not Dan's mental state, they do not pass the false-belief task. As such, children's interpretation of the pragmatic language within false-belief tasks may explain ToM outcomes as opposed to children's ToM understanding (Lewis et al., 2017; Westra & Carruthers, 2017).

These linguistic findings suggest children are likely to develop ToM understanding before they can perform false-belief tasks in part because of the lexical complexity of false-belief tasks. Curenton (2004) provides such an analysis in the interpretation of false-belief performance discrepancies between European American children and African American children through the relation between the oral narratives and false-belief task performance of these groups. African American children failed false-belief tasks at a higher rate than European American children. However, the narratives produced by African American children were much richer with mental-state language than European American children, suggesting the African American children had strong ToM skills, but the false-belief tasks did not capture this. Thus, false-belief tasks may lack construct validity as a measure of ToM, which should be considered as a limitation to research in this area.

Theoretical Support for the Importance of Theory of Mind to Reading Comprehension

The accumulation of theoretical support for ToM's role in reading comprehension stems from its importance to narrative comprehension (Dore et al., 2018). Bruner (1987) defines narrative comprehension as the coordination between two landscapes: action (plot events) and consciousness (mental states of characters). Astington (1990) contends ToM is the mechanism by which children make sense of the landscape of consciousness. Indeed, characters' mental states within narratives follow a predictable causal narrative structure (Lynch et al., 2007). The contemporaneous emergence of ToM, inference making, and narrative comprehension at a

formative time in child development (Dore et al., 2018) likely represents the development of a schema, or a story grammar, for narratives that assists in comprehending stories (Stein & Glenn, 1979). Prior research does suggest that once this development occurs, readers begin to adopt the mental perspective of the main character (Fecica & O'Neill, 2010). Comparisons between the ability of children, adolescents, and adults to attend to characters' mental states suggests this skill increases across development (Pavias et al., 2016).

The contribution of a child's socio-cognitive ToM development is important then because it signifies children's ability to see others, including characters within stories, as intentional actors (Tomasello & Rakoczy, 2003). Their behavior is logically causal, based on a contingency between displays of character behavior prompted by mental states. A character who is sick, for example, wants (mental state) to feel better, so they make themselves a cup of tea (behavior). Authors routinely do not explicitly state these ubiquitous causal relations within stories—if they did, literature would not be so enjoyable. Readers who infer these relations between characters' mental states and succeeding behaviors exhibit a better understanding of narrative structure via tests of narrative recall (Curenton, 2004, 2011; Pelletier & Astington, 2004; Riggo & Cassidy, 2009; Ronfard & Harris, 2014). The social demands of narrative comprehension, however, are more important to narrative recall than simply understanding causality itself (Pavias et al., 2016). In addition to recall, readers who are better able to adopt character perspectives exhibit better comprehension (Dray, 2018), and, importantly, this skill can be taught to struggling readers (Hodges et al., 2018).

As a result, the knowledge needed for comprehending narratives is inherently social in nature *and* must be inferred. Dore et al. (2018) thus posit that ToM is a "hidden" variable in reading comprehension because it supports reading comprehension vis-a-vis inference making.

In essence, a reader with greater social knowledge and the capacity to build such knowledge is better able to generate accurate inferences during narrative reading, which demands specialized (social) content knowledge supported by academic-specific vocabulary understanding (mental states; Cartwright, 2015). In Kim's Direct and Inferential Model of Reading (2017, 2019, 2020) ToM is a measurable way to capture social-inference-making processes, and, because of the link to inference making, ToM is purported in Kim's model under linguistic comprehension.

The Present Synthesis

Despite a comparative wealth of research on ToM's relation to reading comprehension, what is less clear from the literature is whether ToM's relation to reading comprehension offers a meaningful contributor to SVR models, and if and how the relation changes in light of variation in measurement and statistical modeling methods. It is also unknown whether ToM's relation to reading comprehension changes across certain populations and across genres. To my knowledge, no systematic review of the literature exists in this area. My purpose in this chapter is to present a synthesis of the literature that includes ToM in exploratory models of reading comprehension in order to provide clarity on the relation between ToM and reading comprehension under different conditions and with specific samples. To achieve this purpose, my primary research question is, to what degree does ToM correlate with reading comprehension? My secondary research question is whether there a direct effect between ToM and reading comprehension?

To answer these questions, I followed the Cooper and Hedges's protocol (2009) and the procedures set forth in the PRISMA guidelines (Moher et al., 2009. Both describe a systematic process to complete a research synthesis literature review. A description of the literature search, data evaluation, and data analysis in the context of the present synthesis is below.

Method

Search Procedures

I performed a systematic search of the literature to locate articles. Due to the difficulty in systemically accessing dissertations, conference presentations, and technical papers, and the nascent stage of the literature on the relation between ToM and reading comprehension, I excluded these documents from the systematic search and, thus, from inclusion in the synthesis. To begin, I reviewed the keywords across the ToM literature to establish search terms based on the wide use of ToM across literature bases. Next, I consulted with an expert on ToM and reading comprehension to ensure I identified a comprehensive list of appropriate search terms. Because the development of ToM spans pre- and emergent readers, constructs of reading comprehension in this synthesis included reading, story, and listening comprehension.

The final BOOLEAN all text search included the keywords: *theory of mind* OR *ToM* OR *mentalizing* OR *perspective-taking* OR *mental state* OR *social understanding* OR *false belief* OR *false belief understanding* OR *mental state language* OR *socio-cognitive* OR *social cognitive* OR *social cognition* OR *emotional inferences* OR *character perspective* OR *characters' perspective* AND *reading comprehension* OR *narrative comprehension OR text comprehension* OR *listening comprehension* OR *story comprehension* within the EBSCO Database utilizing ERIC, Academic Search Complete, and PyschINFO, resulting in a total of 226 articles.

To ensure the online search did not miss any articles, I then conducted hand searches of three journals that published two or more articles on the relation between ToM and reading comprehension. The three journals searched were: 1) the *Journal of Experimental Child Psychology*, 2) *Frontiers in Psychology*, and 3) *Reading Research Quarterly*. No additional articles met the inclusion criteria for use in this synthesis from the hand searches.

As a first screening, I reviewed the titles, keywords, and abstracts of the 226 identified articles to determine which articles warranted further review based on the following criteria: 1) Studies were published in English between January 1983 and May 2018. The rationale for the search limitation of articles after 1983 is because of the publication of Wimmer and Perner's (1983) seminal piece on ToM, and the rationale for May 2018 is the publication of Dore et al.'s (2018) narrative review of ToM and reading comprehension. 2) Studies were quantitative and empirical. I did not include qualitative work or case studies. 3) Studies did not use brain imaging. I excluded brain-imaging studies because the purpose of the present synthesis does not include a critical discussion of social neuroscience processes. Based on the first screening, 199 articles met the criteria for inclusion in this synthesis.

As a second screening, I reviewed the full text of the remaining 199 identified articles to determine which ones met the following criteria: 1) Samples included children in grades pre-K through twelve. The rationale for this limitation is because the population of interest was students in the pre-K–12 setting. 2) Studies included a measure of ToM and a measure of RC. Many studies indicated the use of general cognitive variables and measures of reading, but did not include specific measures of ToM, RC, or both. Based on the second screening, 13 articles met the criteria for inclusion in this synthesis.

For the third and final stage, I conducted an ancestral search by screening the citations of the 13 identified articles. Two additional studies met the inclusion criteria, making a total of 15 articles. See Chapter 2, Figure 1 for a PRISMA search flow diagram.

Coding

After completing the initial search, I developed a coding matrix for the data evaluation process based on preliminary data analysis. The final established codes fell into four categories: 1) study features, 2) measurement, 3) additional measured variables, and 4) analytic findings.

Study features consisted of codes for publication date, age and grade range, sample characteristics, design, and geographical location. Next, measurement codes included the ToM and reading comprehension measure, as well as the type of assessment used for both variables. I coded each study for ToM according to whether the assessments addressed first order, second order, or advanced ToM. Other measured variables were coded as well, such as measures of EF, linguistic comprehension, additional reading measures, and demographic variables. Lastly, for each study, correlation (r), including partial and semi-partial coefficient values (β); variance explained (r^2); reliability; and validity were analyzed to report on the strength of the relation between ToM and reading comprehension.

I had an outside reading researcher double code 30% of the articles. Coding reliability (α = 94%) was established using point by point agreement. Reliability was calculated by the number of agreements divided by the number of agreements plus disagreements. The two coders discussed all discrepancies in coding until 100% agreement was achieved.

In order to answer my research questions, I first evaluated the included studies by study features, including reported demographics, analysis type, and measures included in statistical models. Next, I evaluated the included articles' analytic findings. I created tables (Chapter 1, Table 1 and Chapter 2, Table 2) to capture and report the results of my evaluation.

Results

A total of 15 studies, reported in peer-reviewed journals between January 1983 and May 2018, met the criteria for inclusion in the present synthesis. All studies were published in

2011 or later. Thirteen total studies examined the relation between ToM and reading comprehension in the general population, one study examined the relation in a sample of students with autism, and the remaining study compared the relation in a sample of students with autism to a typically developing sample. I report the synthesis results by first detailing study features, and second by detailing the analytic findings.

Study Features

The details provided about demographic data varied across studies. All studies reported age and gender. Eight studies reported socio-economic status with a range of 12% to > 90% representing low social economic stratums. Six studies reported race or ethnicity, with three studies reporting results from a homogenous Korean population. Eight studies reported emerging bilingual participants with a range of 0% to approximately 60%. Two studies of a general population reported the presence of students with disabilities in the sample with a range of 0% to 16.7%. The mean age across all studies with data from typically-developing participants is six years and four months with a range of three years and ten months to nine years and five months. I report sample demographics of the included studies in Chapter 2, Table 1.

The geography of the studies' locations varied, including five languages (i.e., English, Korean, German, Dutch, and Spanish). Seven studies took place in the United States, three in South Korea, and one each in Chile, the Netherlands, the United Kingdom, Germany, and Canada. Five (33%) studies were led by one researcher.

All but one study in the present synthesis were non-experimental. Thirteen studies reported analyses from cross-sectional data, with the remaining studies reporting longitudinal results. Eight studies reported results from regression analysis, and seven studies included results

from mediation analysis, including four from structural equation modeling and three from other mediation analyses. One study reported results from a no control, pre- post-test design.

The total number of participants across the 15 studies was 2,024. Sample sizes ranged from 31 to 350 with a mean of 135. It is not possible to report on power for the individual parameter of interest (i.e., ToM to comprehension path), but I do report on power for the studies' overall models. Cohen, et al. (2003) report a large effect as 0.35, medium as 0.15, and small as 0.02. A priori power analysis using G*Power (Faul et al., 2009) for a six-item regression analysis, suggests the minimum number of participants needed to detect a small effect is 485, assuming a power of 0.8 with six predictors and an alpha level of 0.05. The minimum number of participants to detect a medium effect is 68, and to detect a large effect is 32. Based on Cohen et al., 14 studies were adequately powered to detect a large effect, 11 total studies were adequately powered to detect a medium effect. No study was adequately powered to detect a small effect.

Measures

All studies used false-belief measures to assess ToM. Five studies used a composite score of both first and second order false-belief tasks. Two studies included only second-order false-belief tasks and two studies included only first-order false-belief tasks. Two studies used both first- and second-order tasks for separate analyses. One study used first order for listening comprehension and second order for reading comprehension, and the other analyzed effects of first order and second order separately on story comprehension. Five studies included an advanced ToM measure. One study used the ToM scale but adapted the scale into the participants' native language (Spanish) and added an additional task to control for potential celling effects. One study included the RMET in addition to an advanced false-belief measure.

Measures of comprehension varied across studies. Seven studies had only a reading comprehension outcome measure, two studies had only a listening comprehension outcome measure, and three studies had both a reading and listening comprehension outcome measure with separate analyses. One study had a measure of story comprehension (wordless picture book task) and listening comprehension, analyzed separately. Two studies only measured comprehension using a story comprehension measure, both were researcher-developed: one with a wordless picture book and another a fables comprehension task. All studies using regression analytic methods and confirmatory-factor analysis included at least one standardized comprehension score. The five studies using structural equation modeling included latent variables for both listening and reading comprehension, all comprising both narrative and expository components and standardized as well as researcher-developed measures. One researcher conducted all five structural equation model studies.

The inclusion of control variables and covariates varied across studies. Six studies with a reading comprehension outcome measure included SVR variables of decoding and linguistic comprehension. Of those six, five included other known variables related to reading comprehension. Three total studies included a measure of inferencing, eight included a measure of vocabulary (one of which included both vocabulary breadth and depth), and five included a measure of comprehension monitoring. One study included a fluency measure. One study using story comprehension included a standardized early-reading-ability measure as a control (assessing alphabetic and print convention knowledge as well as meaning).

Nine studies included measures of executive function. The most common component of executive function assessed was working memory, with eight studies including a working memory measure. Three studies included a measure of inhibition, two included a measure of

attentional control, and two included a measure of cognitive flexibility. One study utilized a composite executive function score of working memory and a multi-component measure. Language measures varied across studies, with five studies including at least one measure of language ability. One study used a measure of grammar. One study used a measure of syntax. One study used an expressive-only measure. The two remaining studies measured language using a general language measure. Other covariates included age (six studies), gender (three studies), non-verbal ability (two studies), SES (two studies), days in school (one study), IQ (one study), and home literacy environment, print exposure, and school type (one study).

Analytic Findings

Across the 15 studies, 53 correlations (*r*) were reported between ToM and comprehension (range = 0.09–0.69; median = 0.43). Twenty-three correlations examined ToM's relation to reading comprehension (range = 0.09–0.69; median = 0.41). Twenty-seven correlations examined the relation between ToM and listening comprehension (range = 0.24–0.67; median = 0.45). The remaining three correlations examined the relation between ToM and story comprehension (range = 0.18–0.40; median = 0.23). See Chapter 2, Table 2 for each correlation by type of ToM and comprehension assessed.

Thirteen studies reported 19 standardized Beta values, indicating the prediction of comprehension by ToM, ranging from β = -0.17 to β = 0.45 across comprehension outcomes, five of which were not significant at the 0.05 level. One beta value reported the directional relation between reading comprehension to ToM. Unique variance of ToM when predicting reading comprehension was reported in seven studies and ranged from 2–8.9% of comprehension performance beyond the control variables included. One study was experimental, and I calculated an effect size also reported in Chapter 2, Table 2.

Eight studies conducted hierarchical regression analysis and reported ten Beta values across comprehension outcomes. Beta values ranged from β = -0.17 to β = 0.38, four of which were non-significant. Three of the four non-significant findings measured first-order ToM only. The fourth non-significant finding used a composite score of the ToM scale (Wellman & Lui, 2004) plus one second order measure.

Eight studies conducted mediation analysis. Beta values ranged from β = 0.023 to β = 0.45 across comprehension outcomes, one of which was non-significant. Direct mediation occurred between ToM and linguistic comprehension in five studies. A direct relation between ToM and reading comprehension was found in three studies, two of which were conducted with samples of students with autism. Indirect mediation occurred between ToM and reading comprehension through linguistic comprehension in three studies. One study found a direct relation between reading comprehension and ToM. See Chapter 2, Table 2 for a summary of the analytic findings.

Discussion

The studies in my pool varied across methodologies, variables, and assessments. The studies also, then, offered different conclusions. In total, 12 of the 15 analyses of the relation between ToM and reading comprehension found an overall positive, medium relation with a range of small to medium effects in addition to one intervention study finding large effects based on Cohen and colleagues' (2003) guidelines.

Overall, the results suggest the presence of a small, indirect relation between ToM and reading comprehension via linguistic comprehension in a general population that grows in strength across student development. Further, the synthesis results suggest the relation is sensitive to the type of ToM assessed; second order and advanced ToM measures present

stronger relations than first order. Finally, evidence from this review suggests that a more precise interpretation of the results is that ToM, *as measured by false belief understanding*, does not play a direct role in reading comprehension but does play a mediating role through linguistic comprehension.

A wealth of prior research supports listening comprehension's dominance in reading comprehension outcomes (Wolf et al., 2019). In effect, listening comprehension serves as a gatekeeper of comprehension, controlling what and how different variables influence reading comprehension outcomes. In addition, meta-analytic research suggests a bidirectional relation between false-belief understanding and language competence. Current research suggests oral language and listening comprehension in young, emergent readers essentially operates as a single construct due to high multicollinearity (Language and Reading Research Consortium, 2017). Thus, due to the nearly complete use of false belief measures within the synthesis to measure ToM, it stands to reason the mediation effect of listening comprehension would be present.

The reasons for first order ToM to appear non-significant in the synthesis results are less straightforward, but there are a few reasons this finding likely emerged. First order ToM typically develops between the ages of three and five. Participant development could therefore contribute to variance or be a source of constraint of the findings. The mean age of participants in studies that measured a composite ToM score inclusive of first and second order measures was six years and one month. Ceiling effects for first order ToM may mask ToM's significant contribution due to the lack of variance in ToM within the general population for the specific age range studied.

A second reason may derive from the early dominance of other skills in reading comprehension outcomes. During the appropriate 3–5 age range, much of a student's reading

comprehension score depends on word-reading skills subsumed under the decoding domain until they reach automaticity. Only once automaticity in basic decoding is reached, do other, higher-order skills such as ToM explain reading comprehension outcomes. The effect of first-order ToM is thus overpowered by stronger explanatory variables in this age range. Still, even once such higher-order skills do begin to exert an effect on reading comprehension, the effect of first order ToM may be undetectable when other variables such as executive functioning, vocabulary, and inferencing are included, especially given the early developmental ceiling of first order ToM.

A third reason surrounds the structure of the ToM assessments. Correlations reported between comprehension and second order and advanced measures of ToM are consistently stronger than first order ToM across the synthesis results. Second order and advanced measures of ToM more closely resemble narrative reading tasks than first order measures and may represent more valid measures to assess ToM's influence on reading comprehension outcomes.

Studies within the present synthesis examining the relation in students with autism both do and do not help clarify the findings of the synthesis. Both studies of participants with autism (McIntyre et al., 2018; Ricketts, et al., 2013) assessed ToM with advanced measures, controlled for SVR variables, and found a significant relation between ToM and reading comprehension. Further, McIntyre et al.'s findings included a typically-developing sample, which revealed a nonsignificant direct relation between ToM and reading comprehension that precludes generalizing the indirect relation via linguistic comprehension beyond the general population. However, these same findings do not clarify the non-significant findings of first order ToM because due to the age of the participants only second order measures were used.

Work on Theory of Mind and Comprehension Since 2018

Interest in ToM's relation to reading comprehension is proliferating. Since 2018, approximately nine studies have been published with findings that both corroborate and contradict results from the studies included in this section. All eight studies still measured ToM with false-belief measures and used varying comprehension assessments.

Longitudinal findings from these more-recent studies are in line with the findings of the synthesis, further suggesting ToM grows in importance to reading comprehension as participants reach decoding fluency ceilings. Kim (2019) found an indirect effect on reading comprehension of 0.18 in Grade 2, which increased to 0.32 in Grade 4, and a direct effect on listening comprehension of 0.30 in Grade 2, which increased to 0.39 in Grade 4. However, Kim (2020) found a direct effect on listening comprehension in kindergarten of 0.36 (first and second order measures), which decreased to 0.31 (second order only) in Grade 2 with similar covariates, suggesting a u-shaped relation between ToM and listening comprehension, likely due to other variables such as EFs that may be stronger predicters at this earlier age range.

Other longitudinal studies examining ToM's relation to reading comprehension align with and extend findings from the synthesis. In a pair of studies form Ebert (2020a; 2020b), early ToM (first order) did not predict later reading comprehension in adolescents when early language abilities were controlled, but advanced ToM did through listening comprehension. Furthermore, although early ToM did not relate to later listening comprehension, it did relate to later vocabulary (Ebert, 2020a). In a follow up study, Ebert (2020b) demonstrated that mental-state vocabulary mediated the effect of early ToM's relation to listening comprehension. Ebert suggests ToM relates to listening comprehension because it supports children's understanding of mental states and processes and, thus, mental-state language, specifically.

Ebert's (2020b) focus on mental-state language is significant because of the influence of knowledge on reading comprehension outcomes. If a reader has a sufficient level of knowledge to access a text, the relation between knowledge and reading comprehension becomes bi-directional (O'Riley et al., 2019). Mental state language represents academic-specific vocabulary necessary for successful narrative comprehension (Cartwright, 2015), and narratives thus demand specialty knowledge that is, in effect, knowledge of the social world (Pavias et al., 2016). Finally, the most recent findings from Kim et al. (2021) strengthen the claim of mental-state language's importance by reporting that ToM (second order) predicted mental-state talk of children's recall, which strongly predicted (0.77) narrative comprehension and was weakly related to informational text comprehension (0.28). Mental-state language's importance becomes even more prominent considering Kim and Petscher's (2020) research on child, text, and assessment characteristics that found null effects of individual differences in ToM (second order) across narrative and informational texts and literal and inferential question types.

Narratives are also the main mode of engagement for young readers (Pentimonti et al., 2011) and, thereby, support beginning reading comprehension as well as the development of social knowledge and skills (Lynch et al., 2008; Mar & Oately, 2008). Readers with underdeveloped or atypical ToM may experience a Matthew Effect (see Stanovich, 1986), whereby, the gap in overall knowledge between readers with strong ToM and readers with weaker ToM becomes larger due to missing out on crucial learning from texts, further attenuating reading comprehension (Cain & Oakhill, 2011). These occurrences, over time, could result in cascading effects across the lifespan for students with underdeveloped or atypical ToM development. This cascading effect could further explain why Ebert (2020b) found such an effect of mental-state language in adolescence. Nonetheless, based on the linguistic findings and

discussion considered earlier, it remains open to speculation whether a measure of ToM performance on false belief captures such social knowledge.

In fact, Ebert's (2020a; 2020b) recent pair of studies support the presence of a bi-directional relation between ToM and reading comprehension that was found in Cantin et al., (2017). However, Cantin et al. only measured first-order ToM and also found a significant relation to mathematical ability. Conversely, Lecce et al. (2021) found a significant bi-directional relation between advanced ToM and reading comprehension and not to mathematical ability. Other support from Blair and Razza (2007) found false-belief understanding related to mathematic ability, but did not relate to reading skill (decoding; not comprehension) in kindergarten. These findings further support the developmental ToM's importance and suggest first order ToM may better represent a general cognitive skill (like EFs), rather than a socio-cognitive specific skill exclusive to reading, whereas second order and advanced measures may signify more nuanced measures that tap comprehension specific skills and thus are more ecologically-valid measures.

Other recent findings add a new direction of findings than in the present synthesis. Findings from Florit et al. (2019) are directly applicable to 21st-century reading comprehension demands and expand the research examining ToM's relation to reading comprehension by assessing ToM's role in multiple text comprehension. The findings suggest that advanced ToM of children in Grades 4 and 5 predicts multiple text comprehension over and above fluency, vocabulary, and single text comprehension. Florit proposes ToM assists readers in comprehending multiple author perspectives, a finding which could extend to expository comprehension as well. However, Florit et al. do not include a measure of listening comprehension or language ability.

Finally, Tong et al. (2019) compared ToM's relation to literal and non-literal reading comprehension between a typically developing population and population of students with autism, finding that ToM mediated the group effect on non-literal reading comprehension only. The measure used to assess ToM was the ToM inventory (Hutchins et al., 2012), which includes both first order and second order measures, providing some evidence that measures of first order ToM may directly relate to non-literal reading comprehension in samples of children with autism.

Limitations

Limitations of Synthesis Studies

The findings from the synthesis are generally limited due to the correlational nature of the study designs. The studies also lacked participant diversity; a majority of studies, despite being conducted in numerous countries, were comprised of participants that spoke the countries' dominant language, reflected the dominant race, and came from middle- to upper-class socioeconomic strata. ToM was almost solely measured with false-belief tasks and all studies included ToM in addition to other cognitive variables. Most studies did not report the specific direct and indirect effect of ToM, making it difficult to determine its unique contribution. The lack of fluency and decoding measures included in the studies for the majority age-range studied is a particular cause for concern.

The type of measure used to assess ToM is a general limitation across studies. All studies but one measured ToM as a unitary cognitive construct. The one study, the only experimental one, assessed ToM as both (advanced) false-belief understanding (the Social Faux Paus test; Baron-Cohen et al.1999) and emotion recognition as assessed by the RMET. Further, the majority of studies did not include in depth discussions of ToM measure protocols. Although

most ToM measure protocols included a comprehension control question, they did not use control questions for participants' pragmatic inferences of the task, and all studies included measures of ToM with sentential complements. As discussed earlier, such language confounds are concerning as they threaten the validity of the findings from studies within the synthesis, especially for participants from different racial and ethnic groups or low economic social strata.

Finally, it is worth mentioning that all studies within the present synthesis were underpowered to detect a small effect. Evidence from the review suggests the ToM's effect on reading comprehension is small. Therefore, it may be that the studies that found no effect or non-significance simply did not have enough power within the study design to detect its effect.

Limitations of the Present Synthesis

The field is still nascent with limited empirical evidence on the relation between ToM and comprehension. In fact, one researcher conducted 33% of the studies in this synthesis. There is a particular dearth of research examining moderating factors' characteristics, which we know are central to understanding ToM's role. Also, with a limited number of studies and small samples, it is difficult to generalize findings, especially so for sub-populations. The results of the present synthesis do not include sufficient data for special populations such as those with autism, learning disabilities, and other clinical populations; with or at risk for comprehension specific deficits; and emerging bilinguals.

Another potential limitation is my exclusion of other measures of social cognition, (see, for example, LaRusso et al., 2016 for a study on perspective-taking in an adolescent sample). Broadening the search criteria to include other measures of social cognition such as perspective-taking and empathy may have provided an increased sample of studies and could have provided more insight into the role social cognition generally plays in reading

comprehension. This may continue to be a challenge in the field; however, as social cognition is an extensive and chameleon-like construct umbrellaing many terms that are often used interchangeably and measured in disparate ways (Happé et al., 2017), it is difficult to operationalize more broadly for a systemic review. Nonetheless, the present synthesis offers the first of its kind, reporting synthesized findings on the empirical relation between ToM and reading comprehension.

Future Directions

Given the heterogeneity in findings and the aforementioned limitations, the synthesis provides plentiful directions for future research. A first area of need is future conceptual replications with larger samples and disaggregated findings for certain sub-populations with and at risk for deficits in language, such as those disabilities, including, but not exclusive to autism, and emerging bilinguals. At minimum, future replication studies should likely include both a measure of cognitive ToM (false-belief tasks) as well as affective ToM (emotion recognition; RMET), either separately or perhaps as a unitary social understanding measure in addition to known lower-level contributors to reading comprehension, such as decoding and fluency, and supporting cognitive variables such as EFs. The unique contribution of ToM should be reported explicitly. The inclusion of latent measures of ToM could provide a more robust implication or ToM's role in reading comprehension, provided ToM is not a directly-observable skill.

It also may be important for future replications to include different measures and evaluate ToM's relation to expository text comprehension. To date, there is very limited evidence that a relation exists between ToM and expository reading comprehension specifically, and what research exists reports mixed findings (Kim 2017; Kim & Petscher, 2020; Kim et al., 2021). More research is needed to examine the assumption that ToM supports narrative comprehension

to a greater degree than expository. Such endeavors can provide fruitful information for determining ToM's unique role in reading comprehension, distinct from more general reasoning processes, like EFs, and how to best measure ToM for the purpose of reading comprehension.

More broadly, replications are needed with more diverse measures of social cognition. It remains unclear the extent to which ToM ability corresponds with one's use of ToM to comprehend text and what other factors may be at play. For instance, somatic empathy, the phenomenon marked by the presence of physiological responses to another's mental or emotional state, may contribute to reading comprehension. Studies of the reading comprehension performance of adolescents with psychopathic traits, indicating an absence of somatic empathy, demonstrate a reading comprehension deficit (Horan et al., 2016; Vaughn et al., 2011), despite typical ToM development (Shamay-Tsoory et al., 2010).

Prior research indicates students' inference-making can sometimes fail to support comprehension not because students cannot generate inferences, but because they either fail to use the generated inferences appropriately or generate inaccurate ones. ToM is essentially a skill of inference making (Dore et al., 2018). Thus, despite having the ability to do so, adolescents who display psychopathic traits may not use their ToM ability to comprehend the text, while others without psychopathic traits may still use ToM reasoning but do so inaccurately, especially at more complex levels of comprehension. The inclusion of more general socio-cognitive variables spanning cognitive, affective, and somatic markers may represent a more complete picture of social cognition's role in reading comprehension.

Moreover, because skills of social cognition vary in malleability (Happé et al., 2017), such investigations also have the potential to reveal more precisely the extent to which changes in social cognition correspond with changes in reading comprehension outcomes. The

malleability of ToM and the literacy-led training practices, however, provide confidence that ToM-remediation efforts may result in distal improvement of students' reading comprehension. (Hofmann et al., 2016). But we do not yet know if their inclusion supports reading comprehension. Experimental studies in classroom contexts are thus needed in order to better understand implications for classroom practice.

Conclusion

My goal here was to provide the field with a critical overview of the literature on ToM and reading comprehension that supported the publication of ToM and reading comprehension reviews (Dore et al., 2018) and models and frameworks of reading comprehension inclusive of ToM (Duke & Cartwright, 2021; Kim, 2017). To achieve my goal, I aimed to identify more precisely how and what the field can interpret from the research conducted thus far, with an overall aim to support future research in this area.

Indeed, the stagnation of student reading comprehension outcomes (United States Department of Education, 2019) will continue to urge researchers across a variety of domains to consider the role of novel contributors to reading comprehension. One such aspect may focus on social cognition skills. As the present synthesis suggests, there is emerging and growing research that one measure of social cognition, ToM, significantly relates to reading comprehension. It is perhaps most clear, however, that much more research is needed. Such a burgeoning research field offers exciting interdisciplinary opportunities that may highlight more dynamic ways of conceptualizing, assessing, and teaching reading comprehension.

Chapter 3: A Reading Comprehension View of the Prevalence and Purpose of Mental-State Language in Preschool Interactive Read-Aloud Instruction

A wealth of research supports the prominence of high-order literacy skills that develop in the early childhood and elementary years. Such skills go on to support children's success in future employment and post-secondary education (Ju et al., 2012; Kerr et al. 2017), and there is evidence of a negative correlation between students' skill development and chances of dropout and incarceration (Barnett et al., 2021; Christle & Yell, 2008; Hoge et al., 1996). Early literacy skills also support the development of later reading comprehension (Dickinson et al., 2010; Golinkoff et al., 2019; Whitehurst & Lonigan, 1998). Effective early instruction on high-order literacy skills is thus needed to ensure all students can participate, not only in advanced disciplinary literacy learning, but also competitively participate in the 21st-century job market.

The present study's aim was to examine how one high-order reading comprehension skill—ToM—supports early childhood comprehension instruction. A burgeoning and lively research field names ToM as a novel skill that plays a role in students' reading and listening comprehension (Kim, 2021). ToM is a socio-cognitive term that designates a child has developed the aptitude to identify the mental states of others (i.e., others' thoughts, feelings, and intentions). While research on ToM in preschool classroom read-aloud activities exists, to my knowledge, none of it links ToM preschool classroom read-aloud instruction with comprehension of the read-aloud text. My purpose here was to identify and describe for what pedagogical purpose teachers use ToM language, defined as mental-state language, during interactive read-aloud (IRA) instruction. The results reveal how mental-state-language coding for reading and listening comprehension compares to mental-state-language coding for ToM development, and how mental-state-language use supports narrative comprehension in preschool classrooms.

Guiding Theory

The Simple View Reading (SVR) defines reading comprehension as the product of one's mastery of two domains, word decoding and listening comprehension (Hoover & Gough, 1990). Reading comprehension itself is not defined as a literacy skill per se, rather it is the end product of a constellation of skills at work during the process of reading (Tumner & Chapman, 2012). Both domains of the SVR indeed contain subskills that are necessary for reading comprehension, and those associated with listening comprehension are some of the least understood. ToM is one of these subskills and a recent area of interest to comprehension researchers.

ToM is apt for targeted support in early-childhood settings because it begins to develop around age three, approximately the age children begin to attend preschool (Wellman, 2002). The emergence of research on the effect of ToM on children's academic outcomes has spurred investigations in early-childhood education research on the role preschool environments play in supporting ToM development. To date, however, researchers have yet to empirically examine how supporting children's ToM also supports linguistic comprehension. Instead, the research examines the inverse, investigating how supporting linguistic comprehension supports children's emerging ToM. My purpose in this study was to address the literature gap by examining how ToM language, defined as references to characters' mental states, supports read-aloud instruction in preschool classrooms.

Based on the literature I described in Chapter 1 and extend upon in this chapter's conceptual framework below, it is evident that ToM plays some role in children's reading-comprehension development. In the sections that follow, the theoretical framework of the present study extends the discussion on how naturalistic classroom settings provide children with rich discussions about mental states. I maintain that, not only do such discussions support

children's ToM development as the present literature suggests, but that these discussions also support skills pertinent to reading comprehension. Such support is found in the conversations between children and their caregivers and teachers during shared-book reading of narrative texts.

Shared Reading

ToM develops innately along a continuum from the interactions between underlying innate skills and experiences within children's environments (Wellman, 2002). Familial conversations, particularly conversations with older siblings and caregivers, have a direct influence on early ToM development (Tompkins et al., 2018). In particular, the ToM-development literature names conversations with references to the mental states of others (e.g., thoughts, beliefs, feelings, intentions) as a primary environmental driver of ToM development. Conversations about the mental state of characters in stories mirror the same processes involved in inferring other people's mental states (Mar & Oatley, 2008). For example, while reading the classic fable "Little Red Riding Hood," a mother may ask "Why wasn't Little Red Riding Hood scared when she got to her grandmother's house?" In response to the child's answer, the mother might expand or correct as needed, indicating, "Yes, she believed (Little Red Riding's Hood's mental state) the wolf to be her grandmother because he was in disguise, so she wasn't scared!" In this brief exchange, the mother is helping her child understand that the mental state of Little Red Riding Hood (i.e., false belief) explains her behavior (i.e., lack of fear) in the story. In this way, the mother's questioning and feedback help the child understand the contingent relation between mental states and subsequent behaviors, a hallmark of ToM. Prior research shows that such mental state conversations between caregivers and children routinely occur during read-aloud activities and children's greater exposure corresponds with children's ToM-development outcomes (Ziv et al., 2013). At the same time, the mother is supporting the child's understanding

of story structure, and thereby comprehension, by attending to the mental states of Little Red Riding Hood throughout the story. Indeed, prior research also shows the importance of shared reading activities within the home literacy environment to children's later reading comprehension outcomes (Flack et al., 2018; Sénéchal & LeFevre, 2002).

Shared book reading between cargivers and children represents a similar activity to shared reading in early childhood settings (Ziv et al., 2014). As with caregivers and their children, the convergent conclusion from the small but growing research corpus on ToM in preschool is that read-aloud activities in classrooms provide children with rich discussions about the mental states of others through conversations about the characters in children's literature (Martucci, 2016). Likewise, with ToM development, longitudinal studies of early childhood classroom instruction suggest read-aloud activities also support children's later reading comprehension through the development of linguistic comprehension subskills (Lennox, 2013; Lever & Sénéchal, 2011; Pentimonti et al., 2012; Swanson et al., 2011).

Read-aloud activities in early childhood commonly refer to a series of practices that fall under shared reading, an umbrella term that represents instances of teachers reading aloud a text with students. Shared readings vary in type, with some more dialogic in nature than others. For instance, choral or read and respond practices reflect students reading with or repeating after their teacher as the teacher reads segments of text with little to no dialogic interaction. Conversely, IRA and dialogic reading refers to shared reading practices where the teacher reads a text aloud while actively engaging students in the process of meaning-making using the text (Wiseman, 2011; Snow, 2002). During IRAs and dialogic reading, teachers employ explicit instructional practices by stopping at selected points in the story to activate background knowledge, model comprehension strategies, provide independent and group opportunities to

practice target skills, and provide academic feedback to student responses (Lennox, 2013). Dialogic reading differs from IRAs in structure. While IRAs employ instructional practices arbitrarily based on the teacher's discretion to better engage children in the reading of the text, dialogic reading is a structured set of practices akin to explicit instruction that judiciously plan read-aloud instruction with a specific comprehension-focused aim, such as summarization (Baker et al., 2013; Institute for Education Sciences, 2007; Santoro et al., 2016).

Importance of Dialogic Interaction. The conversational aspect of shared book reading found in IRAs and dialogic reading is a critical intervention for ToM development. Research suggests mere exposure to mental states through oral reading is too passive a process for children to garner meaningful learning from the exposure (Devine & Hughes, 2019). In order for children's ToM development to benefit, conversations must occur during reading that help demonstrate for the child the thought process involved in using knowledge of mental states to reason about human behavior (Cutting & Dunn, 1999; Slaughter et al., 2007). The same pattern is found with developing literacy skills. Zucker et al. (2013) report that the frequency of preschool shared readings was not a predictor of children's vocabulary and reading comprehension in kindergarten and Grade 1, but that the rates of teachers' extratextual talk during the shared readings were a predictor of children's vocabulary in kindergarten and trended toward significance in Grade 1 reading comprehension.

Results from observation studies in naturalistic preschool classrooms posit that the support of linguistic comprehension subskills comes from the how and why questions posed by teachers during interactive shared reading practices, which increases teachers' extratextual talk and guides children's inference making during read-aloud activities (Pentimonti et al., 2012). Deshmukh et al. (2019) and Blything et al. (2020) correspondingly report that student responses

to such inferential questions during instruction elicit more linguistically-complex answers than responses to literal questions. Deshmukh and colleagues further report that, although how and why questions occur often during IRAs, such questions remain rare in comparison to other *wh*-questions (who, what, when, where) and simple yes/no questions. Encouragingly however, Blything and colleagues suggest teachers can be trained to include how and why questions at higher rates when provided professional development.

The conversations teachers have with students may be especially important for children's comprehension of characters' mental states because of the natural implicitness of references to mental states found in children's literature. Although explicit mental state references do exist in children's literature (Zunshine, 2019), references are more often implicit, (Dyer et al., 2000). As a result, modeling the ToM inferential process of mental state attribution to characters within the story is an important part of leveraging the affordances of interactive read-aloud practices to support children's ToM development (Devine & Hughes, 2019).

Findings from observations studies conducted in naturalistic preschool classrooms are important to consider because research on teacher read-aloud practice suggests dialogic reading happens less often than IRAs due to a lack of deliberate planning that teachers report when preparing their read-aloud instruction (McCaffrey & Hisrich, 2017). In the present study, the read-aloud activity under investigation represents an IRA practice and not dialogic reading because teachers were not given a prescribed implementation method; rather, teachers were told to conduct a typical classroom read aloud of a researcher developed picture book (*Kingdom of Friends*; Pentimonti & Zucker 2015). As a result, this study's findings provide a blueprint for how to build upon existing shared reading practices to make them more reflective of dialogic reading instructional frameworks.

A Focus on Narrative

Narratives are goal structured in that, at the core of every story, there is an introduction to a main protagonist who is goal oriented but confronted with a central problem in achieving that goal. The central problem of stories is therefore represented by the clash between the protagonist's goal and the antagonist's goal, which are in conflict. The ending of narratives is signaled by the resolution of the central problem. For example, in "Little Red Riding Hood," Little Red Riding Hood sets out to visit her grandmother (protagonist goal), but is confronted with the central problem when she meets the Big Bad Wolf. The Big Bad Wolf wants to eat her (antagonist goal). Little Red Riding Hood is unaware the wolf is trying to trick her by acting as her grandmother. The story is resolved when Little Red Riding Hood realizes the wolf's intentions and escapes. Overtime, repeated exposure to such narrative structure through oral storytelling, television, and text reading produces the development of a narrative schema based on the predictable structure of narratives (e.g., story grammar; Mandler & Johnson, 1977; Stein & Glenn, 1979). Once a child reaches school age, narrative schemas are further reinforced by teacher's story reading and literacy instruction.

Comprehension of mental states requires inference making; ToM represents the skills and knowledge needed to understand the social context of stories from the characters' perspectives (Duke & Cartwright, 2018). ToM serves narrative comprehension by providing the inference-making skills required to glue the narrative structure elements together and form a cohesive understanding of what the story is about (Dore et al., 2018; Graesser et al., 1994; Lynch et al., 2007). Therefore, the ability to use ToM to attend to characters' mental states strengthens one's ability to make inferences pertinent to comprehending narrative story elements (see Chapter 3, Figure 1). Using Little Red Riding Hood again as an example, Little Red Riding

Hood's goal-oriented mental states (*wanting* to visit her grandmother, her *fear* of the wolf, and *not realizing* the wolf is disguised as her grandmother) around the central goal of visiting her grandmother are vital to comprehending the story's structure. If a child, who is listening to or reading this story, does not understand why Little Red Riding Hood is in the forest, fears the wolf, or does not understand her ignorance of the wolf's deception, their comprehension of the story's structure suffers, resulting in a less-cohesive understanding of the story (see Chapter 3, Figure 2).

The mental states of Little Red Riding Hood represent what Bruner (1987) identified as the narrative landscape of consciousness. Successful narrative comprehension, according to Bruner, requires the coordination of the story's action events and the associated consciousness states of the characters (1987). A readers' cognizance of story events and of characters' consciousness, in part, result from a reader's ToM (Astington, 1990). Examinations of children's oral narrations of wordless books suggest such an effect of ToM. Children who exhibited more mature ToM development told oral narrations that were richer in mental state language that were used to tell more cohesive stories (Curenton, 2011; Gamannossi et al., 2014; Peskin & Astington, 2004; Pelletier, 2006; Pelletier & Astington, 2004; Riggio & Cassidy, 2009).

The Present Study

My initial aim in the present study was to explore if and how ToM language supports children's ToM development as well as linguistic comprehension subskills in preschool classroom IRAs. I sought to answer the following research questions: (1) With what frequency do preschool teachers use mental state language during shared reading instruction? and (2) to what degree do different types of mental state language occur during shared reading instruction? The first two research questions provided the data needed to address my third primary research

question, (3) what instructional contexts serve as antecedents to mental-state-language use during IRAs in preschool classrooms and, how, if at all, does mental-state-language-use support ToM and comprehension development? I hypothesized that the instructional context is instructional sequences in which inference-making is the target comprehension skill based on Dore et al.'s ToM and reading comprehension framework that posits ToM supports reading comprehension through inference making. By supporting inference making, children's comprehension skill improves.

Once I began the coding process, however, the a priori coding process I identified for initial use in the study proved too limiting in what it could reveal about how ToM language occurs within interactive read aloud inference-making instruction. Thus, after the initial piloting of an a priori-established coding scheme (Pentimonti et al., 2012; Ruffman et al., 2002), my aim grew twofold. First, to amend the coding scheme that was designed to assess how conversations between mothers and their children support ToM development for use with preschool teachers' IRA instruction. Second, to use this amended coding scheme with a sample of preschool teachers conducting an IRA of a common story. Therefore, here I report the methodology and results in three parts associated with each aim. Part 1 reports results from the initial coding process and the rationale for why I abandoned it to pursue the development of a new coding scheme. Part 2 reports on the development of the amended coding scheme, and Part 3 reveals the findings of my implementation of the amended coding scheme, corresponding to my main research questions.

Study Overview

To examine ToM language occurrence during IRAs, I conducted a mixed-method descriptive analysis of secondary data (Nassaji, 2015). I obtained permission to conduct a secondary analysis of Pentimonti and colleagues' (2012) data from a study of teachers' shared

book reading in preschool classrooms. Secondary research analysis refers to the process by which a researcher uses existing sets of data to answer a novel research question (Glass, 1976). The purpose of primary study was to create a standardized protocol to observe shared reading quality (see Pentimonti et al., 2012). The data set includes professional transcriptions of teachers' self-recorded videos of IRAs using a researcher-developed book, *Kingdom of Friends*. The transcripts report the IRA by utterance, which signifies a single unit of speech (e.g., a sound, word, statement, or question) from a teacher or student. Each IRA transcript is an Excel file with each utterance reported on a separate row. The utterances appear in the order they were spoken during the IRA.

Participants and Setting

The original data were collected during the 2005–2006 school year across state and private preschool settings. The average observation length of the shared reading videos did not exceed thirty minutes. The teacher participants were 95% female with an average of 15 years of teaching experience. More than 70% of teacher participants held at least a Bachelor's degree in education. The majority of teacher participants were white (72%); 20% were Black, 3% were Latinx, and the remainder were multiracial or identified as other race.

There were either 12 or 17 students present in each classroom. The mean age of student participants was 4 years and 4 months, and 54% were male. Half of the student participants were white, 30% were Black, 7% were Latinx, and 6% were multiracial. The majority of student participants (58%) reported an annual family income of $40,000 with a range between $5,000 and $85,000. More detailed information on teacher and student demographics is provided in the original study (Pentimonti et al., 2012).

Book Details

Kingdom of Friends tells the story of Diego's and Petunia's friendship. We are introduced to how Diego and Petunia are the best of friends, but we also find out that Petunia can be a bit bossy toward Diego, which makes Diego really upset. One day, the friends left preschool very upset with one another. Petunia talks to her dad that night, and he tells her that she should apologize to Diego. When they come back to preschool the next day, we see Diego and Petunia reconcile their friendship after Petunia apologizes, and they begin to play together again.

Part 1: Pilot Coding

Part 1 Method

I established my initial purpose based on Dore et al.'s (2018) framework that suggests ToM influences reading comprehension via inference making related to characters' mental states. I initially planned to code the data in three cycles. The first using a coding scheme for mental state language, second using a coding scheme for inference making instruction, and third for instances of overlap between the first and second coding cycles.

Measures. Ruffman et al. (2002) developed a mental state coding scheme to provide a measure of mental state references during mother-child dyads shared reading to assess how mental state reference use is related to children's ToM performance. I used the Ruffman et al. coding scheme to code mental state language. I intended to use the results from the Ruffman et al. coding scheme to answer my first two research questions by reporting the frequency of ToM language occurrence and the breakdown of occurrence by mental state type and, for research question 3, by reporting where the mental state coding co-occurred with inference making instruction.

Pentimonti et al. (2012) developed the Systematic Assessment of Book Reading (SABR) to measure shared reading quality in early childhood classrooms. I planned to measure inference-making instruction using the SABR by identifying a priori a subset of codes from the larger observation protocol (why and how questioning, making predictions, making connections, background knowledge, and causal reasoning; see Chapter 3, Table 1). I intended to use the results from Pentimonti et al.'s coding scheme to answer my third research question by reporting where the inference-making-instruction coding co-occurred with the mental state coding.

Procedures. For the first two coding cycles, I planned to use a protocol coding procedure. Protocol coding is used when there is already an established coding procedure in use for the data under study (Saldaña, 2015). I planned to use a mental state coding scheme from Ruffman et al. (2002) for the first coding cycle. For the last coding cycle, I planned to use a simultaneous coding procedure to identify areas of co-occurrence between the mental state coding scheme and inference making coding scheme. Simultaneous coding is used when two more codes apply to one piece of datum (Saldaña, 2015).

The mental state coding scheme codes each teacher and student utterance in the sequence in which they occur. I first classified each utterance as a mental state utterance occurrence or a non-mental state utterance occurrence and then identified the specific code for each mental state utterance. For example, I classified the teacher utterance, "Let's see if Diego [a story character] *wants* to do what Petunia does" as a mental state utterance with a corresponding mental state type code of 'Desire.' Originally, Ruffman et al. developed the coding scheme for picture analysis. I adapted the coding scheme slightly for use with IRA instruction by adding an additional code as a classroom management utterance code to capture any behavior management

utterances to exclude such utterances from analysis. See Chapter 3, Table 2 for a reprinted table of the original adapted mental state coding scheme.

The Pentimonti et al. coding scheme codes each teacher and student utterance verbatim in the sequence in which they occurred. I analyzed each utterance by type (Child Talk + Teacher Continuations, Teacher Talk, and Extratextual Talk; Chapter 3, Table 1). Extratextual Talk codes can co-occur with other codes from the other two categories. For example, a teacher utterance of "Let's see if Diego [a story character] *wants* to do what Petunia does." is coded as a teacher talk (character reference) with two Extratextual Talk codes (making predictions and desires/preferences). If the utterance does not result in at least one corresponding code, the utterance is identified as not codable. See Chapter 3, Figure 3 for the Part 1 coding procedure.

Data Analysis. During the third coding cycle, I identified utterances that contained co-occurrence between the mental state coding and inference making instruction coding. An utterance is considered to have co-occurrence if the utterance contains at least one code from the inference instruction coding scheme and at least one code from the mental state language coding scheme. For example, I identified the teacher utterance of "How do you *think* Diego *felt* when she called him a scaredy cat?" as an utterance that contained a co-occurrence. This is because the teacher asked a how question (inference making instruction code) with reference to an emotion mental state (felt) and a cognitive mental state (think). Therefore, the utterance would contain a code from the inference instruction coding scheme in addition to two codes from the mental state language coding scheme. I also identified areas of non-concurrence by noting where utterances contained at least one code from one coding scheme and no coding from the other coding scheme. For example, I classified the teacher utterance, "Let's see if Diego *wants* to do what Petunia does" as a mental state code (want) that occurred outside of inference making instruction

because the utterance does not contain an inference instruction code. I identified instances of both co-occurrence and non-concurrence in order to examine how mental-state language is used or not used in the context of inference making instruction.

Part 1 Results

Upon reflection, after piloting the initial three-round coding procedure across ten transcripts, I noticed a growing number of words that could be mental states but were not captured by either coding scheme. I conducted a preliminary analysis of the data I had thus far and realized the process would not result in meaningful findings due to a limitation within the coding system. For example, Chapter 3, Figure 4 shows a dialogic instructional sequence discussion about the word "fight" initiated by the teacher.

Based on the Ruffman et al. coding scheme, I only coded the final utterance as a mental-state utterance because of the use of the emotion-mental-state word *angry*. No utterance within the example qualifies as inference-making instruction under the a priori identified inference-making codes from the SABR Pentimonti et al. (2012) coding scheme. Therefore, none of the utterances within the example co-occur, which would indicate that the use of the word *angry* in this instructional context did not support the teacher's inference-making instruction. The limitation in the instructional sequence of the word *fight* shows its failure within the coding system to capture the sequence in which the mental-state language occurred within the instructional context. In other words, the way I had defined mental-state language and inference making based on the a priori coding systems was too narrow. The restricted coding resulted in the isolation of mental-state language from the larger instructional context, reducing the likelihood of any meaningful analysis as to the pedagogical reason why it had occurred. I noted this limitation by first recognizing ambiguous mental-state words, second by noticing the

proximity in which ambiguous mental-state words appeared to the coding I had completed, and finally by identifying this close proximity occurring within a sequence of instruction, defined as the block of utterances where book reading stopped in the transcript and where it began again. The blocks of utterances in the transcripts between where reading stops and where reading resumed again represents the teacher's dialogic interaction with the students. I concluded that, although mental-state language was somewhat occurring, albeit infrequently, in the context of inference-making instruction, it was mostly used outside inference-making instruction as I had defined it by supporting teacher's instruction on the inference-making process, which I describe in more detail below.

Part 1 Discussion

Despite the null findings emerging from my initial coding method, I reasoned the comprehension of the ambiguous words in the transcripts that gave me pause contained tacit mental-state understanding, which appeared to support the teacher's instruction within instructional sequences. For instance, to comprehend the word *fight* in the aforementioned example, students must understand that the word *fight* can refer to instances of physical altercation as well as verbal arguments. To ensure their students understood this, the teacher began this instructional sequence by stopping their oral reading of the text and prompting students with a sentence stem to identify that Petunia and Diego were in a fight. Subsequently, the teacher then established the understanding that the fight Petunia and Diego were having was not a physical fight by asking a set of questions: "Did they hit each other?" and following up with, "Did they use their words?" Finally, the teacher reinforced how to identify when a verbal argument can be described as a fight by asking, "They did [use their words], but in an angry way, didn't they?" At the conclusion of this instructional sequence, after ensuring students understood

the mental state of the characters and how to describe their physical actions, the teacher continued orally reading the text.

What the example revealed to me was that the teacher supported their students' comprehension of the word *fight*, and therefore the story scene, in two distinct ways that seemed to mirror the two-stage inference making process described by Kendeou (2015) I describe in detail in Chapter 1. First, by activating or providing background knowledge (i.e., mental state vocabulary knowledge; fighting can mean verbal arguments in addition to physical altercations) and second by referencing a specific inferential ability (i.e., perspective taking; the characters were angry which is why they were fighting). This set of skills (background knowledge + perspective taking) *together* assists children in understanding the word *fight* in the story context and is *how* children's ToM skill supports comprehension.

As I describe in detail in Chapter 1, inference making consists of both knowledge activation and integration (Kendeou, 2015), and ToM refers to a child's ability to not only determine others' mental states (perspective take) but to use that information to make sense of others' previous or ensuing social behaviors (Wellman, 2002). If ToM is indeed supporting the comprehension of the word *fight* in this context, it makes sense to hypothesize that the successful comprehension of the word *fight* would mirror the same inferential process of knowledge activation and subsequent integration because ToM is fundamentally a skill of inference making since mental states cannot be seen. To test this hypothesis, consider how inference making can fail. Comprehenders fail to make inferences during reading for three primary reasons: 1) the knowledge needed to generate the inference is not activated due to a lack of prior knowledge about the topic or poor working memory capacity, 2) the knowledge is activated but is incorrect when placed in the context in the story, or 3) the knowledge is activated but is not well

integrated, if integrated at all, into the emerging mental representation (Cain & Oakill, 2011; Cain et al., 2001; Carlson et al., 2014; Oakhill & Cain, 2007; Kendeou & van den Broek, 2007). These same three potential pitfalls can also occur with ToM. First, if a reader has poor mental-state vocabulary (i.e., does not know people can fight with words), the initial knowledge activation stage of inference making will fail and the inference (e.g., the characters are in a fight) will not generate. Second, if a reader has adequate mental state vocabulary but fails to perspective-take (i.e., does not identify the characters as angry), the inference still will not generate. Third, if a reader has adequate mental-state vocabulary and does take the perspective of the characters but is wrong about the thoughts and feelings of the character (e.g., the characters are just play fighting, not real fighting), the inference generated will not support the emerging mental representation. Given this evidence, breaking ToM into the two discrete parts of background knowledge and perspective taking provides a better understanding of how children's ToM skill supports comprehension and where it may interrupt comprehension. In addition, it can also serve as a guide to determine if words such as *fight* are indeed mental state words.

Part 2: Development of Additional Codes

After concluding that the a priori coding system was insufficient to answer my research questions, I decided to altogether abandon the use of the inference-making instruction codes from Pentimonti et al.'s (2012) SABR coding scheme. Instead, I set out to adapt the mental state coding scheme provided by Ruffman et al. (2002) to include other mental state words that required the two-stage inferential process for comprehension. My specific aim in Part 2 of the present study was to develop a broader ToM coding scheme by capturing additional mental state types.

Part 2 Method

Procedures. To develop additional codes for the ToM coding scheme, I first began by consulting another mental state coding scheme put forth by Dyer et al. (2000) that was used to count the frequency with which mental state references occurred in younger and older children's books. When I cross referenced the Ruffman et al. (2002) and Dyer coding schemes, I realized there were mental state types within the Dyer et al. coding scheme that were not in the Ruffman et al. coding scheme. For example, Dyer et al. had additional classifications of mental state word types of moral evaluation (good or bad and associated synonyms) and words of volition (choice). I could not directly transpose these codes onto the Ruffman et al. coding scheme because the blending of codes precluded that action. For example, the Dyer et al. moral evaluation code, in addition to including moral evaluations, included mental state words that Ruffman et al. classified as modulations of assertion (e.g., should, supposed to). Dyer et al.'s volition code included desire mental states that the Ruffman et al. coding scheme already tabulated as a separate mental-state type. Therefore, had I directly transposed the Dyer et al. moral evaluations and volition codes, it would have resulted in double coding utterances that contained modulations of assertions and desires. Therefore, I began the new coding process to identify new mental state type codes I could add to the Ruffman et al. coding scheme to adapt for use in Part 3 of the present study using Dyer et al.'s moral evaluation and volition codes as provisional codes. Provisional codes function as draft codes identified from previous research that are modified throughout the coding process to establish new codes for use in subsequent coding sessions (Saldaña, 2015).

I analyzed the data again through a three-part cyclical deductive coding process using pre-established coding schemes from the literature (Saldaña, 2015). I engaged in analytic memo writing twice during the coding process to include a level of reflexivity to my data analysis

process (Maxwell, 2012). During the analytic memo writing, I reflected on the coding process, theorized preliminary assertions that were emerging from the data, and tracked preliminary data trends to identify if new data provided novel insight, aligned with previous coding, or contradicted previous coding (Saldaña, 2015). If, during the coding process and analytic memo writing, I inductively uncovered codes not captured by the a priori coding criteria, I recoded the transcripts as needed. See Chapter 3 Figure 5 for the Part 2 coding procedure.

Data saturation refers to the point in data collection in which new data does not uniquely inform data analysis (Glaser & Strauss, 1967). Because I identified codes from the literature and pre-determined the criteria by which I was going to determine the suitability for new codes generated from the data, I completed data analysis using an a priori thematic saturation sampling approach (Saunders et al., 2017). Francis et al. (2010) provided a heuristic for identifying the needed sample size to achieve data saturation in deductive qualitative research by classifying a priori an initial random analysis sample of ten and a stopping criterion of three additional samples. This means I needed to code a minimum number of 13 transcripts. Before starting the coding process, I choose a random selection of 13 transcripts. I achieved thematic saturation after coding 13 transcripts.

Coding. I began the new coding process by using an eclectic coding procedure for the second round of coding. Eclectic coding is typically the first step in an iterative coding process that blends the use of two or more coding strategies in order to codify a uniformed coding system for use in subsequent coding sessions (Saldaña, 2015). During this first coding cycle, in addition to the Ruffman et al. coding, when I came across an unclear word, I copied the instructional sequence in which it was used into a new Excel spreadsheet so as to structurally pre-code the data. Pre-coding is a qualitative coding method whereby datum is identified for more specific

and targeted coding in later coding sessions, and structural coding allows for data to be chunked into segments that represent a whole (Saldaña, 2015). I classified an instructional sequence as the block of utterances that occurred between read-only utterances. This means an instructional sequence contained the series of utterances between when the teacher's oral reading of the story stopped and when it resumed.

Once I was done coding ten transcripts, I coded an additional three for a total of 13, in line with the approach outlined by Francis et al. (2010), before stopping to review the first round of coding and moving to the second round. I then engaged in a second round of coding of the pre-coded instructional sequences. I first consulted the provisional codes from Dyer et al. of moral evaluation and volition mental state types to see if and how I could incorporate any of the ambiguous mental-state words identified in the first round of coding into the Ruffman coding through analytic memo writing (Saldaña, 2015). These analytic memos served as a log of my thoughts throughout the coding process.

After coding thirteen transcripts in this way and achieving data saturation, I reviewed my new coding and analytic memos and noticed a pattern in that words of moral evaluation were frequently used to describe characters. However, there also remained a subset of ambiguous words that were not moral evaluations but were still used to describe characters and were evaluations, nonetheless. For example, the word *bossy* was frequently used to describe the character, Petunia, within the story. Bretherton & Beeghly's (1982) early categorization of mental states provides clarity on ambiguous evaluative words like *bossy* distinct from moral evaluations like *good* and *bad*, "The word nice could have been treated as a moral term, since children often use it to refer to moral conformity…the [mental state] categories…should not be taken as absolutes but as organizational aids" (pg. 907). Because neither the Dyer et al. nor

Ruffman et al. coding schemes included a mental state type code that codified words like *bossy*, I engaged in a final round of coding using a hypothesis coding procedure. Hypothesis coding is used when a researcher generates a working hypothesis by which to assess the fitness of the data to the hypothesis. I assessed each remaining ambiguous mental state word against my theoretical hypothesis that the comprehension of mental state words requires a ToM inferential process inclusive of both background knowledge activation and perspective taking.

Trustworthiness. Brantlinger et al. (2005) presents several ways qualitative researchers in special education can establish trustworthiness in their data analysis, a term equated to reliability and validity in quantitative research. I ensured the trustworthiness of the data analysis described above by employing three strategies outlined by Brantlinger et al.: theory triangulation, peer debriefing, and external auditing. First, I analyzed the data across two different theories of child development related to ToM and the skills of reading comprehension (e.g., inference making). Second, throughout the process I shared analytic memos with two experts in developmental psychology and reading comprehension to affirm my interpretations were grounded in the data and made sense in the context of the current literature in the respective fields. Third and finally, an external researcher reviewed the qualitative methodology and data analysis to attest that the assertions I made were grounded in the findings.

Part 2 Results

Character Traits as Mental State Types. When we use words such as *bossy* to describe people and characters in stories, we imbue a connotation that communicates a negative semantic meaning. This evaluative colloquial mechanism signals an attention to and comprehension of intentionality, a hallmark of ToM. Indeed, definitions of the word *bossy* describe someone that is fond of ordering other people around or inclined to domineer. Underlying mental

states (e.g., fond, inclination) and their contingent intentions (e.g., order people around, domineer) are overtly referenced in the definition. Therefore, to accurately describe a character as bossy involves an implicit understanding of their mental state and associated actions (i.e., ToM). Once I established how ToM serves the comprehension of the word *bossy*, I then sought to examine if and how the comprehension of the word bossy required the two-step inferential process to determine if bossy met the criteria of a mental state word.

In the context of stories, bossy is a character trait. To accurately identify a character as bossy, a comprehender must first understand that the character's actions indicate the actions of someone that is bossy, and second perspective take to identify if the characters' intentions (to get their way) align with bossy actions (ordering people around). We know this process can breakdown in three ways based on the inference making pitfalls I previously described. 1) If a reader understands just the concrete actions of the story disassociated from the characters' mental states, a reader will not infer a character trait. Rather, the reader will describe the character's actions when asked to describe the character, such as describing a bossy character as someone that orders people around a lot. If a reader understands the same concrete actions of the story and associates them with a different intention—say, to make a plan or to show confidence—a reader may infer a different character trait; for example, identifying the character as decisive or assertive instead of bossy. Last, a reader may understand the concrete actions of the story and correctly interpret the character intention but still fail to identify the appropriate character trait because they lack the vocabulary to do so; for example, identifying the character as mean instead of bossy.

Behavior Descriptions as Mental State Types. In addition to character-trait words such as bossy, I also determined a set of unclear words that, although related to Dyer et al.'s

classification of volition mental state words, did not fit. For example, the aforementioned ambiguous word *fight* was included in this mental state grouping. I engaged the same process as with bossy to identify if and how I could classify such words for use within my new coding scheme. I showcase this process using the mental state behavioral descriptor word share.

To share generally means to divide something or use something with another person. Understanding the behavioral action of sharing connotes something about the thoughts and feelings of the sharer and potentially about the recipient of the sharing behavior, specifically related to intention. Sometimes we intend to share with others because we want to or because we care about them. Other times, someone else may force us to share. Understanding of such underlying mental states helps readers comprehend the use of the word *share* within the story context, thus resulting in the word share containing implicit mental state understanding. Therefore, ToM serves the comprehension of the word *share*. Next, I examined if and how the comprehension of the word *share* required the two-step inferential process to determine if it met the criteria of a mental-state word.

In the context of stories, share is a description of character behavior. To accurately portray the actions of the character as sharing behavior, a comprehender must first understand that the character's actions indicate the actions of someone that is sharing, and second a comprehender must perspective take to identify if the characters' intentions (including others in an activity) align with sharing actions (by providing access to tangible items or roles). Again, we know this process can breakdown in three ways. 1) If a reader understands just the concrete actions of the story disassociated from the characters mental states, a reader will not infer the appropriate behavioral description. Rather, the reader will describe the character's actions when asked to describe the character, such as describing sharing behaviors exhibited by characters as

giving toys or letting them play with them. If a reader understands the same concrete actions of the story and associates them with a different intention—say, to give away something—a reader may infer a different behavioral description; for example, describing the character behavior as offering instead of sharing. Last, a reader may understand the story's concrete actions and correctly interpret the characters' intention but still fail to accurately describe the characters' behavior due to a lack of vocabulary; for example, describing the behavior as giving instead of sharing.

Non-Examples. After completing the final coding procedure, there remained ambiguous words that did not quite fit as mental-state words but that still seemed to require requisite mental-state understanding. For example, during one part of the story, Petunia is upset and shares with her dad she is upset because her friend, Diego, made her mad. In response, her dad talks to her and tells her she should apologize to her friend. Some teachers discussed this exchange by using the word *advice* to describe the dad's actions. Other teachers used the phrase, "talked to," and still others interpreted the situation as her dad "telling her what to do." The text does say "talked to," which was the most common way teachers described the scenario. These nuanced interpretations may be important to consider because the most precise term for the situation is *advice*, yet still, teachers differed in how they discussed the situation with their students. The word *advice* is the most precise term when we consider the social elements of the story. First, Petunia is a child talking to her father, who tells her she should apologize. When we use the word *advice*, we typically refer to instances of a sometimes wiser other making a recommendation to someone in need of support. In this case, the wiser other is Petunia's father. Second, the key word in the text is *should*, which separates the father's statement from a demand and lends the accurate interpretation of the word *advice*. But advice is not a mental state, rather it is the noun

nominalization of the verb to advise. According to the coding scheme, *advise* would classify as a mental state under the behavioral description category, but *advice* is not a behavioral description—it is the product of engaging in the behavior. Another word like *advice* that occurred in the transcripts was *idea*. The coding scheme does not capture such words. However, such words do, to some extent, require requisite mental state understanding because their use depends upon one's attention to and comprehension of the social context of the story.

Part 2 Discussion

The results from Part 2 reveal similarities and differences when coding mental state language use during IRAs from a lens of reading comprehension instruction as opposed to a lens of ToM development. For example, the addition of character traits and behavioral descriptions as mental-state types to the coding scheme corroborates and expands upon prior research that emphasizes the importance of implicit mental states over explicit references (Peskin & Astington, 2004). My decision to add these two mental-state types to the Ruffman coding scheme is supported by findings detailed in Peskin and Astington's work, which underscores the importance of implicit references for children's ToM development. Peskin and Astington found that participants who were exposed to shared readings with texts designed with implicit mental-state references outperformed control groups on measures of ToM performance who were read the same texts enriched with additional explicit cognitive mental states or a control text. Such a finding is important to consider when examining mental-state language during IRAs because Peskin and Astington contend that the improved performance in ToM of the implicit-references group is explained by the increased conversations that resulted from teachers' explanations of the implicit references. The increase in conversation spurred by implicit mental states is also likely to better support comprehension development.

As previously mentioned, behavioral descriptions such as *sulk* and character traits such as *bossy* are not in and of themselves mental states. Rather, they require an implicit understanding of the associated mental states that prompt the behaviors and traits we identify as *sulk* or *bossy*. Character traits and behavioral descriptions, however, are elements potentially more germane to narrative comprehension than ToM. It is unknown whether the same result pattern for ToM development found in Peskin and Astington's study would extend to mental-state categories.

Part 3: Final Coding and Data Analysis

Part 3 Method

Coding Scheme. The iterative coding process described above in Part 2 resulted in me adapting the Ruffman et al. (2002) coding scheme by adding two additional mental state code types: character traits and behavioral descriptions. The third and final adaption to the coding scheme I made was to classify the mental-state reference as either reading comprehension focused or non-reading comprehension focused. For example, teachers regularly used the cognitive- and desire-state words of *like* and *want* when managing student behavior, asking students to make connections outside the text, and in expressing personal preferences. As examples, teachers would say, "I need you to have a seat." or "I liked that story!" or "In what centers do we pretend?" "What do you think this part of the book is?" I added a code to capture these behavior management instances so that I could adjust the data in order to capture how ToM language supports comprehension instruction specifically, rather than general language learning. See Chapter 3, Table 3 for the final adapted coding scheme with code definitions and examples.

Quantitative Coding Procedure and Analysis. Similar studies within the research literature generally code between 40 and 100 transcripts (Alvarenga et al., 2020; Adrián et al., 2007; Ruffman et al., 2002; Ziv et al., 2013, 2014, 2015). I coded a random selection of

58 transcripts (total available in the dataset, 105) across 49 teachers (ten teachers had two transcripts each). Each transcript contained three types of utterances: teacher utterance, student utterance, and a read-only utterance. The read-only utterance refers to the oral reading of the story by the teacher. I excluded the read-only utterances from the mental state language analysis in order to capture only the mental states referenced by the teacher and students. I then coded each teacher and student utterance verbatim in the instructional sequence of occurrence.

To answer my first two research questions, I first classified each utterance as mental state or non-mental state and then whether or not the utterance was made by a teacher or the student(s). Next, I captured the specific mental state word(s) referenced during the utterance and identified the specific type of each word referenced (e.g., desire, emotion, modulation of assertion, think/know/believe, other, behavior, character trait). Finally, I determined whether the mental state utterance occurred as a part of reading comprehension instruction. After coding the transcripts, I collapsed the data to report the overall descriptive information (e.g., frequency counts, means, percentages, and rates). Once I completed coding all transcriptions, I then applied the same mental state coding scheme to the *Kingdom of Friends* to compare between frequencies of mental states uttered by the teacher and students and the frequencies as mentioned in the text. To complete the coding process, I created a new document with all of the mental state words referenced across all of the utterances in the transcripts to obtain a total amount of words referenced during the shared readings. I also identified the number of unique mental state words by creating yet another sheet with all duplicates removed. I repeated this process with the shared reading text.

Coding Reliability. To establish interrater reliability, I trained a reading researcher at the University of Maryland as a second coder by establishing an interrater reliability of 85% on two

transcripts that I did not use for data analysis. Once I trained the second coder, they double coded 30% of the transcripts (n = 15). I discussed all disagreements with the second coder until we reached 100% agreement. I report average interrater reliability for mental state utterances, reading comprehension focused utterances, and mental state type by dividing the smaller total by the bigger total and multiplying by 100. Reliability for mental state utterances and reading comprehension focused utterances was excellent at 96% and 97%, respectively. Reliability across mental state type ranged from acceptable to excellent: desire-state terms (97%), emotion-state terms (86%), cognitive-state terms (83%), modulations of assertion (66%), character traits (67%), behavior descriptions (65%), and other-defined terms (96%).

Qualitative Coding Procedure and Analysis. To answer the third research question, I employed a secondary coding analysis procedure that included both deductive and inductive coding methods. I used the previously coded data I coded for research questions one and two. I chose the top three transcripts with the most mental state reference based on overall percentage of mental state utterances. The top three transcripts contained reading comprehension focused mental state utterance percentages of 44%, 43%, and 41% (see Chapter 3, Table 4). The structurally coded data from the quantitative data analysis chunked the reading comprehension focused mental state references into instructional sequences based on when the teacher stopped reading the text to ask a question or provide a comment until when they resumed the reading. Therefore, rather than analyzing one utterance at a time, I was able to analyze an entire sequence of teacher read-aloud instruction that contained mental state references, which better allowed me to identify the pedagogical purpose of their mental state language use.

For the first round of coding, I implemented a hypothesis coding procedure comprised of two descriptive codes that I identified a priori from the literature. These two codes referred to

perspective-taking and causality. I allowed for simultaneous coding of the data at this stage because some utterances included both perspective-taking and causality components. After this first round of coding, I then implemented an open descriptive coding procedure to identify any potential themes emerging from the data I had not identified a priori. This round of coding resulted in a code of teacher directed language emerging from the data. I implemented a last coding procedure where I distinguished patterns within the data to identify sub-codes. I began coding the transcript with the highest reading-comprehension focused mental-state-utterance percentage and proceeded with the next transcript and so on. See Chapter 3, Figure 6 for the whole coding procedure in Part 3 and Chapter 3, Table 5 for code and subcode descriptions.

Also at the last coding stage, I implemented magnitude coding. Magnitude coding refers to capturing the magnitude and how often each code occurred relative to the rest (Saldaña, 2015). My magnitude coding occurred in two parts, first by noting each time a code occurred to get an overall tally of its presence in the data. Second, I indicated where the code occurred in the instructional sequence the code (beginning, middle, or end) to situate its presence in the flow of instruction. By including magnitude coding, I was able to capture how often each code occurred compared to other codes, and how the code functioned within the teacher's instruction based on its place within the instructional sequence. Finally, I simultaneously coded the data. Simultaneous coding allows for individual datum to receive more than one code (Saldaña, 2015).

I allowed for this in the coding process because comprehension-focused mental state utterances within the transcripts often served more than one pedagogical purpose that worked together within the teachers' instruction. For example, one teacher asked her students to remember a past experience of seeing another teacher's facial expression, and another teacher asked her students what they could learn from the character's problem solving within the text.

Autobiographical memory retrieval presents as a skill of perspective-taking (Sutin & Robins, 2008), as does self-reflection and awareness of one's own learning (Gerace et al., 2017). Both *remember* and *learn* are mental states that require the student to engage in a metacognitive process of considering their own perspective. Therefore, the task of remembering facial expressions received a perspective taking code with two simultaneous subcodes, one for facial expressions and one for self-orientation. The task of identifying what can be learned from the character's problem solving received simultaneous coding of both causality- and perspective-taking.

Researcher Positionality. Positionality in qualitative research refers to the ways in which the researchers' perspective effects their analysis of the data (Bogdan & Biklen, 2007). The perspective one holds is formed to varying degrees by and in response to ever-changing personal, group, and societal influences (Bogdan & Biklen, 2007). My positionality as a former middle-grade special education teacher and literacy specialist gives me apt insight into teachers' reading comprehension instruction. As a reading comprehension researcher now, I hold a dual scholar-practitioner vantage point, which allows me to witness teacher instruction as both an insider and an outsider. My insider perspective allows me to observe subtleties that others without such background may miss, and my outsider perspective allows me to keenly interrogate teacher practice with a critical and academic lens. I also hold a firm belief in the power of stories to transform people's views about themselves, others, and the world around them, notably by inducing empathy (Djikic et al., 2013). I have witnessed the transformative effect of stories myself as a reader. My professional and personal experiences, as well as my beliefs, are not divorced from my analysis and support my research by offering an increased level of veracity to the findings.

Part 3 Results

Research Questions 1 and 2. I analyzed 11,724 utterances. The total number of mental state utterances that occurred during the IRAs was 3,454, or 29.5% of all utterances, with 4,155 mental state words across the utterances with a mean of 61.12 (Standard Deviation = 42.49). There were 174 unique mental state words. I then removed 559 (4.7%) of mental state utterances that occurred outside of reading comprehension instruction with a range of 0 to 47 utterances removed per transcript. The most common mental state word across both reading comprehension and non-reading comprehension focused utterances was *think*. The mean number of non-reading comprehension focused utterances that were removed across all the IRA transcripts analyzed was 9.64 (SD = 10.13). The final total of mental state utterances with non-reading comprehension focused utterances removed is 2,986 (25.5%) with a final range of 0 to 165. The final mean is 52.21 (SD = 35.53).

Teacher mental state utterances accounted for 2,638 (74%) of all mental state utterances and student mental state utterances accounted for 864 (24%). Teacher mental state utterances occurred with a range of 3 to 140 per transcript, and students' occurred with a range of 0 to 55 per transcript. The mean number of teacher utterances was 45.48 (SD = 33.06) and the mean number of student utterances was 14.90 (SD = 12.57). See Chapter 3, Tables 6 and 7 for overall mental state prevalence and prevalence by type in teachers' IRAs and in the *Kingdom of Friends* text, respectively.

Not all transcripts reported the length of the shared reading instruction; only 37 transcripts (63%) out of the 58 total I analyzed included that information. Of these, the average duration was 12 minutes and 53 seconds (SD = 3 minutes and 46 seconds) with a range of 6 minutes and 28 seconds to 23 minutes and 20 seconds. The rate per second of mental state

utterance across all teacher and student utterances was one mental state utterance every 12.54 (SD = 13.31) seconds, which decreased to one mental state utterance every 15.38 (SD = 17.60) seconds when I removed the non-reading comprehension focused utterances.

Research Question 3. The analysis I conducted for my third research question suggests teachers use mental state language for three main pedagogical purposes, all of which directly or indirectly support comprehension instruction. These purposes, in order of magnitude, are: 1) teachers use mental-state language to call for students to perspective take, 2) guide students' understanding of story structure, and 3) reinforce and teach social skills and awareness. Across the three top three mental state utterance transcripts, instructional sequences routinely revealed instances of teachers referring to the facial expressions of characters in the text illustrations to help students understand what the character was thinking and how the character was feeling and tasking students with self- and other-oriented perspective taking prompts. Teachers also used mental state language for the pedagogical purpose of guiding children's understanding of the story structure and linking character mental states with subsequent displays of behavior through causal related questions and comments. Finally, teachers used mental state language to teach and reinforce social skills and awareness. These three pedagogical purposes worked in concert to provide students with a rich read-aloud dialogue brimming with mental-state language that built connections between student social background knowledge and the text. In the next sections, I review these three pedagogical purposes in each teacher's read aloud by providing a snapshot of each teacher's IRA instruction. See Chapter 3, Table 8 for coding results by prevalence.

***Ms. Taylor*.** In Ms. Taylor's read-aloud transcript, perspective taking mental references most commonly represented over half of all mental state references, followed by causality and then minimal social skills and awareness mental state references. Before reading, Ms. Taylor

asked students, “Remember when we looked at Mrs. G.’s face and when the mouth turns into a smile that looked like a happy face?” The students nodded in response and Ms. Taylor continued, “Now, look at my face” The students responded in unison, “Angry!” Ms. Taylor then used the student’s knowledge of people’s past facial expressions and her current facial expression to activate knowledge for the read aloud by stating “Yeah, there might be some happy faces and there might be some angry faces in here.” Additionally, during reading, Ms. Taylor would frequently ask questions such as “Does Diego look happy in these pictures?” and “How’s he feeling now in the picture?” to position students’ attention to how the illustrations provide a visual aid that show characters’ mental states and how their mental states change throughout the story. Ms. Taylor followed up these perspective taking questions by asking self-oriented and other-oriented questions such as “Let me see your angry faces.” and “[The characters] are angry, aren’t they?” to further reinforce student understanding.

Structurally coding the data for instructional sequence analysis and magnitude coding for positioning within the instructional sequence further revealed that Ms. Taylor used perspective taking questions and comments as a springboard for other pedagogical purposes by initiating many instructional sequences with such mental state language. In effect, in Ms. Taylor’s instruction included perspective-taking mental state instruction that acted as the knowledge activation students needed to engage in further dialogue related to causality within the story and display attention to the social skills and awareness displayed (or not displayed) by the character’s the text. For example, Ms. Taylor started an instructional sequence by asking a perspective-taking facial expression question, “Does Diego look happy in these pictures?” and how is [Petunia] feeling?” She followed student responses of “Mad!” with a causality behavior

question, “You don’t think they’re going to play anymore? Let’s find out.” to conclude the instructional sequence.

Ms. Taylor’s perspective-driven mental state language use also guided student’s attention to causal story structure indirectly. Throughout the read aloud, Ms. Taylor’s instructional sequences included questions such as, “How is the character feeling?” Ms. Taylor then asked various versions of this question repeatedly throughout the read aloud, “How is the character feeling now?” or “Now how is the character feeling?” in response to subsequent story events. Ms. Taylor further supported her student’s understanding of story structure by asking specific causality questions such as, “What do you think will happen in the next part of the story?” and modeling causal behavior predictions of character behavior, “I guess he will have to build it back and she won’t stomp on it again.” Ms. Taylor’s directive to track changes in the characters’ feelings and make predictions related to story structure and character behavior likely assists her students in not only practicing perspective taking, but also in monitor the introduction and resolution of the story’s central problem.

Ms. Taylor only included a few references to social skills and awareness within her read aloud. Amid the main conflict of the story, Ms. Taylor commented, “Oh, they aren’t being nice, are they? Are they being friendly and kind?” and then followed up this social skills and awareness vocabulary support with “I think they forgot the preschool rules” to reinforce the social norms of preschool. This brief social skills and awareness instructional sequence shows Ms. Taylor taking advantage of the fact that the text supplied her with a social scenario in which to buttress other areas of her teaching.

Ms. Hanna. Ms. Hanna’s read aloud contained longer instructional sequence periods than Ms. Taylor and Ms. Martin. Whereas Ms. Taylor and Ms. Martin punctuated their read aloud

with short and pointed instructional sequences, Ms. Hanna's instructional sequences were more complex. For example, in contrast to Ms. Taylor, Ms. Hanna's perspective-taking instruction emerged as a subordinate element within her causality instruction. For example, Ms. Hanna began an instructional sequence by asking her students a behavior causality question and gradually changed the phrasing to an other-oriented perspective-taking question, "[Petunia] said [Diego] was a scaredy cat and then it started to really bother him, right? And then what happened? What did they do? They got really what?" Ms. Hanna then followed up students' responses of "Angry!" with a self-oriented perspective-taking prompt, "Show me your angry faces" to reinforce how to identify someone (including oneself) as angry. She then asked another causal story structure question in the middle of the instructional sequence, "Do you think they did a very good job solving their problem at first?" Ms. Hanna continued alternating back and forth between causality and perspective-taking questions and comments, before concluding the instructional sequence with a self-oriented perspective-taking utterance, "Sometimes when I'm mad at my sister I do my mean face," which provided a model of her connecting her experiences with the text events. Because Ms. Hanna began this instructional with a causal behavior question and returned to causality-focused questions and comments throughout the instructional sequence, she seemed to employ perspective-taking as a means to support her student's understanding of causality within the story rather than as a separate targeted skill of instruction.

All but one of the social skills and awareness utterances in Ms. Hanna's read aloud occurred after reading. The one outlier occurred as a clarifying, seemingly rhetorical question in the middle of her causality-focused instruction based on an initial causal question of, "How did the problem start?" When the student's struggled in identifying how the problem started by not responding, Ms. Hanna then interjected with the social skills and awareness utterance, "[Diego

and Petunia] were playing nicely, right?" before asking a more pointed causal question, "But who was making all the decisions?" at which point the students answered correctly, "Petunia!" Ms. Hanna's interjection served to situate student attention to the fact that a problem between two friends occurs when playing nicely stops because of one or both friends' actions. Ms. Hanna's subsequent question, which referenced the domineering actions of Petunia from the text (Petunia making all the decisions), allowed the students to recognize that the central problem of the story started when Petunia became too bossy when she was playing with Diego.

As a result of the repeated emphasis on causality, Ms. Hanna's read aloud demonstrated the most mental state utterances to causality, in particular to causal story structure. She consistently made explicit references to the story's central problem before, during, and after the read aloud. In the beginning she stated, "So today—what I want you do is—I want you to look for the problem and see if the problem gets solved." During reading she asked her students, "Do you think the problem is starting?" After reading, Ms. Hanna asked, "So, what do you think you can learn from a story like this and the characters?" and a student responded, "That they solve the problem." Ms. Hanna then probed the student response, "And what can you learn from their problem solving?" The student then responded, "Um, you can learn to not be bossy and be nice to each other." Ms. Hanna then reinforced this idea by connecting it to social skills and awareness by stating, "You know in the first place, if we have to realize we're bossing our friends around maybe we shouldn't do that and you know just learn to work together from the beginning" to complete this instructional sequence. Ms. Hanna appeared to use her references to the central problem before and during reading as a way to inculcate social skills and awareness in her students after reading. Ending the read aloud with a dialogue focused on social skill and awareness signaled Ms. Hanna leveraging the opportunity provided by the text to reinforce social

skills and awareness in her students, similar to how Ms. Taylor supplemented her read-aloud instruction with references to social skills and awareness displayed by the characters.

***Ms. Martin*.** Ms. Martin started and ended her read aloud with social skills and awareness mental-state utterances. Before reading, Ms. Martin asked her students what the golden rule was, to which her students dutifully responded, "Treat others how you want to be treated!" suggesting it is a common refrain in her classroom. Then, during reading, around the climax of the story, Ms. Martin asked her students, "How do you think [Diego and Petunia] are feeling right now" to start an instructional sequence. When the students answered "Mad!" Ms. Martin responded, "They both feel mad and how else might they feel?" to which the students respond, "Sad." Then, Ms. Martin inquired, "Do you think Diego forgives [Petunia]?" to end the instructional sequence. Like with Ms. Taylor, Ms. Martin's during-reading instruction shows a pattern of perspective-taking that serves knowledge activation for the causal behavior question related to the resolution of the story's problem.

After reading, Ms. Martin prompted her students to recall, "So how do you think in this story they worked out their problems?" A student's response showed an established understanding of the story structure based on the characters' mental states, "That [Diego and Petunia] were happy friends at first then they were mad in the middle and then he forgived [sic] her and they were friends again." After affirming her student's response, Ms. Martin further activated her student's knowledge by asking a self-oriented perspective taking question, "Do you ever have that happen where sometimes you and your friend don't get along and go home sad about it?" before she returned to social skills and awareness by asking a student, "[Student name], remember what we say to someone when we know we've done something wrong we tell these people this?" to which the student replied, "Sorry." Ms. Martin affirmed the student's

answer with a slight correction, “Yeah. I forgive you.” Ms. Martin then concluded the read aloud by stating, “And when someone tells you they’re sorry for what they did you have to forgive them and you and you can be friends again.” Additionally, what is also emerging from Ms. Martin’s emphasis on social skills and awareness is that her student’s social-skills vocabulary is supported as a part of their social skills and awareness. A similar consequence emerged out of Ms. Taylor’s social skills and awareness mental state utterances as well. However, neither Ms. Taylor, Ms. Hanna, nor Ms. Martin directly addressed mental state vocabulary, while other teachers with less overall mental state utterances did by explaining the meaning of certain words within the text.

By beginning and ending her read aloud with social skills and awareness mental state utterances Ms. Martin signals her use of the text and read-aloud instruction as a deliberate instructional tool and method to teach and reinforce her students’ social skills. Nonetheless, even if students’ comprehension is not the IRA’s main focus, the during-reading instructional sequences exhibited by Ms. Martin suggest her attentiveness to ensuring that her student’s comprehension of the story’s main problem and resolution benefits her student’s comprehension of the story. Here we can see that the inherent social mechanisms of stories serving a dual role. Ms. Martin’s emphasis on social skills and awareness during her IRA instruction resulted in students’ comprehension gaining a secondary benefit and emerging as byproduct of her social skills instruction.

Further, Ms. Martin’s read aloud is somewhat distinctive from both Ms. Taylor and Ms. Hanna in her phrasing of questions. Ms. Martin used cognitive-state terms and the modulation of assertions more often in her read aloud. In doing so, Ms. Martin is tacitly communicating to her

students the inherently ambiguous nature of others' mental states, which, research demonstrates, is important for developing children's ToM (Peskin & Astington, 2004; Leece et al., 2021).

Part 3 Discussion

As previously reviewed throughout this dissertation, the power of text in providing a simulated social world invites comprehenders to employ socio-cognitive processes while comprehending stories (Mar & Oatley, 2008). The results I report in Part 3 reveal that preschool teachers' IRA instruction often includes a focus on this social aspect of stories as evidenced by consistent mental state language use across the IRAs I analyzed. Further, the identification of specific pedagogical purposes of mental state language during IRA instruction underscores the importance of socio-cognitive skills in reading comprehension. However, empirical evidence for such socio-cognitive processes' role in bolstering reading comprehension outcomes is limited.

Character perspective charts (Emery, 1996) and story maps with character perspectives (Shanahan & Shanahan, 1997) are both reading comprehension strategies that utilize a perspective-taking graphic organizer to support students in adopting character perspectives to better make sense of story events and to highlight narrative structure. Hodges et al. (2018) is the only intervention study to my knowledge investigating the impact of perspective-taking specific instruction on reading comprehension outcomes for struggling readers. Hodges and colleagues found that tasking students with taking the perspective of characters within stories to facilitate story map instruction was an effective method by which to teach reading comprehension in Grade 3. However, participant's perspective-taking was not assessed, so it remains unclear what effect participant's perspective-taking ability has on the intervention effectiveness.

Another study aimed to support students in Grades 2 and 3 who were struggling behaviorally in school and in reading comprehension. Lysaker and colleagues (2011) used a relationally oriented method of a dialogic read-aloud intervention with texts with references to character relationships, thoughts, and feelings. During the read alouds, the dialogue centered around personal connections to the text, character perspective taking, and exhibiting empathy for characters. The participants' scores on pre- and post-measures of narrative comprehension and ToM measures demonstrated relationally-oriented reading to be an effective method of intervention for both comprehension and ToM. However, the narrative comprehension measure was a child's oral narration of a wordless picture book, so it remains unknown whether the effectiveness would translate to listening to or reading a story.

The paucity of reading comprehension intervention research utilizing socio-cognitive processes such as ToM and perspective-taking as part of the intervention method could be the result of common reading comprehension assessments used in research and practice. The effect of ToM and perspective-taking in quantitative analyses that account for other known variables of reading comprehension is small (Kelly et al., in progress; LaRusso et al., 2016). Because of the small effect, researchers, like early grade teachers, may be dissuaded from emphasizing such skills in intervention research. However, as the research on ToM and reading comprehension continues to grow and converge, the inclusion of socio-cognitive processes in reading-comprehension-intervention research may become more prevalent in the coming years. As an example, Davidson and Weismer (2018) queried caregivers of children with autism on their child's book genre preferences, revealing a slight preference for non-fiction texts. Including reading motivation and enjoyment measures could shed light on the unique role socio-cognitive processes play in reading comprehension, which could impact reading comprehension

intervention research by underscoring the importance of reporting text genre and other text characteristics. Still, to my knowledge, no such research exists to date on such relations.

Teacher Variation in Mental State Language Use. The results from Part 3 of the present study perhaps most strongly reveal that teachers vary quite significantly in their use of mental state language during shared reading activities, and therefore, may vary in levels of support for reading comprehension. Teachers were not specifically directed to use mental state language nor highlight the mental states of characters, and still the frequency of mental state language use in some teachers' readings was upwards of 50% of all utterances. Other teachers, however, had 0% of utterances containing a mental state reference. As one example, although some teachers discussed the word *sulk* with their students, most teachers did not. In fact, only one of the top three transcripts with the most mental utterances discussed the word *sulk*.

Of the seven mental-state types I coded, the implicit mental-state references of behavioral and character trait codes comprised nearly 50% of mental-state references across the transcripts. Other implicit mental-state references such as emotional mental-state terms and desire-state terms comprised 24% and 13%, respectively. By comparison, the remaining three mental-state types within the coding scheme are all explicit references and together only comprised 33% of all utterances. As stated by Peskin & Astington (2004), perhaps because implicit references are by nature more abstract for children, teachers were more likely to ask questions and facilitate conversation about such textual references to make clear the inferences required to make meaning of the text, resulting in their increased use. However, the findings of the present study extend previous findings on ToM to include other types of implicit mental states perhaps more germane to narrative comprehension: character traits and behavior descriptions.

In fact, character trait and emotion type frequencies of the transcripts, when compared to their frequency in the *Kingdom of Friends* suggests teachers and students do indeed use these mental states with greater occurrence. Although only 9% and 11% of the words in the *Kingdom of Friends* referenced character traits and emotions, the percent of utterances in the transcripts were 27% and 24%, the highest percentages of all mental state types. The reason for the discrepancy in this sample was the result of the references spurring robust discussion.

Teachers also varied in the extent to which they used modulations of assertions. Although modulations of assertions occurred infrequently, some teachers did use these words more frequently. Teachers that did use modulations of assertions more frequent seemed to use words like *maybe* to indicate the potentiality for multiple answers when asking questions that required making an inference, including when asking about characters mental states. Teachers also however, appeared to use cognitive terms like *think* in this way as well. Given past research on the importance of cognitive mental state language (Peskin & Astington, 2004), it is reasonable to speculate using modulations of assertions or cognitive mental state language when making predictions (i.e., "*maybe* Petunia and Diego will be friends again" or "I think Petunia and Diego will be friends again") supports ToM development because it supports children's understanding that the mental states of others are fluid, changing in response to events and one's acquisition of knowledge (Gopnik & Astington, 1988). Such a finding further corroborates past research that suggests that while most preschool teachers acknowledge the importance of supporting their students' social-emotional growth, some teachers support such development more strongly than others during instruction (Zinsser et al., 2014).

Finally, teachers varied in the extent to which they supported their students' mental state vocabulary development. Narrative texts are rich with mental state references (Zunshine, 2019).

As such, narrative texts present students with unfamiliar and sophisticated mental state vocabulary. For example, the *Kingdom of Friends* text used for the shared reading activity includes the word *sulk* to describe a character's behavior. Some teachers, in response to encountering the word *sulk,* paused. A teacher asked, "Do you know what sulk means?" to which the students responded in chorus with, "No!" The teacher then described the word *sulk* and asked the students to act out *sulk,* introducing new vocabulary that describes the physical actions someone may display when they are feeling sad.

Crucially, such vocabulary support even occurred in the absence of the text's explicit use of mental state language. For instance, the story context prompted mental state discussions that expanded children's vocabulary. A teacher asked, "Did that make Diego happy or sad [student name]." In response, a student answers, "mad" to which the teacher replied, "It kinda made him upset at first. Why was he upset? Because he wanted to what?" Interestingly, even though the teacher gave the student two choices here (happy and sad), the student chose a different answer more appropriate for the question. Then the teacher expanded on the student's word by giving an even more precise word (*upset*) and asking for an explanation to the characters mental state. As I detail in Chapter 1, it stands to reason such vocabulary instruction would result in increased story comprehension given the prospect that narratives contain such content-specific vocabulary (Cartwright, 2015).

Importantly, the teachers in the current study were not prompted to include mental state language in their read alouds, and still their mental state language use suggests alignment with supporting socio-cognitive skills important for reading comprehension. However, the diversity in teacher responses supports the need for new research on developing professional development opportunities for teachers on the importance of mental state language. The focus of such

professional development can support teacher practice in moving from unstructured IRAs to dialogic reading practices similar to that implemented by Lysaker et al. (2011), not only so they can better support students' social-emotional needs, but also their students' reading comprehension development.

General Discussion

A wealth of prior research in child development highlights the importance of mental state references to children's ToM development. Research in reading comprehension is also beginning to name ToM as a variable of potential importance under the linguistic comprehension domain of the SVR. The aim of the present study was to examine the instructional contexts in which mental state language occurs during IRA instruction in order to identify and describe how such language use may support comprehension. I achieved this aim by first piloting a coding scheme, which I then amended, and finally using the adapted coding scheme to analyze transcripts of preschool teachers' IRAs of a common story. My final analysis of mental state utterances within instructional sequences revealed that teachers' use of mental state language enriched the dialogic interactions among teachers and students and does support their comprehension instruction in important ways: by providing perspective taking practice, calling attention to casual relations across the text, and teaching and reinforcing social skills and awareness. Such enrichment likely contributes to the development ToM as well as reading comprehension. Therefore, examining references to mental states during read alouds offers a glimpse into the window where ToM and reading comprehension may meet.

Overall, the results indicate that mental state language frequently appears during IRA instruction, and such use supports skills and strategies important for reading comprehension. Thus, teachers' attention to characters' mental states, while supporting children's social

emotional skills, also supports children's comprehension of narratives. The shared reading transcripts I analyzed reveal teachers use mental state language for the pedagogical purpose of providing their students with perspective taking practice, story structure understanding, and social skills and awareness. The results of the present study align with past research that suggests preschool teachers acknowledge the importance of their students' social and emotional development (Zinsser et al., 2014) and highlight narrative texts as rich in opportunities for supporting and teaching social-emotional awareness (Zunshine, 2019), including ToM specifically (Astington, 1990; Martucci, 2016; Tompkins, 2015).

The findings of the present study offer a foundation for IRA instruction that support students' comprehension skills and social-emotional development that is rooted in current teacher practice. Indeed, experimenter-led and caregiver-led intervention studies aimed to specifically promote references to socio-cognitive themes in order to support preschool aged children's ToM development through shared readings show promise (Aram et al., 2013; Bergman-Deitcher et al., 2021; Tompkins, 2015). Such studies, combined with findings from the present study, provide a template for intervention design with the purpose of supporting comprehension while concurrently supporting ToM development.

Implications for Practice

Taken all together, results from all three parts of the present study reveal a clear overlap between comprehension and mental state language in teachers' care in directing students' attention to and comprehension of the contingencies between story events, characters' mental states and subsequent behavioral responses. As I described in Chapter 1, the Common Core State Standards (CCSS) directly call for student mastery in understanding contingent relations between character's mental states and subsequent behaviors in the reading literature standards beginning

in Grade 2. The Reading Literature (RL) Grade 2 standard RL.2.3 states "Describe how characters in a story respond to major events and challenges." (National Governors Association Center for Best Practices, 2010). As a result, the observing of ToM's role in the CCSS reading literature standards reinforces the role of ToM in reading comprehension, and the results of the present study emphasize early childhood literacy practices as important to instructing students how to use their emerging ToM to achieve mastery with such standards.

Moreover, when considering student responses to such instruction, we see students using ToM reasoning. Take, for example, teacher questions related to story structure and characters' mental states that are more specific, "Why do you think the character has that big tear in her eye?" or "Why did the character want to do it that way?" Varied student responses to such questions signal distinct ToM developmental stages present across the students within each class. Some students responded with accurate references to the character's mental states, "The character is *sad* and *worried* about their friend." In contrast, other students were more action-oriented, "Because the character is crying." As Astington (1990) argues, ToM supports narrative comprehension by assisting in the coordination of Bruner's (1987) two landscapes of narratives—the landscapes of action and consciousness. The former student's response coordinates the two landscapes, while the latter student's response remains squarely within the landscape of action. Interrogating such student responses may offer teachers an indication of their students' ToM, and, at the same time, offer a way to monitor their comprehension of the story at a deeper level. Consequently, ToM-informed interpretations of student responses provide teachers with the opportunity to provide pertinent feedback to students' comprehension of the story events and performance on Common Core reading literature standards, again indicating an

overlap between ToM and reading comprehension instruction. Future professional development in this area should include specific ToM instruction for teachers.

Limitations and Future Research Areas

There are limitations to this study. First, the generalizability is limited to the specific text studied, which is not authentic children's literature, but rather researcher developed. The text is also not expository text, which constrains the findings to the narrative genre. Future research of shared readings utilizing the same coding procedure with authentic children's literature that includes both narrative and expository texts is needed.

A second limitation related to generalizability is that most of the data does not contain multiple transcriptions of shared readings conducted by the same teacher, nor does it contain samples from natural teacher instruction. I also was unable to conduct member checks. As such, the data captured represents more of a moment of time in shared reading instruction and may not be a true representation of individual teacher's shared reading practices within their classrooms. The data do, however, include a substantial number of teacher participants, so this limitation did not affect data saturation. Nonetheless, future studies that collect data with a smaller number of teachers over a longer period of time and during their natural instruction are needed in order to present a more stable illustration of the teachers' shared reading practices. Such data is needed to determine to what degree mental state language prevalence benefits students and if there is a threshold at which the number of mental state references does not benefit students' comprehension.

Third, I do not have student data. The inclusion of student outcome data could provide more robust claims as to how ToM language use by preschool teachers during shared reading instruction contributes to reading comprehension outcomes. Other data pertaining to students,

such as motivation for perspective-taking and overall reading motivation and engagement could play a moderating or mediating role in the effect of mental state language on comprehension (Wigfield et al., 2008; Kim et al., 2018). Controlling for other common individual differences, such as executive functions and word recognition, could further inform the field of mental state language's unique role in reading comprehension outcomes (Kim, 2021). Additionally, I do not have access to teacher data. Measuring teacher's ToM may be important in future studies to determine the effect individual differences teacher ToM has on the use of mental state language across teachers, as is other teacher level demographics. Therefore, future research examining the relation between teachers' mental-state language use and students' reading comprehension performance that includes such individual differences across both teachers and students is needed.

Conclusion

The complexity of reading comprehension continues to motivate innovative studies to examine the multitude of variables that contribute to student reading comprehension outcomes. In the present study, the piloting, amendment, and use of a mental-state coding scheme to examine mental-state language occurrence during IRAs in preschool classrooms revealed how one such variable of interest, ToM, may support comprehension instruction in preschool classrooms. The present study's results help move the field out of theoretical and empirical models of ToM's role in reading comprehension and into exploring the ways in which ToM can help inform the establishment of dialogic reading instructional frameworks and, thus, early childhood reading comprehension. Although teachers may already reference mental states and use mental-state language naturally during IRAs, the deliberate and systematic use of such an instructional framework would support early childhood literacy instruction in better serving

students' reading comprehension as well as their social emotional development. In short, by having access to such instructional practices, teachers' instruction can better meet the needs of the whole child.

Chapter 4: Teaching Theme Using Theory of Mind and Inference Making During Interactive Read Alouds of Authentic Children's Literature

Without a doubt, reading comprehension is an integral, if not the most prominent, skill students need to master for post-secondary success. Students who do not develop strong reading comprehension continue to struggle in post-secondary educational settings and employment (Kerr et al., 2017; Ju et al., 2012; Williams, 2008). The Simple View of Reading (SVR) defines reading comprehension as the product of word reading (e.g., phonics, phonemic awareness, encoding, fluency) and listening comprehension (e.g., understanding texts that are read aloud; Hoover & Gough, 1990). Both sets of skills develop over time as a student learns to read but vary in rate of development, with the decoding domain representing stronger explanatory power in reading comprehension during the early reading stages (Foorman et al., 2018). The listening comprehension domain only becomes dominant once a reader obtains automaticity with decoding sills, typically around Grade 2

The apparent lack of developmental importance of the listening comprehension domain in early grades is misleading because the component skills of listening comprehension begin to develop during pre-, emergent, and beginning literacy activities, but do not exert an influence in measurement models until the student can read text independently (Oakhill & Cain, 2012; Tompkins et al., 2013). Pre-literacy activities refer to children's engagement with storytelling outside of formal reading activities such as a child "reading" a wordless picture book (Paris & Paris, 2003), watching cartoons and movies (Nathanson et al., 2013; van den Broek et al., 1996), or listening to a story told aloud (Paris & Paris, 2003). Pre-literacy activities serve comprehension because they expose children to predictable narrative schemas that they begin to encode to help them make sense of stories they encounter in the future (Lynch et al., 2008;

Rumelhart, 1980). Once children begin to read picture books with others and on their own during emergent and beginning literacy, their developing narrative schemas continue to strengthen as children learn to formally read. Theory of Mind (ToM) is one listening comprehension component skill that is emerging as a novel potential contributor to reading comprehension, which develops, in part, in response to visual stimuli such as movies and cartoons, "reading" wordless pictures, and read aloud activities, such as shared readings between children, their caregivers, and teachers (Dore et al., 2018; Kim, 2020; Mar et al., 2010; Ziv et al., 2014).

ToM is a skill of social cognition, referring to one's ability to identify, or infer, another person's mental state (i.e., thoughts, feelings, beliefs, intentions) and use that knowledge to predict and explain the behavior of others (Wellman, 2002; Wimmer & Perner, 1983). Because mental states occur in one's mind, ToM is inherently a skill of inference making. Indeed, ToM emerges in child development around age three when component skills of the listening comprehension domain are also developing. Therefore, educators of pre-, emergent, and beginning readers have reason to attend to such skills. Yet, the evidence for ToM's role in reading comprehension to date is theoretical and correlational. No studies, to my knowledge, have examined the effects of implementing a comprehension intervention informed by what the field knows about ToM. The aim of the present study was to address this research gap by developing and assessing the effectiveness of a ToM-informed listening comprehension intervention with beginning readers. In the sections that follow, I extend upon my theoretical framework in Chapter 1 to provide support for the development of such an intervention from the current ToM and reading comprehension literature base.

Theoretical Framework

In everyday life, people with adequate social knowledge successfully navigate social interactions (Fiske & Taylor, 2013). Narrative stories mirror everyday life experiences by simulating social interactions and perspective-taking processes like those required by ToM (Mar & Oatley, 2008). Similar inferential processes that undergird ToM (e.g., inferring another's motive to predict another's behavior) apply to the socio-cognitive inferential processes necessary for comprehending narratives (Heyes & Frith, 2014). As such, narrative texts are a great vehicle to display the interplay between socio-cognitive development and comprehension.

To generate an accurate inference is a two-stage process, representing an initial stage of knowledge activation and a subsequent stage of knowledge integration (Kendeou, 2015). Once knowledge is activated, the second and final stage of inference making is integrating the activated knowledge with implicit information in the text. Socio-cognitive inferential processes, such as ToM, assist in comprehending narratives because a narrative's structure is driven by the characters' mental states (Lynch et al., 2008). Although references to mental states are prevalent in children's literature (Zunshine, 2019), they are often implicit. Readers of narrative texts, therefore, must use their ToM skill to interpret the implicit social aspects of the story from the characters' perspectives by making inferences (Duke & Cartwright, 2018). As I describe in depth in Chapter 1, comprehenders with stronger ToM skill make more accurate inferences that cohere text structure resulting in greater comprehension of the text's macrostructure (Dore et al., 2018). Such macrostructure comprehension assists theme identification by providing a social world microcosm (Mar & Oatley, 2008) that comprehenders can then generalize from in order to make sense of larger universal concepts (i.e., identifying theme; Graesser, 2007). In line with this theoretical frame, prior research on theme identification highlights the importance of targeting narrative text structure when teaching theme identification (Williams, 2005).

Within a narrative comprehension context, this ToM inferential process results from the child's ability to first, take the perspective of the character and accurately determine their mental state (i.e., activate knowledge), and second, use that knowledge to explain or predict the character's behavior within the story (i.e., integration). In this way, comprehenders who engage their ToM skill during reading or listening to a narrative do so by making knowledge-based socio-cognitive inferences about the story's social context. First, comprehenders' knowledge is activated through references to characters' mental states, which provides the prior knowledge necessary to generate an inference. One known pitfall of accurate inference making is the activation of irrelevant or incorrect knowledge, which then leads to inaccurate integration and an overall inaccurate inference. Comprehenders that recruit their perspective-taking skill more efficiently likely catch and fix irrelevant and incorrect knowledge as a form of comprehension monitoring before completing the inference by integrating the activated knowledge to their emerging mental representation of the text.

Comprehension, ToM, and Inference Making

Comprehension of text is a latent skill, meaning it cannot be directly measured. As a result, readers must show their comprehension in some manner to provide a proxy that serves as a measurable product of comprehension. One common proxy for *deep* reading comprehension is the accurate identification of the main idea or theme. The identification of narrative themes, however, is one of the most difficult challenges readers face when comprehending stories (Graesser et al., 2002). In contrast to text-based, local inference making, such as anaphoric inference making linking pronouns with appropriate referents, theme identification is not done spontaneously during reading and, therefore, is effortful and initiated by the reader (Zwaan and Radvansky, 1998). In addition, narrative themes are rarely, if ever, explicitly

stated in a text, rendering their identification a product of inference making. This type of skill is a global inference, which requires a reader to make sense of how textual knowledge fits with their background knowledge and, therefore, represents situational level comprehension, the deepest level of comprehension where learning is purported to occur (Graesser, 2015; Kintsch, 1998). Because themes are implicit and identified with effort by the reader, theme identification must be explicitly taught.

Students at risk for reading deficits display peculiar difficulty with theme identification (Gersten et al., 2001). The process of producing a theme statement tasks students with demonstrating their knowledge of a text in a more cognitively-challenging way than basic recall (Krathwohl, 2002). Samples of adolescent and elementary students at risk for or with learning disabilities showed theme awareness and an understanding of basic theme concepts (e.g., what a theme is) as measured by open-ended researcher-developed questions that were scaffolded with prompts as necessary. The participants were unable to produce a coherent theme statement, however, despite explicit instruction and sentence starters (Williams et al., 1994). The discrepancy between theme awareness and the production of a theme statement is similar to a more general reading difficulty such students face in strategically using reading comprehension strategies (Gersten et al., 2001). Nevertheless, prior research suggests teaching theme identification is malleable in response to intervention when it is explicitly taught using a structured approach (Wilder & Williams, 2001; Williams, 1994; *Cohen's d* = 1.41 to 5.93 as reported in Edmonds et al., 2009). Further efforts from Williams et al. (2002), using a similar structured system of questioning and explicit instruction, found a similar pattern with students in Grades 2 and 3. Participants showed theme awareness and basic theme concept understanding, but were unable to produce a coherent theme statement.

The successful generation of accurate global, thematic inferences occurs through comprehension of the story's macro structure (Graesser, 2007). As stated earlier, children begin to encode narrative structure before formal schooling in response to familial shared reading practices and also in response to television or other forms of oral storytelling (Nathanson et al., 2013). Once children begin formal schooling, narrative structure is explicitly taught. Educators, for example, outline narrative structure as consisting of three basic elements: (1) an introduction to the protagonist and main goal, (2) the protagonist's struggles to achieve the main goal, and (3) the protagonist's end to the journey in achieving the main goal, (Chapter 3, Figure 1). The glue that links these three distinct elements of narrative structure into an overall understanding of what the story is generally about (i.e., theme) is inference making (Cain, 2010; Kintsch, 1988; van den Broek et al., 2015). As an example, in the classic fable "Little Red Riding Hood," a reader's inference-making abilities assist them in linking the ignorance of Little Red Riding Hood to her subsequent behavior (i.e., lack of fear) when encountering the disguised wolf once she arrives at her grandmother's house (Astington, 1990).

As students' ToM emerges, they can begin to attend to characters' mental states by inferring their thoughts, feelings, and intentions, which improves their understanding of the overall macro narrative structure (Astington, 1990; Bruner, 1987; Duke & Cartwright, 2018; Dore et al., 2018; Kintsch, 1988; Lynch et al., 2008). Without the ability to attend to macrostructure, students will likely incorrectly identify the theme of the story. Narrative themes are often nominalized mental states referenced or ascribed to characters throughout the story. For example, *The Diary of Anne Frank* invokes themes of resiliency and hope in the face of insurmountable hardship. Clearly, attention to Anne's mental states (i.e., referring to her as resilient or hopeful) and many other characters as the story progresses, as well as monitoring

temporal changes in characters' mental states in response to story events (i.e., Anne's mental state changing from hopeless to hopeful), contributes to accurately extracting such themes from the story. As a result, comprehenders with weaker ToM skill are likely less attentive to characters' mental states, resulting in weaker and less-frequent inference making, which contributes to weaker understanding of macrostructure, reducing the likelihood of accurate theme identification. Making ToM processes (i.e., social inference making) overt through explicit questioning and feedback and explicit instruction on teaching theme statement production, together, is likely to improve theme-identification performance.

Indeed, prior research suggests that weak perspective-taking abilities (a characterization of impaired ToM) appear to contribute to lower reading comprehension performance (Donahue, 2014) and that explicitly guiding children to take the perspective of characters does increase reading comprehension performance (Hodges et al., 2018). However, even if readers readily take the perspective of characters, they may lack the mental state vocabulary, thus resulting in a breakdown of the ToM inference-making process due to lack of or incorrect knowledge activation. To further compound this issue, even if readers readily take the characters' perspective and have adequate mental state vocabulary, the strategic use of such knowledge in the form of successful integration of the generated inference presents as a particular area of weakness for less-skilled readers. In a study conducted by Carlson et al. (2014), skilled and less-skilled comprehenders did not differ on text-based inference making but did significantly differ in knowledge-based inferencing, especially so in response to knowledge-based inference making demands that required connecting knowledge to related text-based information. While a general lack of knowledge likely plays a role in such a finding, other causes could be the result of deficits in lower-order skills critical to successful

comprehension, such as executive functions (Bowyer-Crane & Snowling, 2005; Butterfuss & Kendeou, 2018; Kim, 2017; 2019; 2020). Therefore, ToM inference-making training for narrative comprehension must attend to all three potential pitfalls by supporting children's, mental state vocabulary, perspective-taking, and inference integration, and must do so using explicit instruction techniques to accommodate potential effects of individual differences in executive functions (Siregar, 2021).

In the subsections below, I detail how I accomplished meeting these components in the development of my intervention using literature from ToM and inference-making trainings. Prior research indicates the instructional practice of interactive read aloud (IRA), a shared reading process that involves the teacher reading aloud a text and interacting with a group of children, using a corresponding graphic organizer is an especially appropriate context for teaching such skills (Haagar & Vaughn, 2013; Lennox, 2013; Swanson et al., 2011).

Theory of Mind Training: A Blueprint for Knowledge Activation

For children needing targeted ToM support, ToM interventions demonstrate promising results (Hofmann et al., 2016). From a meta-analysis conducted by Hofmann et al., the majority of training studies utilized narrative passage excerpts to provide perspective-taking scenarios from which to ask questions about characters' mental states in the stories and correct children's mental state attributions during the task. Such trainings, which aimed to remediate ToM weaknesses, provided me with a starting point for developing a comprehension-focused intervention by offering a blueprint for how to activate and teach social background knowledge.

For example, Guajardo & Watson (2002) conducted two ToM training studies by engaging students in a dialogic reading of a narrative text with an adult instructed to focus on asking questions about and explaining the mental state (in this case, false belief) elements within

the story. In the first study, the dialogic reading took place with a small group of students and in the second, the training took place with individual students. The individual training effect size was $g = 0.537$ in comparison to $g = 0.183$ for the small group training. Tompkins (2015) implemented a similar one-on-one ToM training study with a sample of children from low socio-economic strata using children's literature. In contrast to Guajardo & Watson (2002), Tompkins included questions and conversations using emotional states in addition to false belief conversations. Children in the experimental group outperformed those in the control group as well as another storybook comparison group that read the same stories but without the ToM discussion component. This effect held for all outcome measures (i.e., false belief understanding, social competence, and emotion understanding).

Another set of ToM trainings focused on mental state attribution accuracy. Lecce, et al. (2014) conducted a training study that asked students to infer characters' mental states from researcher-developed narrative passages. The training included the trainer reading the passage with a small group of students between nine- and ten-years-old and asking questions related to the characters' mental states. After asking questions and eliciting student responses, the trainer gave the participants feedback on the accuracy of their mental-state attributions. The training group significantly outperformed the control group post-test and maintained better performance over time. Bianco and Lecce (2016) next tested the efficacy of Lecce et al.'s training in naturalistic classroom settings with classroom teachers and students between eight- and nine-years-old. Teachers and small groups of students read narratives together. During the read aloud, teachers would stop and ask the students to accurately identify characters' mental states and provide corrective or affirmative feedback in response to student replies. The results indicated that training children's ToM could be effective in naturalistic classroom settings and

when conducted alongside typical instructional practices. In yet another follow-up study by Bianco and colleagues (2019), the experimenter-led trainings in classrooms given to students in small groups resulted in positive ToM growth in even younger children, aged seven- to eight-years-old. However, in all studies mentioned, including the author's, researcher-developed narrative passages were used rather than children's literature.

Taken together, this brief review of ToM trainings using narratives provides important takeaways for comprehension-focused intervention development targeting ToM-directed inference making. First, one-on-one settings using children's literature as the basis for ToM interventions may lead to better outcomes than small or whole group settings. Second, explicit instruction geared toward explaining ToM concepts, including modeling mental state attribution and practice opportunities with corrective and affirmative feedback is key. Third, such instruction has the potential to transfer to classroom settings with the teacher as the implementer. By implementing ToM intervention components, I fulfill the requirement of supporting participants' mental state vocabulary and perspective-taking.

Inference Training: How to Integrate Activated Knowledge

Meta-analytic evidence suggests less-skilled readers benefit more from direct inference instruction than skilled readers on both literal (less skilled reader $d = 0.97$ and skilled reader $d = 0.06$) and non-literal comprehension measures (less skilled reader $d = 0.8$ and skilled reader $d = 0.55$; Elleman, 2017), and inference training may be an important skill for children with learning disabilities in particular (Hall, 2016). Because of inference making's integral role in successful reading comprehension (Kendeou, 2015), explicitly teaching inference making is likely beneficial for both general and inferential reading comprehension. Across the Hall and Elleman reviews, text genres used in the interventions reviewed varied across the studies.

General procedures used to teach inference making also varied, consisting of activating background knowledge, asking why and how questions, delivering corrective feedback, summarizing, and using graphic organizers. In general, the amount of instruction was minimal, with an average of 10 hours reported in Elleman's meta-analysis, suggesting brief interventions can result in meaningful gains.

McMaster and Espin (2017) conducted a more robust review of inference making intervention procedures based on two broad effective categories of inference making instruction identified by McKewon et al. (2009)—content-focused instruction and strategy instruction. Content-focused instruction included background knowledge activation where interventionists pointed out important text parts and words, pre-taught vocabulary, asked inferential questions, and provided structured feedback. Strategy instruction methods were more explicit in linking parts of text together and text with background knowledge, teaching metacognition (e.g., thinking aloud, self-questioning) and the use of graphic organizers. However, neither method appeared more beneficial than any other. McMaster and Espin report five elements as the most important for teaching inference making that reflect the core components of explicit instruction (Archer & Hughes, 2010).

Across the inference making intervention studies included in the reviews discussed above (Ellman, 2017; Hall, 2016; McMaster & Espin, 2017), few focused on beginning readers, and none included beginning reader participants. McMaster and colleagues, however, are developing web-based interactive software designed to teach inference making for comprehension in Grades 1 and 2 using visual narratives that show initial promise from field experiments (Technology-Based Early Language Comprehension Intervention [TeLCI]; Kendeou et al., 2017). McMaster et al. (2019) extends specific guidance on inference making

intervention development in recounting their efforts toward the development of TeLCI. Based on the results of the McMaster and Espin (2017) and McMaster et al. (2019) guidance documents, I used the design highlighted across both to structure TIER so that it aligns with best practices in the literature for teaching inference making to early readers. I provide a table of the intervention components I include that correspond to each element highlighted by McMaster and colleagues in Chapter 4, Table 1.

Theory of Change

The purpose of this study was to examine the effects of a dialogic read aloud on narrative listening comprehension for early elementary students. I used read alouds of authentic children's literature to implement TIER (Theme and Inference making for Emerging Readers). Due to the COVID-19 pandemic, I executed TIER online using Amazon's Kindle Cloud Reader and Zoom. TIER is a 1:1 read aloud intervention that utilizes an explicit instruction approach with strategic inference making questioning and feedback. I designed the TIER framework from ToM and inference making intervention literature bases to supporting beginning readers' theme identification and statement formation. I posit in the theory of change that strategically using ToM inference making questioning and feedback during read alouds to support explicit instruction on theme identification and statement formation will improve the target students' narrative listening comprehension. See Chapter 4 Figure 1 for the theory of change. Specifically, this study sought to answer the following research question: to what degree does TIER improve the listening comprehension outcomes of beginning readers who are less-skilled comprehenders as measured by their theme statement accuracy?

Method

To answer my research question, I conducted an experimental study using a multiprobe multiple baseline (MPMB) single case research design (SCRD) across participants. The purpose of using an SCRD is to establish a functional relation in behavior change (i.e., theme identification) between baseline and intervention sessions. SCRD is a widely-used experimental research design whereby a functional relation is established through the manipulation of the independent variable across distinct phases, resulting in the interpretation of a causal relationship between dependent and independent variable(s). In the present study, I included three distinct phases across three participants resulting in within- and across-participant experimental control. (Ledford & Gast, 2018). MPMB is an appropriate research design given the evidence from prior studies that suggest one-on-one ToM training is most effective (Hofmann et al., 2016), and given the nature of the intervention content, instructional effects cannot be removed through a reversal or withdrawal design.

Participants and Setting

I planned the intervention with the WWC SCD standards in mind. WWC SCD design standards requires at least three manipulations of the independent variable and three participants (WWC, 2020). I gave each caregiver a participant demographic survey to complete (see Appendix A, Figure 1). One male and one female student, both eight-years-old and in Grade 2, and a nine-year-old female in Grade 3 participated in the study for a total of three participants. My rationale for selecting children in Grades 2 and 3 was that the Common Core State Standards first introduce the skill of theme identification beginning in Grade 2, and ToM is adequately established in typical populations by this age range (Wellman & Liu, 2004). I excluded emergent bilinguals and students receiving special education services. All participants

attended public schools. All names used here are pseudonyms to protect the anonymity of the participants. Kai was an eight-year-old, second grade male student of native Hawaiian descent who received free or reduced-price lunch. Kai demonstrated slightly-delayed ToM based on his Theory of Mind Task Battery (ToMTB). Kai passed all the early and basic ToM tasks, but one. Kai failed four advanced tasks, two due to failing the control questions. His score when compared to an age-matched sample (Chapman et al., 2006) on the RMET was above the mean by approximately four questions. Kelani was an eight-year-old, second grade female student of Hawaiian descent who received free or reduced-price lunch and was enrolled in a dual language, Hawaiian and English, school program. Kelani showed delayed ToM based on the ToMTB. Kelani passed all the early ToM questions, but failed two basic ToM questions. One fail in basic ToM was due to Kelani failing the control question. Kelani passed one advanced ToM task, failed another task, and failed three due to failing the control questions. Kelani's score on the RMET was below the aged-matched sample by approximately two questions. Kelsey was a nine-year-old third grade female white student with a Section 504 plan in school. Kelsey demonstrated developmentally appropriate ToM based on his ToMTB, passing all the early and basic ToM tasks, but failing two advanced tasks, one due to failing the control question. His score when compared to an age-matched sample on the RMET is above the mean by approximately four questions. See Chapter 4 Table 2 for participants' pre-test scores.

I initially sought to recruit participants at local elementary schools, but due to the ongoing COVID-19 pandemic, I resorted to recruiting study participants through word of mouth in my community and then communicating via email to caregivers who expressed interest in acquiring comprehension support for their children. Before starting the screening process, I gathered parent consent, and before every session, including pre-assessment and intervention, I

also obtained verbal student assent. I held all sessions on Zoom. I used Amazon's Kindle Cloud Reader service to obtain the texts used for the intervention and to conduct the read alouds. For a list of instructional books see Appendix A, Figure 2.

Selection Process and Measures

I initially aimed to include only participants with comprehension-specific deficits. In most cases, struggling readers' deficits coincide with deficits in decoding (Gough & Tunmer, 1986), but approximately 10% of all children have comprehension-specific deficits defined as deficits in linguistic comprehension without deficits in decoding (Landi & Ryherd, 2017). Such students are identified as scoring below the 25th percentile in listening comprehension and above the 40th percentile in word recognition and fluency (Catts et al, 2006). I aimed to only include this subgroup because such students present peculiar weaknesses in the oral and cognitive sub-skills within the listening comprehension domain (Clarke et al., 2010; Petscher et al., 2018). One specific area of linguistic comprehension weakness for students with comprehension specific deficits is inference making (Cain, 2010; Carlson et al., 2014; Gersten et al., 2001). As a result, readers with comprehension-specific deficits present as prime candidates for a listening comprehension intervention. However, due to the ongoing COVID-19 pandemic, recruitment was difficult because I could not access schools. Therefore, I modified my selection procedures to include less-skilled comprehenders, defined as readers who scored below the 50th percentile in reading comprehension both with and without commensurate deficits in fluency. As shown by participants' pre-test scores in Chapter 4, Table 2, there was one participant, Kelani, who demonstrated comprehension-specific deficits.

To identify less-skilled comprehenders for the study, I gave three pre-assessments to each participant. Each assessment lasted approximately 15 minutes, resulting in approximately

45 minutes of pre-assessment time per student. First, I administered the Woodcock Johnson-IV test of Achievement (WJIV) oral comprehension subtest (Schrank, et at., 2014) to identify participants with less-skilled comprehension, as measured by scores at or below the 50th percentile. The WJIV oral comprehension subtest is a standardized, individual-administered measure that includes cloze sentences read aloud to the student by a recording. Participants listen to a sentence or passage and identify the missing word or phrase. Median split-half reliability in the examiner's manual for ages 5–19 is reported as 0.82 (Mather & Wendling, 2014).

I then administered the final pre-testing measures, the ToM Task Battery (ToMTB; Hutchins & Prelock, 2010), to assess children's cognitive ToM, and the Reading the Mind in the Eyes (RMET; Baron-Cohen et al., 2001) to assess children's affective ToM. The ToMTB is a standardized individual-administered measure that includes a series of social scenarios described with words and pictures followed by associated questions, including both control questions that assess the student's basic comprehension of the scenario and target ToM assessment questions. The ToMTB consists of five early ToM tasks (developmental age around 1–3.5 years of age), five basic ToM (developmental age around 3.5–6.5 years), and five advanced ToM tasks (around 6.5–8 years). Internal consistency in the ToMTB technical manual for a typically developing sample, ages 2–9, is reported as 0.94, and criterion-related validity with the Theory of Mind Inventory (a caregiver report measure of ToM) is $r = 0.66$ (Hutchins et al., 2012). The RMET is a measure of ToM that predominantly assesses affective ToM. The RMET presents images of individuals' eyes to assess how well participants can discern the mental state of the person depicted, given images of their eyes. In the child version of the test, participants are shown 28 black and white picture of eyes and given four possible answer choices (e.g., angry, surprised, hate, relaxed). The participant is asked to identify which emotion is shown in the eye picture.

Previous studies with participants aged 8–10 report internal consistency of the RMET child version as 0.68 (Vogindroukas et al. 2014).

I also included a researcher-developed measure of social validity to give to each participant at the end of their involvement (see Appendix A, Figure 3).

Intervention Procedures

The independent variable in this study is the interactive read aloud instruction. The intervention procedures consisted of the interventionist (myself) reading a narrative children's picture book using the Kindle app over Zoom with participants in a one-on-one setting. I began each session by asking, "Are you ready to work with me today" to obtain verbal assent before starting. I then shared my screen with the participant that showed the Kindle book. Before reading, I reviewed three key vocabulary words (mental state vocabulary) and set the purpose for reading (to identify theme). I also reviewed the background knowledge necessary for comprehension of the story. For example, a main character in one of the books used is described as an engineer. Before reading the story, I explained what an engineer was to the participant. As I read the stories with the target participant, I paused at planned key points in the text to ask the participant questions. I asked questions during the IRA because Freed and Cain (2017) identified that students benefitted more from asking questions during IRAs than after.

For each key point in the story that I identified prior to the lesson, I included two questions. The first question was designed to activate the participant's background knowledge by asking the participant to attribute an appropriate mental state to a character or explain the meaning of a specific reference to a characters' mental state. The second question was causal, which was intended to integrate the activated background knowledge into the story by making sense of the characters' behaviors. I deliberately sequenced the questions in this way to scaffold

the two-stage inference-making process. I provided corrective feedback or reinforcement of the correct answer in reply to participant's accurate or inaccurate responses.

After reading, I filled out a graphic organizer (See Appendix B: Sample Lesson) with the participant in order summarize the text and have the participant identify the theme. I first asked the participant what goes into each space. If the participant gave a correct response, I typed out their answer in the corresponding space on the graphic organizer. If the participant did not know or gave an incorrect response, I filled in the correct answer and verbalized the correct answer to the student. If the participant gave an incomplete response, I filled in the rest of the response and verbalized the correct answer to the student. Graphic organizers are visual representations of ideas and structures within a narrative and expository texts (Dexter & Hughes, 2011). The graphic organizer I used for this study visually represents how inferences made during the text reading support understanding the typical structure of narratives (e.g., introduction, the main character, event, the central problem, the central problem's resolution). In addition, the graphic organizer visually represents how the inferences made during reading and the overall structure help identify the story's theme. I scripted all lessons before implementation (see Appendix B: Sample Lesson).

Dependent Variable and Measure

I used a narrative protocol by Peskin and Astington (2004), to measure the distal listening comprehension dependent variable of theme identification. The protocol assesses narrative comprehension by considering how children use ToM to integrate the landscapes of consciousness and actions in stories based on Bruner's (1987) framework of narrative comprehension, which posits narrative comprehension as the product of a reader's ability to integrate the action events in the story with characters' mental states. Student responses are

scored as one of three categories: Action-Only, Consciousness-Only, or Coordinated Action-Consciousness. Peskin and Astington (2004) reported inter-rater reliability for each category: Action-Only, $\kappa = 0.91$; Consciousness-Only, $\kappa = 0.76$; Coordinated Action-Consciousness, $\kappa = 0.98$. In the present study, I used the code for Coordinated Action-Consciousness to assess children's theme statements. This means that children must first identify a mental state word (1 point; 0.5 for inaccurate mental state word) and an action from the story (1 point; 0.5 for an inaccurate action). To produce a theme statement, the mental state word and action must be linked through the child explaining how the mental state assists in completing the action (1 point; 0.5 for linking the inaccurate statements together). A participant earns a minimum score of zero if they do not identify any part of the theme statement and a maximum score of three if they identify all parts. Astington (1990) extends Bruner's theory by reasoning that ToM is the method by which children begin to link the two landscapes to comprehend a narrative. Astington's claim has considerable empirical evidence (Curenton, 2004, 2011; Riggio & Cassidy, 2009; Ronfard & Harris, 2014) that lends support to the use of Peskin and Astington's narrative comprehension protocol (see Appendix A, Figure 4 for Peskin and Astington's narrative comprehension protocol). Here is an example theme statement that would earn a score of three: Hope (mental state) is important because it helps us keep going (action) even if something is hard (story integration).

Book Selection

At minimum, I needed 19 books to complete the intervention, but because I gave a book-familiarity survey of the 19 pre-identified books and discovered participants were familiar with some instructional books, I added five more as alternates for a total of 24 books (see Appendix A, Figure 5 for the book familiarity survey). I chose books by contacting authors of

ToM training studies and identifying ones used in prior studies. I also consulted ToM literature for books rich with mental state themes and lessons, such as folktales (Astington, 1990). I included the Lexile ranges for all 24 books to establish stability in text difficulty (Grade 2 Lexile range = 420–800; Williamson et al., 2012). Lexile scores classify books by difficulty, frequently using the modifier AD meaning "Adult Directed" before the Lexile level. An adult-directed book means the book is better read by an adult with a child than by the child independently. Most books chosen for the intervention contained the AD modifier.

Experimental Design

To enact a multiprobe multiple baseline SCD design across participants, I randomly assigned students to an order that determined whether they received the intervention first, second, or third. Then, I sequentially introduced the intervention across assigned participants. I collected at least three baseline data points for all participants prior to beginning the intervention to establish a stable baseline for each participant. I gathered the first three baseline data points for all participants on the same three consecutive days before introducing the intervention to the first student. After day three, I collected one more baseline item for participants two and three at the start of participant one's intervention. I then continued collecting baseline data for participant two for two more sessions before introducing the intervention. I repeated the same process with participant three when participant two started the intervention phase.

Baseline. There were a total of 16 baseline sessions across participants, which ranged from 3 minutes and 48 seconds to 15 minutes and 40 seconds with a mean of 9 minutes and 31 seconds. The time varied due to different text lengths and student response times to question prompts. I did not give treatment instruction to participants during the baseline phase. Before reading, I stated the purpose for the read aloud and provided necessary background knowledge,

but did not review mental-state vocabulary. During reading, I stopped and asked the inference making questions, but did not provide explicit feedback to the participants. After reading, I asked the participants to identify the theme of the story, but did not fill out a graphic organizer nor provide a summary for the participants.

Explicit Instruction Session. After I established a stable baseline with each participant, I provided one treatment-only session inclusive of the ToM inference questions with explicit feedback, where I explicitly modeled how to write a theme statement for the participant. There were a total of three explicit instruction sessions across participants, one for each participant. Before reading, I stated the purpose of the read aloud, provided necessary background knowledge, and reviewed the three key mental state vocabulary words for the text. While reading, I stopped and asked the inference-making questions and provided explicit feedback to the participants' responses. After reading, I filled out a graphic organizer and provided a summary for the participants. I concluded the sessions by asking the participants to identify the story's theme. I did not score participant responses during the explicit instruction session.

Treatment + assessment probe sessions. There were a total of 15 treatment + assessment probe sessions across participants, five per participant. Invention sessions ranged from 10 minutes and 29 seconds to 23 minutes and 5 seconds, with a mean of 17 minutes and 15 seconds. Intervention sessions across participants ranged from 3 minutes and 48 seconds to 13 minutes with a mean of 9 minutes and 31 seconds. Following the treatment only session, there were a total at least five treatment + assessment probes sessions. During the treatment + assessment probe sessions, I conducted the IRA. I asked questions at the pre-established stopping points, noted students' accuracy of response, and provided feedback. Finally, I finished with providing a summary of the text using the graphic organizer referenced. I then asked the

participant to produce a theme statement, which the participants did orally. I asked the participants to identify the theme of the story by forming a theme statement through the 3-question theme statement prompt.

Post-Intervention. One week after the final treatment + assessment probe session, I administered a final assessment probe to determine if participants maintained the target skill of the theme statement formation in the absence of the intervention. Each participant received one post-intervention assessment probe for a total of three post-intervention sessions, one for each participant.

Procedural Fidelity, Reliability, and Interobserver Agreement

Procedural fidelity data was collected across 30% of baseline, intervention, and post-intervention sessions. I trained one outside researcher, an assistant professor, to assess my fidelity of implementation on all intervention lessons using a fidelity checklist (see Appendix A, Figures 6 and 7 for intervention and baseline fidelity checklists). No treatment session scored below a 3 (<80% of fidelity items complete) in my data analysis. Implementation fidelity ranged from 96–91% for the baseline and 81–92% for intervention. Fidelity was 100% for the one instruction only session and 91% for post-intervention.

I trained another assistant professor with a specialty in narrative comprehension to serve as a second rater for theme statement coding of the selected lessons to obtain interobserver agreement for reliability before I completed data analysis, in line with SCD design standards (Kratochwill et al., 2013). I established interobserver agreement by dividing the number of agreements by the total number of agreements and disagreements and multiplying by 100. I did not continue data analysis until interobserver agreement reached a threshold of 80%,

as identified in the What Works Clearinghouse (WWC) standards (WWC, 2020). The IOA was 88% for theme identification.

Data Analysis

To evaluate the effects of interactive read aloud intervention, I measured narrative listening comprehension in three less-skilled readers. The data I obtained is presented in Chapter 4, Figure 2. The median scores for baseline and intervention and post-intervention phases for each participant are displayed in Chapter 4, Table 3. I used visual analysis to determine whether there was a functional relation between the intervention and students' theme identification by examining the percentage of non-overlapping data as well as the percentage of data points exceeding the mean (Parker et al., 2011). I also evaluated the graphed data for level, trend, variability, immediacy of the effect, and overlap within each phase and across phases.

Results

Kai

During the baseline phase, Kai demonstrated a stable low level of theme statement (Median = 0.5, range 0.5–1). After the intervention was introduced, his theme statement accuracy score showed an increase in trend with an immediate effect that was slightly variable throughout the intervention phases but remained above the baseline median score. In the intervention session, his theme statement accuracy score increased to a median of 2.50 with a range of 2–3. In three of out of the six intervention sessions I conducted with Kai, he scored a 3, which is the maximum score. In three other sessions, he scored a 2. This indicates no overlap (0%) between his baseline and intervention sessions and 100% of intervention sessions exceeding the median. It appears his theme statement accuracy was maintained at a stable level after the week-long intervention absence as indicated by his post-intervention probe session in which he scored a 2.

Kelani

During the baseline phase, Kelani demonstrated a stable low level of theme statement accuracy (Median = 0.25, range 0–0.5). After the intervention was introduced, her theme statement accuracy score showed an immediate effect that was overall stable but with a slight decrease in trend. In the intervention sessions, her theme statement accuracy score increased to a median of 2 with a range of 1–3. In one of Kelani's six intervention sessions, she scored a 3, which is the maximum score. In three sessions, she scored a 2, and in one session, she scored a 1. This indicates no overlap (0%) between her baseline and intervention sessions and 100% of intervention sessions exceeding the median. It appears her theme statement accuracy was maintained at a stable level after the week-long intervention absence as indicated by her post-intervention probe session, in which she scored a 1.5.

Kelsey

During the baseline phase, Kelsey demonstrated a low to moderate level of theme statement accuracy with some variability (Median= .5, range 0.5–1). After the intervention was introduced, her theme statement accuracy score showed an immediate effect and an increasing trend that was stable throughout the intervention phases. In the intervention session, her theme statement accuracy score increased to a median of 2 with a range of 1–3. In one out Kelsey's six intervention sessions, she scored a 3, which is the maximum score. In four sessions, she scored a 2, and in one session she scored a 1. This indicates an overlap of one data points (20%) between her baseline and intervention sessions and 100% of intervention sessions exceeding the median. It appears her theme statement accuracy was maintained after the week-long intervention absence as indicated by her post-intervention probe session, in which she scored a 2.

Social Validity

Intervention participants completed the social validity measure by telling me which picture represented their feelings best. All participants reported that they enjoyed reading with adults and that, when they are in school, reading with adults is their favorite part of the day. Additionally, all participants reported that reading with adults helps them better understand the story and become a better reader. Kelani and Kelsey reported that they feel they learn a lot when they read with adults. Kai reported he felt neutral about his learning when he reads with adults. No participant reported a negative response.

Discussion

The purpose of the present study was to evaluate the effects of a dialogic reading intervention on narrative listening comprehension with three participants who present as less-skilled comprehenders. The findings revealed that all participants demonstrated improvements in narrative listening comprehension as measured by their scores in theme statement accuracy between baseline and intervention sessions. This finding suggests a functional relation between the dialogic reading intervention and theme statement accuracy. In addition, all three participants maintained their improved theme identification and statement formation when assessed after a weeklong absence of instruction. Overall, this study demonstrates that less-skilled comprehenders who are learning to read can form adequate theme statements when given explicit instruction during dialogic reading activities aimed to support their socio-cognitive knowledge of the text.

Explicit Instruction

The dialogic reading intervention involved numerous components that may explain the improvements in theme statement accuracy of the participants. Overarchingly, the principal benefit of the intervention is its alignment with the five essential components of explicit

instruction: 1) segmentation, 2) models, 3) scaffolds, 4) practice opportunities, and 5) feedback (Archer & Hughes, 2010; Hughes et al., 2017).

Segmentation (1) refers to breaking down complex tasks into component parts. I segmented the task of forming a theme statement by breaking the theme statement down into the three main parts of identifying the most important mental-state word to the story, explaining why it is important, and connecting it to the story's context. Modeling (2) refers to a clear demonstration of the internal thought processes that occur when working through a complex task. I included one day of explicit modeling by the interventionist without participant assessment. In addition, the remainder of assessment days included explicit modeling of correct theme statements in the form of corrective feedback. Scaffolding (3) refers to the systematic use of prompts and other guides as a method to support student's in achieving independent mastery of specific skills (Archer & Hughes, 2010). I embedded scaffolding within the intervention in four main ways. First, I specifically sequenced the inference-making questions to reflect knowledge activation and then integration (Kendeou, 2015). Second, I provided the participants with a sentence frame to assist in completing the target task of forming a complete theme statement. Sentence frames are a specific type of sentence stem that provides students with the syntactical structure of the sentence so as to lessen the task's cognitive load (Rodriguez-Mojica et al., 2018). In doing so, participants were able to better focus on the theme statement's content and not its syntactic structure. Third, in addition to the sentence frame, I further scaffolded the participant's theme statement construction by asking three questions that corresponded with each part of the theme statement to form the whole theme statement, instead of asking them to form the entire theme statement at once. Fourth, and finally, I used a graphic organizer to aid students in in summarizing the text after reading. Feedback (5) refers to specific corrective and

affirmative teacher responses to student errors. I provided such feedback to participant responses to the inference making questions during reading and when summarizing after reading. In addition, I provided specific corrective and affirmative feedback when supporting participants' theme statement formation.

Theme and Background Knowledge

As stated in Chapter 1, the role of background knowledge in reading comprehension is well-established in the literature. This research also highlights the background knowledge's important role in theme identification (Goldman et al., 2015; Graesser, 2007). Still, because theme is seldom used in comparison to other comprehension measures across the field, less is known about the exact background knowledge students need to identify narrative themes. For instance, we know it is important for students to have topic background knowledge for any text (McCarthy et al., 2018). If a student does not know what baseball is, they may have a hard time comprehending a narrative text that contains references to the sport. The question of domain knowledge for narrative texts, in contrast, remains elusive with scant mention in the literature. One mention comes from Cartwright's (2015) assertation that social cognition, a construct that represents one's knowledge about the social world (Fiske & Taylor, 2013), represents the domain knowledge students need to comprehend narrative texts. The present study adopted the ToM construct as the method to clarify the socio-cognitive knowledge necessary to generate accurate inferences essential to text comprehension. The positive effects of the intervention method suggest that supporting socio-cognitive knowledge in this way does indeed support narrative comprehension. Narratives then likely do contain domain knowledge much in the same way as expository texts, and domain knowledge is knowledge of the social world. The present study suggests that the ToM skill may serve as one measurable way to capture social knowledge

because it represents a proxy for the comprehender's capacity for generating socio-cognitive knowledge-based inferences. Providing and correcting such knowledge through feedback on their ToM performance during reading (i.e., mental state attribution of character's mental state and explanations of character's behaviors) may benefit comprehenders with deficits in decoding, specifically by compensating for such deficits (Miller & Keenan, 2009).

However, other boundary conditions may decrease the potency of including a social background knowledge component to comprehension interventions. Kaefer (2020) studied the effect of activating versus providing background knowledge during read aloud instruction. Activating knowledge refers to bringing the knowledge students already have into working memory (Hambrek & Engel, 2002), while provided knowledge is presenting students with completely new knowledge. Ninety-two students across three conditions were provided with either activated knowledge, provided knowledge, or no knowledge before engaging in a read aloud activity. The results suggested that activated knowledge was more effective in aiding deep comprehension than providing background knowledge, meaning that students must already have a base level of knowledge to benefit from including a discussion of background knowledge during before-reading activities (see also O'Reilly et al., 2019). This presents a challenge to teachers of students who may have atypical socio-cognitive development or low levels of social cognition when reading narrative texts. The present study addressed this potential challenge by not only activating students' background knowledge before reading, but also by further providing knowledge and correcting knowledge during and after reading through the inference making questions and summary graphic organizer activity. In doing so, I was better able to ensure participants had the accurate knowledge needed to identify the story's theme and create an accurate theme statement. While the ToM skill may be an important reading skill, its

importance may be due to the knowledge that is attained by engaging such a skill during reading. Therefore, general socio-cognitive knowledge mediates the impact ToM skill has on reading comprehension outcomes.

Limitations and Future Research Directions

There are limitations to the current study. First, a small number of participants limits the generalization of the current study's findings. I cannot generalize the current study's findings to other populations of readers. Relatedly, I cannot generalize the current study's findings to typical in-person classroom instruction given the online nature of the intervention method due to the COVID-19 pandemic. However, given the nascent stage of the literature on ToM and reading comprehension, the findings of the proposed study will inform the potential utility of conducting future experimental and quasi-experimental group design studies during typical in-person classroom instruction in schools.

Second, I served as the interventionist for the study, collected data for all students, and was aware of each participant's conditions when testing the students' instruction. However, the use of a common instructional practice in today's classrooms (i.e., IRAs) strengthens the ecological validity of the proposed intervention (Ledford et al., 2018). And, as stated previously, the findings of the proposed intervention will inform future group design studies, which will better explore the practicality and sustainability of the intervention with teachers as interventionists (Fuchs et al., 2001).

Third, the present study cannot conclude that the intervention led to increased ToM performance due to the small sample and lack of ToM outcome data. The conversations had during the intervention may result in distal effects on both ToM and comprehension outcomes (Tompkins et al., 2019). Future experimental work in this area with larger samples

should aim to include both a comprehension outcome measure as well as a ToM measure to assess distal transfer effects across both constructs. Moreover, second order ToM may be a stronger contributor to comprehension than first order ToM (Kelly et al. in progress). It is possible that ceiling effects in first-order (basic) ToM were present in the target students given their age and performance on the ToMTB. However, it remains unknown if this effect aided or disrupted the potential effectiveness of the intervention. It is also unknown what the effects of the study would be on a different population of readers. I chose the population purposefully to support their comprehension due to evidenced difficulties with comprehension specifically, which likely contributed to the effective findings. It may be that TIER has null effects for populations of readers who are adequate or skilled comprehenders. Relatedly, I did not measure ToM or inferencing ability during the intervention to examine what effect TIER had on those discrete skills. Future studies should focus on parsing out the differential impact of student ability, especially in middle childhood, across ToM construct components (including both cognitive and affective) on listening and reading comprehension outcomes. Future studies should include both standardized measures as well as practice-oriented measures, such as theme. Such investigations could reveal the effect of TIER on discrete reading comprehension skills such as inference making and ToM as well as narrative, reading, and listening comprehension generally.

Fourth, there is also a chance the Lexile scores of the books I chose for instruction and assessment do not accurately predict theme-identification difficulty (Allington et al., 2015). Some books may present a harder challenge for some students, which may make it more difficult to identify an intervention effect. However. the inclusion of additional dependent variables and the randomization of books to different intervention phases mitigated this potential internal threat to internal validity.

Fifth, I did not include a measure of student's language ability. Common narrative themes, such as happiness, represent the nominalized noun of a mental state verb (to feel happy). The identification of narrative themes using such language then represent a form of verb to noun nominalization. Such an ability develops out of syntactic bootstrapping, which describes the process by which children learn the meanings of words, specifically verbs, by recognizing words as different parts of speech (Fisher et al., 2010). The syntactic ability to nominalize a verb as a noun via such a process to portray a theme statement may present a significant challenge for students with low syntactic abilities (Goswami & Bryant, 2016). My use of prompts, explicitly modeling, and sentence frames alleviated constraints of syntax, but future work in this area should consider controlling for student's language ability and executive functions.

Implications for Practice

The main implication from the present study is that students in early elementary classrooms will benefit if educators make clear the socio-cognitive elements of the story during read aloud activities. Additionally, the presented framework explains in concrete terms the knowledge and vocabulary students need to comprehend narratives at a deeper level. Prior research suggests upwards of 50% of early elementary educators do not plan their read aloud instruction. Educators should consider using the presented intervention framework to plan for and support their read aloud instruction.

Educators should also consider the cultural familiarity of the text to their students. Reading, particularly narrative reading, is in effect, a cultural act (Tsai, 2007). When engaged in the reading of stories, students are thrust into a cultural world more or less similar to or different from their own (Brooks & Browne, 2012; Galda & Beach, 2001). The degree of congruence between the cultural world of the text and the cultural funds of knowledge brought to the text by

the reader, exerts an effect on reading comprehension outcomes, including the reader's ability to perspective-take (Dray & Selman, 2010). Because of such interconnectedness between social and cultural understanding of a text, and that reading comprehension is defined as a meaning-making process impacted by the text, reader, and activity, educators should strive to choose texts that are culturally relevant better support their students' comprehension.

Other implications extend to the content areas. History and science texts can follow the basic narrative structure of character and setting introduction, problem, and solution (i.e., narrative nonfiction; Dahlstrom, 2014). In fact, evidence suggests presenting information in the form of narratives can increase student knowledge in content areas that then supports deeper learning (Browning, 2021; Flynn & Hardman, 2019). Therefore, the intervention framework presented in the current study can support students' comprehension of narrative texts and, at the same time, build content area knowledge. Finally, the intervention framework can also support students in secondary settings. As narratives grow in complexity, the main structure of character and setting introduction, problem, and solution remains intact with potentially numerous, smaller structures subordinate to the main event, and sometimes told in a non-sequential temporal order. The basic framework presented in the current study can extend to benefit secondary students' comprehension of such complex narratives.

Conclusion

By engaging participants less-skilled in comprehension in a strategic dialogic reading activity, I was able to examine the impact of TIER on participants' listening comprehension of authentic children's literature. The findings of the present study suggest the intervention framework is effective in strengthening narrative listening comprehension as evidenced by a demonstrated increase in participant's theme statement accuracy. This finding is the first of its

kind to examine a causal link between ToM-informed instruction and improvements in comprehension. Thus, teachers' implementation of dialogic reading activities purposefully designed to attend to the socio-cognitive knowledge of narrative texts through ToM inference-making questioning and feedback has the potential to support students' comprehension in addition to supporting their socio-cognitive development.

Chapter 5: Conclusion

Reading comprehension is a highly complex constellation of skills that become more or less important based on reader, text, and task characteristics (Snow, 2002). A significant objective of research in reading comprehension then aims to make sense of and explain how this constellation of skills functions during reading for successful comprehension across individual readers as well as profiles of readers, varying reading tasks, and when reading different text types. A novel skill in this research in vogue of late is the socio-cognitive skill ToM (Kim, 2020). My overall aim in the present dissertation was to examine how ToM contributed to reading comprehension outcomes—an active and lively research field. As I engaged in my dissertation work, I encountered a preponderance of research that converged to suggest ToM contributes to reading comprehension via the listening comprehension domain of the Simple View of Reading (SVR; Kim, 2017) as a form of inference making (Dore et al., 2018). Therefore, my aim in the dissertation was to specifically examine how ToM contributes to reading comprehension as a socio-cognitive, knowledge-based inference-making process through a series of three papers.

First in the series (Chapter 2), I report a synthesis of literature designed to examine the relation between ToM and comprehension generally defined. The synthesis uncovered scant applied and observational research examining ToM in classrooms when compared to correlational research. The second paper (Chapter 3) addresses this literature gap. I report the results of a mixed methods analysis of secondary observational data aimed at examining the prevalence of mental-state vocabulary use (e.g., think, believe, want, intend) by preschool teachers during classroom read alouds and the pedagogical purpose of such language for comprehension instruction. In the third paper in the series (Chapter 4), I report a single-case

study intervention I designed from the ToM and reading comprehension research and the observational findings from the second paper. Here (Chapter 5), I summarize in more detail the three studies (Chapters 2, 3, and 4) that comprise the present dissertation and review their findings as it pertains to the overall aim. I will also discuss the future research implications the dissertation highlights and conclude with implications for classroom practice.

Studies Review

My goal in conducting the research synthesis was to analyze the burgeoning quantitative research on the relation between ToM and comprehension in order to better understand the boundaries and constraints of the relationship. Specifically, I examined how the relation changed in response to different ToM and comprehension and measures, construct dimensionsm and across different populations. A total of 15 studies qualified for inclusion in the synthesis. The synthesized findings suggest a mediated relation between ToM and reading comprehension exists mainly for second order and advanced ToM measures and through the SVR listening comprehension domain. In special populations, such as those with autism, however, the relation may emerge as direct, over and above both listening comprehension and decoding SVR domains, but more research is needed. One limitation of the literature corpus, as identified from the synthesis, is the paucity of research examining the role of ToM in classroom instruction geared toward comprehension skills and strategies. As a result of the lack of practice-based research, while there exists a relation between ToM and reading comprehension, the present literature cannot determine the utility of leveraging the relation for classroom instruction. Therefore, I conducted two studies to address this gap: first, an observational study to ascertain teachers' current practice surrounding ToM skills during comprehension instruction, and second, a study

to evaluate the efficacy of a ToM-structured intervention I designed to support early elementary students' listening comprehension.

My initial plan for the observational study of the present dissertation was to observe classroom instruction myself and collect original data to analyze. Unfortunately, the COVID-19 pandemic required I obtain this data from a secondary source. I reached out to Dr. Jill Pentimonti after reading some of her papers on preschool classroom instruction, and, graciously, Dr. Pentimonti and her research team loaned me observational data of preschool classroom read aloud instruction to analyze for my first study. I conducted an observational study as my first study in the present dissertation with the foresight of designing an ecologically valid intervention to assess in Study 2 that not only included best practices from the reading comprehension literature and ToM literature, but that also included and built upon current classroom practice. As such, the findings from Study 2 supported the crafting of the intervention in Study 3.

The second study consisted of three separate parts. In the first part, I report a pilot of the data evaluation method I initially set out to use for data analysis, which consisted of two coding schemes reported in the literature, one for comprehension instruction (Pentimonti et al., 2012) and one for ToM (Ruffman et al., 2002). After piloting the data evaluation method however, I noticed a need to adapt the ToM coding-scheme protocol to increase its sensitivity to ToM mental-state language occurrence during comprehension instruction. Additionally, the comprehension instruction coding protocol proved too restrictive in what it captured about ToM language during comprehension instruction. As a result of these limitations across the two coding schemes, I needed to reformulate my data evaluation method principally by altering my coding approach. I then abandoned the comprehension instruction coding scheme entirely and amended the ToM coding scheme for use with read aloud comprehension instruction, specifically.

Therefore, the second part of the study reports the results of the qualitative process by which I identified the adaptations I made to the ToM coding scheme I utilized to analyze the data in Part 3 of Study 2. In Part 3 of Study 2, I report the findings resulting from my new coding using the adapted coding scheme from Part 2.

Chapter 2 is a descriptive mixed method study. Therefore, the findings presented are both quantitative and qualitative. The quantitative findings indicated that across the teacher IRAs I analyzed, reading comprehension focused mental state utterances occurred in 24.7% of all utterances with teacher utterances occurring at a rate of 4:1 when compared to student mental state utterances. One hundred-seventy-four unique mental-state words occurred during the read aloud in comparison to 88 unique mental-state words in the text. Character traits were the most common mental states referenced (27%), followed by emotion-state terms (24%), cognitive-state terms (22%), behavioral descriptions (20%), desire-state terms (13%), other-defined terms (7%), and finally, modulations of assertion (4%). Last, teacher variation in mental-state utterances showed significant variation with a range of 0% to 50%.

The qualitative findings demonstrated how teachers' mental-state utterances supported student's comprehension of the text during the IRA. As expected, teachers' utterances directed students to take the perspective of the characters in the story, specifically in reference to characters' feelings throughout the book. Research suggests narrative texts do indeed invite such perspective-taking practice, and recent research reveals that advanced comprehension of the social aspects of stories, in particular, are more important to comprehension than understanding the causal story structure itself. Further, and again in line with prior research on narrative reading comprehension, the perspective driven questions, statements, and feedback by the teachers also appeared as a form of comprehension monitoring. Teachers appeared to explicitly ask character

perspective questions to direct students' attention to how the characters' mental state changed throughout the story and, in doing so, moved the story forward. When teacher's provided student feedback on character perspective taking, teachers structured their feedback to consider how the mental states of characters functioned within the story's context in this causal way. Finally, another common use of mental-state references by teachers was in response to explaining vocabulary. Sometimes the vocabulary came straight from the text. For example, the behavioral description word *sulk* occurred in the story and was a common word teachers explained during the IRA. Other vocabulary was introduced and clarified by teachers to describe the thoughts feelings, and actions of the characters not explicitly stated in the text.

For the final chapter in the series (Chapter 4), I designed a listening comprehension intervention that focused on theme identification using a ToM inference making structure using a dialogic reading format. Again, my initial plan was to work in Grade 2 classrooms in a community school, but COVID-19 had other plans in store for me. I partnered with three caregivers interested in obtaining extra comprehension support for their child and implemented the intervention via Zoom. All three participants presented as readers less-skilled in comprehension. Following the synthesis and observational study, my aim in Chapter 4 was to assess the efficacy of a ToM-informed listening comprehension intervention on students' theme-statement accuracy.

I conducted a multiple probe across participants' single case research design to assess the efficacy of the intervention I developed. The intervention procedures consisted of activities coordinated before, during, and after reading. In the before reading activities, I set the purpose for reading (to identify theme), introduced three mental state vocabulary words highlighted per text by defining them with the participant, and reviewed any needed background knowledge for

the text. Next, during reading, I asked pre-planned questions structured to reflect the two-step inference making process of knowledge action + integration at strategic points in the story. I designed each set of inference making questions to correspond with three critical plot points within each text corresponding to 1) the introduction of the main character and story setting, 2) the main story problem, and 3) the main story problem resolution. The mental state vocabulary highlighted before reading corresponded to comprehending the character(s) mental state at these critical plot points. Finally, after reading, I summarized the story with the participant using a graphic organizer and asked the participant to identify the most important mental-state vocabulary word to the story (the theme), why that mental state word is important, and how it connected to the story. These three questions together formed the theme statement.

All three participants per visual analysis demonstrated significant improvement in their theme statement accuracy. In addition, the intervention read aloud method demonstrated strong social validity per participant survey results. Fidelity across all intervention implemented was sufficient (range = 91%–100%).

Summary

Returning to the overarching question guiding my dissertation work of *how* ToM supports reading comprehension, the results as a whole help clarify how ToM functions within the linguistic comprehension domain as an inference-making skill in conjunction with other known variables. First, as previously outlined in Chapter 1, inference making entails knowledge activation and integration with the comprehenders' emerging mental representation of the text (Kendeou, 2015). Knowledge activation within a ToM inferential process is in response to mental state references within the text. Once a comprehender encounters mental state references, the social knowledge communicated from such references is activated in working memory. Such

activation occurs either automatically by the reactivation of previously activated knowledge currently stored in working memory or with effort by the comprehender retrieving information from their long-term memory (i.e., background knowledge; van den Broek, 2015), which supplies the necessary knowledge to make an inference. The former activation routinely occurs online when the topic is familiar to the reader but is more effortful when the topic is unfamiliar (Noordman et al., 2015).

Next, a reader then may engage in perspective taking as a form of comprehension monitoring to assess the accuracy and relevancy of the activated knowledge to the story context. Previous research demonstrates skilled comprehenders display such comprehension monitoring when presented with incongruent information in a story (van Moort et al., 2021). Once the activated knowledge is verified, it is integrated into the comprehender's overall emerging mental representation of the text (Perfetti & Stafura, 2014). The ToM inferential process concludes with such integration occurring when the activated knowledge from the mental state references explains the character(s) past or current behaviors or allows for a comprehender to predict the character(s) future behavior.

Let's consider an example sentence: "A *smile* crept across Sarah's face as she *tried* to temper her show of ***elation*** when her name was a read aloud for the student of the year award." The words *smile* and *tried* are in italics to represent indirect references to Sarah's mental state. *Smile* and *tried* are indirect in that they both represent underlying mental states. For example, one *smiles* when one is **happy**, and one *tries* when one has a **desire** or **need** to achieve a goal. ***Elation***, however, is in italics and bolded to represent a direct relation. Elation is a direct reference to Sarah's mental state; Sarah is feeling extreme happiness.

When a reader encounters these words in the text, knowledge of Sarah's mental state is automatically activated if one has the background knowledge (i.e., knows what the word *elation* means; knows that *smiling* means someone is happy, etc.) or if previous knowledge from the story is stored in working memory. For instance, if the story contained information about how Sarah wanted to be student of the year, a reader may still infer Sarah's mental state of happiness because such previous information would compensate for lack of mental state understanding. This is also why working memory capacity supports inference making, especially for those with inadequate background knowledge (Smith et al., 2021). Such a process may require effort if a reader must supply background knowledge to understand Sarah's mental state because they do not have the mental state understanding to do so nor do they have access to previous textual knowledge stored in working memory. A reader may then ascribe the same mental state to Sarah by identifying how they or someone would feel receiving a student of the year award. Once the activated knowledge is verified, it is integrated into the reader's emerging mental representation. In the above example, Sarah receiving her award may signal the conclusion of the story with Sarah achieving her goal.

Past research identifies that understanding the socio-cognitive causal relations in narrative text, as displayed above with Sarah, supports comprehension (Pavias et al., 2016). ToM provides a measurable way to assess comprehender's capacity for this skill. Perhaps more importantly, the research on ToM provides a structure of instruction to support students in developing this skill. Research demonstrates relations between ToM and an array of other socio-cognitive constructs (Imuta et al., 2016; Weimer et al., 2021). Therefore, putting this guidance into practice not only has the potential to better students' reading comprehension

outcomes, but it can also support them in exhibiting more skillful social understanding, displaying empathy toward others, and developing closer relationships with those around them.

Implications for Classroom Practice

Examining the conceptualization of ToM as a socio-cognitive, knowledge-based inference-making process necessary for narrative reading comprehension across the three studies in the present dissertation illuminate connections to classroom practice. One main implication is detailing how to guide students' attention to characters' mental states. It may not be enough for teachers to simply ask how a character is feeling. Rather, perspective-taking questions should be couched within an inference-making question set that asks students to first identify character mental states and then explain how such knowledge informs the story. This higher-level analysis is exemplified in the Common Core Reading Anchor Standard 3 CCSS.ELA-LITERACY.CCRA.R.3: Analyze how and why individuals, events, or ideas develop and interact over the course of a text (National Governors Association Center for Best Practices, 2010). For students who are less-skilled in comprehension, modeling this process via a think-aloud procedure during read-aloud activities is an example of explicit instruction that benefits students at risk for reading failure. Professional development for teachers that centers on explaining ToM as an inferential process and linking it with current practice in reading comprehension instruction could greatly benefit teachers' read-aloud instruction.

Furthermore, my focus on theme as the dependent variable in Chapter 4 was a purposeful choice to for two main reasons. First, there exist inherent limits within common standardized reading comprehension measures in capturing more complex reading comprehension processes (Ellman & Oslund, 2019). Such limitations directly contribute to the gap between research and practice. Even more specific measures that capture individual skills, such as

inference making, rarely, if ever, task participants with theme-statement formation. I stated in Chapter 1 that I cannot assume that changes in students' discrete reading comprehension skills such as inference making will result in changes to theme identification. Therefore, I did not assess inference making or ToM skill as a dependent variable in Chapter 4 because I deliberately focused on theme identification—the reading task the participants in my study will be assessed on in their schools. Teachers now have a research-based dialogic reading instructional framework to use in future to teach theme identification.

Future Research Directions

The three studies in the present dissertation point to numerous avenues of future inquiry. The synthesis especially highlights the need to assess the differential relations between ToM and reading comprehension resulting from varying ToM measures and varying reading comprehension assessments, including expository measures. More broadly, the construct of social cognition as a whole, of which ToM is but one part, is intricately complex and interrelated (Happé et al., 2017). As such, investigations should also assess the differential impact of varied socio-cognitive measures, such as cognitive ToM versus affective ToM measures or physiological measures of empathy, such as skin conductance, heart rate monitoring, and pupil dilation, among other conceptualizations of the construct (Sesso et al., 2021). More generally, it may also benefit the field to assess how the relation changes over time using developmentally appropriate measures to see if the contribution of socio-cognitive skill changes in strength across student development. Finally, disaggregation by different student populations is needed to examine the generalizability of findings. One population routinely involved in research on ToM is students with autism. However, given evidence ToM's confluence with language competence, ToM may play a significant and direct role in reading comprehension of narrative texts in other

clinical populations with characteristic deficits in language, such as those with learning disabilities and those who exhibit traits of psychopathy.

An initial step to expand the findings in the second study is to explore mental state references across a range of texts, including expository, to examine the extent to which text type results in differences and similarities in mental state references, particularly when the cultural knowledge of the text is dissimilar or similar to the cultural identities of the reader (see, for example, Dray & Selman, 2010). Additionally, encompassing a wider range of grades would reveal if and how mental state references change across early elementary grades. Lecce et al. (2021) conducted a study of teachers' mental state talk on 8–11-year-old Italian students' ToM performance on an advanced measure of ToM. Lecce and colleagues reported teachers' mental state references positively correlated with students' ToM performance. A similar study that includes students' reading comprehension outcome data is needed to examine if variation in teachers' mental state references during read-aloud activities results in differential student outcomes on reading comprehension.

Last, the course of research for Study 3 is to continue to scale up the intervention and assess its effectiveness across text genres and with different student populations. The expansion should include standardized reading comprehension measures alongside ecologically and socially-valid researcher-developed measures of theme identification to strengthen the overall validity of intervention findings. Additionally, researchers should also include valid and reliable ToM measures to assess changes in participants' ToM to see if commensurate changes in both reading comprehension and ToM occur in response to the read-aloud intervention. Included in these scale-up studies should be research on the effectiveness of providing teacher's professional development on including mental state language in their read-aloud activities and having

teachers serve as the intervention implementers. Finally, the importance of mental state language demonstrated in the present dissertation and current ToM and reading comprehension research (Ebert, 2020; Kim et al., 2021) presents an opportunity for the development of a measure that assesses children's mental state understanding. Such a measure can support efforts aimed to bolster background knowledge for narrative comprehension and serve as a control measure for background knowledge when assessing reading comprehension intervention outcomes with narrative texts.

Conclusion

The present dissertation began with a problem statement centered on the Grade 4 slump phenomenon, which represents the difficulty some students face when transitioning from learning to read to reading to learn (Chall et al., 1990). Despite the incredible strides we have made in the last few decades on teaching students how to accurately and fluently read the words on a page, the Grade 4 slump persists (United States Department of Education, 2019). There remains, then, an obstinate and pressing concern regarding how to teach students to craft and gain meaning from what they read *and* show that they have done so on meaningful tasks.

The role of knowledge in reading comprehension is emerging as potentially critical in this endeavor as we enter the early post-common-core era (Catts, 2021). Given the known benefits of early intervention (Whitehurst & Lonigan, 1998), examining how preschool and early elementary students use knowledge to make meaning of a text as they learn to read is vital to addressing the role knowledge plays in reading comprehension outcomes across upper elementary, secondary settings, and beyond. My dissertation aligns with this avenue of inquiry in two interconnected ways: first, by suggesting that mental state references and vocabulary provide a foundation for understanding the social background knowledge necessary to comprehend

narrative texts (Cartwright, 2015), and second, by naming and describing a ToM inferencing making process that supports students in making sense of their own social knowledge and the social knowledge inherent within narrative texts—the key to learning from texts (Goldman et al., 2015). It is my intention that this dissertation offers a small but significant fragment to the reading comprehension research mosaic. I hope that I and others can continue to build upon this fragment to add depth and dimension to the ever-evolving mosaic in the many years to come.

Tables

Chapter 1 Table 1

Theory of Mind Analysis of National Assessment of Education Progress Literary Skill Items

Score	Reading Skill Assessed
Basic (Cut Score: 208)	
206	Recognize the meaning of a word with multiple meanings as it is used in a story—Integrate/Interpret (MC)
203	Recognize description of character's action explicitly stated in a story—Locate/Recall (MC)
203	*Recognize a generalization about how a story character feels based on text description—Integrate/Interpret (MC)*
Proficient (Cut Score: 238)	
212	*Form an opinion about a character's behavior and provide an example from the story to support opinion—Critique/Evaluate (CR)*
213	*Recognize a reason for a character's action that is implied in a section of a story—Locate/Recall (MC)*
229	*Describe explicitly stated motive of character in story—Integrate/Interpret (CR)*
231	*Recognize the word that describes the way the speaker feels in a stanza of a poem—Integrate/Interpret (MC)*
Advanced (Cut Score: 268)	
245	Recognize the best description of what the speaker of a poem is suggesting in a specific stanza—Integrate/Interpret (MC)
250	Sort events referred to or described in a story according to when they occur—Integrate/Interpret (SR)
250	Recognize best description of the way an author organizes information in a literary non-fiction article—Critique/Evaluate (MC)
266	*Evaluate story events to form an opinion about a character's behavior and support opinion with information from the story—Critique/Evaluate (CR)*

Note. Skills in italics indicate scores that require Theory of Mind.

MC = Multiple Choice Response; CR = Constructed Response; SR = Selected Response

Chapter 1 Table 2

Narrative Inference Classification Examples in Little Red Riding Hood

	Local Inference	Global Inference
Text-Based Inference	Definition: Making connections between propositions and references. Text Example: "The door was unlocked, so the wolf opened it quietly and crept inside. He tiptoed into the grandmother's bedroom and found her asleep in bed." Necessary Inferences: Associating the pronoun *he* in the second sentence with the subject *the wolf* in the previous sentence, and the pronoun *she* with the subject *grandmother* in the second sentence. Associating 'the door' with 'it' in the first sentence.	Definition: Using past textual information to make sense of novel text information. Text Example: "Little Red Riding Hood knocked on the front door... 'Come in, my dear,' said her grandmother. Little Red Riding Hood went into the bedroom." Necessary Inference: She wasn't scared because she didn't know the wolf was masquerading as her grandmother even though she met the wolf earlier.
Knowledge-Based Inference	Definition: Using relevant background knowledge to make sense of implicit information within the text. Text Example: "Little Red Riding Hood went into the bedroom. She stared at her grandmother lying in bed." Necessary Inference: Associating a mental state of the character to an action: A little girl staring at her grandmother suggests she is confused rather than excited or happy to see her grandmother.	Definition: Making connections between information from the text and broader universal concepts. Text Example: "Well, I think you've learned your lesson,' said her grandmother." Necessary Inference: Understanding that stories often teach us a valuable lesson requires students to infer how knowledge gained from the story can apply to other settings. (e.g., theme identification). The theme of Little Red Riding Hood is we be cautious around strangers because they might try and trick us.

Chapter 2 Table 1

Study Features

			Sample Demographics			
Study (Year)	Location	Analytic Sample Size	%Low SES	%Race	%Sped	%SLL
Atkinson et al. (2017)	United Kingdom	N = 62	12%	NR	NR	0%
Boerma et al. (2017)	Netherlands	N = 117	NR	NR	NR	0%
Cantin et al. (2016)	USA	N = 93	22%	AA: 1% Asian: 3% Hispanic: 2% Other: 12%	NR	NR
Guajardo & Cartwright (2016)	USA	N = 31	22.5%	32% Minority	NR	NR
Kim & Phillips (2014)	USA	N = 156	~90%	AA: 38% Hispanic: 5% Multi: 14% Other: 2%	16.7%	<5%
Kim (2015a)	South Korea	N = 130	NR	NR	0%	0%
Kim (2015b)	South Korea	N = 145	NR	NR	0%	0%
Kim (2016)	South Korea	N = 201	NR	NR	NR	NR
Kim (2017)	USA	N = 350	74%	AA: 33% Hispanic: 6% Multi: 4%	2%	1.8%
Lockl et al. (2017)	Germany	N = 57	< 50%	NR	NR	17.4%

		Sample Demographics				
Study (Year)	Location	Analytic Sample Size	%Low SES	%Race	%Sped	%SLL
McIntyre et al. (2018)	USA			AA: 0% Asian: 4%	AA: 0% Asian: 0%	
HFASD		N = 70	~25%	Hispanic: 9%	Hispanic: 6%	NR
TD		N = 35	~25%	Multi: 13% Other: 4%	Multi: 11% Other: 0%	NR
Pelletier & Beatty (2015)	Canada	N = 186	~30%	~30%	NR	~60%
Ricketts et al. (2013)	USA	N = 100	NR	NR	100%	NR
Strasser & Del Rio (2013)	Chile	N = 257	NR	NR	NR	NR

Note. USA = United States of America; NR = Not reported; SES = Socioeconomic Status; Sped = Special Education; SLL = Second Language Learner; AA = African American; Multi = Multi-race; HFSA = High Functioning Autism Spectrum Disorder; TD = Typically Developing.

~ = Approximate value reported.

Chapter 2 Table 2

Description of Theory of Mind Tasks

Study (Year)	Mean Age	(Number of Tasks) ToM Task Type	Possible Score	Mean Score	SD	α	Pearson Correlation with Comprehension (RC type)
Atkinson et al. (2017)	3:10	(2) First Order	0–5	2.26	1.90	1.00	0.44* (LC)
† Boerma et al. (2017)	9:5	(8) Advanced	0–16	9.98	2.85	0.60	0.29* (RC)
Cantin et al. (2016)	9:0	(7) Advanced	0–7	4.47	1.66	0.97	0.46* (RC)
Guajardo & Cartwright (2016)							
T1	4:4	(5) First Order	0–5	2.58	1.72	0.84	0.31 (RC)
T2	8:1	(2) Second Order	0–4	2.74	1.12	0.94	0.55* (RC)
Kim & Phillips (2014)	6:2	(3) Second order	0–6	1.49	1.65	0.73	0.26–0.57* (LC)
			0–6	1.36	1.51	0.73	
			0–6	1.31	1.52	0.71	
Kim (2015a)	6:1	(4) First order (2) Second order	2–17	10.04	3.07	0.72	0.46–0.55* (LC) 0.43–0.52* (RC)
Kim (2015b)	6:1	(4) First order (2) Second order	2–17	10.02	3.09	0.74	0.45–0.57* (LC) 0.41–0.50* (RC)
Kim (2016)	6:8	(1) First order (2) Second order	0–15	5.24	2.91	0.74	0.29–0.55* (LC)
Kim (2017)	7:5	(3) Second order	0–17	7.79	3.80	0.71	0.51* (LC) 0.47* (ELC)

Study (Year)	Mean Age	(Number of Tasks) ToM Task Type	Possible Score	Mean Score	SD	α	Pearson Correlation with Comprehension (RC type)
							0.34–0.36* (RC)
Lockl et al. (2017)	4:2	(3) First Order	0–5	2.24	1.9	0.64	0.09 (Grade 1 RC) 0.16 (Grade 2 RC)
McIntyre et al. (2018) HFASD TD	12:6 12:8	(2) Advanced	0–22	6.74 7.99	2.08 1.54	0.97	0.65* (RC) 0.32 (RC)
Pelletier & Beatty (2015)	NR	(4) First order	0–16	JK: 6.29 SK: 7.40	JK: 2.42 SK: 2.33	NR	0.21 (RC) 0.23 (SC)
		(2) Second order		JK: 6.23 SK: 7.05	JK: 2.44 SK: 2.21		0.38* (RC) 0.40* (SC)
Ricketts et al. (2013)	15:6	(4) Advanced	0–2	0.85	0.53	NR	NR
†Strasser & Del Rio (2014)	5:5	(5) First Order (1) Second order	0–8	5.21	1.86	0.64 1	0.18 (SC)

Note. Age is reported years: months; SD = standard deviation; α = interrater agreement reliability reported as Cronbach's alpha; LC = Listening Comprehension; RC = Reading Comprehension; T1 = Time One; T2 = Time 2; ELC = Expository Listening Comprehension; HFSA = High Functioning Autism Spectrum Disorder; TD = Typically Developing; JK = Junior Kindergarten; SK = Senior Kindergarten; SC = Story Comprehension.

† = Adapted tasks into native language of participants.

* = $p < 0.05$.

Chapter 2 Table 3

Regression Analytic Findings

Study (Year)	Theory of Mind Type	Comprehension Type	Statistical Model	Standardized Regression Coefficient (β)	Unique Variance Explained (r^2)
Atkinson, et al. (2017)	First order	Reading Comprehension (E & N) (N)	Age Nonverbal ability Decoding Linguistic comprehension Executive function	0.38*	0.08
Guajardo & Cartwright (2016)	First order	Reading Comprehension (E & N) (N)	Age Vocabulary SES Phrase and sentence comprehension Decoding Working memory Counterfactual reasoning Cognitive flexibility	0.154	NR
	Second order			0.078*	0.089
Lockl et al. 2017	First Order	Reading Comprehension (E & N)	SES Gender Nonverbal ability Working memory Language	-0.17 (Grade 1) -0.04 (Grade 2)	NR
McIntyre et al. (2018) HFASD	Advanced	Reading Comprehension (E & N)	IQ Word Reading Language	0.30*	0.06
TD				0.20	0.03

Study (Year)	Theory of Mind Type	Comprehension Type	Statistical Model	Standardized Regression Coefficient (β)	Unique Variance Explained (r^2)
Pelletier & Beatty (2015)	(4) First order (2) Second Order	Story comprehension	Age Vocabulary Early reading ability	NR	NS 0.042
Ricketts, et al. (2013)	Advanced	Reading Comprehension (E & N)	Word recognition Linguistic comprehension	0.24*	0.05
			Word recognition Linguistic comprehension	0.23*	0.04
Strasser & Del Rio (2014)	Composite	Story comprehension	Days in school Age Gender Vocabulary breadth Vocabulary depth Working memory Inhibition Attention Inference making Comprehension monitoring	0.005	NR

Note. Composite refers to a ToM score that is a composite of first and second order tasks. E & N = comprehension measure that comprises both expository and narrative passages; N = narrative reading comprehension measure; NR = not reported; HFSA = High Functioning Autism Spectrum Disorder; TD = Typically Developing.

* = $p < 0.05$.

Chapter 2 Table 4

Mediation Analytic Findings

Study (Year)	Theory of Mind Type	Analytic Method	Comprehension Type	Other Variables in Statistical Model (covariates)	Main Findings	Standardized Regression Coefficient (β)	Effect Type
Atkinson et al. (2017)	First-order	Bootstrap	Reading Comprehension (E & N)	NA	ToM is directly related to reading comprehension and indirectly related via listening comprehension	2.51 (*b*) 0.81 (*b*)	Direct Indirect
Boerma, et al. (2017)	Advanced	Confirmatory Path Analysis	Reading Comprehension (E & N)	Expressive verbal ability Print exposure HLE (Sex, Age, School Type)	ToM was directly related to reading comprehension and mediated the relation between home literacy environment and reading comprehension. Differences were not found across sex, age, and school. type on ToM	0.19	Direct
Cantin et al. (2017)	Advanced	Confirmatory Path Analysis	Reading Comprehension (E & N)	Age Mathematics Performance Working Memory Inhibition Flexibility	No path was reported for ToM to RC. The path from RC to ToM was .39 and was significant.	NR	NR

Study (Year)	Theory of Mind Type	Analytic Method	Comprehension Type	Other Variables in Statistical Model (covariates)	Main Findings	Standardized Regression Coefficient (β)	Effect Type
Kim & Phillips (2014)	Second-order	Structural Equation Model	Listening Comprehension (N)	Inhibitory Control Vocabulary Comprehension-monitoring	ToM, comprehension monitoring, and inhibitory control were all independently related to listening comprehension over and above vocabulary and age.	0.45	Direct
Kim (2015a)	Composite	Structural Equation Model	Listening Comprehension (N)	Syntax Working memory Comprehension monitoring Vocabulary	ToM was independently related to listening comprehension.	0.30	Direct
			Reading Comprehension (N)	Syntax Working memory Word reading Comprehension monitoring Vocabulary Listening Comprehension	Listening comprehension and word reading mediated the relation between all variables and reading comprehension	NR	Indirect

Study (Year)	Theory of Mind Type	Analytic Method	Comprehension Type	Other Variables in Statistical Model (covariates)	Main Findings	Standardized Regression Coefficient (β)	Effect Type
Kim (2015b)	Composite	Structural Equation Model	Reading Comprehension (E & N)	Vocabulary Grammatical Knowledge Working Memory Comprehension Monitoring Text reading Fluency	ToM, vocabulary, and grammatical knowledge were all independently related to reading comprehension	0.20	Direct
Kim (2016)	Composite	Structural Equation Model	Listening Comprehension (E & N) (N)	Working memory Attention Syntax Vocabulary Inferencing Comprehension monitoring	ToM, inference, grammatical knowledge, and working memory were all directly related to reading listening comprehension. ToM had some mediating effect between all other variables	0.45 0.07	Direct Indirect
Kim (2017)	Second order	Structural Equation Model	Listening Comprehension (N) (N & E) (E)	Inference Comprehension Monitoring Vocabulary Grammatical Knowledge Working Memory	ToM directly related to listening comprehension. No indirect relation was found via other variables within the statistical model.	0.37	Direct

Study (Year)	Theory of Mind Type	Analytic Method	Comprehension Type	Other Variables in Statistical Model (covariates)	Main Findings	Standardized Regression Coefficient (β)	Effect Type
			Reading Comprehension (E) (N) (E & N)	Inference Comprehension Monitoring Vocabulary Grammatical Knowledge Working Memory Listening comprehension Word reading	Listening comprehension and word reading mediated the relation between all variables and reading comprehension	0.18	Indirect
Strasser & Del Rio (2014)	Composite	Regression	Story Comprehension	Working memory	ToM did not mediate the effect of working memory on reading comprehension	0.023 (NS)	Indirect

Note. Composite refers to a ToM score that is a composite of first and second order tasks; NA = Not applicable; (*b*) indicates nonstandardized beta coefficient; E & N = reading comprehension measure that comprises both expository and narrative passages; NR = Not reported; E = expository reading comprehension measure; N = narrative reading comprehension measure; NS = Not significant.

Chapter 3 Table 1

Pentimonti (2012) SABR Coding Scheme

CHILD TALK + TEACHER CONTINUATIONS	
CHILD TALK	
Child Comment	Definition: Child declarative sentence form used to convey information or imperative sentence used to request action from the listener.
Child Questions	Definition: Child interrogative sentence form designed to elicit a response from the listener. – *Why doesn't Diego get a turn?* Keywords: *Who, What, When, Where, Which, Why, How, Have (Has, Had, Having), Can, Could, Do (Does, Did), Will, Would, Is, Are, Was, Were, Being, Been, May, Might, Must, Need, Shall, Should*
TEACHER TALK Triggered by Child Talk	
Repeat/Recast/Extend	Definition: Elaborating on child utterances by repeating, recasting or expanding the topic in a way that uses at least one word from the child's previous utterance. These continuations of child topics/utterances must use at least one word from the child's previous utterance(s); this includes root word that are carried over from C's talk (C: They read. T: They're reading). This must include teacher talk that immediately follows a child utterance.
TEACHER TALK	
Auxiliary-Fronted Yes/No Question	Definition: These questions are designed to elicit a yes/no response and contain an auxiliary ("helper") verb. *Note: Auxiliary verb may not be the very first word in the question, but should still be coded here. – *Do you see that?* / *When they were playing, was she being nice?* Keywords/Auxiliary verbs: *Have, Has, Had, Having, Can, Could, Do, Does, Did, Will, Would, Is, Are, Was, Were, Being, Been, May, Might, Must, Need, Shall, Should*
Yes/No Question	Definition: A question that can be answered with yes/no question response but does NOT have an auxiliary verb at beginning – *See that?* May also include tag questions (rhetorical questions) – *She's bossy, huh?*
Basic Wh- Question	Definition: Who, what, when, where, which (but not why) + interrogative sentence form/question – *What happened?*
Why Question	**Definition: Why + interrogative sentence form/question – *Why are they mad?***
How Question	**Definition: *How* + interrogative sentence form/question – *How did they build the tower so high?***

Turn-Taking Question	Definition: Where the teacher uses a question form, but it does not fit the other Q categories and is designed to give the child a turn to speak (*Yes, Felicia?*)
Redirection/Reminders	Definition: Positive and negative behavior management related talk (redirection and reminders) Key phrases: *Raise your hand, just a minute, hold on, I am looking for…., I like how…, shhh, sit down, wait, wait your turn, stop, be quiet*; *Look, see* (Do not double count attention-directing words if they are embedded in a longer behavior-focused utterance, e.g., *Sit criss-cross applesauce and look up here!)*
Book and Print Conventions	Definition: Talks about how to use books, book parts or rules and conventions that English print requires. Keywords/phrases*: Title, title page, page, cover, spine, end pages, dedication page, turn page, left/ right, top/bottom, "this is called" + title of book, genre*
Letters/Words/Writing	Definition: Talks about letter names, letter sounds, alphabetical order or letter features; identifies whole words in print; models writing of words; talks about how to write, invented spelling, or writing notes to friends Keywords: *Letters, uppercase, lowercase, letter sound, starts with + letter name, rhyming word, write, writing center*
Character References	Definition: Names specific characters with proper noun. Keywords: Proper nouns that name characters in text (e.g., Petunia, Diego)
Cognition	Definition: Indicates cognitive processes in the brain. Keywords: *Think, know, believe, make believe, remember/remind, pretend, idea, plan, find out, learn, understand, imagine/imagination, wonder, guess, decide/decision, choose/chose/choice, pick*
Feelings/Emotions	Definition: Captures feeling/emotions. Common keywords: *Mad, sad, happy, angry, sorry, miss, scared*
Define/Elaborate on Vocabulary	Definition: Asks for or providing a word's definition, or elaborating on word meaning. Key phrases: *"What does __ mean?"; "The word __ means …"*
Act Out/Pretend	Definition: Promotes dramatization of the book or other pretend role-play. Key phrases: "*Let's pretend…"; "Let's imagine…"; make believe, (dramatic) play*
EXTRATEXTURAL TALK	
Author/Illustrator reference	Definition: This code involves teacher talk that name the author/illustrator or discusses the role of the author/illustrator. Example*: authors write words, illustrators draw pictures*

Making Predictions	**Definition: This code includes inferences on a forecasted causal chain into the future such as predicting new plans for the character or asking what event will happen next.** **Example: *"What will happen…?"; "Guess what's next?"; "I expect…"***
Making Connections	**Definition: This code involves modeling the implicit link or explicit comparison between text and personal experiences.** **Example: *"Remember when we…"; "Reminds me of our…"***
Background Knowledge	**Definition: This code involves modeling connections between the text and background information/facts.** **Example: *"What do you know about…?"***
Causal Reasoning/Problem Solving	**Definition: These are inferences on a causal chain between the current, explicit event/state and previous text information. Causal effects reference antecedents or consequences/effects of text events or physical states/objects.** **Example: *because, why?, since/so, problem, solve, trouble***
Desires/Preferences	Definition: This code involves naming or inferring desires, wants, preferences or other volition terms. Example: *like, dislike, want, hate, decide, favorite*
Judgments/Opinions	Definition: This code includes opinions, attitudes and assertions that express character/self/others' judgments about the quality of something, traits/identity of someone or other attitudes about stimulus/state. Example: *mean/nice, bossy, fair, smart, friendly, bully, agree, good/bad*
	NOT CODEABLE
Inaudible Talk	Example: *If you cannot hear what someone says*
Filler Words	Example: *Ummmm, uhhh, well*
Manners	Example: *Please, Thank you*
Yes/No Responses	Example: *Teacher replies to student question or response with yes or no; Student replies to teacher question with yes or no*
Overlapping Speech	Example: *Teacher and students talk at the same time. More than one student talks at a time.*
Teacher Praise	Example: *Good job! Well done!*
Exact Repetitions	Example: *Teacher repeats what the student says exactly; teacher tells students, "repeat after me."*
Talk in a language other than English	Example: *Foreign language use such as Spanish spoken with a child*

Note. Items in bold represent codes used in Part 1 of the present study.

Chapter 3 Table 2

Part 1 Adapted Ruffman et al. (2002) Coding Scheme

Category	Examples
Mental State Utterances	
Desire	Want, like, love, hope, wish, dream, prefer, keen on
Emotion	Happy, sad, unhappy, feel, cross, angry, grumpy
Modulations of Assertions	Might, maybe, perhaps, possibly, probably, could be, must, certainly,
Think, Know, Believe	Do you know what that is? She knows that's going to happen. They're thinking hard. Let me think. I think it's lovely.
Other Mental States	We went to France, remember? I understand that.
Behavior	*Play, share,*
Character Traits	*Bossy, Nice, Mean*
Non-Mental State Utterances	
Factual Talk	A stethoscope is for listening to the heart.
Descriptions	She's riding a bicycle.
Link to past learning or experiences	We did that when we looked outside for bugs to catch.
Repetitions	Techer repeats student utterance.
Self-repetitions	Teacher repeats own utterance
Orienting utterances	Look, what's that?
Class Management	*Directives related to student behavior. [Student name] take a seat.*

Note. Item in italics represents additional codes not included in Ruffman's et al. original coding scheme.

Chapter 3 Table 3

Part 3 Final Coding Scheme

Category	Code	Definition	Examples
Desire	D	Refers to mental states that reflect a person's want or aspiration but also how they feel about a certain thing or event.	Want, like, love, hope, wish, dream, prefer, keen on, dislike, avoid, need
Emotion	E	Refers to mental states that reflect any emotional state of a person.	Happy, sad, unhappy, feel, cross, angry, grumpy, scared, joy, excited,
Modulations of Assertions	M	Refers to mental states that indicate the probability that the mental state of oneself or others could be wrong.	Might, maybe, perhaps, possibly, probably, could be, must, certainly, definitely, sure, guess, figure, reckon, certain, suppose, wonder, expect, curious, bet
Think, Know, Believe	T	Refers to mental states that reflect cognitive actions of a person.	Think, Know, Believe, Idea, Pretend, Predict, Make believe, A say
Behavioral descriptions	*B*	*Refers to actions that reveal and require the comprehension of underlying mental states of the characters that drive character behavior.*	*Play, Share, Fix it, Protect, Get along, Work together, Decide, Sorry/apologize (can also be a feeling), choose/choice, Plan*
Character Traits	*C*	*Refers to descriptions of characters within the story that require the comprehension of their mental states that drive their behavior.*	*Bossy, Nice, Mean, Determined, Hard-worker, Friend, Family*
Other Mental States	O	Refers to mental states that describe any other mental state not captured within the other categories.	Remember, Understand, Learn, Listen, let's see, find out

Note. Item in italics represents additional codes not included in Ruffman's et al. original coding scheme identified in Part 2 of the present study.

Chapter 3 Table 4

Breakdown of Teachers' Mental State Language Use in Top three Transcripts with Most Mental State Language

Utterance Totals	Ms. Taylor	Ms. Hanna	Ms. Martin	Sample Median
Total	108	230	89	206
Total Read Only	18	3	10	25
Total Analyzed	95 (83)	227 (97)	79 (89)	170 (83)
Total Mental State	45 (50)	100 (44)	33 (41)	52 (29)
Total Non-Reading Comprehension	5 (11)	3 (3)	1 (1)	5 (13)
Total Reading Comprehension Instruction Mental State	40 (44)	97 (43)	32 (41)	45 (26)
Mental State Type				
Total Desire	0 (0)	20 (20)	5 (15)	6 (11)
Total Emotion	21 (47)	15 (15)	15 (45)	12 (23)
Total Cognitive	10 (22)	30 (30)	12 (36)	10 (23)
Total Modulations of Assertion	0 (0)	10 (10)	5 (15)	1 (2)
Total Character Traits	9 (20)	31 (31)	14 (42)	13 (25)
Total Behavior	2 (4)	21 (21)	5 (15)	8 (18)
Total Other	7 (16)	13 (13)	1 (3)	3 (7)

Note. Items in parentheses are percentages based on total utterances.

Chapter 3 Table 5

Part 3 Codes

Code Sub-Code	Definition
Perspective Taking	Teacher explanations of or tasking students with identifying the mental states of characters within the story.
Non-verbal communication	Teacher directing students to perspective-take by referencing character facial expressions, body language, or actions conveyed in the book illustrations.
Self-Oriented	Teacher modeling how to or directing students to perspective-take by asking what they would think and feel.
Other-Oriented	Teacher modeling how to or directing students to perspective-take by asking what others would think and feel.
Causality	Teacher explanations of or tasking students with identifying contingent relations between character mental states and subsequent or past behaviors.
Behavior Causality	Teacher modeling how or directing students to identify contingent relations between character mental states and subsequent or past behaviors at a local text level (i.e., sentence level or immediate story elements)
Causal Story Structure	Teacher modeling how or directing students to identify contingent relations between character mental states and subsequent or past behaviors at a global text level (i.e., central problem and its resolution).
Social Skills and Awareness	Teacher commenting on or asking questions about social aspects in the social with the effect of reinforcing or teaching social skills and understanding
Vocabulary	Teacher explanations of mental state vocabulary within the text or read aloud dialogue

Note. Items indented represent subcodes.

Chapter 3 Table 6

Mental State References by Type

Mental State Type	Frequency Count	Mean per transcript (SD)	Range	Percentage of all mental state utterances	Rate per second of instruction
Desire	460	7.93 (7.43)	0–86	13	62.19
Emotion	848	14.62 (14.00)	0–74	24	33.73
Cognitive	787	13.57 (13.86)	0–72	22	36.35
Modulations of Assertion	139	2.40 (4.03)	0–20	4	205.80
Character Traits	954	16.45 (12.41)	0–48	27	29.99
Behavior Descriptions	708	12.21 (11.97)	0–55	20	40.40
Other	259	4.47 (4.5)	0–20	7	110.45

Note. SD = Standard Deviation

Chapter 3 Table 7

Mental State References by Type in Kingdom of Friends

Mental State Type	Frequency Count (Total = 88)	Percentage of all mental state utterances
Desire	21	24
Emotion	10	11
Cognitive	26	29
Modulations of Assertion	1	1
Character Traits	8	9
Behavior	22	25
Other	0	0

Chapter 3 Table 8

Breakdown of the median rates of code and subcode prevalence by teacher and overall

Code Sub-Code	Ms. Taylor	Ms. Hanna	Ms. Martin	Total Utterance Median
Total Reading Comprehension Utterances	40	97	32	45
Perspective Taking	21 (53)	19 (20)	21 (66)	61 (36)
Non-verbal communication	10 (25)	2 (2)	0 (0)	12 (7)
Self-Oriented	2 (5)	13 (13)	8 (25)	23 (14)
Other-Oriented	9 (23)	6 (6)	13 (41)	28 (17)
Causality	16 (40)	52 (58)	4 (13)	72 (43)
Behavior Causality	8 (20)	19 (20)	4 (13)	31 (18)
Causal Story Structure	8 (20)	30 (33)	0 (0)	38 (22)
Social Skills and Awareness	3 (8)	26 (29)	9 (0)	38 (22)
Vocabulary	2 (5)	1 (1)	7 (22)	9 (5)

Note. Items indented represent subcodes.

Items in parentheses are the percentages of based on total reading comprehension focused mental state utterances.

Chapter 4 Table 1

McMaster & Espin (2017) and McMaster et al. (2019) Guidance

Explicit Instruction Component	Included Intervention Component
Explicitly model the process	Before assessment, I included one treatment only day of explicit modeling without assessment. In addition, the remainder of assessment days included explicit instruction treatment.
Time questions purposefully	I developed questions in the sample lesson purposefully based on the narrative structure framework I designed from my theoretical framework.
Provide scaffolds	I provided explicit instruction on Day 1 of the treatment phase before beginning data collection. I further provided explicit instruction by scaffolding instruction and providing corrective feedback in response to participants' performance of target skills.
Sequence inference making instruction purposefully	I purposefully sequenced inference making instruction in two ways. First, I asked a question to activate background knowledge and provided scaffolding as needed through corrective or reinforcing feedback before asking the inference making question. Second, I sequenced inference making instruction across the text based on the narrative structure framework I designed from my theoretical framework.
Ask questions and provide structured feedback	I included scaffolded feedback in the sample lesson in response to correct and incorrect student answers to ToM and inference making prompts. I also scaffolded the inference making process by first asking a question about the characters' mental states before asking inference making questions.
Provide students with multiple practice opportunities	Over the course of the intervention, I provided students multiple practice opportunities with different texts.
Gradually release responsibility to the student	I achieved a gradual release of responsibility through the explicit instructional design of the intervention.
Instruction Content and Design Guidance	
Inference making instruction should transfer to different comprehension mediums	I examined the transference of the inference making intervention on varying texts.
Activate background knowledge	Questions specifically aimed at clarifying mental state verbs and concepts provided within the text serve as background knowledge activation before inference making.

Pre-teach important vocabulary and concepts	I included a pre-teaching component in the sample lesson before reading section where I reviewed important vocabulary words (mental states) for the lesson.
Read texts together	I conducted the intervention using the instructional method of IRA which is a dialogic shared reading procedure.

Chapter 4 Table 2

Pre-Assessment Raw Scores

	Listening Comprehension[a]	Oral Reading Fluency[b]	Affective Theory of Mind	Theory of Mind	Passage Comprehension[c]
Kai	19 (45)	83 (49)	19	10	28 (40)
Kelani	9 (2)	108 (69)	13	9	19 (6)
Kelsey	17 (21)	40 (5)	18	13	22 (9)

Note. Items in parentheses represent percentile ranks

[a] Percentile ranks obtained from the online Riverside scoring system*:* https://riversidescore.com

[b] Percentile ranks obtained from https://dibels.uoregon.edu/sites/dibels1.uoregon.edu/files/DIBELS8-Percentiles-2020.pdf

[c] Percentile ranks obtained from the online Riverside scoring system: https://riversidescore.com

Chapter 4 Table 3

Median Theme Statement Scores Baseline to Intervention and Post-Intervention

Participant	Baseline	Intervention and Post-Intervention[a]
Kai	0.5	2.5
Kelani	0.25	2
Kelsey	.5	2

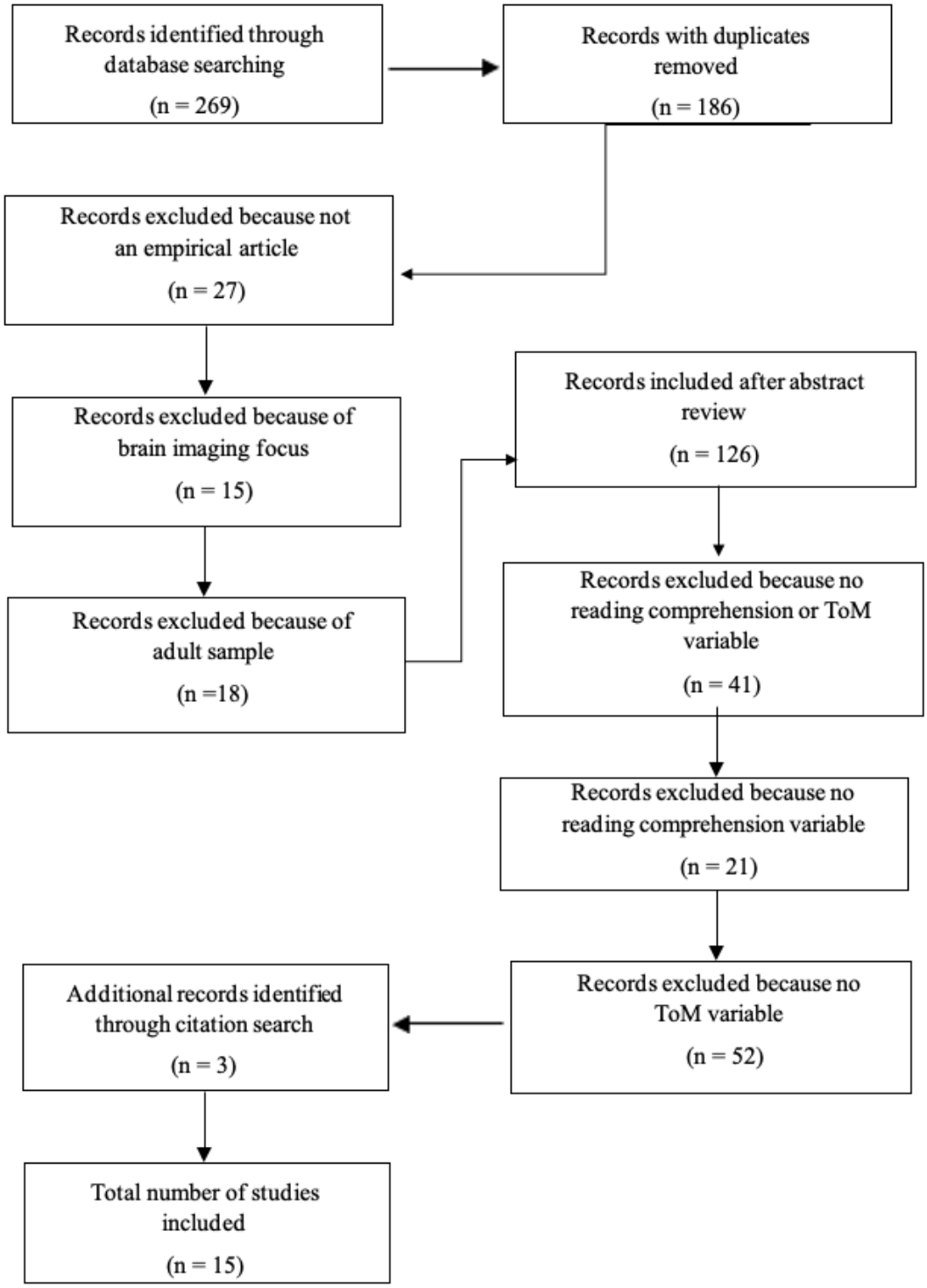
Note. CI = Confidence Interval.

[a] Includes post-intervention probe in median score.

Figures

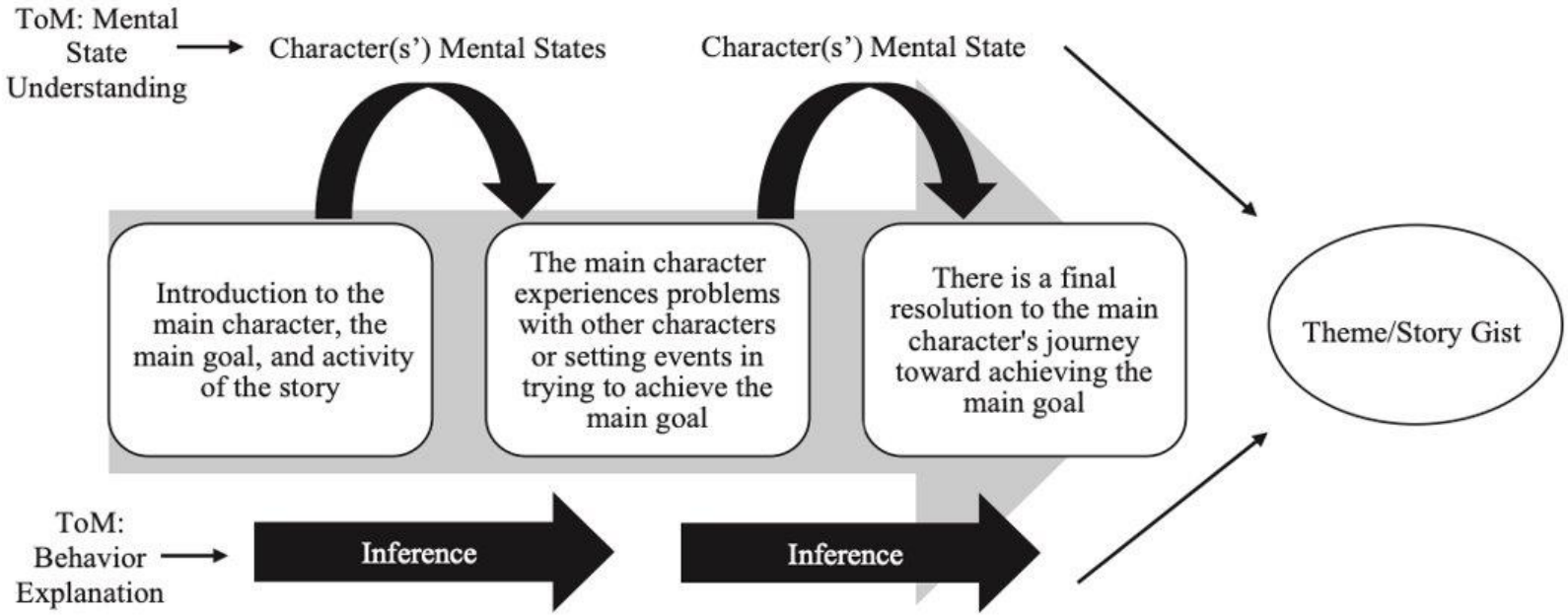
Chapter 2 Figure 1

Search Flow Diagram



Chapter 3 Figure 1

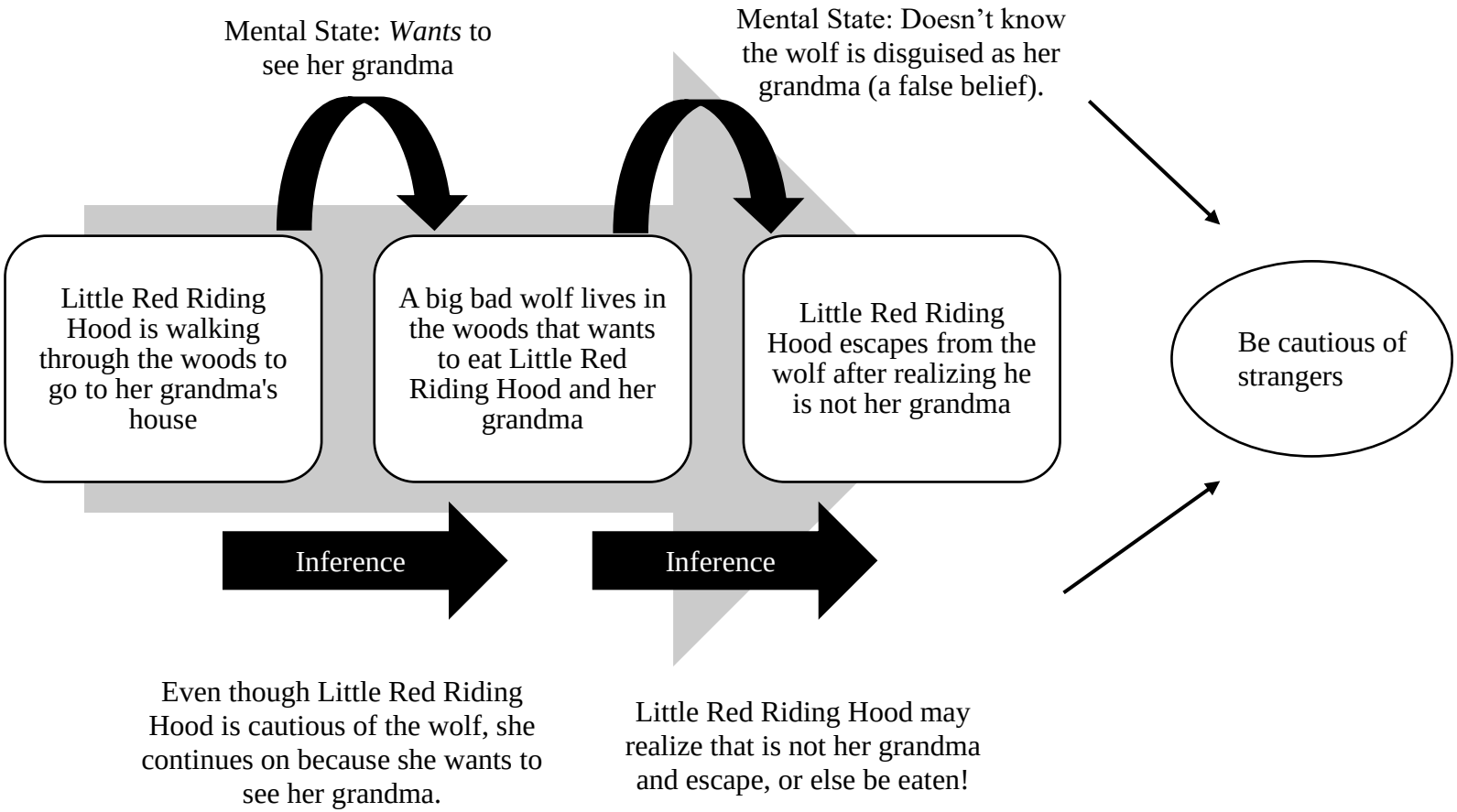
Theory of Mind and Narrative Structure



Chapter 3 Figure 2

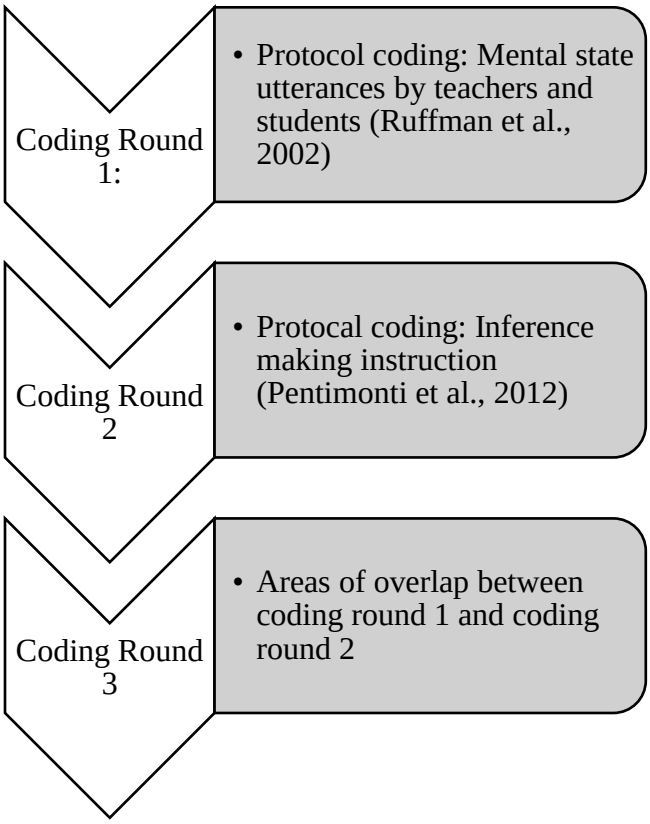
Narrative Structure with Theory of Mind

Little Red Riding Hood Narrative Structure with Theory of Mind



Chapter 3 Figure 3

Part 1 Pilot Protocol Coding Procedure



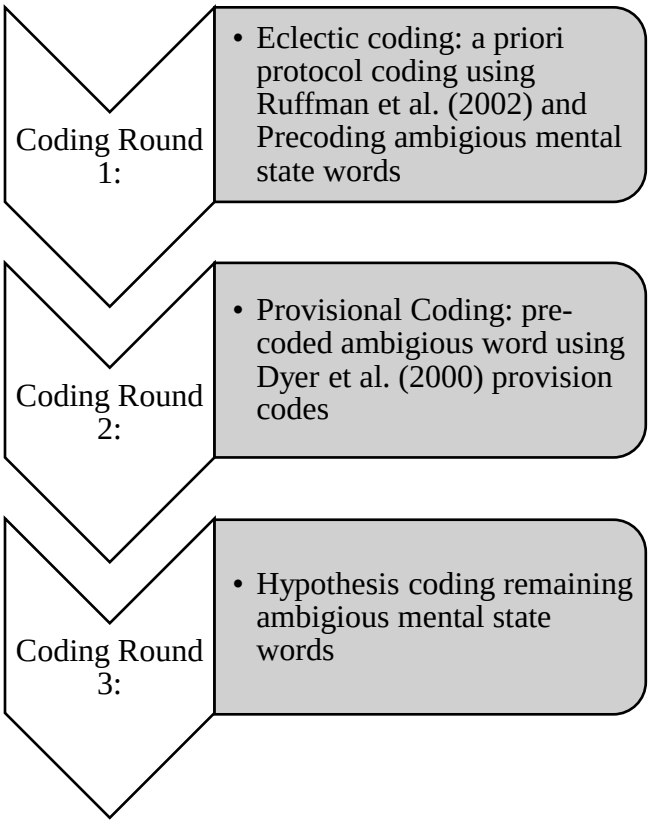
Chapter 4 Figure 4

Teacher and Student Instructional Sequence Dialogue

Speaker	Utterance
Teacher:	They had a big ……
Students:	Fight.
Teacher:	Fight.
Teacher:	Did they hit each other?
Students:	No.
Teacher:	No.
Teacher:	Did they use their words?
Students:	Yeah.
Teacher:	They did, but in an angry way, didn't they?

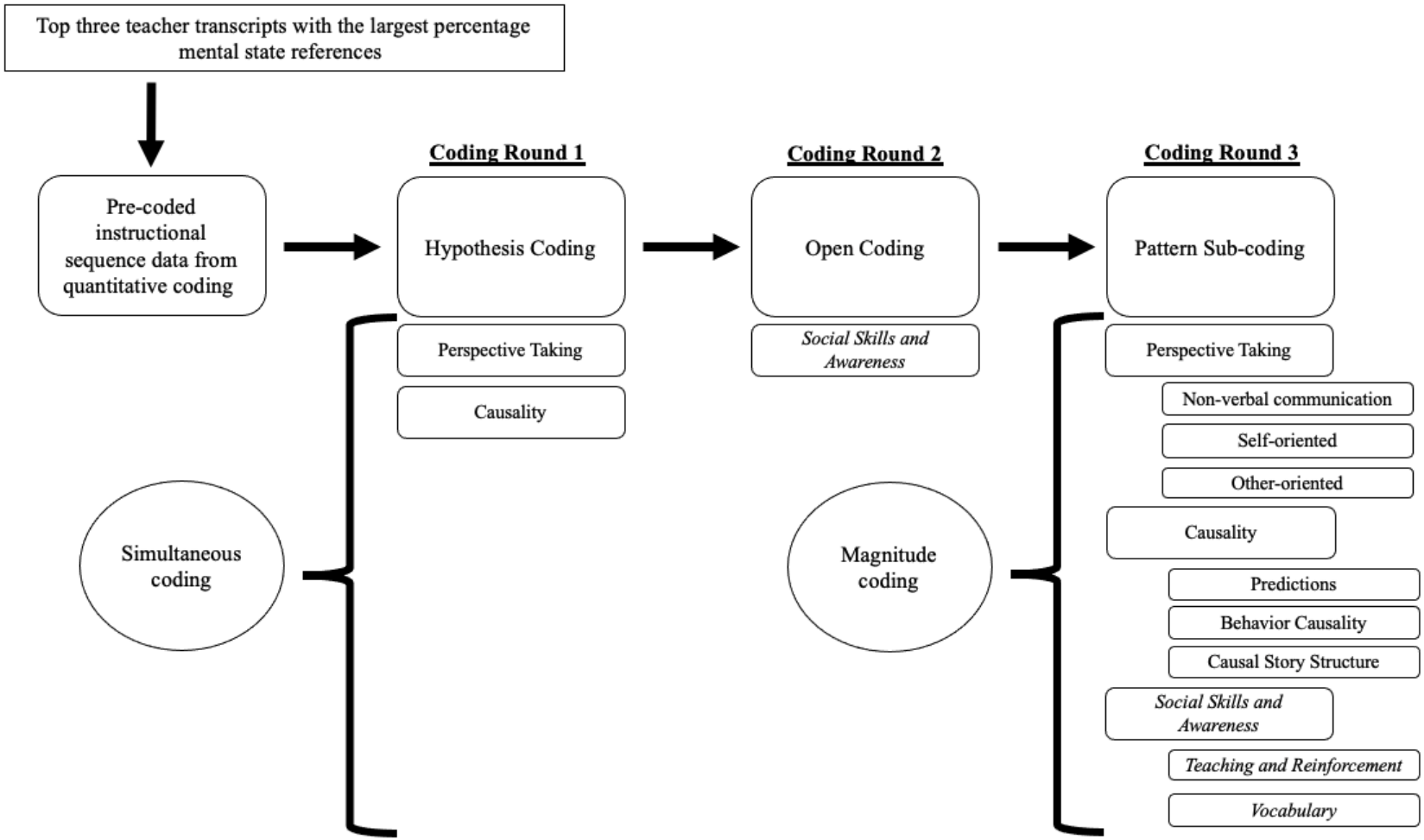
Chapter 3 Figure 5

Part 2 Coding Procedure



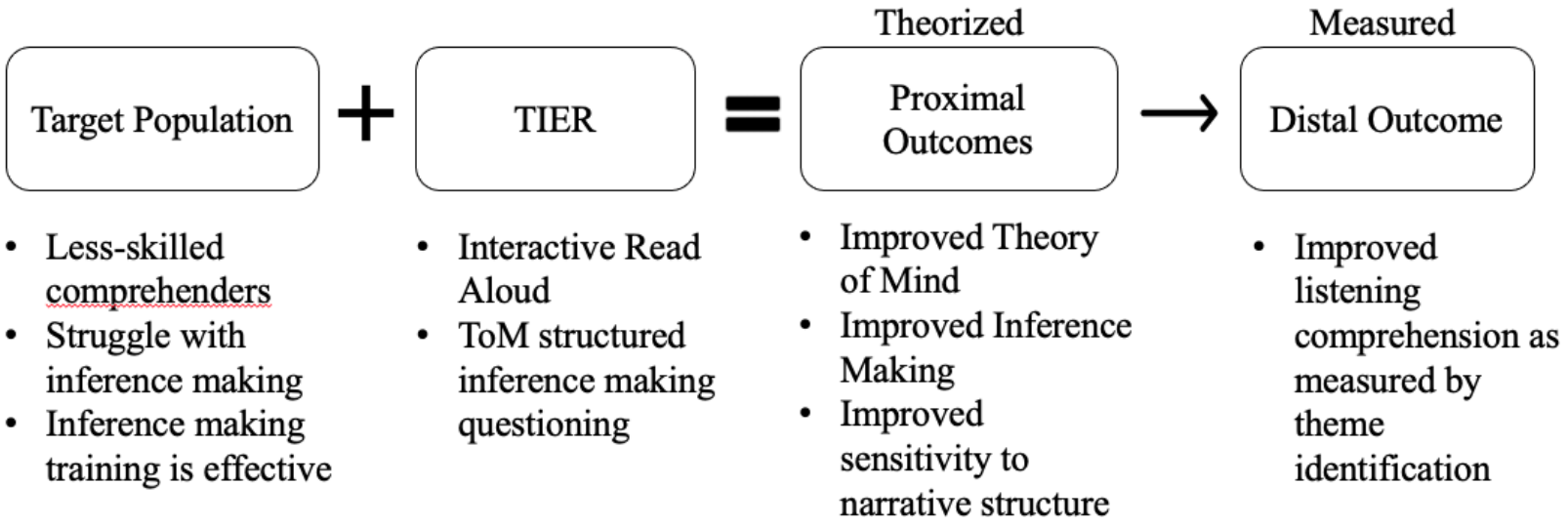
Chapter 3 Figure 6

Part 3 Coding Procedure



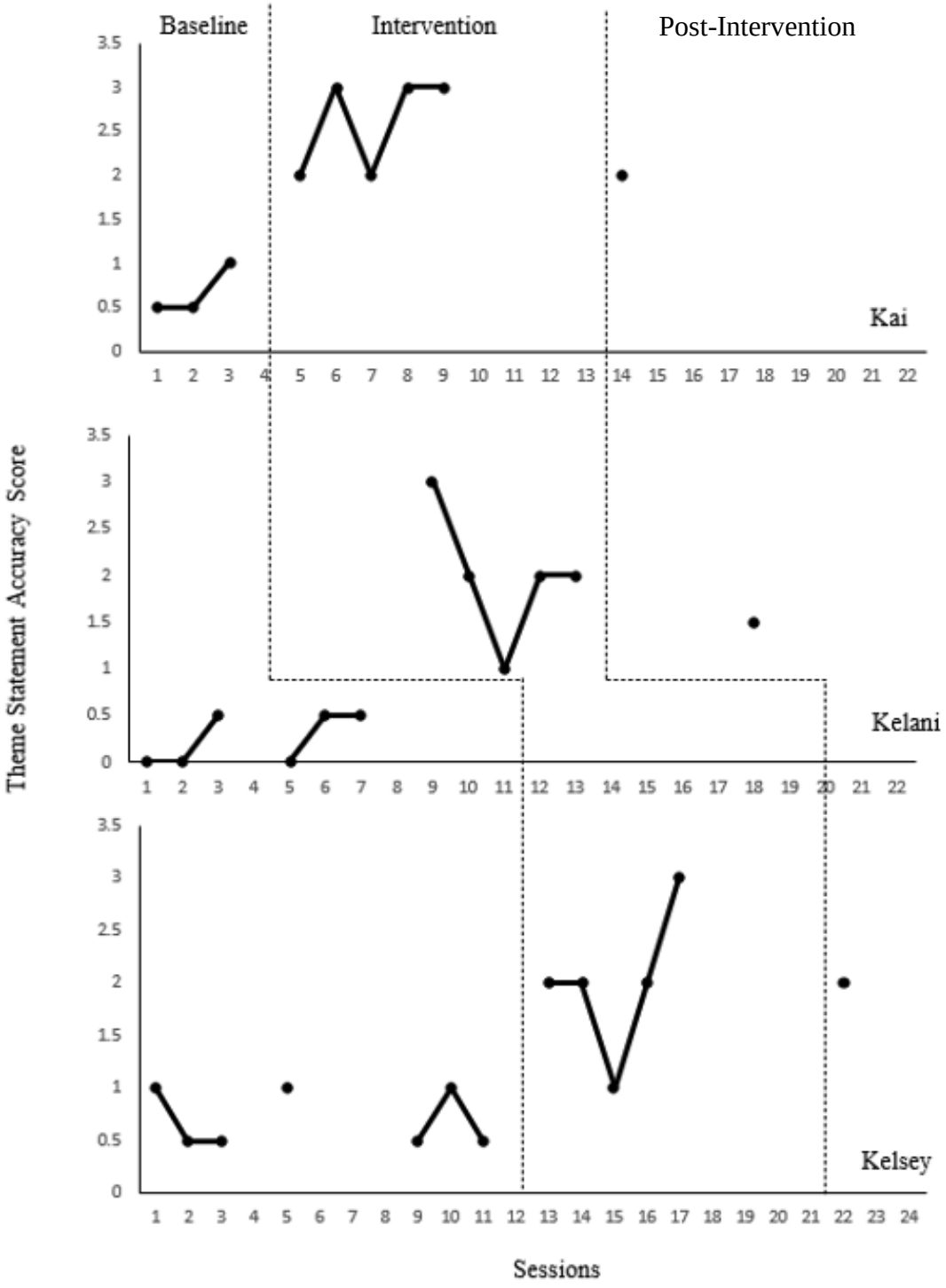
Chapter 4 Figure 1

Intervention Theory of Change



Chapter 4 Figure 2

Results of TIER on Participant's Theme Statement Accuracy



Appendix A: Intervention Materials

Appendix A Figure 1
Student Demographic Form for Caregiver

1. Student Name:

2. Date of Birth:

3. Gender:
   - Male
   - Female
   - Non-Binary

4. Race
   - American Indian or Alaska Native
   - Asian
   - Black or African American
   - Native Hawaiian or Other Pacific Islander
   - White

5. Ethnicity
   - Hispanic or Latinx or Spanish Origin
   - Not Hispanic or Latinx or Spanish Origin

6. Does your student qualify for free or reduced-price meals?
   - Yes
   - No

7. What is the primary language spoken in the home?
   - Yes
   - No

8. What is your child's primary language?

9. Does your child receive any outside of school educational services (e.g., speech or language services)?
   - Yes
   - No

10. If yes, what services does your child receive?

11. Does your child receive any inside of school educational services beyond typical instruction from their classroom teacher (e.g., speech or language services)?
   - Yes
   - No

Appendix A Figure 2

Instructional Books

Number	Title	Publication Date	Author	Lexile Level	# of Pages
1.	Jabari Jumps	2020	Gaia Cornwall	AD670L	32
2.	Morris Micklewhite and the Tangerine Dress	2020	Christine Baldacchino	AD530L	
3.	Pancho Rabbit and the Coyote: A Migrant's Tale	2013	Duncan Tonatiuh	AD590L	32
4.	Those Shoes	2009	Maribeth Boelts	AD550L	40
5.	Adrian Simcox Does NOT Have a Horse	2018	Marcy Campbell	AD600L	40
6.	The Invisible Boy	2013	Trudy Ludwig	AD680L	40
7.	The Proudest Blue	2019	Ibtihaj Muhammad	500L	40
8.	My Two Blankets	2015	Irena Kobald	AD480L	32
9.	Rosie Revere, Engineer	2013	Andrea Beaty	AD780L	32
10.	The Day the Crayons Quit	2013	Drew Daywalt	AD730L	40
11.	Giraffe Problems	2018	Jory John	AD530L	42
12.	When We Were Alone	2016	David Robertson	600L	32
13.	Life on Mars	2017	Jon Agee	AD470L	32
14.	The Name Jar	2003	Yangsook Choi	AD590L	40
15.	Owl Moon	1987	Jane Yolen	550L	32

Appendix A Figure 3

Student Social Validity Survey

The following statements will be read to you.

Circle only one face that matches what you think of each statement. The very happy face means you strongly agree, and the very sad face means strongly disagree. If you don't know, you can choose the face in the middle.

Question	Agree	Neutral	Disagree
1. I like reading with my adults.			
2. I learn a lot when I read with adults.			
3. Reading with adults helps me understand what I read better.			
4. Reading with adults help me become a better reader.			
5. Reading time is one of my favorite times during the day.			
6. Reading with an adult is one of my favorite activities during the school day.			

Appendix A Figure 4

Pelletier & Astington (2004) Coding Scheme

(1) Action-Only, (2) Consciousness Only, and (3) Coordinated Action-Consciousness, as follows:

1) Action-Only: verb phrases which referred to simple description of setting, events or actions, that is, the landscape of action; for example, “Here is a picture of Bunny Rabbit”, “That’s them going back to look”, “Rabbit is taking him to his friend Crocodile”, “And they’re running away”, “Brown Squirrel is taking the nut...and putting it in the flower pot”, “Grey Squirrel is looking in there again.”

2) Consciousness-Only: verb phrases which referred only to the picture within the thought-bubble without reference to the character who held the thought. From our perspective, this is the landscape of consciousness-only; for example, when shown a picture of Scott thinking about the squirrel looking under the bush, the Consciousness-Only score would be given for: “Squirrel looks under the bush.”

3) Coordinated Action-Consciousness: verb phrases which referred to characters’ thought, mentioning both the character and the thought content, that is, integrating the landscapes of action and consciousness; for example, “The boy thought the Grey one could find it in the plant” (picture of Scott thinking of Grey Squirrel looking under the bush), “The boy got an idea...he put the nut out of the flower pot... and he put it under the bush” (picture of Scott thinking about moving the nut from the flower pot to the bush), “Then the wolf had an idea how he could get that bunny in the pot” (picture of Wolf thinking about Rabbit in a stew pot), “And the rabbit is thinking about a crocodile...I want Wolf to meet my friend Crocodile” (picture of Rabbit thinking about Alligator while holding out his hand to Wolf).

Appendix A Figure 5

Book Familiarity Survey for Caregiver and Teacher

Please indicate the level of familiarity your child has with the book by circling one of the given choices (1 = not familiar, 2 = some familiarity, 3 = very familiar).

1. *I Will Surprise My Friend* by Mo Williams

    1 = not familiar 2 = some familiarity 3 = very familiar

2. *The Gruffalo* by Julia Donaldson

    1 = not familiar 2 = some familiarity 3 = very familiar

3. *One Cool Friend* by Toni Buzzeo

    1 = not familiar 2 = some familiarity 3 = very familiar

4. *Life on Mars* by Jon Agee

    1 = not familiar 2 = some familiarity 3 = very familiar

5. *The Day you Begin* by Jacqueline Woodson

    1 = not familiar 2 = some familiarity 3 = very familiar

6. *Thunder Boy Jr.* by Sherman Alexie

    1 = not familiar 2 = some familiarity 3 = very familiar

7. *The Name Jar* by Yangsook Choi

    1 = not familiar 2 = some familiarity 3 = very familiar

8. *Mango, Abuela, and Me* by Meg Medina

    1 = not familiar 2 = some familiarity 3 = very familiar

9. *Where's Jamela* by Niki Daly

    1 = not familiar 2 = some familiarity 3 = very familiar

10. *Grumpy Monkey* by Suzanne Lang

1 = not familiar 2 = some familiarity 3 = very familiar

11. *Owl Moon* by Jane Yolen

1 = not familiar 2 = some familiarity 3 = very familiar

12. *Exclamation Mark* by Amy Krouse Rosenthal

1 = not familiar 2 = some familiarity 3 = very familiar

13. *Hair Love* by Matthew Cherry

1 = not familiar 2 = some familiarity 3 = very familiar

14. *Grow Happy* by Jon Lasser

1 = not familiar 2 = some familiarity 3 = very familiar

15. *Small in the City* by Sydney Smith

1 = not familiar 2 = some familiarity 3 = very familiar

16. *The Day the Crayons Quit* by Drew Daywalt

1 = not familiar 2 = some familiarity 3 = very familiar

17. *Smokey Night* by Eve Bunting

1 = not familiar 2 = some familiarity 3 = very familiar

18. *Barkus, Sly, and the Golden Egg* by Angela McAllister

1 = not familiar 2 = some familiarity 3 = very familiar

19. *Pancho Rabbit and the Coyote* by Duncan Tonatiuh

1 = not familiar 2 = some familiarity 3 = very familiar

20. *The Invisible Boy* by Trudy Ludwig

1 = not familiar 2 = some familiarity 3 = very familiar

Appendix A Figure 6

Fidelity Checklist: Intervention Phases

<table>
<tr><th>Interactive Read Aloud Lesson</th><th colspan="3">Fidelity Rating</th></tr>
<tr><td rowspan="2">Before Reading
☐ T gains student assent
☐ T provides an introduction to the book
☐ T sets a purpose for reading
☐ T reviews relevant topic background knowledge
☐ T highlights key mental state vocabulary before reading
During Reading
☐ T asks questions at the pre-identified key points during interactive read aloud
☐ T asks inference making questions in lesson plan order
☐ T consistently provides corrective or reinforcing academic feedback on all inference making questions
☐ T models fluent reading
☐ T responds to spontaneous student questions during the interactive read aloud
After Reading
☐ T fills in graphic organizer with student's help
☐ T provides a summary of the story
☐ T scaffolds theme statement formation by asking scaffolded questions for students
☐ T's theme statement feedback includes a mental state word, an action, and story integration</td><td>3 = High</td><td>2 = Medium</td><td>1 = Low</td></tr>
<tr><td>Completes > 80%of the expected elements and procedures.</td><td>Completes between 80% and 50% of the quality indicators</td><td>Completes < 50% of the required elements and procedures</td></tr>
<tr><td colspan="4">Notes:</td></tr>
</table>

Appendix A Figure 7

Fidelity Checklist: Baseline

<table>
<tr><th>Interactive Read Aloud Lesson</th><th colspan="3">Fidelity Rating</th></tr>
<tr><td rowspan="2">Before Reading
☐ T gains student assent
☐ T provides an introduction to the book
☐ T sets a purpose for reading
☐ T reviews relevant topic background knowledge
During Reading
☐ T asks questions at the pre-identified key points during interactive read aloud
☐ T asks inference making questions in order
☐ T models fluent reading
☐ T responds to spontaneous student questions during the interactive read aloud
After Reading
☐ T asks the participant to identify the theme of the story</td><td>3 = High</td><td>2 = Medium</td><td>1 = Low</td></tr>
<tr><td>Completes <u>> 80%</u>of the expected elements and procedures.</td><td>Completes <u>between 80% and 50%</u> of the quality indicators</td><td>Completes <u>< 50%</u> of the required elements and procedures</td></tr>
<tr><td colspan="4">Notes:</td></tr>
</table>

Appendix B: Sample Lesson

Before Reading Script: Book Introduction

T gains student assent: Are you ready to read today?

T sets a purpose for reading: We will identify the theme of the story.

T reviews relevant topic background knowledge: Owling is when you go out at night to try and spot an owl

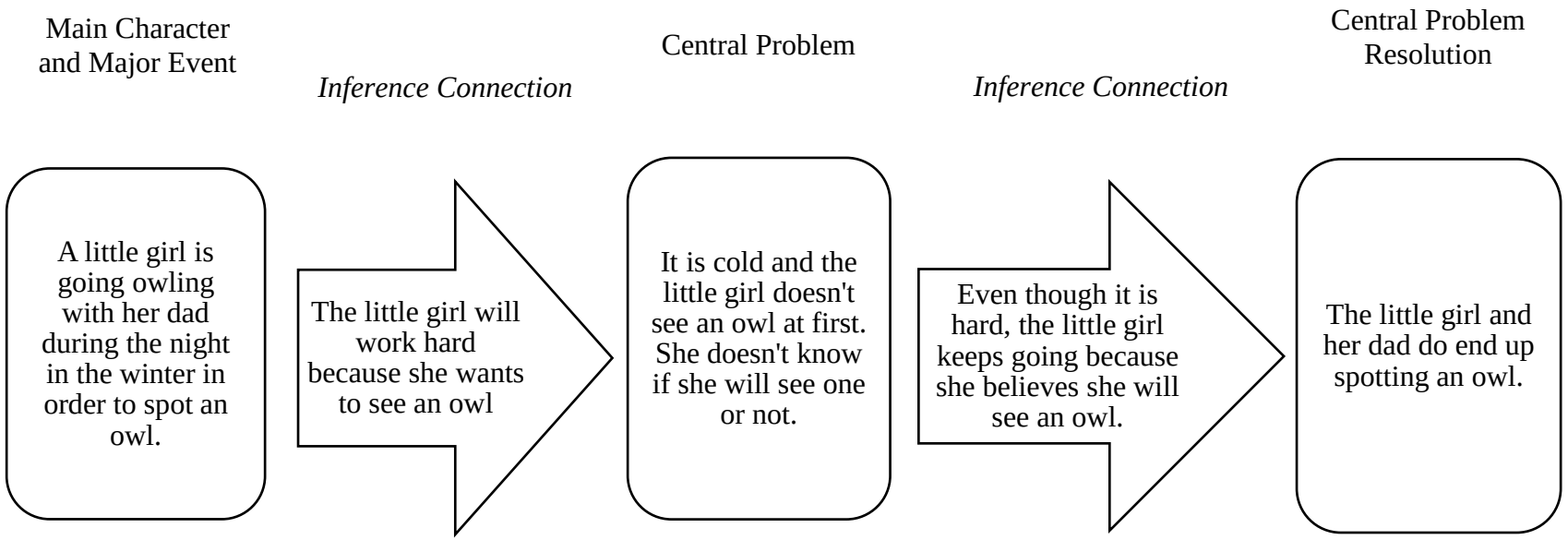
T highlights key mental state vocabulary before reading: [Ask student what words mean and correct or reinforce in response to student replies]

Believe
Doubt
Unaware

During Reading: [Read aloud the story and ask the questions at the pre-identified points in the story. Provide student's responses feedback to inference responses and fill in the graphic organizer.]

After Reading: Before we identify the theme of the story, let's review what the story is about by looking back at our graphic organizer [Provide a summary of text]. The story we read was about a girl who went owling her dad. It was cold and the little girl didn't see an owl at first, but then her and her dad finally saw an owl. See Figure 1. Now, let's write out the theme statement by using the thought word hope. In our inference connections I see the action of working hard and not giving up. I can use these actions in my theme statement. [Write out theme statement: Hope helps us not give up if something is hard.]

Narrative Structure Graphic Organizer



Theme statement: *Hope* (word that explains a feeling or a thought) helps us *keep going* (action) even if something is hard.

*Optional Follow-up Question if incorrect theme response: Can you think of a time when hope helped you do something hard?

Interactive Read Aloud Theory of Mind Question and Answer Plan for Owl Moon

Question	Purpose	Book Passage	Theory of Mind feedback for incorrect response: Modeling a Correct Response	Theory of Mind reinforcement for correct response
1. How do you think the little girl is feeling about going owling?	Introduction to central character and event	"I had been waiting to go owling with Pa for a long, long time."	I think she feels **excited** because the text says she had been waiting a long time which means she has **wanted** to go owling for a long time, and when I wait a long time for something, I'm **excited** when it finally happens.	Right! When I wait a long time for something, I'm **excited** when it finally happens. She must feel **excited** because she's waited for so long to go owling.
2. Why do you think the little girl says she isn't disappointed?	Problem identification and resolution	"Again, he called out. And then again. After each call he was silent and for a moment we both listened. But there was no answer. Pa shrugged. I shrugged. I was not **disappointed**. My brothers all said sometimes there's an owl and sometimes there isn't."	I think she says that she isn't **disappointed** because in the sentence before it says, "there was no answer." This means that no owls responded to Pa's call. The little girl could be **disappointed** because something she **wanted** didn't happen, but her brothers said this might happen, so she was ready for it, and wasn't **disappointed** it hadn't happened yet. She isn't giving up.	Right! She **wanted** to hear an owl respond to the call from her Pa and it didn't happen, so she could have been **disappointed.** But her brothers said this might happen, so she was ready for it, and wasn't **disappointed** it hadn't happened yet. She isn't giving up.
3. Why does the little girl say you only need hope when you go owling?	Theme Identification	"When you go owling, you don't need words or warmth, or anything but **hope**. That's what Pa says. The kind of **hope** that flies on silent wings under a shining Owl Moon."	I think she says that you only need **hope** because she **believed** they would see an owl even though she felt cold and maybe **wanted** to go home. She was brave and patient and, in the end, she was able to see an owl because she never gave up **hope. Hope** helped her keep going.	Right! The little girl can't **know** she is going to see an owl. If she wants to see an owl, she has to **believe** that she could see one because that belief keeps her going. It gives her **hope**. **Hope** helped her keep going.

Note. Adapted from Bianco and Lecce (2016).
Italic items in bold indicate the use of mental-state verbs in teacher-student interactions.

Interactive Read Aloud Inference Connection Question and Answer Plan for Owl Moon

Question	**Purpose**	**Inference feedback for incorrect response: Modeling a Correct Response**	**Inference reinforcement for correct response**
1. How does the little girl's excitement help her work hard while owling in the cold?	Establish a link between the central character and activity to the main problem	The little girl's excitement helps because it gives her the strength to her work hard even though it is cold.	Right! When I am excited to do something, I really want, I will work hard to do it. The little girl will work hard to see an owl.
2. Why is it important that the little girl isn't disappointed?	Establish a link between the main problem to the problem's resolution.	It is important that the little girl isn't disappointed because it means she will not give up trying to spot an owl.	Right! She isn't disappointed which means she won't give up in trying to spot an owl!
3. What is the theme of Owl Moon?	Identify Theme	The little girl had hope, which helped her keep going to see an owl even though it was hard because it was cold, at night, and took a long time. The theme is that hope helps us keep going if something is hard.	Right! The theme is that hope helps us keep going if something is hard.

Note. Adapted from Bianco and Lecce (2016).

References

Adlof, S.M., Catts, H.W., & Lee, J. (2010). Kindergarten predictors of second vs. eighth grade reading comprehension impairments. *Journal of Learning Disabilities, 43*(4), 332–345. https://doi.org/10.1177/0022219410369067

Adrián, J. E., Clemente, R. A., & Villanueva, L. (2007). Mother's use of cognitive state verbs in picture-book reading and the development of children's understanding of mind: a longitudinal study. *Child Development, 78*(4), 1052–1067. http://dx.doi.org/10.1111/j.1467-8624.2007.01052.x

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher, 61*(5), 364–373. https://doi.org/10.1598/RT.61.5.1

Alexander, P. A., & The Disciplined Reading and Learning Research Laboratory. (2012). Reading into the future: Competence for the 21st century. *Educational Psychologist, 47*(4), 259–280. https://doi.org/10.1080/00461520.2012.722511

Allington, R. L., McCuiston, K., & Billen, M. (2015). What research says about text complexity and learning to read. *The Reading Teacher, 68*(7), 491–501. https://doi.org/10.1002/trtr.1280

Alvarenga, P., Zucker, T. A., Tambyraja, S., & Justice, L. (2020). Contingency in teacher-child emotional state talk during shared book reading in early childhood classrooms. *Early Education and Development, 31*(8), 1187–1205. https://doi.org/10.1080/10409289.2020.1722786

Apperly, I. (2010). Mindreaders: *The cognitive basis of theory of mind.* Psychology Press.

Aram, D., Fine, Y., & Ziv, M. (2013). Enhancing parent-child shared book reading interactions: Promoting references to the book's plot and socio-cognitive themes. *Early Childhood Research Quarterly, 28*(1), 111–122. https://doi.org/10.1016/j.ecresq.2012.03.005

Archer, A. L., & Hughes, C. A. (2010). *Explicit instruction: Effective and efficient teaching*. Guilford Press.

Astington, J. W. (1990). Narrative and the child's theory of mind. In B. Britton & A. D. Pellegrini (Eds.), *Narrative thought and narrative language* (pp. 151–171). Erlbaum.

Astington, J. W. (2003). Sometimes necessary, never sufficient: False-belief understanding and social competence. In B. Repacholi & V. Slaughter (Eds.). *Individual differences in theory of mind: Implications for typical and atypical development* (p. 13–38). Psychology Press.

Astington, J. W., & Baird, J. A. (Eds.). (2005). *Why language matters for theory of mind*. Oxford University Press.

Atkinson, L., Slade, L., Powell, D., & Levy, J. P. (2017). Theory of mind in emerging reading comprehension: A longitudinal study of early indirect and direct effects. *Journal of Experimental Child Psychology, 164*, 225–238. https://doi.org/10.1016/j.jecp.2017.04.007

Baker, S. K., Santoro, L. E., Chard, D. J., Fien, H., Park, Y., & Otterstedt, J. (2013). An evaluation of an explicit read aloud intervention taught in whole-classroom formats in

first grade. *The Elementary School Journal, 113*(3), 331–358. https://doi.org/10.1086/668503

Barnert, E. S., Perry, R., Shetgiri, R., Steers, N., Dudovitz, R., Heard-Garris, N. J., ... & Chung, P. J. (2021). Adolescent protective and risk factors for incarceration through early adulthood. Journal of Child and Family Studies, 30(6), 1428–1440. https://doi.org/10.1007/s10826-021-01954-y

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind?" *Cognition, 21*(1), 37–46. https://doi.org/10.1016/0010-0277(85)90022-8

Baron-Cohen, S., O'riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders,* 29(5), 407–418. https://doi.org/10.1023/A:1023035012436

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42*(2), 241–251. https://doi.org/10.1111/1469-7610.00715

Beaudoin, C., Leblanc, É., Gagner, C., & Beauchamp, M. H. (2020). Systematic review and inventory of theory of mind measures for young children. *Frontiers in psychology*, 10, 2905. https://doi.org/10.3389/fpsyg.2019.02905

Bergman Deitcher, D., Aram, D., Khalaily-Shahadi, M., & Dwairy, M. (2021). Promoting preschoolers' mental-emotional conceptualization and social understanding: A shared

book-reading study. *Early Education and Development, 32*(4), 501–515. https://doi.org/10.1080/10409289.2020.1772662

Best, R. M., Floyd, R. G., & McNamara, D. S. (2008). Differential competencies contributing to children's comprehension of narrative and expository texts. *Reading Psychology, 29*(2), 137–164. https://doi.org/10.1080/02702710801963951

Bianco, F., & Lecce, S. (2016). Translating child development research into practice: Can teachers foster children's theory of mind in primary school? *British Journal of Educational Psychology*, *86*(4), 592–605. https://doi.org/10.1111/bjep.12125

Bianco, F., Lombardi, E., Massaro, D., Castelli, I., Valle, A., Marchetti, A., & Lecce, S. (2019). Enhancing advanced Theory of Mind skills in primary school: A training study with 7- to 8-year-olds. *Infant and Child Development, 28*(6), e2155. https://doi.org/10.1002/icd.2155

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*(2), 647–663. https://doi.org/10.1111/j.1467-8624.2007.01019.x

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false-belief task as a test of theory of mind. *Cognition, 77*, 25–31. https://doi.org/10.1016/S0010-0277(00)00096-2

Blything, L. P., Hardie, A., & Cain, K. (2020). Question asking during reading comprehension instruction: A corpus study of how question type influences the linguistic complexity of primary school students' responses. *Reading Research Quarterly*, *55*(3), 443–472. https://doi.org/10.1002/rrq.279

Bodgan, R., & Biklin, S. K. (2007). Qualitative research for education: An introduction to theory and methods (5th ed.). Pearson.

Boerma, I. E., Mol, S. E., & Jolles, J. (2017). The role of home literacy environment, mentalizing, expressive verbal ability, and print exposure in third and fourth graders' reading comprehension. *Scientific Studies of Reading*, *21*(3), 179–193. https://doi.org/10.1080/10888438.2016.1277727

Bohn-Gettler, C. M., & Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology*, *39*(3), 206–219. https://doi.org/10.1016/j.cedpsych.2014.05.003

Bowyer-Crane, C., & Snowling, M. J. (2005). Assessing children's inference generation: What do tests of reading comprehension measure?. *British Journal of Educational Psychology, 75*(2), 189–201. https://doi.org/10.1348/000709904X22674

Brantlinger, E., Jimenez, R., Klingner, J., Pugach, M., & Richardson, V. (2005). Qualitative studies in special education. *Exceptional Children, 71*(2), 195–207. https://doi.org/10.1177/001440290507100205

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology, 18*(6), 906. https://doi.org/10.1037/0012-1649.18.6.906

Brooks, W., & Browne, S. (2012). Towards a culturally situated reader response theory. *Children's Literature in Education, 43*(1), 74–85. https://doi.org/10.1007/s10583-011-9154-z

Brown, H. M., Oram-Cardy, J., & Johnson, A. (2013). A meta-analysis of the reading comprehension skills of individuals on the autism spectrum. *Journal of Autism and Developmental Disorders, 43*(4), 932–955. https://doi.org/10.1007/s10803-012-1638-1

Browning, E. (2021). *Bridging the Gap Between Past and Present: Narrative Nonfiction in the Primary History Classroom* (Doctoral dissertation, King's College London).

Bruner, J. (1987). Life as narrative. *Social Research, 54*(1), 305–314.

Butterfuss, R., & Kendeou, P. (2018). The role of executive functions in reading comprehension. *Educational Psychology Review, 30*, 801–826. https://doi.org/10.1007/s10648-017-9422-6

Cain, K, Oakhill, J., Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The-influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology*, *96*(4), 671–681. https://doi.org/10.1037/0022-0663.96.4.671

Cain, K. (2010). *Reading development and difficulties: An introduction*. Wiley-Blackwell.

Cain, K. E., & Oakhill, J. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing, 11*(5-6), 489–503. https://doi.org/10.1023/A:1008084120205

Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology, 76*(4), 683-696. https://doi.org/10.1348/000709905X67610

Cain, K., & Oakhill, J. (2011). Matthew effects in young readers: Reading comprehension and reading experience aid vocabulary development. *Journal of Learning Disabilities*, *44*(5), 431–443. https://doi.org/10.1177/0022219411410042

Cain, K., & Oakhill, J. (2014). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension?. *L'Annee Psychologique, 114*(4), 647–662. https://doi.org/10.4074/S0003503314004035

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition, 29*(6), 850–859. https://doi.org/10.3758/BF03196414

Cantin, R. H., Gnaedinger, E. K., Gallaway, K. C., Hesson-McInnis, M. S., & Hund, A. M. (2016). Executive functioning predicts reading, mathematics, and theory of mind during the elementary years. *Journal of Experimental Child Psychology*, *146*, 66–78. https://doi.org/10.1016/j.jecp.2016.01.014

Carlson, S. E., van den Broek, P., McMaster, K., Rapp, D. N., Bohn-Gettler, C. M., Kendeou, P., & White, M. J. (2014). Effects of comprehension skill on inference generation during reading. *International Journal of Disability, Development and Education*, *61*(3), 258–274. https://doi.org/10.1080/1034912X.2014.934004

Carpenter, R. D., & Paris, S. G. (2005). Issues of validity and reliability in early reading assessments. In S. G. Paris and S. A. Stahl (Eds.), *Children's reading comprehension and assessment*, (pp. 279–304). Lawrence Erlbaum.

Cartwright, K. B. (2015). *Executive skills and reading comprehension: A guide for educators*. Guilford Publications.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest, 19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Catts, H. W. (2021). Rethinking How to Promote Reading Comprehension. *American Educator, Winter, 2022*.

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*. *49*(2), 278–93. https://doi.org/10.1044/1092-4388(2006/023)

Catts, H. W., Hogan, T. P., & Fey, M. E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities*, *36*(2), 151–164. https://doi.org/10.1177/002221940303600208

Catts, H. W., Nielsen, D. C., Bridges, M. S., & Liu, Y. S. (2016). Early identification of reading comprehension difficulties. *Journal of Learning Disabilities, 49*(5), 451–465. https://doi.org/10.1177/0022219414556121

Catts, H. W., Nielsen, D. C., Bridges, M. S., Liu, Y. S., & Bontempo, D. E. (2015). Early identification of reading disabilities within an RTI framework. *Journal of Learning Disabilities, 48*(3), 281–297. https://doi.org/10.1177/0022219413498115

Chall, J. (1983). *Stages of reading development*. McGraw-Hill.

Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis: Why poor children fall behind*. Harvard University Press.

Chapman, E., Baron-Cohen, S., Auyeung, B., Knickmeyer, R., Taylor, K., & Hackett, G. (2006). Fetal testosterone and empathy: evidence from the empathy quotient (EQ) and the "reading the mind in the eyes" test. *Social Neuroscience, 1*(2), 135–148. https://doi.org/10.1080/17470910600992239

Cho, E., Capin, P., Roberts, G., Roberts, G. J., & Vaughn, S. (2019). Examining sources and mechanisms of reading comprehension difficulties: Comparing English learners and non-English learners within the simple view of reading. *Journal of Educational Psychology, 111*(6), 982–1000. https://doi.org/10.1037/edu0000332

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science, 21*(8), 1106–1116. https://doi.org/10.1177/0956797610375449

Cohen, C., & Cohen, R. West & Aiken (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (5th ed.). Erlbaum

Connor, C. M., Phillips, B. M., Kaschak, M., Apel, K., Kim, Y. S., Al Otaiba, S., ... & Lonigan, C. J. (2014). Comprehension tools for teachers: Reading for understanding from prekindergarten through fourth grade. *Educational Psychology Review, 26*(3), 379–401. https://doi.org/10.1007/s10648-014-9267-1

Cook, A. E., & O'Brien, E. J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes, 51*(1–2), 26–49. https://doi.org/10.1080/0163853X.2013.855107

Cook, A. E., & O'Brien, E. J. (2017). Fundamentals of inferencing during reading. *Language and Linguistics Compass, 11*(7), e12246. https://doi.org/10.1111/lnc3.12246

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). Russell Sage Foundation.

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *99*(2), 311. https://doi.org/10.1037/0022-0663.99.2.311

Curenton, S. M. (2004). The association between narratives and theory of mind for low-income preschoolers. *Early Education and Development*, *15*(2), 124–146. https://doi.org/10.1207/s15566935eed1502_1

Curenton, S. M. (2011). Understanding the landscapes of stories: The association between preschoolers' narrative comprehension and production skills and cognitive abilities. *Early Child Development and Care, 181*(6), 791–808. https://doi.org/10.1080/03004430.2010.490946

Currie, N. K., & Cain, K. (2015). Children's inference generation: The role of vocabulary and working memory. *Journal of Experimental Child Psychology, 137*, 57–75. https://doi.org/10.1016/j.jecp.2015.03.005

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

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*, *10*(3), 277–299. https://doi.org/10.1207/s1532799xssr1003_5

Dahlstrom, M. F. (2014). Using narratives and storytelling to communicate science with nonexpert audiences. Proceedings of the National Academy of Sciences, 111(Supplement 4), 13614–13620. https://doi.org/10.1073/pnas.1320645111

Davidson, M. M., & Weismer, S. (2018). A preliminary investigation of parent-reported fiction versus non-fiction book preferences of school-age children with autism spectrum disorder. *Autism & Developmental Language Impairments, 3,* 1–12. https://doi.org/10.1177/2396941518806109

Deshmukh, R. S., Zucker, T. A., Tambyraja, S. R., Pentimonti, J. M., Bowles, R. P., & Justice, L. M. (2019). Teachers' use of questions during shared book reading: Relations to child responses. *Early Childhood Research Quarterly, 49*, 59–68. https://doi.org/10.1016/j.ecresq.2019.05.006

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

Devine, R. T., & Hughes, C. (2019). Let's talk: Parents' mental talk (not mind-mindedness or mindreading capacity) predicts children's false belief understanding. *Child Development*, *90*(4), 1236–1253. https://doi.org/10.1111/cdev.12990

Dexter, D. D., & Hughes, C. A. (2011). Graphic organizers and students with learning disabilities: A meta-analysis. *Learning Disability Quarterly, 34*(1), 51–72. https://doi.org/10.1177/073194871103400104

Diakidoy, I. A. N., Stylianou, P., Karefillidou, C., & Papageorgiou, P. (2005). The relationship between listening and reading comprehension of different types of text at increasing

grade levels. *Reading Psychology*, *26*(1), 55–80. https://doi.org/10.1080/02702710590910584

Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language: Why language is central to reading development. *Educational Researcher, 39*(4), 305–310. https://doi.org/10.3102/0013189X10370204

Djikic, M., Oatley, K., & Moldoveanu, M. C. (2013). Reading other minds: Effects of literature on empathy. Scientific study of literature, 3(1), 28–47. https://doi.org/10.1075/ssol.3.1.06dji

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

Donahue, M. (2014). Perspective-taking and reading comprehension of narratives: Lessons learned from The Bean. In C. A. Stone, E. Silliman, B. Ehren, & G. Wallach (Eds.), *Handbook of language and literacy development and disorders* (2nd ed., pp 323–338). Guilford Press.

Dore, R. A., Amendum, S. J., Golinkoff, R. M., & Hirsh-Pasek, K. (2018). Theory of Mind: a Hidden Factor in Reading Comprehension? *Educational Psychology Review*, *30*(3), 1–23. https://doi.org/10.1007/s10648-018-9443-9

Dray, A. J. (2018). Perspective taking in fiction: A comparison of two texts. *The Elementary School Journal, 118*(3), 452–476. https://doi.org/10.1086/696148

Dray, A. J., & Selman, R. L. (2010). Culture and comprehension: A mixed methods study of children's responses to a fictional story about interracial conflict. *Reading & Writing Quarterly, 27*(1-2), 48–74. https://doi.org/10.1080/10573569.2011.532714

Duke, N. K., & Cartwright, K. B. (2018). The DRIVE model of reading: Deploying reading in varied environments. In D.E. Alvermann, N.J. Unrau, M. Sailors, and R.B. Ruddell (Eds.), *Theoretical models and processes of literacy* (7th Ed., pp. 134–151). Routledge.

Duke, N. K., & Cartwright, K. B. (2021). The science of reading progresses: Communicating advances beyond the simple view of reading. *Reading Research Quarterly, 56,* S25–S44. https://doi.org/10.1002/rrq.411

Dvash, J., & Shamay-Tsoory, S. G. (2014). Theory of mind and empathy as multidimensional constructs: Neurological foundations. *Topics in Language Disorders, 34*(4), 282–295. https://doi.org/10.1097/TLD.0000000000000040

Dyer, J. R., Shatz, M., & Wellman, H. M. (2000). Young children's storybooks as a source of mental state information. *Cognitive Development, 15*(1), 17–37. https://doi.org/10.1016/S0885-2014(00)00017-4

Eason, S. H., Goldberg, L. F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader-text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. *Journal of Educational Psychology, 104*(3), 515. https://doi.org/10.1037/a0027182

Ebert, S. (2020a). Early language competencies and advanced measures of mental state understanding are differently related to listening and reading comprehension in early adolescence. *Frontiers in Psychology*, *11*, 952. https://doi.org/10.3389/fpsyg.2020.00952

Ebert, S. (2020b). Theory of mind, language, and reading: Developmental relations from early childhood to early adolescence. *Journal of Experimental Child Psychology*, *191*, 104739. https://doi.org/10.1016/j.jecp.2019.104739

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Klingler Tackett, K., & Wick Schnakenberg, J. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research, 79*(1), 262–300. https://doi.org/10.3102/0034654308325998

Elleman, A. M. (2017). Examining the impact of inference instruction on the literal and inferential comprehension of skilled and less skilled readers: A meta-analytic review. *Journal of Educational Psychology*, *109*(6), 761. https://doi.org/10.1037/edu0000180

Elleman, A. M., & Oslund, E. L. (2019). Reading comprehension research: Implications for practice and policy. *Policy Insights from the Behavioral and Brain Sciences, 6*(1), 3–11. https://doi.org/10.1177/2372732218816339

Emery, D. W. (1996). Helping readers comprehend stories from the characters' perspective. *The Reading Teacher, 49*(7), 534–541

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fecica, A. M., & O'Neill, D. K. (2010). A step at a time: Preliterate children's simulation of narrative movement during story comprehension. *Cognition, 116*(3), 368–381. https://doi.org/10.1016/j.cognition.2010.05.014

Fisher, C., Gertner, Y., Scott, R. M., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(2), 143–149. https://doi.org/10.1002/wcs.17

Fiske, S., & Taylor, S. E. (2013). *Social cognition* (2nd Ed.). Mcgraw-Hill.

Flack, Z. M., Field, A. P., & Horst, J. S. (2018). The effects of shared storybook reading on word learning: A meta-analysis. *Developmental Psychology, 54*(7), 1334–1346. https://doi.org/10.1037/dev0000512

Flavell, J. (1963). *The developmental psychology of Jean Piage*t. D Van Nostrand. http://dx.doi.org/10.1037/11449-000

Florit, E., Cain, K., & Mason, L. (2019). Going beyond children's single-text comprehension: The role of fundamental and higher-level skills in 4th graders' multiple-document comprehension. *British Journal of Educational Psychology*. Advanced online publication. https://doi.org/10.1111/bjep.12288

Flynn, S., & Hardman, M. (2019). The use of interactive fiction to promote conceptual change in science. *Science & Education, 28*(1), 127–152. https://doi.org/10.1007/s11191-019-00032-6

Foorman, B. R., Petscher, Y., & Herrera, S. (2018). Unique and common effects of decoding and language factors in predicting reading comprehension in grades 1–10. *Learning and Individual Differences, 63,* 12–23. https://doi.org/10.1016/j.lindif.2018.02.011

Francis, D. J., Kulesz, P. A., & Benoit, J. S. (2018). Extending the simple view of reading to account for variation within readers and across texts: The complete view of reading (CVRi). *Remedial and Special Education, 39*(5), 274–288. https://doi.org/10.1177/0741932518772904

Francis, J. J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M. P., & Grimshaw, J. M. (2010). What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychology and Health, 25*(10), 1229–1245. https://doi.org/10.1080/08870440903194015

Freed, J., & Cain, K. (2017). Assessing school-aged children's inference-making: the effect of story test format in listening comprehension. *International Journal of Language & Communication Disorders, 52*(1), 95–105. https://doi.org/10.1111/1460-6984.12260

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*(4), 483–527. https://doi.org/10.1016/0885-2014(95)90024-1

Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., Yang, N. J., Braun, M., O'Connor R. E. (2001). Is reading important in reading-readiness programs? A randomized field trial with teachers as program implementers. *Journal of Educational Psychology, 93*(2), 251–267. https://doi.org/10.1037/0022-0663.93.2.251

Galda, L., & Beach, R. (2001). Response to literature as a cultural activity. *Reading Research Quarterly, 36*(1), 64–73. https://doi.org/10.1598/RRQ.36.1.4

Gerace, A., Day, A., Casey, S., & Mohr, P. (2017). 'I think, you think': understanding the importance of self-reflection to the taking of another person's perspective. *Journal of Relationships Research, 8*. https://doi.org/10.1017/jrr.2017.8

Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research, 71*(2), 279–320. https://doi.org/10.3102/00346543071002279

Geva, E., & Massey-Garrison, A. (2013). A comparison of the language skills of ELLs and monolinguals who are poor decoders, poor comprehenders, or normal readers. *Journal of Learning Disabilities, 46*(5), 387–401. https://doi.org/10.1177/0022219412466651

Glaser, B. G. and Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research.* Aldine Publishing

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*(10), 3–8. https://doi.org/10.3102/0013189X005010003

Goldman, S. R., & Pellegrino, J. W. (2015a). Research on learning and instruction: Implications for curriculum, instruction, and assessment. *Policy Insights from the Behavioral and Brain Sciences, 2*(1), 33–41.https://doi.org/10.1177/2372732215601866

Goldman, S. R., McCarthy, K. S., & Burkett, C. (2015b). Interpretive inferences in literature. In E. O'Brien, A. Cook, & R. Lorch (Eds.), *Inferences during reading* (pp. 386–415). Cambridge University Press.

Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2019). Language matters: Denying the existence of the 30-million-word gap has serious consequences. *Child Development, 90*(3), 985–992. https://doi.org/10.1111/cdev.13128

Good, R.H., & Kaminski, R. A. (Eds.) (2002). Dynamic Indicators of Basic Early Literacy Skills (6th ed.): Institute for the Development of Educational Achievement. Available: https://dibels.uoregon.edu/

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development,* 26–37. https://doi.org/10.2307/1130386

Goswami, U., & Bryant, P. (2016). *Phonological skills and learning to read*. Psychology Press.

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

Graesser, A. C. (2007). An introduction to strategic reading comprehension. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 3–26). Lawrence Erlbaum.

Graesser, A. C. (2015). Deeper learning with advances in discourse science and technology. *Policy Insights from the Behavioral and Brain Sciences, 2*(1), 42–50. https://doi.org/10.1177/2372732215600888

Graesser, A. C., & Kreuz, R. J. (1993). A theory of inference generation during text comprehension. *Discourse Processes, 16*, 145–160. https://doi.org/10.1080/01638539309544833

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational researcher, 40*(5), 223–234. https://doi.org/10.3102/0013189X11413260

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36*(2), 193–202. https://doi.org/10.3758/BF03195564

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48*(1), 163–189. https://doi.org/10.1146/annurev.psych.48.1.163

Graesser, A. C., Person, N. K., & Hu, X. (2002). Improving Comprehension through Discourse Processing. *New Directions for Teaching and Learning*, 89, 33–44. https://doi.org/10.1002/tl.45

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*(3), 371–395. https://doi.org/10.1037/0033-295X.101.3.371

Guajardo, N. R., & Cartwright, K. B. (2016). The contribution of theory of mind, counterfactual reasoning, and executive function to pre-readers' language comprehension and later reading awareness and comprehension in elementary school. *Journal of Experimental Child Psychology, 144*, 27–45. https://doi.org/10.1016/j.jecp.2015.11.004

Guajardo, N. R., & Turley-Ames, K. J. (2004). Preschoolers' generation of different types of counterfactual statements and theory of mind understanding. *Cognitive Development, 19*(1), 53–80. https://doi.org/10.1016/j.cogdev.2003.09.002

Guajardo, N. R., & Watson, A. C. (2002). Narrative discourse and theory of mind development. *The Journal of Genetic Psychology*, *163*(3), 305–325. https://doi.org/10.1080/00221320209598686

Guajardo, N. R., Parker, J., & Turley-Ames, K. (2009). Associations among false belief understanding, counterfactual reasoning, and executive function. *British Journal of Developmental Psychology, 27*(3), 681–702. https://doi.org/10.1348/026151008X357886

Gygax, P., & Gillioz, C. (2015). Emotion inferences during reading: Going beyond the tip of the iceberg. l. In E. J. O'Brien, A. E. Cook, & R. F. Lorch, Jr (Eds.), *Inferences during reading* (pp. 122–139). Cambridge University Press.

Gygax, P., Tapiero, I., & Carruzzo, E. (2007). Emotion inferences during reading comprehension: What evidence can the self-pace reading paradigm provide?. *Discourse processes*, *44*(1), 33–50. https://doi.org/10.1080/01638530701285564

Haager, D., & Vaughn, S. (2013). The common core state standards and reading: Interpretations and implications for elementary students with learning disabilities. *Learning Disabilities Research & Practice*, *28*(1), 5–16. https://doi.org/10.1111/ldrp.12000

Hall, C. S. (2016). Inference instruction for struggling readers: A synthesis of intervention research. *Educational Psychology Review*, *28*(1), 1–22. https://doi.org/10.1007/s10648-014-9295-x

Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive psychology*, *44*(4), 339–387. https://doi.org/10.1006/cogp.2001.0769

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

Happé, F., Cook, J. L., & Bird, G. (2017). The structure of social cognition: In(ter)dependence of sociocognitive processes. *Annual Review of Psychology, 68*, 243–267. https://doi.org/10.1146/annurev-psych-010416-044046

Hayward, E. O., & Homer, B. D. (2017). Reliability and validity of advanced theory-of-mind measures in middle childhood and adolescence. *British Journal of Developmental Psychology*, *35*(3), 454–462. https://doi.org/10.1111/bjdp.12186

Heyes, C. M. & Frith, C. D. (2014). The cultural evolution of mind reading. *Science, 344*(6190), 1243091. https://doi.org/10.1126/science.1243091

Hjetland, H. N., Lervåg, A., Lyster, S. A. H., Hagtvet, B. E., Hulme, C., & Melby-Lervåg, M. (2019). Pathways to reading comprehension: A longitudinal study from 4 to 9 years of age. *Journal of Educational Psychology, 111*(5), 751. https://doi.org/10.1037/edu0000321

Hodges, T. S., McTigue, E., Wright, K. L., Franks, A. D., & Matthews, S. D. (2018). Transacting with characters: Teaching children perspective taking with authentic literature. *Journal of Research in Childhood Education, 32*(3), 343–362. https://doi.org/10.1080/02568543.2018.1464529

Hofmann, S. G., Doan, S. N., Sprung, M., Wilson, A., Ebesutani, C., Andrews, L. A., Curtiss, J., & Harris, P. L. (2016). Training children's theory-of-mind: A meta-analysis of controlled studies. *Cognition*, *150*, 200–212. https://doi.org/10.1016/j.cognition.2016.01.006

Hoge, R. D., Andrews, D. A., & Leschied, A. W. (1996). An investigation of risk and protective factors in a sample of youthful offenders. Journal of Child Psychology and Psychiatry and Allied Disciplines, 37(4), 419–424. https://doi.org/10.1111/j.1469-7610.1996.tb01422.x

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

Horan, J. M., Brown, J. L., Jones, S. M., & Aber, J. L. (2016). The influence of conduct problems and callous-unemotional traits on academic development among youth. *Journal of Youth and Adolescence*, *45*(6), 1245–1260. https://doi.org/10.1007/s10964-015-0349-2

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

Hutchins, T. L., Prelock, P. A., & Bonazinga, L. (2012). Psychometric evaluation of the Theory of Mind Inventory (ToMI): A study of typically developing children and children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *42*(3), 327–341. https://doi.org/10.1007/s10803-011-1244-7

Hutchins, T.L., Prelock, P.A., (2010). The theory of mind task battery. Available at: theoryofmindinventory.com.

Imuta, K., Henry, J. D., Slaughter, V., Selcuk, B., & Ruffman, T. (2016). Theory of mind and prosocial behavior in childhood: A meta-analytic review. *Developmental Psychology, 52*(8), 1192–1205. https://doi.org/10.1037/dev0000140

Institute of Education Sciences. (2007). What Works Clearinghouse. Accessed February 16, 2021, from http://ies.ed.gov/ncee/wwc/

Joshi, R. M., & Aaron, P. G. (2000). The component model of reading: Simple view of reading made a little more complex. *Reading Psychology, 21*(2), 85–97. https://doi.org/10.1080/02702710050084428

Ju, S., Zhang, D., & Pacha, J. (2012). Employability skills valued by employers as important for entry-level employees with and without disabilities. *Career Development and Transition for Exceptional Individuals, 35*(1), 29–38. https://doi.org/10.1177/0885728811419167

Kaefer, T. (2020). When did you learn it? How background knowledge impacts attention and comprehension in read-aloud activities. *Reading Research Quarterly, 55*, S173–S183. https://doi.org/10.1002/rrq.344

Kelly, J., H., Cummings, K., D., & Cartwright, K. (in progress). A theoretical overview and empirical research synthesis on the relation between theory of mind and reading comprehension. To be submitted to *Reading and Writing*.

Kendeou, P. & O'Brien, E. J. (2018). Theories of text processing: A view from the top-down. In M. F. Schober and D. N. Rapp (Eds.), *Handbook of discourse processes* (2nd Ed., pp. 7–21). Routledge. https://doi.org/10.4324/9781315687384-2

Kendeou, P. (2015). A general inference skill. In E. J. O'Brien, A. E. Cook, & R. F. Lorch, Jr (Eds.), *Inferences during reading* (pp. 160–181). Cambridge University Press.

Kendeou, P., & Van Den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition, 35*(7), 1567–1577. https://doi.org/10.3758/BF03193491

Kendeou, P., McMaster, K. L., Butterfuss, R., Kim, J., Bresina, B., & Wagner, K. (2020). *Topics in Cognitive Science, 12*(1), 256–273. https://doi.org/10.1111/tops.12457

Kendeou, P., Papadopoulos, T. C., & Spanoudis, G. (2012). Processing demands of reading comprehension tests in young readers. *Learning and Instruction, 22*(5), 354–367. https://doi.org/10.1016/j.learninstruc.2012.02.001

Kendeou, P., Smith, E. R., & O'Brien, E. J. (2013). Updating during reading comprehension: Why causality matters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 854. https://doi.org/10.1037/a0029468

Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology, 101*(4), 765. https://doi.org/10.1037/a0015956

Kerr, M. M., & Frese, K. M. (2017). Reading to learn or learning to read? Engaging college students in course readings. *College Teaching, 65*(1), 28–31. https://doi.org/10.1080/87567555.2016.1222577

Kim, H. Y., LaRusso, M. D., Hsin, L. B., Harbaugh, A. G., Selman, R. L., & Snow, C. E. (2018). Social perspective-taking performance: Construct, measurement, and relations with academic performance and engagement. *Journal of Applied Developmental Psychology,* 57, 24–41. https://doi.org/10.1016/j.appdev.2018.05.005

Kim, Y. S. G. & Phillips, B. (2014). Cognitive correlates of listening comprehension. *Reading Research Quarterly, 49*(3), 269–281. https://doi.org/10.1002/rrq.74

Kim, Y. S. G. (2015a). Developmental, component-based model of reading fluency: An investigation of predictors of word-reading fluency, text-reading fluency, and reading comprehension. *Reading Research Quarterly, 50*(4), 459–481. https://doi.org/10.1002/rrq.107

Kim, Y. S. G. (2015b). Language and cognitive predictors of text comprehension: Evidence from multivariate analysis. *Child Development, 86*(1), 128–144. https://doi.org/10.1111/cdev.12293

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

Kim, Y. S. G. (2017). Why the simple view of reading is not simplistic: Unpacking component skills of reading using a direct and indirect effect model of reading (DIER). *Scientific Studies of Reading, 21*(4), 310–333. https://doi.org/10.1080/10888438.2017.1291643

Kim, Y. S. G. (2019). Hierarchical and dynamic relations of language and cognitive skills to reading comprehension: Testing the direct and indirect effects model of

reading (DIER). *Journal of Educational Psychology*. Advanced online publication. https://doi.org/10.1037/edu0000407

Kim, Y. S. G. (2020). Theory of mind mediates the relations of language and domain-general cognitions to discourse comprehension. *Journal of Experimental Child Psychology*, *194*, Advanced Online Publication. 104813. https://doi.org/10.1016/j.jecp.2020.104813

Kim, Y. S. G., & Petscher, Y. (2020). Influences of individual, text, and assessment factors on text/discourse comprehension in oral language (listening comprehension). *Annals of Dyslexia*, 1–20. https://doi.org/10.1007/s11881-020-00208-8

Kim, Y. S. G., Dore, R., Cho, M., Golinkoff, R., & Amendum, S. J. (2021). Theory of mind, mental state talk, and discourse comprehension: Theory of mind process is more important for narrative comprehension than for informational text comprehension. *Journal of Experimental Child Psychology*, 209, 105181. https://doi.org/10.1016/j.jecp.2021.105181

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163. https://doi.org/10.1037/0033-295X.95.2.163

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*(4), 294–303. https://doi.org/10.1037/0003-066X.49.4.294

Kintsch, W. (1995). How readers construct situation models for stories: The role of syntactic cues and causal inferences. In M. A. Gernsbacher and T. Givón (Eds.), *Coherence in spontaneous text*, (pp. 139–160.). John Benjamins. https://doi.org/10.1075/tsl.31.06kin

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363–394. https://doi.org/10.1037/0033-295X.85.5.363

Kispal, A. (2008). *Effective teaching of inference skills for reading: Literature review* (DCSF Research Report 031). London: DCSF

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice, 41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*(1), 26–38. https://doi.org/10.1177/0741932512452794

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293–323. https://doi.org/10.1016/0010-0285(74)90015-2

Landi, N., & Ryherd, K. (2017). Understanding specific reading comprehension deficit: A review. *Language and Linguistics Compass, 11*(2), e12234. https://doi.org/10.1111/lnc3.12234

Language and Reading Research Consortium & Chiu, Y. D. (2018). The simple view of reading across development: Prediction of grade 3 reading comprehension from prekindergarten skills. *Remedial and Special Education, 39*(5), 289–303. https://doi.org/10.1177/0741932518762055

Language and Reading Research Consortium, Currie, N. K., & Muijselaar, M. M. L. (2019). Inference making in young children: The concurrent and longitudinal contributions of verbal working memory and vocabulary. *Journal of Educational Psychology, 111*(8), 1416–1431. https://doi.org/10.1037/edu0000342

Language and Reading Research Consortium. (2017). Oral language and listening comprehension: Same or different constructs?. *Journal of Speech, Language, and Hearing Research, 60*(5), 1273–1284. https://doi.org/10.1044/2017_JSLHR-L-16-0039

LaRusso, M., Kim, H. Y., Selman, R., Uccelli, P., Dawson, T., Jones, S., & Snow, C. (2016). Contributions of academic language, perspective taking, and complex reasoning to deep reading comprehension. *Journal of Research on Educational Effectiveness*, *9*(2), 201–222. https://doi.org/10.1080/19345747.2015.1116035

Lecce, S., Bianco, F., & Hughes, C. (2021a). Reading minds and reading texts: Evidence for independent and specific associations. *Cognitive Development, 57*, 101010. https://doi.org/10.1016/j.cogdev.2021.101010

Lecce, S., Bianco, F., Devine, R., Hughes, C., & Banerjee, R. (2014). Promoting theory of mind in middle childhood: A training program. *Journal of Experimental Child Psychology*, 126, 52–67. https://doi.org/10.1016/j.jecp.2014.03.002

Lecce, S., Ronchi, L., & Devine, R. T. (2021b). Mind what teacher says: Teachers' propensity for mental-state language and children's theory of mind in middle childhood. *Social Development.* https://doi.org/10.1111/sode.12552

Ledford, J. R., & Gast, D. L. (Eds.). (2018). *Single case research methodology: Applications in special education and behavioral sciences*. (3rd ed.). Routledge.

Ledford, J. R., Lane, J. D., & Tate, R. (2018). Evaluating quality and rigor in single case research. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (3rd ed., pp. 365–392). Routledge.

Lee, H. (1960). *To Kill a Mockingbird*. Random House.

Lehr, S. (1988). The child's developing sense of theme as a response to literature. *Reading Research Quarterly*, 337–357. https://doi.org/10.2307/748046

Lennox, S. (2013). Interactive read-alouds-An avenue for enhancing children's language for thinking and understanding: A review of recent research. *Early Childhood Education Journal, 41*(5), 381–389. https://doi.org/10.1007/s10643-013-0578-5

Lever, R., & Sénéchal, M. (2011). Discussing stories: How a dialogic reading intervention improves kindergarteners' oral narrative construction. *Journal of Experimental Child Psychology, 108*(1), 1–24. https://doi.org/10.1016/j.jecp.2010.07.002

Lewis, S., Hacquard, V., & Lidz, J. (2017). "Think" Pragmatically: Children's Interpretation of belief reports. *Language Learning and Development*, *13*(4), 395–417. https://doi.org/10.1080/15475441.2017.1296768

Linderholm, T. (2002). Predictive inference generation as a function of working memory capacity and causal text constraints. *Discourse Processes, 34*(3), 259–280. https://doi.org/10.1207/S15326950DP3403_2

Lockl, K., Ebert, S., & Weinert, S. (2017). Predicting school achievement from early theory of mind: Differential effects on achievement tests and teacher ratings. *Learning and Individual Differences, 53,* 93–102. https://doi.org/10.1016/j.lindif.2016.11.007

Lonigan, C. J., Burgess, S. R., & Schatschneider, C. (2018). Examining the simple view of reading with elementary school children: Still simple after all these years. *Remedial and Special Education*, *39*(5), 260–273. https://doi.org/10.1177/0741932518764833

Loukusa, S., & Moilanen, I. (2009). Pragmatic inference abilities in individuals with Asperger syndrome or high-functioning autism. A review. *Research in Autism Spectrum Disorders, 3*(4), 890–904. https://doi.org/10.1016/j.rasd.2009.05.002

Lynch, J. S., & van den Broek, P. (2007). Understanding the glue of narrative structure: Children's on- and off-line inferences about characters' goals. *Cognitive Development, 22*(3), 323–340. https://doi.org/10.1016/j.cogdev.2007.02.002

Lynch, J. S., van den Broek, P., Kremer, K. E., Kendeou, P., White, M. J., & Lorch, E. P. (2008). The development of narrative comprehension and its relation to other early reading skills. *Reading Psychology*, *29*(4), 327–365. https://doi.org/10.1080/02702710802165416

Lysaker, J. T., Tonge, C., Gauson, D., & Miller, A. (2011). Reading and social imagination: What relationally oriented reading instruction can do for children. *Reading Psychology, 32*(6), 520–566.

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

Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science, 3*(3), 173–192. https://doi.org/10.1111/j.1745-6924.2008.00073.x

Mar, R. A., Tackett, J. L., & Moore, C. (2010). Exposure to media and theory-of-mind development in preschoolers. *Cognitive Development, 25*(1), 69–78. https://doi.org/10.1016/j.cogdev.2009.11.002

Martucci, K. (2016). Shared storybook reading in the preschool setting and considerations for young children's theory of mind development. *Journal of Early Childhood Research, 14*(1), 55–68. https://doi.org/10.1177/1476718X14523750

Mather, N., & Wendling, B. J. (2014). Examiner's manual. *Wookcock-Johnson IV tests of achievement.* Riverside

Maxwell, J. A. (2012). *Qualitative research design: An interactive approach*. (3rd Edition). Sage.

McArthur, G. M., Hogben, J. H., Edwards, V. T., Heath, S. M., & Mengler, E. D. (2000). On the "specifics" of specific reading disability and specific language impairment. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 41*(7), 869–874. https://doi.org/10.1111/1469-7610.00674

McCaffrey, M., & Hisrich, K. E. (2017). Read-alouds in the classroom: A pilot study of teachers' self-reporting practices. *Reading Improvement, 54*(3), 93–100.

McCarthy, K. S. & Goldman, S. R. (2015). Comprehension of short stories: Effects of task instructions on literary interpretation. *Discourse Processes, 52*(7), 585–608. https://doi.org/10.1080/0163853X.2014.967610

McCarthy, K. S. (2015). Reading beyond the lines: A critical review of cognitive approaches to literary interpretation and comprehension. *Scientific Study of Literature, 5*(1), 99–128. https://doi.org/10.1075/ssol.5.1.05mcc

McCarthy, K. S., & Goldman, S. R. (2019). Constructing interpretive inferences about literary text: The role of domain-specific knowledge. *Learning and Instruction, 60*, 245–251. https://doi.org/10.1016/j.learninstruc.2017.12.004

McCarthy, K. S., Guerrero, T. A., Kent, K. M., Allen, L. K., McNamara, D. S., Chao, S. F., ... & Sabatini, J. (2018). Comprehension in a Scenario-Based Assessment: Domain and Topic-Specific Background Knowledge. *Discourse Processes, 55*(5-6), 510–524. https://doi.org/10.1080/0163853X.2018.1460159

McIntyre, N. S., Oswald, T. M., Solari, E. J., Zajic, M. C., Lerro, L. E., Hughes, C., Devin, R. T., & Mundy, P. C. (2018). Social cognition and reading comprehension in children and adolescents with autism spectrum disorders or typical development. *Research in Autism Spectrum Disorders, 54*, 9–20. https://doi.org/10.1016/j.rasd.2018.06.004

McIntyre, N. S., Solari, E. J., Gonzales, J. E., Solomon, M., Lerro, L. E., Novotny, S., ... & Mundy, P. C. (2017). The scope and nature of reading comprehension impairments in school-aged children with higher-functioning autism spectrum disorder. *Journal of Autism and Developmental Disorders, 47*(9), 2838–2860. https://doi.org/10.1007/s10803-017-3209-y

McKeown, M. G., Beck, I. L., & Blake, R. G. K. (2009). Rethinking reading comprehension instruction: A comparison of instruction for strategies and content approaches. *Reading Research Quarterly*, 44(3), 218–253. https://doi.org/10.1598/RRQ.44.3.1

McMaster, K. L., & Espin, C. A. (2017). Reading comprehension instruction and intervention: Promoting inference making. D. Compton, R. Partial, & K. Cain (Eds.). *Theories of reading development*. John Benjamin's. https://doi.org/10.1075/swll.15.25mcm

McMaster, K.L., Kendeou, P., Bresina, B., Slater, S., Wagner, White, M.J., Butterfuss, R., Kim, J., & Umana, C. (2019). Developing an interactive software application to support young children's inference-making. *L1- Educational Studies in Language and Literature, 19*, 1–30. https://doi.org/10.17239/L1ESLL-2019.19.04.04

McNamara, D. S., & Kendeou, P. (2017). Translating advances in reading comprehension research to educational practice. *International Electronic Journal of Elementary Education, 4*(1), 33–46

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 51, pp. 297–384). Elsevier Science. https://doi.org/10.1016/S0079-7421(09)51009-2

Miller, A. C., & Keenan, J. M. (2009). How word decoding skill impacts text memory: The centrality deficit and how domain knowledge can compensate. *Annals of dyslexia, 59*(2), 99–113. https://doi.org/10.1007/s11881-009-0025-x

Miller, S. A. (2012). *Theory of mind: Beyond the preschool years*. Psychology Press.

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78*(2), 622–646. https://doi.org/10.1111/j.1467-8624.2007.01018.x

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses:* The PRISMA Statement. *PLoS Med 6*(7): e1000097. https://doi.org/10.1371/journal.pmed.1000097

Nassaji, H. (2015). Qualitative and descriptive research: Data type versus data analysis. *Language Teaching Research*, 19(2), 129–132. https://doi.org/10.1177/1362168815572747

Nathanson, A. I., Sharp, M. L., Aladé, F., Rasmussen, E. E., & Christy, K. (2013). The relation between television exposure and theory of mind among preschoolers. *Journal of Communication, 63*(6), 1088–1108. https://doi.org/10.1111/jcom.12062

Nation, K., Cocksey, J., Taylor, J. S., & Bishop, D. V. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *Journal of Child Psychology and Psychiatry, 51*(9), 1031–1039. https://doi.org/10.1111/j.1469-7610.2010.02254.x

National Governors Association Center for Best Practices (2010). *Common Core State Standards: English Language Arts*. Washington, D.C.: National Governors Association Center for Best Practices, Council of Chief State School Officers. Retrieved from http://www.corestandards.org/ELA-Literacy/

Nilsson, K. K., & de Lopez, K. J. (2016). Theory of mind in children with specific language impairment: A systematic review and meta-analysis. *Child Development, 87*(1), 143–153. https://doi.org/10.1111/cdev.12462

Noordman, L., Vonk, W., Cozijn, R., & Frank, S. (2015). Causal inferences and world knowledge. In E. J. O'Brien, A. E. Cook, & J. R. F. Lorch (Eds.), *Inferences during Reading* (pp. 260–289). Cambridge University Press.

O'Reilly, T., Wang, Z., & Sabatini, J. (2019). How much knowledge is too little? When a lack of knowledge becomes a barrier to comprehension. *Psychological Science*, *30*(9), 1344–1351. https://doi.org/10.1177/0956797619862276

Oakhill J., Cain K., Elbro C. (2019). Reading comprehension and reading comprehension difficulties. In D. Kilpatrick, R. Joshi, and R. Wagner (Eds.), *Reading development and difficulties* (pp. 83–115). Springer. https://doi.org/10.1007/978-3-030-26550-2_5

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading, 16*(2), 91–121. https://doi.org/10.1080/10888438.2010.529219

Oakhill, J., & Cain, K. (2007). Issues of causality in children's reading comprehension. In D. McNamara (Ed.), *Reading comprehension strategies. Theories, interventions, and technologies* (pp. 47–71). Erlbaum.

Oakhill, J., Cain, K., McCarthy, D. (2015). Inference processing in children: The contributions of depth and breadth of vocabulary knowledge. In E. J. O'Brien, A. E. Cook, & J. R. F. Lorch, (Eds.). *Inferences during reading* (pp. 140–159). Cambridge University Press

Osterhaus, C., Koerber, S., & Sodian, B. (2016). Scaling of advanced theory-of-mind tasks. *Child development, 87*(6), 1971–1991. https://doi.org/10.1111/cdev.12566

Paris, A. H., & Paris, S. G. (2003). Assessing narrative comprehension in young children. *Reading Research Quarterly, 38*(1), 36–76. https://doi.org/10.1598/RRQ.38.1.3

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40*(2), 184–202. https://doi.org/10.1598/RRQ.40.2.3

Paris, S. G., & Paris, A. H. (2006). The influence of developmental skill trajectories on assessments of children's early reading. In W. Damon, R. Lerner, K. A. Renninger, & I. E. Siegel (Eds.), *Handbook of Child Psychology*, Volume 4: Child Psychology in Practice (6th Ed., pp. 48–74). John Wiley & Sons. https://doi.org/10.1002/9780470147658.chpsy0402

Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior modification, 35*(4), 303–322. https://doi.org/10.1177/0145445511399147

Pavias, M., van den Broek, P., Hickendorff, M., Beker, K., & Van Leijenhorst, L. (2016). Effects of social-cognitive processing demands and structural importance on narrative recall: Differences between children, adolescents, and adults. *Discourse Processes, 53*(5–6), 488–512. https://doi.org/10.1080/0163853X.2016.1171070

Pelletier, J. (2006). Relations among theory of mind, metacognitive language, reading skills and story comprehension in L1 and L2 learners. In A. Antonietti (Ed), *Theory of mind and language in developmental contexts* (pp. 77–92). Springer US.

Pelletier, J., & Astington, J. (2004). Action, consciousness and theory of mind: Children's ability to coordinate story characters' actions and thoughts. *Early Education and Development, 15*(1), 5–22. https://doi.org/10.1207/s15566935eed1501_1

Pelletier, J., & Beatty, R. (2015). Children's understanding of Aesop's fables: relations to reading comprehension and theory of mind. *Frontiers in Psychology, 6*, 1448. https://doi.org/10.3389/fpsyg.2015.01448

Peng, P., Fuchs, D., Fuchs, L. S., Elleman, A. M., Kearns, D. M., Gilbert, J. K., ... & Patton III, S. (2019). A longitudinal analysis of the trajectories and predictors of word reading and reading comprehension development among at-risk readers. *Journal of Learning Disabilities, 52*(3), 195–208. https://doi.org/10.1177/0022219418809080

Pentimonti, J. M., & Zucker, T. A. (2015). *Kingdom of friends*. University of Texas Health Science Center Houston

Pentimonti, J. M., Zucker, T. A., & Justice, L. M. (2011). What are preschool teachers reading in their classrooms? *Reading Psychology*, *32*(3), 197–236. https://doi.org/10.1080/02702711003604484

Pentimonti, J. M., Zucker, T. A., Justice, L. M., Petscher, Y., Piasta, S. B., & Kaderavek, J. N. (2012). A standardized tool for assessing the quality of classroom-based shared reading: Systematic Assessment of Book Reading (SABR). *Early Childhood Research Quarterly, 27*(3), 512–528. https://doi.org/10.1016/j.ecresq.2011.12.007

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22–37. https://doi.org/10.1080/10888438.2013.827687

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that…" attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, *39*(3), 437–471. https://doi.org/10.1016/0022-0965(85)90051-7

Peskin, J., & Astington, J. W. (2004). The effects of adding metacognitive language to story texts. *Cognitive Development*, *19*(2), 253–273. https://doi.org/10.1016/j.cogdev.2004.01.003

Petscher, Y., Cabell, S. Q., Catts, H. W., Compton, D. L., Foorman, B. R., Hart, S. A., ... & Wagner, R. K. (2020). How the Science of Reading Informs 21st-Century Education. *Reading Research Quarterly*, 55, S267–S282. https://doi.org/10.1002/rrq.352

Petscher, Y., Justice, L. M., & Hogan, T. (2018). Modeling the early language trajectory of language development when the measures change and its relation to poor reading comprehension. *Child development*, *89*(6), 2136–2156. https://doi.org/10.1111/cdev.12880

Quesque, F., & Rossetti, Y. (2020). What do theory-of-mind tasks actually measure? Theory and practice. *Perspectives on Psychological Science, 15*(2), 384–396. https://doi.org/10.1177/1745691619896607

Rapp, D. N., Broek, P. V. D., McMaster, K. L., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading*, *11*(4), 289–312. https://doi.org/10.1080/10888430701530417

Reese, E., Suggate, S., Long, J., & Schaughency, E. (2010). Children's oral narrative and reading skills in the first 3 years of reading instruction. *Reading and Writing, 23*(6), 627–644. https://doi.org/10.1007/s11145-009-9175-9

Ricketts, J., Jones, C. R., Happé, F., & Charman, T. (2013). Reading comprehension in autism spectrum disorders: The role of oral language and social functioning. *Journal of Autism*

and Developmental Disorders, 43(4), 807–816. https://doi.org/10.1007/s10803-012-1619-4

Riggio, M. M., & Cassidy, K. W. (2009). Preschoolers' processing of false beliefs within the context of picture book reading. *Early Education and Development, 20*(6), 992–1015. https://doi.org/10.1080/10409280903375685

Roberts, C. (2012). Information structure: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics, 5*, 1–69. https://doi.org/10.3765/sp.5.6

Rodriguez-Mojica, C., & Briceño, A. (2018). Sentence stems that support reading comprehension. *The Reading Teacher, 72*(3), 398–402. https://doi.org/10.1002/trtr.1705

Ronfard, S., & Harris, P. L. (2014). When will Little Red Riding Hood become scared? Children's attribution of mental states to a story character. *Developmental Psychology, 50*(1), 283. https://doi.org/10.1037/a0032970

Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development, 73*(3), 734–751. https://doi.org/10.1111/1467-8624.00435

Rumelhart, D. (1980). Schemata: The building blocks of cognition. In: R. J. Spiro, B. C. Bruce and W. F. Brewer. (Eds.), *Theoretical issues in reading comprehension*. Erlbaum

Sáenz, L. M., & Fuchs, L. S. (2002). Examining the reading difficulty of secondary students with learning disabilities: Expository versus narrative text. *Remedial and Special Education*, *23*(1), 31–41. https://doi.org/10.1177/074193250202300105

Saldaña, J. (2015). *The coding manual for qualitative researchers*. (3rd Edition). Sage.

Santoro, L. E., Baker, S. K., Fien, H., Smith, J. L. M., & Chard, D. J. (2016). Using read-alouds to help struggling readers access and comprehend complex, informational text. *Teaching Exceptional Children, 48*(6), 282–292. https://doi.org/10.1177/0040059916650634

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2017). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity: International Journal of Methodology, 52*(4), 1893–1907. https://doi.org/10.1007/s11135-017-0574-8

Schrank, F. A., McGrew, K. S., & Mather, N. (2014). *Woodcock–Johnson IV*. Riverside

Semingson, P., & Kerns, W. (2021). Where is the evidence? Looking back to Jeanne Chall and enduring debates about the science of reading. *Reading Research Quarterly*, 56, S157–S169. https://doi.org/10.1002/rrq.405

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

Sesso, G., Brancati, G. E., Fantozzi, P., Inguaggiato, E., Milone, A., & Masi, G. (2021). Measures of empathy in children and adolescents: A systematic review of questionnaires. *World Journal of Psychiatry, 11*(10), 876. https://doi.org/10.5498/wjp.v11.i10.876

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

Shanahan, T., & Shanahan, S. (1997). Character Perspective Charting: Helping children to develop a more complete conception of story. *The Reading Teacher, 50*(8), 668–677.

Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C., & Torgesen, J. (2010). Improving reading comprehension in kindergarten through 3rd grade: A practice guide (NCEE 2010-4038). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from whatworks.ed.gov/publications/practiceguides.

Silverman, R. D., Speece, D. L., Harring, J. R., & Ritchey, K. D. (2013). Fluency has a role in the simple view of reading. *Scientific Studies of Reading, 17*(2), 108–133. https://doi.org/10.1080/10888438.2011.618153

Siregar, N. R. (2021). Explicit instruction and executive functioning capacity: a new direction in cognitive load theory. *Journal of Education*. Advance Online Publication. https://doi.org/10.1177/00220574211033256

Slaughter, V., Peterson, C. C., & Mackintosh, E. (2007). Mind what mother says: Narrative input and theory of mind in typical children and those on the autism spectrum. *Child Development, 78*(3), 839–858. https://doi.org/10.1111/j.1467-8624.2007.01036.x

Smith, R., Snow, P., Serry, T., & Hammond, L. (2021). The role of background knowledge in reading comprehension: A critical review. *Reading Psychology, 42*(3), 214–240. https://doi.org/10.1080/02702711.2021.1888348

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Rand Corporation.

Snow, C. E. (2018). Simple and not-so-simple views of reading. *Remedial and Special Education, 39*(5), 313–316. https://doi.org/10.1177/0741932518770288

Spencer, M., Gilmour, A. F., Miller, A. C., Emerson, A. M., Saha, N. M., & Cutting, L. E. (2019). Understanding the influence of text complexity and question type on reading outcomes. *Reading and Writing, 32*(3), 603–637. https://doi.org/10.1007/s11145-018-9883-0

Spencer, M., Quinn, J. M., & Wagner, R. K. (2014). Specific reading comprehension disability: Major problem, myth, or misnomer?. *Learning Disabilities Research & Practice, 29*(1), 3–9. https://doi.org/10.1111/ldrp.12024

Stanley, C. T., Petscher, Y., & Catts, H. (2018). A longitudinal investigation of direct and indirect links between reading skills in kindergarten and reading comprehension in tenth grade. *Reading and Writing, 31*(1), 133–153. https://doi.org/10.1007/s11145-017-9777-6

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*, 32–71. https://doi.org/10.2307/747348

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–367. https://doi.org/10.1598/RRQ.21.4.1

Stein, N., & Glenn, C. (1979). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *New directions in discourse processing* (pp. 53–120). Ablex.

Strasser, K., & Río, F. D. (2014). The role of comprehension monitoring, theory of mind, and vocabulary depth in predicting story comprehension and recall of kindergarten children. *Reading Research Quarterly*, *49*(2), 169–187. https://doi.org/10.1002/rrq.68

Sutin, A. R., & Robins, R. W. (2008). When the "I" looks at the "Me": Autobiographical memory, visual perspective, and the self. *Consciousness and Cognition*, *17*(4), 1386–1397. https://doi.org/10.1016/j.concog.2008.09.001

Swanson, E., Vaughn, S., Wanzek, J., Petscher, Y., Heckert, J., Cavanaugh, C., ... & Tackett, K. (2011). A synthesis of read-aloud interventions on early reading outcomes among preschool through third graders at risk for reading difficulties. *Journal of Learning Disabilities*, *44*(3), 258–275. https://doi.org/10.1177/0022219410378444

Tilstra, J., McMaster, K., van den Broek, P., Kendeou, P., & Rapp, D. (2009). Simple but complex: Components of the simple view of reading across grade levels. *Journal of Research in Reading*, *32*(4), 383–401. https://doi.org/10.1111/j.1467-9817.2009.01401.x

Tomasello, M., & Rakoczy, H. (2003). What makes human cognition unique? From individual to shared to collective intentionality. *Mind & Language*, *18*(2), 121–147. https://doi.org/10.1111/1468-0017.00217

Tompkins, V. (2015). Improving low-income preschoolers' theory of mind: A training study. *Cognitive Development*, *36*, 1–19. https://doi.org/10.1016/j.cogdev.2015.07.001

Tompkins, V., Benigno, J. P., Kiger Lee, B., & Wright, B. M. (2018). The relation between parents' mental state talk and children's social understanding: A meta-analysis. *Social Development*, *27*(2), 223–246. https://doi.org/10.1111/sode.12280

Tompkins, V., Farrar, M., & Montgomery, D. E. (2019). Speaking your mind: Language and narrative in young children's Theory of Mind development. *Advances in Child Development and Behavior, 56,* 109–140. https://doi.org/10.1016/bs.acdb.2018.11.003

Tompkins, V., Guo, Y., & Justice, L. M. (2013). Inference generation, story comprehension, and language skills in the preschool years. *Reading and Writing, 26*(3), 403–429. https://doi.org/10.1007/s11145-012-9374-7

Tong, S. X., Wong, R. W. Y., Kwan, J. L. Y., & Arciuli, J. (2019). Theory of mind as a mediator of reading comprehension differences between Chinese school-age children with autism and typically developing peers. *Scientific Studies of Reading*. Advanced online publication. https://doi.org/10.1080/10888438.2019.1666133

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, *24*(5), 612–630. https://doi.org/10.1016/0749-596X(85)90049-X

Trabasso, T., & Wiley, J. (2005). Goal plans of action and inferences during comprehension of narratives. *Discourse processes, 39*(2-3), 129–164. https://doi.org/10.1207/s15326950dp3902&3_3

Trabasso, T., van den Broek, P. W., & Suh, S. Y. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes*, *12*(1), 1–25. https://doi.org/10.1080/01638538909544717

Tsai, J. L. (2007). Ideal Affect: Cultural Causes and Behavioral Consequences. *Perspectives on Psychological Science*, *2*(3), 242–259. https://doi.org/10.1111/j.1745-6916.2007.00043.x

Tunmer, W. E., & Chapman, J. W. (2012). The simple view of reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities, 45*(5), 453–466. https://doi.org/10.1177/0022219411432685

United States. Department of Education, National Assessment for Educational Progress. (2019b). *The nations report card*. Author. Retrieved from https://www.nationsreportcard.gov/

United States. Department of Education, National Assessment for Educational Progress. (2019a). 2015 *Student questionnaire results*. Author. Retrieved from https://www.nationsreportcard.gov/sq_classroom/ - /reading

van den Broek, P. (1990). The causal inference maker: Towards a process model of inference generation in text comprehension. In D. A. Balota, F. D'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 423–446). Lawrence Erlbaum.

van den Broek, P. (1994). Comprehension and memory of narrative texts: Inferences and coherence. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 539–588). Academic Press.

van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes, 54*(5–6), 360–372. https://doi.org/10.1080/0163853X.2017.1306677

van den Broek, P., Beker, K., & Oudega, M. (2015). Inference generation in text comprehension: Automatic and strategic processes in the construction of a mental representation. In E. J. O'Brien, A. E. Cook, & J. R. F. Lorch, (Eds.). *Inferences during reading* (pp. 94–122). Cambridge University Press.

van den Broek, P., Bohn-Gettler, C. M., Kendeou, P., Carlson, S., & White, M. J. (2011). When a reader meets a text: The role of standards of coherence in reading comprehension. In M. McCrudden, J. Magliano, and G. Schraw (Eds.), *Text relevance and learning from text* (pp. 123–140). Information Age Publishing.

van den Broek, P., Rapp, D.N., & Kendeou, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes, 39*(2–3), 299–316. https://doi.org/10.1080/0163853X.2005.9651685

van den Broek, P., Risden, K., & Husebye-Hartmann, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In R. F. Lorch & E. J. O'Brien (Eds.), *Sources of coherence in text comprehension* (pp. 353–373). Erlbaum

van den Broek, P., Helder, A., & Van Leijenhorst, L. (2013). Sensitivity to structural centrality: Developmental and individual differences in reading comprehension skills. In M. A. Britt, S. R. Goldman, & J.-F. Rouet (Eds.), *Reading: From words to multiple texts* (pp. 132–146). Routledge

van den Broek, P., Lorch, R., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition, 29*(8), 1081–1087. https://doi.org/10.3758/BF03206376

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* Academic Press

van Moort, M. L., Koornneef, A., & van den Broek, P. W. (2021). Differentiating text-based and knowledge-based validation processes during reading: Evidence from eye movements. *Discourse Processes, 58*(1), 22–41. https://doi.org/10.1080/0163853X.2020.1727683

Vaughn, M. G., DeLisi, M., Beaver, K. M., Wexler, J., Barth, A., & Fletcher, J. (2011). Juvenile psychopathic personality traits are associated with poor reading achievement. *Psychiatric Quarterly, 82*(3), 177–190. https://doi.org/10.1007/s11126-010-9162-y

Vogindroukas, I., Chelas, E. N., & Petridis, N. E. (2014). Reading the Mind in the Eyes Test (children's version): a comparison study between children with typical development, children with high-functioning autism and typically developed adults. *Folia Phoniatrica et Logopaedica*, *66*(1-2), 18–24. https://doi.org/10.1159/000363697

Warnell, K. R., & Redcay, E. (2019). Minimal coherence among varied theory of mind measures in childhood and adulthood. *Cognition, 191*, 103997. https://doi.org/10.1016/j.cognition.2019.06.009

Weimer, A. A., Warnell, K. R., Ettekal, I., Cartwright, K. B., Guajardo, N. R., & Liew, J. (2021). Correlates and antecedents of theory of mind development during middle childhood and adolescence: An integrated model. *Developmental Review*, 59, 100945. https://doi.org/10.1016/j.dr.2020.100945

Wellman, H. M. (2002). Understanding the psychological world: Developing a theory of mind. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp.167–187). Blackwell.

Wellman, H. M. (2017). The development of theory of mind: Historical reflections. *Child Development Perspectives*, *11*(3), 207–214. https://doi.org/10.1111/cdep.12236

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*(2), 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655–684. https://doi.org/10.1111/1467-8624.00304

Wellman, H. M., Fang, F., & Peterson, C. C. (2011). Sequential progressions in a theory-of-mind scale: Longitudinal perspectives. *Child Development*, *82*(3), 780–792. https://doi.org/10.1111/j.1467-8624.2011.01583.x

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

Westra, E., & Carruthers, P. (2017). Pragmatic development explains the theory-of-mind scale. *Cognition*, *15*, 165–176. https://doi.org/10.1016/j.cognition.2016.10.021

What Works Clearinghouse. (2020). *What Works Clearinghouse standards handbook* (Version 4.1). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf

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

Wigfield, A., Guthrie, J. T., Perencevich, K. C., Taboada, A., Klauda, S. L., McRae, A., & Barbosa, P. (2008). Role of reading engagement in mediating the effects of reading comprehension instruction on reading outcomes. *Psychology in the Schools*, 45, 432–445. https://doi.org/10.1002/pits.20307.

Wilder, A. A., & Williams, J. P. (2001). Students with severe learning disabilities can learn higher order comprehension skills. *Journal of Educational Psychology*, *93*(2), 268–278. https://doi.org/10.1037/0022-0663.93.2.268

Williams, J. P. (2005). Instruction in reading comprehension for primary-grade students: A focus on text structure. *The Journal of Special Education*, *39*(1), 6–18. https://doi.org/10.1177/00224669050390010201

Williams, J. P., Brown, L. G., Silverstein, A. K., & de Cani, J. S. (1994). An instructional program in comprehension of narrative themes for adolescents with learning disabilities. *Learning Disability Quarterly*, *17*(3), 205–221. https://doi.org/10.2307/1511074

Williams, J. P., Lauer, K. D., Hall, K. M., Lord, K. M., Gugga, S. S., Bak, S. J., Jacobs, P. R., & deCani, J. S. (2002). Teaching elementary school students to identify story themes. *Journal of Educational Psychology*, *94*(2), 235–248. https://doi.org/10.1037/0022-0663.94.2.235

Williamson, G. L. (2008). A text readability continuum for postsecondary readiness. *Journal of Advanced Academics*, *19*(4), 602–632. https://doi.org/10.4219/jaa-2008-832

Williamson, G.L., Koons, H., Sandvik, T., & Sanford-Moore, E. (2012). The text complexity continuum in grades 1–12. (Metametrics Research Brief), MetaMetrics Inc.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128. https://doi.org/10.1016/0010-0277(83)90004-5

Wiseman, A. (2011). Interactive read alouds: Teachers and students constructing knowledge and literacy together. *Early Childhood Education Journal, 38*(6), 431–438. https://doi.org/10.1007/s10643-010-0426-9

Wolf, M. C., Muijselaar, M. M., Boonstra, A. M., & de Bree, E. H. (2019). The relationship between reading and listening comprehension: shared and modality-specific components. *Reading and Writing, 32*(7), 1747–1767. https://doi.org/10.1007/s11145-018-9924-8

Yeh, Z. T., Tsai, M. C., Tsai, M. D., Lo, C. Y., & Wang, K. C. (2017). The relationship between theory of mind and the executive functions: Evidence from patients with frontal lobe damage. *Applied Neuropsychology: Adult, 24*, 342–349. https://doi.org/10.1080/23279095.2016.1185425

Zinsser, K. M., Shewark, E. A., Denham, S. A., & Curby, T. W. (2014). A mixed-method examination of preschool teacher beliefs about social-emotional learning and relations to observed emotional support. *Infant and Child Development, 23*(5), 471–493. https://doi.org/10.1002/icd.1843

Ziv, M., Smadja, M. L., & Aram, D. (2013). Mothers' mental-state discourse with preschoolers during storybook reading and wordless storybook telling. *Early Childhood Research Quarterly, 28*(1), 177–186. https://doi.org/10.1016/j.ecresq.2012.05.005

Ziv, M., Smadja, M. L., & Aram, D. (2014). Mothers' and teachers' mental-state discourse with preschoolers during storybook reading. *Journal of Cognitive Education and Psychology, 13*(1), 103–119. https://doi.org/10.1891/1945-8959.13.1.103

Ziv, M., Smadja, M. L., & Aram, D. (2015). Preschool teachers' reference to theory of mind topics in three storybook contexts: Reading, reconstruction and telling. *Teaching and Teacher Education, 45*, 14–24. https://doi.org/10.1016/j.tate.2014.08.009

Zucker, T. A., Cabell, S. Q., Justice, L. M., Pentimonti, J. M., & Kaderavek, J. N. (2013). The role of frequent, interactive prekindergarten shared reading in the longitudinal development of language and literacy skills. *Developmental Psychology, 49*(8), 1425. https://doi.org/10.1037/a0030347

Zunshine, L. (2019). What Mary Poppins knew: Theory of Mind, children's literature, history. *Narrative*, *27*(1), 1–29. https://doi.org/10.1353/nar.2019.0000

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*(2), 162. https://doi.org/10.1037/0033-2909.123.2.162