FACILITATING SMALL-GROUP DISCUSSIONS:
EFFECTS OF TEACHER DISCOURSE MOVES ACROSS FOUR GROUP COMPOSITIONS IN FOURTH AND FIFTH GRADE

A Thesis in
Educational Psychology
by
Rachel Miriam Vriend Croninger

© 2019 Rachel M. V. Croninger

Submitted in Partial Fulfillment
of the Requirements
for the Degree of
Master of Science

May 2019
The thesis of Rachel Miriam Vriend Croninger was reviewed and approved* by the following:

P. Karen Murphy  
Distinguished Professor in Education (Educational Psychology)  
Harry and Marion Eberly Faculty Fellow  
Thesis Advisor

Peggy N. Van Meter  
Associate Professor of Education (Educational Psychology)  
Director of Undergraduate and Graduate Studies

Pui-Wa Lei  
Professor of Education (Educational Psychology)  
Professor in Charge, Educational Psychology

*Signatures are on file in the Graduate School
ABSTRACT

Engaging in small-group, text-based discussions can promote students’ high-level comprehension and ability to think critically about, around, and with text and content. During productive discussions students participate in extended episodes of talk as they exchange, critique, and evaluate each other’s ideas and reasoning. Although numerous studies on text-based discussion approaches have been conducted, the role of the teacher in facilitating the type of talk proven to promote high-level comprehension of text and content has not been investigated fully. As a facilitator, teachers utilize targeted actions (i.e., teacher moves) to guild students’ talk, but it is unclear how specific pedagogical decisions or contextual factors such as group composition influence the relationship between teachers’ facilitation practices and students’ productive talk.

To extend our understanding of teachers’ role as a facilitating during small-group, text-based discussions, in the current study I examined the turn-by-turn effect of teachers’ targeted facilitation practices on students’ productive talk across four small group composition formats (i.e., homogenous above-average, homogeneous average, homogeneous below-average, or heterogenous) in the context of an empirically-supported, teacher-facilitated discussion approach (i.e., Quality Talk). Fourth- and fifth-grade teachers (n = 4) and their students (n = 62) participated in a year-long implementation of Quality Talk for which students were randomly assigned to either a homogeneous or heterogenous ability discussion group based on a measure of oral reading fluency at the beginning of the school year. Six researchers coded 108 discussion videos for instances of teachers’ facilitation practices (i.e., teacher moves) and students productive talk (i.e., elaborated explanations).

Lag-sequential analysis revealed that group composition influences the relationship between teachers’ facilitation practices and students’ productive talk. Specifically, group composition influenced the strength of the relationship between individual teacher moves and
students’ elaborated explanations as well as how teachers responded to students elaborated explanations during discussions. Results indicated that prompting gave way to elaborated explanations, regardless of group composition, but that prompting was essential for the homogenous below-average group. Furthermore, teachers were also more likely to mark students’ discourse or challenge students’ elaborated explanations in the homogeneous below-average groups than other group composition formats. Finally, results also indicated that students responded to each other differently across group compositions.

The current study is significant in that it extends our understanding of the teacher’s role as a facilitator during small-group, text-based discussions by demonstrating that the turn-by-turn effect of teachers’ targeted facilitation moves during discussion can vary as a function of group composition. Moreover, the present study provides evidence that group composition may influence the extent to which students emulate each other’s discursive practices. The present study also has practical implications for teacher professional development aimed at supporting teachers’ implementing small-group discussions approaches in their classrooms.
### TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................... vii

LIST OF TABLES .......................................................................................................... viii

ACKNOWLEDGEMENTS .............................................................................................. ix

Chapter 1 INTRODUCTION ....................................................................................... 1

Chapter 2 REVIEW OF LITERATURE ......................................................................... 5
  Text-Based Classroom Discussions ................................................................. 5
  Theoretical and empirical underpinnings of discussion-based pedagogy .......... 5
  Text-based discussion approaches ................................................................. 10
  Discourse elements ......................................................................................... 13
  Small-Group Composition ............................................................................ 16
  Teacher Facilitation Practices ..................................................................... 21

Chapter 3 METHOD .................................................................................................... 26
  Participants ........................................................................................................ 26
  Grouping ......................................................................................................... 27
  Intervention .................................................................................................... 28
    Professional development ........................................................................... 29
    Quality Talk mini-lessons ........................................................................ 30
    Quality Talk literacy journals .................................................................... 31
    Quality Talk discussions .......................................................................... 31
  Measures ........................................................................................................ 32
    Oral reading fluency ................................................................................... 32
    Discourse coding ....................................................................................... 33
  Procedures and analysis ............................................................................... 34

Chapter 4 RESULTS .................................................................................................. 37
  Fourth-Grade Results ..................................................................................... 38
    Descriptive statistics .................................................................................. 38
    Relationship between teacher moves and group composition ................. 42
    Influence of teachers’ scaffolding moves on students’ productive talk .... 44
    Influence of teachers’ scaffolding moves across group composition ....... 58
  Fifth-Grade Results ......................................................................................... 60
    Descriptive statistics .................................................................................. 60
    Relationship between teacher moves and group composition ................. 64
    Influence of teachers’ scaffolding moves on students’ productive talk ....... 65

Chapter 5 DISCUSSION ............................................................................................... 70

References ............................................................................................................... 75
LIST OF FIGURES

Figure 3-1: Sample of a coded sequence of talk.................................................................35

Figure 4-1: Discourse example from a fourth-grade heterogeneous ability group of prompting eliciting an elaborated explanation.................................................................50

Figure 4-2: Discourse example from a fourth-grade heterogeneous ability group of an elaborated explanation eliciting an elaborated explanation.................................................................50

Figure 4-3: Discourse example from the fourth-grade homogenous average ability group of prompting eliciting an elaborated explanation.................................................................51

Figure 4-4: Discourse example from the fourth-grade homogenous average ability group of an elaborated explanation eliciting marking.................................................................52

Figure 4-5: Discourse example from the fourth-grade homogenous average ability group of an elaborated explanation eliciting an elaborated explanation.................................................................53

Figure 4-6: Discourse example from the fourth-grade homogenous below average ability group of prompting eliciting an elaborated explanation.................................................................54

Figure 4-7: Discourse example from the fourth-grade homogenous below average ability group of an elaborated explanation eliciting marking.................................................................55

Figure 4-8: Discourse example from the fourth-grade homogenous below average ability group of an elaborated explanation eliciting challenging.................................................................57

Figure 4-9: Discourse example from the fifth-grade homogenous below-average ability group of prompting eliciting an elaborated explanation.................................................................68

Figure 4-10: Discourse example from the fifth-grade homogenous below-average ability group of an elaborated explanation eliciting marking.................................................................68
LIST OF TABLES

Table 3-1:  List of Quality Talk Mini-lessons. .................................................................30

Table 4-1:  Frequency of type of turns by group composition in fourth grade. ...............41

Table 4-2:  Frequency of scaffolding moves by group composition in fourth grade. ........41

Table 4-3:  Crosstabulation of teacher move category and group composition in fourth
grade. ...............................................................................................................................43

Table 4-4:  Crosstabulation of teacher move category and ability level in fourth grade. ....44

Table 4-5:  Likelihood ratio tests for nonstationarity, independence, and heterogeneity in
dquarters grade...................................................................................................................47

Table 4-6:  Yule’s Q associated with sequential patterns across group composition in
fourth grade.........................................................................................................................59

Table 4-7:  Frequency of type of turns by group composition in fifth grade. ...................63

Table 4-8:  Frequency of scaffolding moves by group composition in fifth grade. ..........63

Table 4-9:  Crosstabulation of teacher move category and group composition in fifth
grade. .................................................................................................................................64

Table 4-10: Crosstabulation of teacher move category and ability level in fifth grade. .......65

Table 4-11:  Likelihood ratio tests for nonstationarity, independence, and heterogeneity in
fifth grade .........................................................................................................................66
ACKNOWLEDGEMENTS

This thesis was a lesson in resilience. I was never alone, however, throughout the entire process. During every moment of joy, excitement, surprise, desperation, anguish, and panic, I enjoyed the intellectual and emotional support of my mentors, peers, friends, and family without whom this endeavor would not have been possible.

First and foremost, I would like to show my deepest gratitude to P. Karen Murphy for her unwavering support and guidance over the course of this research. You have proven to be the greatest coach I have ever had, continually offering invaluable mentorship, dedication, and kindness. I will forever appreciate your honesty and confidence in me not only as a developing scholar, but also as an individual. I could not be more excited to see what we can accomplish together moving forward. I am also immensely grateful to your family, John, Lauren, and Paige Alexander, for their encouragement, patience, and understanding throughout this process. I must also thank Peggy Van Meter for helping me conceptualize my research as well as her helpful advice on how I can continue to build this line of inquiry.

I would also like to extend my sincere thanks to my fellow Quality Talk team members. I am truly thankful to Carla Firetto and Ana Butler for their leadership and example as I began my graduate studies. I am deeply indebted to Sara Baszczewski, Chelsea Cameron, Heather Parks Shedlock, Megan Huffstickler, and Emily Chukusky for the time they dedicated to coding these data. I would also like to say a special thank you to Mengyi Li, Liwei Wei, and Christa Oeder who were with me through every struggle and whose insights, ideas, and friendship will forever be a contributor to my successes and achievements. I am honored to be a part of this team, thank you for your contributions and for making 108 Rackley more than a place where we show up to work.
To my father, Robert G. Croninger, my mother, Ruth Vriend, and my brother, Nicholas Croninger, your love and support cannot be overstated; it is because of your efforts and encouragement in all aspects of my life that I was able to push through. From page proofs, to analysis, to meal prep, you have always been patient, present, and practical and I could not be more grateful. Finally, I would like to express my deepest appreciation to my grandmother, Margret Vriend, who we lost shortly before I began this project and whose steadfast belief in me sparked my interest in educational research.
Chapter 1

INTRODUCTION

In recent years the educational community has reconceptualized the role of the teacher. In addition to teaching students an abundance of facts, teachers are now expected to instill students with the necessary skills to judge the quality of the information they encounter in text. Indeed, these skills broaden students’ understanding of the ideas and concepts they encounter while reading which can lead to greater personal opportunities and societal growth (Resnick, Asterhan, & Clarke, 2015). Reading comprehension instruction geared toward generating a basic understanding of text is insufficient. Students must also be able recognize patterns, discern an authors’ motivations for writing, and interpret, consider, and weigh various points of view within and across texts (National Governors Association Center for Best Practices [NGA Center], 2010). Furthermore, students must also be equipped to utilize these skills across domains as well as in and out of the classroom. In essence, teachers are now expected to train students to not only learn about a text, but also think around and with it. I refer to this type of comprehension as high-level comprehension.

In accordance with the shifting landscape of expectations for teachers and students, there has been a steady increase in students’ reading and comprehension scores on national measures of student achievement over the last 30 years (National Center for Educational Statistics, 2018). The majority of students, however, are still demonstrating basic reading and comprehension skills. In 2017 68% of fourth-grade students preformed at or above Basic in reading on the National Assessment of Education Progress (NAEP) demonstrating the ability to locate a main idea and supporting details and make simple inferences. Only nine percent of fourth graders performed at the Advanced level in 2017. Students reading and comprehending at the Advanced
level demonstrate the ability to make complex inferences and judgments about characters’ motivations and perspectives and also use reasons and evidence to justify their thinking or critique of a text. With the goal of pushing students beyond basic reading and comprehension skills, many researchers have focused their attention on how engaging in text-based discussions can foster students’ high-level comprehension of text (Wilkinson & Son, 2010).

The data from the studies on classroom discussions provides evidence that discussion-based pedagogy can foster reading comprehension gains (Kamil, Dorman, Dole, Salinger, & Torgesen, 2008; Murphy, Wilkinson, Soter, Hennessy, & Alexander, 2009; Nystrand, 1997; Shanahan et al., 2010; Soter et al., 2008), help students clarify misunderstandings and learn to engage complex problem solving (Chi, 2000; Cooper, 1999; Rosecoco & Chi, 2007), retain knowledge (Adey & Shayer, 2015), and bolster students’ achievement on standardized tests (Fall, Webb, & Chudowsky, 2000). Indeed, the National Center for Education Statistics (2012) found that students who engaged in more frequent discussions outperformed their peers on the NAEP reading assessment. A set of basic assumptions about the type of discourse that promotes high-level comprehension has emerged from the growing body of literature on text-based discussion (Resnick et al., 2015). Students think out loud during productive discussion by asking questions, expressing their ideas, and engaging in argumentation as they develop and redevelop their understanding (Soter et al., 2008). They also make claims and counterclaims which are challenged and built on by their peers with reasons and evidence (Mercer, 2002). Finally, it is assumed that discussions leading to high-level comprehension do not spontaneously occur, but are facilitated and guided by a teacher (Wilkinson, Murphy, & Binici, 2015).

Researchers have identified a handful of popular teacher moves that teachers use to facilitate text-based discussions. Teacher moves are specific, targeted actions (e.g., prompting, modeling, clarifying) that teachers employ to promote productive talk (Dwyer, Kelcey, Berebitsky, & Carlisle, 2016; Michaels, O’Connor, & Resnick, 2008; Wei, Murphy, & Firetto,
2018). For example, a teacher may prompt a student to justify their thinking by asking, “Why do you think that?” after a student makes a claim with no supporting reasons or evidence. Though a large number of studies have been conducted about text-based discussions more generally, only a limited number of studies have been conducted on the influence of teacher moves during text-based discussions. For instance, Jadallah et al. (2011) illustrated that some teacher moves had delayed or immediate effects on students’ use of evidence during student-led, text-based discussions. There is also evidence that teacher moves impact the type of thinking students engage in during discussion (Lin et al., 2015). The influence of the teacher moves has not, however, been investigated across multiple contexts.

We do not know if the influence of individual teacher moves on student talk is mediated by various grouping factors such as ability grouping – the practice of assigning students to small groups based on some measure of proficiency. While controversial, data from the National Assessment of Educational Progress show that ability grouping has regained popularity in recent years with 71% of fourth graders in 2009 experiencing ability grouping for reading instruction compared to 28% in 1998 (Loveless, 2013). Thus, the purpose of this study is to investigate the relationship between teachers’ facilitation practices and group composition through an examination of teachers’ use of teacher moves and the influence of those moves during small-group, text-based discussions between homogenously and heterogeneously grouped students. In the following chapter, I review the literature on small-group, text-based discussions, group composition, and teacher moves in more detail.

This study address three primary research questions:

RQ1: What is the relationship between teacher’s facilitation practices and group composition during text-based discussions?
RQ2: To what extent do teachers’ facilitation practices influence students’ productive talk during text-based discussions?

RQ3: To what extent does the relationship between teachers’ facilitation practices and students’ productive talk vary as a function of group composition?
Chapter 2

REVIEW OF LITERATURE

In this chapter, I overview the extant literature pertaining to small-group, text-based discussions, within class group composition, and teacher facilitation practices that serves to undergird the central aim of this study. As such, within each section I overview these areas as to their influence on student talk during discussion as well as student reading comprehension. After briefly overviewing the theoretical (i.e., social constructivist, sociocognitive, cognitive, and dialogic) and empirical foundations of discussion-based pedagogy, I introduce the literature on individual discussion approaches as well as specific discursive behaviors shown to promote students’ high-level comprehension. I begin the subsequent section by summarizing the theoretical arguments for and against within class ability grouping before moving into a review of empirical findings regarding its effects on student talk, achievement, and reading comprehension. Finally, I overview the literature on the effect of teacher facilitation practices during small-group, text-based discussions. Taken together, these three sections describe the state of the literature as well as areas of the literature that need further investigation.

Text-Based Classroom Discussions

Theoretical and empirical underpinnings of discussion-based pedagogy

The use of discussion-based pedagogy to improve students’ reading comprehension is grounded in social constructivist, sociocognitive, cognitive, and dialogic perspectives on teaching and learning. Social constructivist Lev Vygotsky (1978) proposed that complex mental functions
(e.g., high-level comprehension and critical-analytic thinking) develop gradually during social interactions. During discussion students’ thinking is externalized as talk, allowing them to think together or to interthink. Thus, students are able to build a deeper understanding of the text or content than they would be able to without each other’s support. As students discuss, they begin to internalize the psychological processes necessary to engage in similar thinking individually. Vygotsky (1978) also stressed the importance of interacting with more knowledgeable others, such as the teacher or a more advanced peer, who can scaffold the discussion in such a way that students perform just beyond their current ability.

Alternatively, Bandura (1977) proposed a sociocognitive theory which states that learning occurs within the context of a reciprocally deterministic relationship between learners, their behavior, and their environment. In other words, a learner’s cognitive processes affect her behavior, which in turn influences her environment, and the environment ultimately influences her cognitive processes and vice versa. Bandura emphasized that students learn by way of observing others’ behavior as well as how those behaviors are subsequently rewarded or punished. Consider, for example, a discussion in which a teacher praises a student’s use of evidence. From this perspective, the student internalizes that providing evidence is good discussion behavior and her peers begin to mimic that behavior so they may too be rewarded. As such, students’ environments influence their cognitive processes and over time, they develop their comprehension and critical-analytic thinking abilities.

Cognitive theories of learning focus on how participating in discussions can promote active engagement in meaning making from text. Specifically, these theories are interested in how discussions effect student knowledge acquisition in terms of the restructuring, maintenance, and use of said knowledge. Kintsch’s Construction-Integration Model (1988) proposes that learners construct mental representations of content as they read and discuss text. These representations vary in complexity, and therefore, result in different levels of comprehension (i.e.,
surface model, textbase model, and situation model). The surface model represents the lowest level of complexity, consisting of simple decoding of words or phrases which forms a literal representation of the text in working memory. The next level of complexity, the textbase, consists of a semantic representation of the text. At this level, connections are made between propositions and ideas within the text, but not with ideas beyond the text. Students with a textbase understanding would likely be able to answer low-level or basic comprehension questions, however, they would be unable to think critically or analytically because they have not integrated their mental representation of the text with their larger knowledge network. The situation model, the most complex representation, is constructed by doing just that – integrating the content of the text with one’s prior knowledge.

Finally, the dialogic perspective distinguishes between monologic and dialogic discourse environments (Bakhtin, 1984). A monologic environment is one in which discussion is highly controlled and scripted, typically by a teacher who asks low-level, recall questions; a dialogic environment is one in which students build on what each other has said and work together to construct a deeper understanding of a text. Bakhtin (1984) argued that monologic environments only promote low-level learning because students are unable to truly participate in the discussion due to the power differential between teacher and student. Moreover, he proposed that students must hold equal rights as the teacher to fully participate in a discussion. Dialogic environments, on the other hand, afford students those rights and thus, can foster meaningful learning through the give and take of true, open discourse.

There is considerable consensus amongst scholars working from these various theoretical foundations that classroom discussions can promote reading comprehension (e.g., Abrami et al., 2015; Almasi, 1995; Brown & Palincsar, 1989; Johnson, Johnson, Stanne, & Garibaldi, 1990; Nystrand, 2006; Nystrand & Gamoran, 1991; Nystrand, Wu, Gamoran, Zeiser, & Long, 2003). Working from a dialogic perspective, Nystrand (1997) and colleagues observed the instructional
discourse of eighth- and ninth-grade teachers in 112 classrooms. Their analysis included 451 observations over a 2-year period in both English and social studies classes. Their results revealed that productive discussions do more than prompt students to recite facts or recall what others think or have said. Instead, classroom discourse is only productive to the extent that it required students to think for themselves. The extent to which discussions required students to think was evidence by the frequency of authentic questions, uptake, and the degree to which teachers incorporated students’ ideas into the conversation. Particularly striking was the degree to which even short discussions were associated with increases in student reading comprehension.

Using the same theoretical foundation, Applebee, Langer, Nystrand, & Gamoran (2003) largely replicated the findings in Nystrand (1997). Applebee and colleagues observed 974 students from 64 middle- and high-school English classrooms located in five states over the course of a year. The findings reported indicate that discussion is an effective means of promoting literature achievement and reading comprehension over a range of academic contexts (i.e., grade levels, urban vs suburban schools, ability levels). Langer’s (2001) multi-group study reached a similar conclusion. Incorporating both dialogic and constructivist views of learning and teaching, Langer attempted to identify instructional practices associated with student achievement in reading, writing, and English by comparing the instructional practices of high performing schools with those of average performing schools. The results indicated that across multiple settings students in high performing schools participate in engaging, whole-class as well as small-group discussions.

In an effort to begin to understand how environmental factors (i.e., who holds primary control of the discussion) influence classroom discussions, Almasi (1995) conducted a quasi-experimental study involving 97 fourth-grade students and their six teachers. Teachers and students participated in either peer- or teacher-led discussion for 11 weeks. After which, they completed a cognitive conflict scenario transfer task and took part in semi-structured interviews.
Cognitive conflict refers to what readers experience when incoming information conflicts with their current understanding of the text or prior knowledge. Successful readers recognize and reconcile such conflicts. Cognitive conflict is a tenant of sociocognitive and cognitive perspectives of learning. According to Piaget’s (1932) theory of cognitive development, students learn to reason by reconciling conflicting points of view as well as justifying their own perspective during social interaction. Working with a sociocognitive lens, Doise and Mugny (1984), found that experiencing social conflict prompted students to productively reorganize and elaborate their mental representations. These theoretical notions provide insight into how successful readers form mental models of text as they read and incorporate new information with prior knowledge.

Three primary findings emerged from Almasi’s (1995) analysis. First, peer-led groups were better able to recognize and resolve cognitive conflict on a transfer task compared to students in teacher-led discussion groups. Second, who controlled the discussion influenced the type of conflict students experienced. Specifically, students in peer-led groups recognized and initiated conversation about their own conflicts in an attempt to reconcile them. Students in teacher-led groups did not seem to recognize their own conflicts until they were pointed out by the teacher or peers. Finally, the nature of the talk was different across peer- and teacher-led groups. Students in peer-led groups discussed more alternative points of view, asked more questions, and responded with more elaborated, complex talk. The structure of talk in teacher-led groups followed a traditional, procedural pattern (i.e., the teacher voiced the majority of questions and evaluate student responses).

In sum, the potential of discussion-based pedagogy is undergirded by various theories of teaching and learning as well as empirical findings. However, the discussed works investigated the effects of discussion in naturalistic settings and say little about specific environmental factors (i.e., specific parameters of an instructional frame, teacher’s facilitation practices) that may
facilitate or inhibit the positive effects of discussion on student reading comprehension. The following section reviews empirical work regarding individual text-based discussion approaches as well as how they differentially effect student reading comprehension.

**Text-based discussion approaches**

Scholars have devoted concerted efforts to developing specific discussion approaches aimed at promoting student reading comprehension. While these approaches have a lot in common, they can be categorized in terms of their instructional frame (Chinn, Anderson, & Waggoner, 2001). An approach's instructional frame helps to set the boundaries for what happens during discussions. In reviewing prominent approaches to classroom discussion, Wilkinson and Son (2010) identified several parameters that can be used to describe an approach's instructional frame. Those parameters include but are not limited to: stance adapted toward text; who controls turns and text; who leads the discussion; who holds interpretive authority; group composition; as well as pre- and post-discussion activities. Taken together, the individual approaches can largely be categorized by their primary stance toward text and the extent to which the teacher controls the discussion (e.g., controlling turns and holding interpretive authority).

The concept of stance originated as an explanation as to why readers who have read the same text can walk away with strikingly different interpretations (Roseneblatt, 1978), suggesting that how the reader approaches the text determines what the reader takes away from the text. According to Roseneblatt (1978), readers comprehend or process information from the text differently depending on the nature of their relationship with the text. In other words, how readers comprehend the text is largely influenced by the purpose and goals they set forth for any particular reading event. Murphy et al. (2009) identified three common stances adopted by text-
based discussion models (i.e., expressive, efferent, and critical-analytic). The expressive stance encourages readers to live through the text (Soter et al., 2008). The efferent stance encourages students to extract as much information, such as facts, concepts, or details, from a text as possible (Rosenblatt, 1978). Finally, a critical-analytic stance encourages students to consider multiple perspectives within the text as well as evaluate its underlying arguments (Wade, Thompson, & Watkins, 1994).

Teachers typically exert the least control over discussion approaches espousing an expressive stance. As such, students have the flexibility to evaluate their emotional reactions to the text features (Rosenblatt, 1978) and build connections between the text and their personal lives (Jakobson, 1987; Soter, Wilkinson, Connors, Murphy, & Shen, 2010). Teachers generally have the most control over discussion approaches espousing an efferent stance. Such discussions include traditional question-answer classroom discussions during which the teacher asks fact-based questions and has control of who talks and when, as well as what is considered ‘right.’ Teachers and students, however, share control of discussion approaches characterized by the critical-analytic stance.

Shared control provides structure and flexibility for students to respond to any unsolved issues or problems they see in those arguments with reasoned, text-based arguments.

Whereas some discussion approaches are aimed at promoting such high-level thinking, others by comparison are aimed at promoting students’ basic comprehension or fostering students’ personal connections to the text. In an effort to better understand the effect of individual discussion approaches on high-level comprehension, Murphy et al. (2009) conducted a comprehensive meta-analysis examining the results from 42 studies documenting the effects of nine popular discussion models (i.e., Collaborative Reasoning, Paedia Seminar, Philosophy for Children, Book Club, Grand Conversations, Literature Circles, Instructional Conversations, Questioning the Author, and Junior Great Books). The findings reported indicate that some
approaches are better at promoting high-level comprehension and critical-analytic thinking than others. Specifically, their results showed that stance accounted for the most variance in terms of efficacy of approaches for promoting high-level comprehension. Discussion approaches holding an efferent stance were particularly good at promoting basic comprehension. Alternatively, the critical-analytic stance better supported high-level comprehension. Notably, Murphy and colleagues found that increasing the quantity of student talk was necessary but not sufficient for enhancing high-level comprehension. Instead, their findings indicated that it was essential that discussion approaches aimed at high-level comprehension encouraged students to engage in specific discursive behaviors.

Some researchers have worked to identify the discourse elements characteristic of the kind of talk that leads to high-level comprehension. Several studies show that productive discussions are marked by high levels of student and teacher authentic questions (Nystrand, 1997; Nystrand & Gamoran, 1991), as well as episodes of individual and collaborative oral argumentation (Schwarz & Linchevski, 2007; Schwarz, Neuman, & Biezuner, 2000). Among them, Soter and colleagues (2008) closely examined the nature of the talk representative of the nine discussion approaches included in Murphy et al. (2009). Importantly, Soter and colleagues operationalized episodes of individual argumentation as elaborated explanations and episodes of collective argumentation as exploratory talk. The data for their analysis consisted of 36 discussion transcripts (i.e., four per discussion model) which took place in classrooms spanning grades three through nine.

Their results indicated that the frequency of authentic questions as well as individual and collective oral argumentation varied by stance. Specifically, expressive approaches were marked by high frequencies of student authentic questions and exploratory talk, but fewer episodes of elaborated explanations. Efferent approaches were indicative of relatively high levels of teacher test questions and elaborated explanations, but few episodes of expository talk. Approaches
taking a critical-analytic stance were marked by high levels of teacher and student authentic questions, as well as high levels of both elaborated explanations and exploratory talk. Soter and colleagues (2008) hypothesized that the high levels of all three indicators was due to the shared control characteristic of critical-analytic discussions allowing for more teacher modeling and scaffolding. These findings further suggested that the discourse patterns indicative of each stance co-occurred with particular reading comprehension outcomes.

Taken together, these findings suggest that discussions likely to result in high-level comprehension are structured, yet flexible, and facilitate an open exchange of ideas about, around, and with text. Discussion approaches fitting this description typically take a critical-analytic stance, encourage the teacher and students to share responsibility, and emphasize the quality of student talk over quantity. As such, discussions that promote high-level comprehension are marked by specific discussion elements (i.e., authentic questions, elaborated explanations, and exploratory talk). Given their relationship with high-level comprehension, the section that follows briefly defines and reviews the extant literature regarding three specific discussion elements--authentic questions, elaborated explanations, and exploratory talk.

**Discourse elements**

Authentic questions are questions for which the questioner does not know the answer or genuinely wants to know how others will respond (Nystrand, 1997; Nystrand & Gamoran, 1991; Nystrand et al., 2003). As such, authentic questions do not have pre-specified answers. The lack of a pre-specified answer allows students to respond with their own thinking and actively participate in constructing knowledge. Nystrand (1997) argued that the use of authentic questions signals underlying classroom epistemology. That is, the extent to which classrooms value the exploration of ideas. Thus, asking authentic questions reflects Baktin’s notion of a dialogic
environment while monologic classrooms incorporate high levels of test questions which serve to test students’ memory. Moreover, authentic questions foster reading comprehension as they typically elicit generalizations, analysis, or speculation. They also invite students to generate extratextual (i.e., intertextual, shared knowledge, affective) connections which further aid in comprehension (Allington & Johnson 2002; Applebee et al., 2003; Bloome & Egan-Robertson, 1993; Edwards & Mercer, 1987; Taylor, Pearson, Peterson, & Rodriguez, 2003).

As students discuss, authentic questions elicit episodes of individual and collaborative argumentation. Argumentation can be conceptualized as supporting claims with reasons and evidence as well as critically evaluating others’ arguments and ideas (Allen, Montalbano, & Duke, 2017). Engaging in oral argumentation during discussion can increase students’ content knowledge (Chinn, O’donnel, & Jinks, 2000), stimulate conceptual change (Asterhan & Schwarz, 2009), and raise students’ awareness of their misconceptions as well as knowledge gaps (Webb & Pal, 1996). As such, discussions leading to high-level comprehension are marked by frequent instances of both individual and collective oral argumentation (Soter et al., 2008). Specifically, during productive discussions, students respond to their peers’ authentic questions with elaborated explanations (i.e., individual argumentation) and participate in exploratory talk (i.e., collective argumentation).

Elaborated explanations go beyond providing a correct answer to a problem or question; they include an explanation or justification of one’s thinking (Webb, 1984; 1991). Therefore, elaborated explanations are ones in which students provide more than one reason or piece of evidence to support or justify their claims. These explanations are considered elaborated because they contain multiple pieces (e.g., a claim, two connective reasons, a reason and evidence).

Imagine, for instance, that a student asks, “Do you think Harry Potter is a good role model for kids?” and another student responds, “No, because he didn’t finish school and he is kinda grouchy with his friends sometimes.” This is an elaborated explanation because it has a
claim (i.e., no) and two pieces of support (i.e., didn’t finish school and he is grouchy). However, it would not be an elaborated explanation if the student simply said, “No,” or “No because he didn’t finish school.”

Collectively constructing and deconstructing arguments as a group can also foster the ability to critically analyze new ideas and information. For example, Mercer, Wegerif, and Dawes (1999) examined the effects of exploratory talk on fourth and fifth graders’ reasoning in a quasi-experimental study. Episodes of exploratory talk are those in which students co-construct knowledge and reason together (Mercer, 2002). Students do this by critically engaging with each other’s ideas and building toward some level of agreement. For example, one student may offer a claim that is supported by reasons and evidence and another student may challenge that claim or the validity of the reasons and evidence. Episodes of exploratory talk can be conceptualized as a dialogic game of chutes and ladders. Students build up to some shared understanding, but also encounter setbacks along the way. Mercer and colleagues found that, on average, participating in episodes of exploratory talk improved students’ scores on a standardized reasoning measures compared to a comparison group.

These studies examined the types of talk characteristic of discussions holding a critical-analytic stance. However, the discussed findings do not shed light on how additional aspects of the instructional frame (e.g., group composition) may influence student talk. Further, these studies do not address how a teacher’s specific discourse moves may interact with the instructional frame to influence student talk. As such, the present study explores the interaction of these two factors on student talk. Specifically, the study considers the extent to which the relationship between a teacher’s facilitation practices and instances of students’ elaborated explanations vary as a function of group composition. The following sections introduce the literature on group composition and teacher facilitation practices and serve to set the context for the current study.
Small-Group Composition

While research findings have firmly established the benefits of small-group compared to whole-class instruction (Kulik, 1992; Lou et al., 1996, Slavin, 1987), there remains little consensus regarding how to organize the groups. The most controversial issue concerning group composition is whether or not students should be placed in groups with students of similar (i.e., homogeneous) or dissimilar (heterogeneous) ability. It is worth noting that in this context ability is a measure of student proficiency at a particular point in time. Thus, ability can change over time based on student experience (Dewey, 1938).

Advocates of homogenous grouping argue that it allows teachers to differentiate instruction (Esposito, 1973; Froman, 1981; Li, 2018; Oakes, 1985; Wilkinson & Fung, 2002). From this perspective, homogeneous grouping allows the teacher to scaffold instruction within each student’s zone of proximal development. Vygotsky (1978) refers to the Zone of Proximal Development (ZPD) which represents a student’s learning potential when instruction is challenging, yet attainable with appropriate levels of assistance. As such, advocates of ability grouping argue that high performing students will be held back if they are placed in a group with average or low performing students, and that ability grouping allows for any necessary remediation. Alternatively, the primary argument for heterogenous grouping is that students of varying abilities can help each other (Slavin, 1987).

Research findings as to the effect of homogeneous grouping compared to heterogeneous grouping on student achievement, however, are not consistent across ability level (i.e., low, average, high). Lou et al. (1996) conducted a meta-analysis comparing the effects of homogeneous and heterogeneous grouping on student achievement in populations ranging from first grade to college. Their analysis included 103 effect sizes from 51 studies investigating the overall impact of small-group instruction. Results indicated that, on average, small-group
instruction had a small, but significant effect on student achievement. Lou et al. (1996) also compared 20 effect sizes from the 12 studies in their sample that directly compared homogeneous and heterogeneous group compositions. Their findings indicated a slight preference for homogeneous grouping in reading overall. However, the superiority of homogeneous grouping was not consistent across ability level. Specifically, their findings indicated that low-ability students benefit more from heterogeneous groupings, average-ability students learned more in homogeneous groups, and high-ability students performed well regardless of group composition. These findings were largely replicated in Lou, Abramim, and Spence (2000).

There is some evidence that student talk may be the key to understanding the differential effect group composition has on students of varying ability levels. Wilkinson and Fung (2002) hypothesized that group composition indirectly effects student achievement by influencing student participation as well as teacher facilitation practices. In essence, group composition impacts the nature of student discourse by shaping the normative social expectations of the group. Saleh, Lazonder, and De Jong (2005) investigated how ability grouping influences student achievement, interaction, and motivation in a biology course. A hundred and four fourth-graders were randomly assigned to a homogenous or heterogeneous group using a standardized science achievement measure. Students participated in 16 plant biology lessons consisting of brief whole-class introduction followed by small-group collaboration. In regards to overall achievement, Salah and colleagues’ findings support those of Lou et al. (1996). Low ability students fared best in heterogeneous groups, average ability students fared best in homogeneous groups, and high ability students did well regardless of group composition. However, their findings also indicated that student talk was related to group composition. Specifically, homogeneous ability-groups produced more collaborative explanations (i.e., explanations constructed by multiple students), on average, than their heterogeneously grouped peers who tended to produce more individual explanations (i.e., explanations constructed by one student).
Evidence supporting this hypothesis is particularly strong for low-ability students. Research on group processes indicates that low-ability students form a teacher-student relationship with their high-ability peers in heterogeneous groups (Salah et al., 2005; Webb, 1982, 1991). As such, low-ability students ask more questions and, thus, receive more explanations in heterogeneous groups (Azmitia 1988; Tudge, 1989; Webb, 1980). Furthermore, low-ability students in homogeneous groups tend to produce fewer explanations overall and have a tendency to provide incorrect information. High-ability students tend to address these misconceptions in heterogeneous groups. While there is evidence that generating elaborated explanations is generally more beneficial than receiving them (Webb, 1992), Webb and Farivar (1999) found that receiving elaborated explanations was related to achievement among students who possess content-related misconceptions. Decreasing the diversity of student ability within the group effectively decreases the number of peers to whom low-ability students can ask questions and receive elaborated explanations. Webb, Farivar, and Mastergeorge (2002) also indicated a variety of conditions that are necessary for help-giving and help-receiving to be effective. Explanations must be relevant to the students’ needs, timely, accurate, and detailed. It may be that these conditions are satisfied in heterogeneous groups and not homogeneous groups.

Alternatively, empirical evidence suggests that average-ability students benefit from homogenous groupings (Lou et al., 1996; Saleh et al., 2005; Webb, 1991). The research is scant, but average-ability students seem to be left out of the teacher-student relationship formed by low- and high-ability students in heterogeneous groups (Salah, Lazonder, & De Jong, 2007; Webb & Palincsar, 1996). As such, heterogeneous groups do not afford them the opportunity to ask questions, generate elaborated explanations, or receive explanations that are timely and relevant to their needs.

Research findings concerning the effect of group composition on high-ability student talk is mixed. Azmitia (1988) concluded that high-ability students engage in similar talk regardless of
group composition. Others, however, argue that high-ability students benefit from filling the teacher role in heterogeneous groups (Johnson, Skon, & Johnson, 1980; Webb, 1980, 1991). While they ask fewer questions in heterogeneous groups, they also provide almost all of the explanations (Salah et al., 2005) and engage in more complex reasoning (Johnson et al., 1980). Alternatively, there is evidence that high-ability students generate more explanations in homogeneous groups as they try to reconcile conflicting views. In a study involving 662 seventh- and eighth-graders, Webb, Nemer, Chizhikm, and Sugue (1998) examined the effects of group composition on science achievement. Students completed a hands-on test individually or in a group of three. Groups were designed as either heterogeneous or homogenous ability and instructed to ask and answer questions as they solved test items. Their results indicated that high-ability students in homogeneous groups were better able to build on each other’s ideas and engaged in more collaborative reasoning.

In essence, there is some indication that group composition affects the quality of student talk but the findings are relatively limited. Notably, none of the reviewed studies occurred within the context of a specific discussion approach. Recently however, Murphy, Greene, et al. (2017) examined the extent to which group composition was associated with variation in students’ basic and high-level comprehension as well as the nature of student and teacher talk during Quality Talk discussions across three time points. Quality Talk (QT) is a flexible, teacher-facilitated small-group discussion approach aimed at promoting high-level comprehension and critical-analytic thinking and reasoning about, around, and with text and content (Murphy & Firetto, 2018). The QT model espouses a primarily critical-analytic stance while also incorporating elements of efferent and expressive stances. Sixty-two fourth- and fifth-graders were assigned within-grade to a homogenous or heterogeneous discussion group on the basis of oral reading fluency. Over the course of a year, each group participated in a weekly discussion about a story they read as part of their normal language arts curriculum. Murphy, Greene, et al.’s findings
indicated that while group composition did not have an effect on students’ basic comprehension, it did have an effect on students’ high-level comprehension. Specifically, heterogeneous groups were, on average, superior to homogeneous groups. Moreover, low-ability students in homogeneous groups did not demonstrate the same growth in high-level comprehension as did their peers. Low-ability students in homogeneous groups did, however, demonstrate the largest gains in basic comprehension.

Murphy, Greene, et al. (2017) also examined the extent to which group composition was associated with the quality of student and teacher talk during Quality Talk discussions. Their results indicated three interesting findings. First, group composition seemed to affect the nature of student discourse over time. While all groups posed comparable amounts of authentic questions between Time 1 and Time 2, the homogenous low group did not demonstrate the same growth in oral argumentation between Time 2 and Time 3 as the others. Next, ability level was associated with the quality of student discourse regardless of group composition. High-ability student questions elicited more generalizations, analyses, and speculations. Further, high-ability students were more likely to build on each other’s arguments as well as challenge what others said. Finally, the degree to which the teacher released control of the discussion over time differed by group composition. Specifically, the teacher continued to use discussion scaffolding and modeling moves for a longer period of time with the homogeneously grouped low-ability students than either the homogeneously grouped high-ability students or the heterogeneously grouped students.

In sum, the use of within-class ability grouping is controversial in part because the impact of group composition is not consistent across ability levels. Some scholars hypothesize that the inconsistent findings may be due to the fact that ability grouping mediates the nature of student talk and engagement. Indeed, there is evidence that average-ability students produce and receive more elaborated explanations in homogenous groups than in heterogeneous groups, while low-
ability students benefit from receiving more elaborated explanations in heterogeneous settings. However, there is also evidence that the effect of within-class ability grouping on the nature of student talk can be mediated by the teacher’s facilitation practices (Saleh et al., 2007). As such, the following section reviews the extant literature on the role of the teacher during discussion.

**Teacher Facilitation Practices**

While less is known about the role of the teacher, there is considerable consensus amongst scholars that the teacher has a powerful effect on the nature of classroom discussions (Brown, Palincsar, & Armbruster, 1984; Michaels & O’Connor, 2015; Nystrand, 1997; Webb, Franke, Turrou, & Ing, 2015; Wei et al., 2018). The teacher makes a handful of decisions that set the boundaries of the instructional frame and, subsequently, the nature of student talk and learning. As such, the teacher ultimately determines who controls the topic and turn taking as well as who poses interpretive authority. During traditional question-and-answer style classroom discussions, for example, the teacher holds primary control over all aspects of the instructional frame. Students do not have opportunities to ask questions, engage in oral argumentation, or generate explanations. Alternatively, during productive discussions, teachers gradually release control of the discussion to their students, providing greater opportunities for students to engage in higher-quality talk.

In an effort to better understand how discourse unfolds in classrooms, Nystrand et al. (2003) conducted a secondary analysis of the data presented by Nystrand (1997). Their analysis of 33,904 questions shed light on the kind of teacher and student talk that prompts and sustains meaningful discussion. Their findings indicated that teacher authentic and uptake questions as well as student questions prompted productive discussions. Student questions proved to be particularly powerful. One student question posed within the last five questions raised the
probability of a productive discussion by 200% compared to sequences of five questions including one teacher question. Importantly however, their results also showed that teachers can facilitate productive discussions in direct and indirect ways. Teachers can prompt discussions by asking authentic questions, eliciting student thinking, and encouraging students to respond to each other instead of evaluating student answers themselves.

Nystrand et al. (2003) also examined whether the same talk patterns held in low-track and high-track classrooms. Their results provided further context for the finding presented by Nystrand (1997). Namely, that while the same number of authentic questions are asked in low- and high-track classrooms, instances of authentic questions cluster differently by track. Authentic questions are asked in closer succession in high-track classrooms as compared to low-track classrooms where instances of authentic questions are more widely distributed. Thus, students in low-track classrooms experience fewer productive discussions which build on what was said previously. Further, students in low-track classrooms posed fewer questions than their peers in high-track classrooms and when they did, those questions tended to be off topic (Nystrand, 1997).

Based on their findings, Nystrand and colleagues (2003) argued that how teachers and students interact with each other drives the extent to which discussions are productive. As such, it is more important to consider how teachers and students interact than what each party does individually. This sentiment is consistent with the findings reported in Soter et al. (2008). In particular, that teachers and students share control of discussions likely to promote high-level comprehension. By comparison, teachers typically control discussions shown to promote basic comprehension. Unfortunately, a regular message found in the literature is that teachers struggle to release control because doing so requires radically changing classroom norms (Alvermann & Hayes, 1989; Alevermann, O’Brien, & Dillon, 1990; Billings & Fitzgerald, 2002; Kucan, 2009; Michaels & O’Connor, 2015).
As such, researchers have identified specific behaviors (i.e., teacher moves) that teachers can use to facilitate productive discussions (e.g., Beed, Hawkins, & Roller, 1991; Gaskins, Rauch, Gensemer, & Cunicelli, 1997; Gillimore & Tharp, 1990; Wei et al., 2018). More recently, Wei et al. produced an integrated taxonomy of teacher moves through a systematic review and analysis of the literature on the various discussion approaches aimed at promoting reading comprehension. Their review included two articles representing each of the nine discussion approaches included in Murphy et al. (2009) (i.e., 18 articles total). Twenty-eight descriptions of teacher moves were extracted from the included articles. Seven discourse coders and four teachers with experience facilitating discussions participated in a three-phased card sort. These procedures resulted in a final, 12-move taxonomy (i.e., backchanneling, challenging, checking, clarifying, debriefing, instructing, marking, modeling, procedural, prompting, reading, and summarizing). See the appendix for definitions and discourse examples of each move.

Teacher moves are tools teachers can use to encourage students to share their thinking, listen and respond to each other’s ideas, and challenge students to think more deeply during discussion (Michaels & O’Connor, 2015). Research findings concerning the effectiveness of teacher moves are rare. Of the few studies that have been conducted, most have examined the aggregate effectiveness of teacher moves (e.g., Beck, McKeown, Sandora, Kucan, & Worthy, 1996; Biskin, Hoskinsson, & Modlin, 1976; Chesser, Gellalty, & Hale, 1997; Chinn et al., 2001; Croninger & Croninger, 2016; Kong & Kitch, 2003; McGee, Courtney, & Lomaz, 1994; Webb et al., 2015). Among them, Webb et al. (2015) found that the degree to which students working in pairs generated complete and correct explanations was associated with the degree to which teachers prompted for details. Furthermore, their results also indicated that teacher’s use of instructing and procedural moves were associated with whether both students generated explanations. Gillies and Khan (2008) showed that the frequency of explanations as well as student performance on a reasoning and problem-solving task were associated with teacher’s use
of teacher moves during discussion (e.g., prompting, clarifying, challenging). While these studies provide some indication that teacher moves facilitate discussions and promote meaningful teacher-student interaction, they do not shed light on the power of each move individually.

Only two studies have addressed the influence of individual teacher moves on the turn-by-turn sequence of events unfolding in a discussion. Both studies were conducted in the context of Collaborative Reasoning discussions. Collaborative Reasoning is a text-based discussion approach with a critical-analytic stance in which the teacher facilitates discussions and gradually release control to the students. First, a study conducted by Jadallah et al. (2011) examined the influence of teacher moves on student’s use of evidence. Their analysis included 30 discussion transcripts of three, fourth-grade discussion groups. Sequential analysis was used to determine the effect of three teacher moves (i.e., prompting and marking student use of evidence, clarifying, and challenging) on the talk patterns taking place during discussion. Their results showed that teacher moves can influence not only the immediate student turn, but several subsequent student turns as well. Prompting for evidence increased the likelihood that the next two student turns would include text-based evidence. Notably, evidence typically built on or challenged what had been said previously. Jadallah et al. also found that prompting for evidence was more effective than praise. Finally, students were more likely to collectively construct an argument if the teacher clarified or challenged a student’s contribution to the discussion.

In a second study involving 124 fourth-grade students, Lin et al. (2015) examined the influence of teacher moves on students’ relational thinking over the course of ten discussions. Thirty-seven teacher moves were grouped into four major categories (i.e., prompting, modeling, marking, and managing). A total of 57 discussions (i.e., three discussions per group) were coded for instances of relational thinking and teacher moves. Lin and colleagues also used sequential analysis to determine the relationship between the teacher moves and instances of relational
thinking. Findings indicated that even minimal teacher involvement served as a catalyst for students to engage in relational thinking.

Overall, empirical finding suggest that teachers play a key role in determining the nature of classroom discussions. For instance, teachers can use teacher moves while students discuss to facilitate productive talk about, around, and with the text. Based on the previously reviewed studies, however, it seems likely that the effect of individual teacher moves may vary by group composition. Yet, few studies have examined the extent to which student talk, group composition, and teacher moves experience a reciprocal relationship. The current study addresses this gap through a systematic investigation of the turn-by-turn nature of text-based, small-group discussions in the context of Quality Talk.
Chapter 3

METHOD

Data for the current investigation were collected as part of a large, multi-year project funded by the Institutes of Educational Science (IES) entitled “Quality Talk: Developing students’ discourse to promote critical-analytic thinking, epistemic cognition, and high-level comprehension” (Grant number: R305A130031). The present study examined the influence of specific teacher moves on students’ elaborated explanations across various ability group compositions during Quality Talk (QT) discussions. This chapter begins with a description of the participants included in the study and an outline of the grouping procedures. Next, a brief description of the QT intervention is presented followed by a description of the measures and coding procedures that were employed in the study. Finally, the chapter concludes with an outline of how the study was conducted and a brief introduction to the analysis used to address each research question.

Participants

Study participants included fourth and fifth-grade teachers (n = 4) and their students from a private elementary school in a small, Midwestern city. The racial make-up of the school’s student population was predominantly White (86%), and approximately 30% received free or reduced lunch. The average reading proficiency was 76% across grades four and five. Students who participated in the study were demographically representative of the school.

All participating teachers were female and possessed 1 – 17 years of teaching experience. A total of 63 students had their parents’ permission to and agreed to participate in the research.
One student moved during the school year, however, so the sample was reduced to 62. Among the sampled students, 28 were in fourth grade (15 female, 13 male) and 34 were in fifth grade (19 female, 15 male). Three of the teachers and approximately three fourths of the fifth-grade students had a year of experience participating Quality Talk (QT) prior to the study.

**Grouping**

In consultation with participating teachers, students were grouped across class, within grade to ensure equivalent exposure to within grade teachers during discussions. Students were divided into six discussion groups per grade (i.e., twelve groups total) based on their oral fluency level (i.e., number of words read correctly in one minute). Oral fluency was used as the grouping variable because a number of studies have indicated it is the most valid indicator of student reading comprehension ability (Fuchs, Fuchs, Hosp, & Jenkins, 2001; Goffreda & DiPerna, 2010; Johnson, Jenkins, Petscher, & Catts, 2009).

Students were divided into six grade level groups using a two-phase process. The aim of the first phase was to randomly assign students to either a homogeneous or heterogenous group. First, a member of the research team measured students oral reading fluency (i.e., words read correctly per minute) using AIMSweb. Students were then ranked within grade based on their oral reading fluency score and paired with a comparable, grade-level peer. Next, each student was randomly assigned a number using Excel. The student with a higher number in each pair was assigned to one of three within grade homogeneous groups (i.e., above average ability, average ability, below average ability). The student with the lower number was placed in one of three within grade heterogenous groups in alternating order so that each heterogenous group contained students with above-average, average, and below-average oral reading fluency.
The aim of the second phase was to assess group characteristics and finalize group membership. Group fluency means and standard deviations were calculated to ensure that the procedures outlined above resulted in a high degree of homogeneity across the three homogeneous groups as well as a high degree of heterogeneity across the three heterogeneous groups in each grade. Larger standard deviations indicated more heterogeneity among students oral reading ability while smaller stand deviations indicated more homogeneity. Adjustments were made to ensure that each group had a mix of students from each grade-level class and at least one male student. The resulting lists were collaboratively reviewed and finalized with the teachers. The grouping procedures resulted in six ability groups in each grade (i.e., a homogenous below average group, a homogenous average group, a homogenous above average group, and three heterogeneous groups). The fourth-grade groups were comprised of four to five students while the fifth-grade groups were comprised of five to six students.

**Intervention**

Quality Talk (QT) is a teacher-facilitated, small-group discussion model designed to enhance students’ high-level comprehension and epistemic cognition about, around, and with text and content (Murphy, 2018). In order to implement QT with fidelity, teachers must possess a thorough understanding of the four components included in the QT model (i.e., an ideal instructional frame, discourse elements, teacher modeling and scaffolding, and a set of pedagogical principles) and students need to possess knowledge about how to engage in productive talk. Thus, teachers participated in initial and ongoing professional development. Students learned how to engage in productive discussions through a set of discourse mini-lessons, completed pre- and post- discussion activities in their QT literacy journals, and engaged in weekly, teacher facilitated small-group discussions about a text included in their regular language
arts curriculum. The following subsections provide a comprehensive explanation of the QT professional development, mini-lessons, literacy journals, and discussions.

**Professional development**

Teachers participated in an initial two-day professional development workshop prior to implementing QT in their classrooms. The purpose of the workshop was to overview the four components of the QT model (i.e., an ideal instructional frame, discourse elements, teacher modeling and scaffolding, and a set of pedagogical principles) to the teachers. As such, the teachers were introduced to QT’s ideal instruction frame including QT’s stance toward text as well as the different roles teachers and students would hold during QT discussions. They were also taught to identify productive discourse elements (e.g., authentic questions and elaborated explanations), employ a set of teacher moves (i.e., prompting, marking, modeling, summarizing, and challenging), and effectively facilitate QT discussions. Finally, teachers were introduced to the QT mini-lessons.

In addition to the initial profession development workshop, teachers met one-on-one with a discourse coach for approximately one hour twice a month throughout the school year. The purpose of these coaching sessions was to offer ongoing support to the teachers as well as to debrief about recent discussions and QT mini-lessons. Before each coaching session, the teacher and their discourse coach independently coded a ten-minute segment from one of the previous weeks discussions using the Discourse Reflection Inventory for Teachers (DRIFT). During coaching, the teacher and their coach compared their coded segments. Comparing the coded segments provided an opportunity to analyze the turn pattern, discuss any abundance or absence of discourse elements in the discussions, and set goals for future discussions. Furthermore, the
teacher and their coach talked about what was going well overall and how to build on those successes during future implementation.

**Quality Talk mini-lessons**

The teachers implemented twelve, researcher developed, mini-lessons (see Table 3-1) which were organized into three broad categories: questioning, argumentation, and writing. The questioning and argumentation mini-lessons aimed to enhance students’ critical-analytic thinking and reasoning through explicit instruction of specific discourse skills. Students were taught how to ask different types of authentic questions including uptake questions, high-level thinking questions, and connection questions. They were also taught how to respond to different types of questions with reasoned arguments comprised of claims supported by multiple reasons or pieces of evidence from the text as well as challenge their peers’ arguments with additional reasoning and evidence. With each mini-lesson, students were introduced to a new discourse element and given an opportunity to practice identifying and creating the new question type or piece of an argument.

<table>
<thead>
<tr>
<th>Category</th>
<th>Mini-Lesson One</th>
<th>Mini-Lesson Two</th>
<th>Mini-Lesson Three</th>
<th>Mini-Lesson Four</th>
<th>Mini-Lesson Five</th>
<th>Mini-Lesson Six</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question Types</td>
<td>Road Trip</td>
<td>Authentic Questions, Test Questions &amp; Uptake</td>
<td>Inter-textual Questions</td>
<td>Generalization Questions &amp; Analysis Questions</td>
<td>Speculation Questions</td>
<td>Affective Questions</td>
</tr>
<tr>
<td>Argumentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3-1: List of Quality Talk Mini-lessons.
The writing mini-lessons aimed to enhancing student’s persuasive and compare and contrast essay writing through explicit instruction of a generic writing strategy as well as several genre-specific strategies. Students were taught to plan for writing with several mnemonics and a graphic organizer. In addition, students learned how to transition their notes from the graphic organizer to writing as well as how to evaluate and revise their writing. Finally, they also learned a set of transition words to enhance their writing.

**Quality Talk literacy journals**

Each student was provided with a Quality Talk literacy journal which contained before, during, and after reading activities for each text in the curriculum. Journal activities were rooted in the extant literature on reading comprehension but tailored to the routines of participating teachers’ instruction and curriculum (e.g., Powerful words). Before-reading activities prompted students to consider their prior knowledge related to the text, text-genre, and new vocabulary as well as the author’s purpose. During-reading activities encouraged students to consider text-genre as well as main idea and supporting details. After-reading activates included nine Likert scale items that elicited student’s perceptions about the text, generating different types of authentic questions, and recording individual and group discussion goals. Teachers provided students with feedback on their questions and goals prior to the discussion. Students brought their journal to each discussion so that they could reference their questions.

**Quality Talk discussions**

Teachers facilitated weekly QT discussions centered around a text from the school’s regular Language Arts curriculum (i.e., *Reading Street*) throughout the school year. Each week,
the teachers facilitated three within grade discussion groups, including at least one homogeneous and heterogenous group. After facilitating discussions with the same three groups for approximately three to four weeks the teachers switched groups within grade and continued to switch groups every three to four weeks for the remainder for the school year.

Teachers began each discussion by reviewing the normative expectations for productive discussions as well as a group’s individual discussion goals. After the review, teachers and students asked and answered authentic questions about, around, and with the text. The talk was characterized by an open participation pattern where the students managed the flow of the discussion by responding to each other naturally, calling on each other, and challenging each other. Each discussion lasted approximately 15 to 20 minutes and ended with a short debrief about how the discussion went as well as goal setting. Discussions were audio and video-recorded for later coding.

**Measures**

**Oral reading fluency**

Student’s oral reading fluency was assessed using AIMSweb’s Reading Curriculum Based Measurement (R-CBM; Shinn & Shinn, 2002). This measure was selected for two reasons. First, the participating school was using AIMSweb to assess student’s oral reading at the beginning of the school year. Second, scores on the AIMSweb measure have been deemed reliable and valid (SEM = .97) (Daniel, 2010; Shinn, Good, Knutson, Tilly, & Collins, 1992). Student’s oral reading fluency was assessed individually by a trained member of the research team. Each student read three grade-level passages aloud for one minute. After which, the research team member who administered the assessment scored each passage for number of
words read correctly. The median score was used to assign students to either a homogeneous or heterogeneous ability group prior to the intervention. The national average on AIMSweb in fourth grade is 121.5 words per minute (SD = 25.3); the national average in fifth grade is 132.1 words per minute (SD = 29.1).

**Discourse coding**

Six trained discourse coders analyzed the middle 10-minutes of 108 QT discussion videos in accordance with the Quality Talk Discourse Coding Manual (Murphy, Firetto, Greene, & Butler, 2017) and the taxonomy of teach moves included in Wei et al. (2018). The middle 10-minutes were used to ensure a representative sample of the discourse and to control for the varied length of the discussions. The 108 discussions were chosen to achieve a balanced sample of discussions across students’ exposure to the QT mini-lesson content and to allow for teachers’ gradual release of control. As such, the sample included nine discussions from each of the twelve discussion groups across fourth and fifth grade: three discussions that occurred during implementation of the questioning mini-lessons (i.e., week 2, week 4, week 6), three discussions that occurred during implementation of the argumentation mini-lessons (i.e., week 8, week 10, week 12), and three discussions that occurred during implantation of the writing mini-lessons (i.e., week 15, week 17, week 19).

Using StudioCode (version 5.8.3), coders identified instances of (1) teacher’s discourse moves and (2) students’ responses according to the procurers delineated in Murphy, Firetto, et al. (2017). Specifically, coders analyzed each turn in the 108 discussions to identify instances of challenging, checking, clarifying, instructing, marking, modeling, prompting, reading, summarizing, procedural moves, and elaborated explanations in order to establish a sequence of meaningful utterances for each individual discussion. A turn was operationalized as a time
segment during which someone holds the floor of the discussion. For example, the boundaries of a turn were often signed by alternating speakers. A miscellaneous code was assigned to participation turns which did not include any teacher moves or elaborated explanations. Each discussion was coded by double coded and reconciled. See Figure 3-1 for a sample of a coded sequence of talk from a fourth-grade discussion.

**Procedures and analysis**

Teachers, parents, and students were consented or assented as appropriate before the beginning of the study. QT was implemented in all four classrooms immediately following baseline discussions. Within grade, students read a story, completed literacy journal activities, and discussed a text each week. Teachers facilitated 19 QT discussions over the course of the intervention. Each week teacher facilitated three within grade discussions; rotating which discussion groups they facilitated approximately every three weeks. Throughout the school year, each teacher also implemented all of the questioning, argumentation, and writing lessons. All QT activates were video and audio recorded.
Figure 3-1: Sample of a coded sequence of talk.
Chapter 4

RESULTS

The results presented in this chapter are organized within grade to preserve clarity. Within grade, I present descriptive statistics and analyses addressing the three research questions I posed in Chapter 1. The descriptive statistics provide a picture of the distribution of turns by group composition. Specifically, the overall talk pattern that emerged for the homogeneous and heterogeneous groups in terms of the frequency and percentage of turns each discourse indicator (i.e., ten teacher moves and elaborated explanations) accounted for within group. Next, I describe the analysis and findings for my first research question which pertained to the relationship between teacher moves and group composition. In order to fully address this question, I present an analysis of the association between teacher moves and group composition overall (i.e., homogenous groups and heterogeneous groups) as well as an analysis of the association between teacher moves and group composition across the levels of the homogeneous groups (i.e., below-average, average, and above-average). Then, I present the findings associated with my second research question pertaining to the sequential patterns imbedded in the discourse over the course of the yearlong intervention. I present the results of the lag-sequential analysis for the pooled heterogeneous groups as well as the homogeneous below-average group, average group, and above-average group individually. Finally, I address my third research question via a comparison of the effect sizes associated the sequential patterns presented in the third subsection.
Fourth-Grade Results

Descriptive statistics

Trained graduate students coded a total of 6,074 turns across the 54 fourth-grade discussions for the teacher moves exhibited by the two fourth-grade teachers as well as students’ elaborated explanations. We also coded turns that did not represent either a teacher move or an elaborated explanation as participation. As shown in Table 4-1, the fourth-grade students produced the majority of turns during discussions over the course of the year. Consequently, the fourth-grade teachers contributed only a small percentage of the total turns across the 54 discussions. Notably, however, the majority of teachers’ turns represented one of the ten teacher moves presented by Wei et al. (2018).

The ten teacher moves were divided into two categories (see Table 4-1) for the analysis. The first category of teacher moves consisted of the five scaffolding moves likely to directly induce students’ productive talk, while the second category of teacher moves consisted of the five-discourse management moves likely to indirectly induce students’ productive talk (Murphy, Firetto, et al., 2017). Scaffolding moves included challenging, marking, modeling, prompting, and summarizing; discourse management moves included checking, clarifying, instructing, procedural, and reading. Over the course of the intervention the two fourth-grade teachers employed more discourse management moves than scaffolding moves. Within group frequencies indicate that the teachers employed more discourse management moves with the homogenous below-average ability group than they did while facilitating other discussion groups (see Table 4-1). There did not, however, seem to be an appreciable amount of variance in the proportion of total turns teachers’ management moves accounted for across the six discussion groups.
The two fourth-grade teachers employed 247 scaffolding moves over the course of the 54 discussions included in the analysis, accounting for 19.2% of their total turns (see Table 4-2). Across the six discussion groups, the teachers employed 41.167 scaffolding moves on average, with a range of 23 scaffolding moves with the homogeneous above-average ability group to 48 scaffolding moves with a heterogeneous ability group (Mdn = 45.5). The teachers utilized prompting more frequently than the other four scaffolding moves, accounting for 11.1% of their total scaffolding moves, followed by marking, challenging, modeling, and summarizing respectively.

To compare teacher’s use of each individual (i.e., prompting, marking, challenging, modeling, and summarizing) scaffolding move across groups, within group I calculated the percentage of total scaffolding moves accounted for by each individual move. Prompting, for example, accounted for a higher percentage of the scaffoldings moves employed during the homogeneous below-average ability group’s discussions (37%) than during the homogeneous above-average ability group’s discussions (30%). In contrast, across groups, marking accounted for the greatest percentage of teachers’ scaffolding moves in the homogeneous above-average ability group (34%) and the lowest percentage of teachers’ scaffolding moves in the homogeneous below-average ability group (15%).

The teachers challenged students a total of 36 times, accounting for 4% of their overall total scaffolding moves (see Table 4-2). Across the six discussion groups, challenging accounted for the highest percentage of scaffolding moves exhibited by the two fourth-grade teachers within the homogeneous above-average group (22%) and the lowest percentage (10%) of scaffolding moves within one of the heterogeneous ability groups. Teachers’ use of modeling was fairly limited, ranging from three instances with the homogeneous above-average and homogeneous below-average ability groups respectively to seven instances with the homogeneous average and a heterogeneous ability groups. With this in mind, within group, marking accounted for 15% of
teachers scaffolding moves utilized with the homogeneous below-average ability group compared to 34% of the scaffolding moves utilized with the homogeneous above-average ability group. Finally, the fourth-grade teachers employed summarizing the least frequently of the five scaffolding moves. Summarizing, however, accounted for approximately 22% of the scaffolding moves employed by the fourth-grade teachers in the homogeneous below-average ability group which was notably more than the remaining five fourth-grade discussion groups (see Table 4-2).

The fourth-grade students produced 415 elaborated explanations over the course of the intervention (see Table 4-1). In particular, students in the homogeneous average group produced the highest frequency of elaborated explanations, followed by two of the heterogeneous groups. The homogenous below-average and homogenous above-average groups produced a similar number of elaborated explanations. Importantly, however, Murphy, Greene, et al. (2017) reported that the elaborated explanations generated by the above-average ability and below-average ability students differed in duration as well as overall quality. Specifically, students in the homogenous above-average group offered longer and more complex elaborated explanations as they progressed through the intervention. Murphy, Greene, et al. noted that this pattern was mirrored in heterogeneous groups discourse as well. By comparison, students in the homogenous below-average group did not demonstrate the same level growth overtime. Below-average ability students, instead, produced elaborated explanations representative of the early stages of Quality Talk throughout the entire school year.
Table 4-1: Frequency of type of turns by group composition in fourth grade.

<table>
<thead>
<tr>
<th>Type of Turn</th>
<th>Homogeneous Ability Groups</th>
<th>Heterogeneous Ability Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Above Average</td>
<td>Average</td>
</tr>
<tr>
<td></td>
<td>% of Turn</td>
<td>% of Turn</td>
</tr>
<tr>
<td>Teacher</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffolding Moves</td>
<td>23</td>
<td>46</td>
</tr>
<tr>
<td>Management Moves</td>
<td>92</td>
<td>12</td>
</tr>
<tr>
<td>Participation</td>
<td>79</td>
<td>54</td>
</tr>
<tr>
<td>Total</td>
<td>194</td>
<td>212</td>
</tr>
<tr>
<td>Student</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elaborated Explanations</td>
<td>68</td>
<td>89</td>
</tr>
<tr>
<td>Participation</td>
<td>866</td>
<td>566</td>
</tr>
<tr>
<td>Total</td>
<td>934</td>
<td>655</td>
</tr>
<tr>
<td>Grand Total</td>
<td>1128</td>
<td>867</td>
</tr>
</tbody>
</table>

Table 4-2: Frequency of scaffolding moves by group composition in fourth grade.

<table>
<thead>
<tr>
<th>Type of Scaffolding Move</th>
<th>Homogeneous Ability Groups</th>
<th>Heterogeneous Ability Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Above Average</td>
<td>Average</td>
</tr>
<tr>
<td>Challenging</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Marking</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>Modeling</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Prompting</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>Summarizing</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>46</td>
</tr>
</tbody>
</table>
**Relationship between teacher moves and group composition**

My first research question pertained to the relationship between teachers’ facilitation practices and group composition. To examine the relationship between category of teacher move and group composition, I utilized two chi-square tests of independence. First, I conducted a chi-square test of independence between teacher move category and the degree of heterogeneity in group composition to determine if frequency of scaffolding and management moves employed by the teachers were distributed differently for homogeneous and heterogeneous discussion groups. Both variables were considered dichotomously. Category of teacher move had two levels (i.e., scaffolding moves and management moves) and degree of heterogeneity in group composition had two levels as well (i.e., homogeneous and heterogeneous composition). As such, the frequency of scaffolding and management moves were collapsed within overall group composition type. All expected cell frequencies were greater than five.

Results suggested that the observed frequency of management moves across the three heterogeneous discussion groups was lower than expected but higher than expected in the homogeneous discussion groups. In contrast, the observed frequency of scaffolding moves was higher than expected in the three heterogeneous groups and lower than expected in the homogeneous discussion groups. There was a statistically significant association between category of teacher move and group composition, $\chi^2 (1) = 7.251$, $p = .007$ (see Table 4-3). The association between the two variables, however, was very small (Cohen, 1988), $\varphi = -.090$.

Adjusted standardized residuals were used to further characterize the nature of the association between category of teacher move and degree of heterogeneity in group composition. Positive standardized residuals indicate that there are more observed events than expected for a given cell; negative adjusted standardized residual indicates that there were fewer observed
events than expected. The larger the absolute value of the adjusted standardized residual the more a cell can be considered to be contributing to the overall chi-square value. Adjusted standardized residuals with absolute values exceeding two are considered deviating significantly from the null hypothesis of a particular cell (Agresti, 2019). All cells were associated with meaningful adjusted standardized residuals indicating that the fourth-grade teachers employed fewer scaffolding moves than expected while facilitating discussions with homogeneous groups and used fewer discourse management moves than expected while facilitating heterogeneous groups discussions (see Table 4-3). The adjusted standardized residuals indicated, however, that each cell deviated from the null hypothesis equally. No cell contributed more than another to the associated between category of teacher move and degree of heterogeneity in group composition.

Next, I utilized a second chi-square test of independence to between category of teacher move and category of homogeneous group composition to determine if the frequency of scaffolding and management moves were distributed differently across the three homogeneous discussion groups. Category of teacher move was treated as a dichotomous variable (i.e., scaffolding moves and management moves); category of homogenous group composition had three levels (i.e., below-average, average, and above-average). As such, the frequency of scaffolding moves and management moves were not collapsed for this analysis, but representative of what students participating in the three individual homogenous groups were exposed to throughout the intervention. All expected cell frequencies were greater than five.

<table>
<thead>
<tr>
<th>Teacher Move Category</th>
<th>Group Composition</th>
<th>( X^2 )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Homogeneous</td>
<td>Heterogenous</td>
<td></td>
</tr>
<tr>
<td>Scaffolding Moves</td>
<td>114</td>
<td>133</td>
<td>7.251*</td>
</tr>
<tr>
<td></td>
<td>(-2.7)</td>
<td>(2.7)</td>
<td></td>
</tr>
<tr>
<td>Management Moves</td>
<td>368</td>
<td>287</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.7)</td>
<td>(-2.7)</td>
<td></td>
</tr>
</tbody>
</table>

Note. * = \( p < .05 \). Adjusted standardized residuals appear in parentheses below group frequencies.
Observed frequencies suggested that teachers employed slightly more management moves in the homogeneous below-average and above-average groups but employed slightly fewer scaffolding moves. This pattern is consistent with the findings of the previous analysis of the association between category of teacher move and degree of heterogeneity in group composition. The observed frequencies for the homogeneous average group, however, deviated from this pattern. Observed frequencies indicated that teachers employed fewer discourse management moves and more scaffolding moves while facilitating the homogeneous average group’s discussions. There was not, however, a statistically significant association between category of teacher move and category of homogeneous group composition, \( \chi^2 (2) = 4.136, p = .126 \) (see Table 4-4). These findings suggest that while the two teachers used fewer scaffolding moves and more discourse management moves when facilitating homogenous groups, their use of those moves was not different than expected.

Table 4-4: Crosstabulation of teacher move category and ability level in fourth grade.

<table>
<thead>
<tr>
<th>Teacher Move Category</th>
<th>Ability Level</th>
<th>( \chi^2 )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Above Average</td>
<td>Average</td>
<td>Below Average</td>
</tr>
<tr>
<td>Scaffolding Moves</td>
<td>23</td>
<td>46</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>(-1.1)</td>
<td>(2.0)</td>
<td>(-1.0)</td>
</tr>
<tr>
<td>Management Moves</td>
<td>92</td>
<td>111</td>
<td>164</td>
</tr>
<tr>
<td></td>
<td>(1.1)</td>
<td>(-2.0)</td>
<td>(1.0)</td>
</tr>
</tbody>
</table>

**Influence of teachers’ scaffolding moves on students’ productive talk**

My second research question pertained to the sequential, turn-by-turn effect of teachers’ facilitation practices on students’ productive talk. Turns were operationalized as segments of time during which a teacher or student held the floor of the discussion (Murphy, Firetto, et al., 2017). Every turn could represent either teachers’ facilitation practices, students’ productive talk, or participation. Teacher’s facilitation practices were operationalized as the five scaffolding
moves (i.e., Challenging, marking, modeling, prompting, and summarizing); student's productive talk was operationalized as elaborated explanations. Turns consisting of teacher or student talk not representative of a scaffolding move or an elaborated explanation were considered participation. I utilized a lag-sequential analysis (Bakeman & Gottman, 1997) to analysis the sequential, turn-by-turn patterns in the discourse and address my second research question.

Lag-sequential analysis compares the probability of a random sequence of events to that of a systematic sequence of events. Consider, for example, that a teacher prompts a student to elaborate on something they said (e.g., Why do you think that?) and the student responds with an elaborated explanation. Lag-sequential analysis assesses the transitional probability between the criterion event (e.g., the teacher’s prompt) and the target event (e.g., the student’s elaborated explanation). A statistically significant transitional probability provides evidence that the criterion event increases the likelihood of the target event. In terms of our example, that a prompt increases the likelihood that a student will provide an elaborated explanation.

Lag-sequential analysis requires a large number of observations (i.e., turns) be represented in the data. Specifically, the number of turns included in the analysis must exceed 5KL+1. Where K represents the number of codes (i.e., scaffolding moves, elaborated explanations, and participation) in the data and L represents the number of lags examined. Lags refers to the number of turns between the criterion event and the target event. Our previous example, in which the target event occurred immediately after the criterion event, evaluates the transitional probability at Lag 1. Including two lags in the analysis would allow you to investigate the transitional probability of the criterion event and the target event with one intervening event between them. To ensure a sufficient number of turns, only the most frequently used scaffolding moves (i.e., challenging, marking, and prompting) at Lag 1 were included in the analysis for fourth grade. As such, at least 125 turns (5x51+1 = 5x52) were needed. This criterion is met when discussions are pooled within group.
Assumptions

The data must meet three assumptions in order to meaningfully interpret the results of a lag-sequential analysis: dependence, stationarity, and homogeneity. Lag-sequential analysis is predicated on the idea that observations are, by nature, dependent (Bakeman & Gottman, 1997). Therefore, unlike most statistical analysis which relay on independent observations, lag-sequential analysis requires dependence in the sequential observations in order to examine sequential patterns. While all other groups met the dependence assumption, the homogeneous above average group failed to meet the dependence assumption and was therefore, dropped from further analysis (O’Connor, 1999).

The null hypothesis of the dependence assumption states that there is no difference between the conditional probability of the target event (i.e., elaborated explanations) given the criterion event (i.e., scaffolding move) and the unconditional probability of the target event (Allison & Liker, 1982). The dependence assumption is a comparison between the proportion of times elaborated explanations and specific scaffolding moves occur together and the proportion of instances they occur independently. As such, the lack of dependence in the homogeneous above-average group suggests that students were just as likely to produce elaborated explanations spontaneously as they were to produce them in response to any initiating event. Furthermore, teachers were also just as likely to employ scaffolding moves in response to elaborated explanations as they were to employ scaffolding moves independently of those events.

Using the SEQGROUPS program (O’Connor, 1999), I assessed the degree to which the data from the homogeneous average group, the homogeneous below-average group, and the three heterogenous groups met the stationarity assumption. Stationarity refers to the stability of the sequential pattern within group across time. In other words, stationarity represents the degree to which the sequential patterns in the data during the first discussion continued throughout the
intervention within group. No statistically significant results were found indicating that the sequential patterns within each group were stable across time (see Table 4-5). This suggests that within group, students responded to prompting, challenging, and marking the same way at the beginning of the year as they did at the end of the year and that teachers responded the same way to their students across time as well. Given the stability of the sequential patterns, I pooled the sequential data within group for the remainder of the analysis.

Table 4-5: Likelihood ratio tests for nonstationarity, independence, and heterogeneity in fourth grade

<table>
<thead>
<tr>
<th></th>
<th>Nonstationarity</th>
<th>Independence</th>
<th>Heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LRx  df  p</td>
<td>LRx  df  p</td>
<td>LRx  df  p</td>
</tr>
<tr>
<td><strong>Homogenous Groups</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above Average</td>
<td>71.37 160 1.000</td>
<td>14.14 16 .5886</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>95.35 160 1.000</td>
<td>33.05 16 .0073</td>
<td></td>
</tr>
<tr>
<td>Below Average</td>
<td>106.22 160 .9997</td>
<td>49.09 16 .0000</td>
<td></td>
</tr>
<tr>
<td><strong>Heterogenous Groups</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>67.66 160 1.000</td>
<td>31.66 16 .0111</td>
<td>36.37 40 .6330</td>
</tr>
<tr>
<td>Group 2</td>
<td>78.91 160 1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 3</td>
<td>118.90 160 .9937</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Finally, I assessed the degree to which the three heterogenous groups met the homogeneity assumption. While stationarity refers to the stability of sequential patterns across time, homogeneity refers to the stability of sequential patterns across groups. Homogeneity of sequential patterns indicates that the sequential patterns are similar across groups; heterogeneity of the sequential patterns indicates that the sequential patterns are different across groups. The likelihood ratio chi-square test for homogeneity was not statistically significant, indicating that the sequential patterns were stable across the three heterogeneous groups $X^2 (40) = 36.37, p = .633$ (see Table 4-5). Therefore, I pooled the data across the heterogeneous groups’ discussions allowing me to compare the heterogenous groups as an overall group type to the homogeneous average group and homogeneous below-average group. Interestingly, the homogenous below
average and average groups did not meet the assumption of homogeneity suggesting that the sequential patterns were inconsistent across these groups. In the following sections, I present the results of the lag-sequential analysis for the pooled heterogeneous groups followed by the results of the lag-sequential analysis for the homogenous average and homogenous below average groups respectfully.

**Fourth-grade heterogeneous ability groups**

One teacher scaffolding move was associated with students’ productive talk at Lag 1 in the heterogeneous ability groups. Prompting predicted students’ elaborated explanations during the heterogeneous ability group’s discussions (see Figure 4-1). When a teacher prompted a student, there was a 20.0% probability (p < .001) that the student would respond with an elaborated explanation in the adjacent turn. In the following discourse example students were discussing *Lost City: The Discovery of Machu Picchu* (Lewin, 2013; see Figure 4-1). In turn 5, a student claims that they would not have wanted to journey to Machu Picchu and offers one reason to why they would not have wanted to go (i.e., I do not like hiking). The teacher prompts the student for evidence (i.e., how do you know that?) of the student not liking to hike in the following turn and the student replies with an additional reasons and evidence about a time they went hiking in Shenandoah National Park (i.e., My feet are not meant to hike [reason]. I hiked up Old Rag and I didn’t like it [evidence]). Marking and challenging were not statistically associated with students’ productive talk at Lag 1 during the heterogeneous group’s discussions, indicating that students were not more likely to generate elaborated explanations in the turn following a teacher’s use of marking or challenging. Furthermore, students’ productive talk did not increase the likelihood of teachers employing marking or challenging in the following turn.
The analysis indicated one additional meaningful sequence in the heterogenous groups’ talk. Specifically, the findings suggest that students’ elaborated explanations snowballed at Lag 1 (Anderson et al., 2001). When a student produced an elaborated explanation, there was 11.8% probability (p < .001) that another student would generate an elaborated explanation in the following turn. For example, as shown in the excerpt in Figure 4-2, students were discussing if they would make the same decision as the main character in *Cliff Hanger* (George & Minor, 2013), a story about a boy who rescues his dog during a thunderstorm. In turn 4, a student claims that she would do the same thing as the character and supports that claim with two conjunctive reasons (i.e., the lighting during a storm could come down further [reason 1] and you could just be electrocuted [reason 2]). A second student builds on the same claim in the following turn by adding that lighting strikes high places not lower places [reason 1] and therefore, they would jump to Monkey Ledge to stay safe [reason 2]. This sequence of events is consistent with Anderson and colleagues (2001) finding that students’ often replicate each other’s discursive behaviors during small group discussions. Similarly, Jadallah et al. (2011) found that teachers could initiate the snowball phenomenon during Collaborative Reasoning discussions by praising a student’s use of evidence. While Anderson and colleagues (2001) and Jadallah et al.’s (2011) findings help to guild an interpretation of students’ elaborated explanations snowballing, I can
only hypothesize that the teacher may be able to elicit this phenomenon with prompting given that the analysis was limited to Lag 1.

Figure 4-2: Discourse example from a fourth-grade heterogeneous ability group of an elaborated explanation eliciting an elaborated explanation.

Figure 4-1: Discourse example from a fourth-grade heterogeneous ability group of prompting eliciting an elaborated explanation.
Fourth-grade homogeneous average ability group

Three statistically significant sequential patterns at Lag 1 emerged from the homogenous average ability group’s discussions. Similar to the heterogenous groups in fourth grade, teacher prompting had an immediate effect on students’ productive talk. There was a 31.2% probability (p = .006) that students would respond to a teachers’ prompts with an elaborated explanation at Lag 1. See Figure 4-3 for a sample of the sequential relationship between prompting and elaborated explanations in the homogeneous average group during the fourth-grade homogeneous average ability group’s discussion of Adelina’s Whales (Sobol, 2013). The second statistically significant sequential pattern concerned teachers’ use of marking in response to students’ productive talk. Results indicated that students’ productive talk increased the likelihood of teachers utilizing marking. When students generated an elaborated explanation, there was 4.5% probability (p =.027) that the teacher would mark the student’s elaborated explanations.
Figure 4-4 shows an excerpt from the homogeneous average ability group’s discussion of *Cliff Hanger* (George & Minor, 2013). Instead of discussing if they would jump from one hiking ledge to another, however, they talk about whether they would disobey their father to rescue their dog like the main character, Axle, did in the story. In her response in turn 6, a student supports her claim that she would disobey her father with two conjunctive reasons and a piece of evidence and the teacher explicitly praises her use of evidence in the adjacent turn. In contrast, marking was not associated with an increased probability of elaborated explanations in the following turn, suggesting that while teachers were more likely to mark elaborated explanations, marking did not stimulate further generation of elaborated explanations at Lag 1. There were also no statistically significant sequential patterns associated with challenging in the homogeneous average group.

The third statistically significant transitional probability in the homogeneous average group’s discussion was representative of the snowball phenomena. As seen in Figure 4-5, there
was a 22.4% probability (p > .001) that an elaborated explanation would lead to another elaborated explanation indicated that the snowball phenomenon was also present in the homogeneous average group. The following discourse example is from the homogeneous average group’s discussion of *Lost City: The Discovery of Machu Picchu* (Lewin, 2013) during which they discuss if they would like to go on a journey similar to the one in the story. In turn 8, one student claims that he would like to go on a similar journey because it is a once in a life time opportunity [reason 1] and he would be grateful to achieve one of his goals, particularly as a young boy [reason 2]. Another student challenges this idea in the following turn by saying that she would not be excited [claim] because there could be venomous snakes [reason 1] and they could bite you [reason 2].
Fourth-grade homogeneous below average ability group

All three teacher scaffolding moves including in the analysis (i.e., prompting, challenging, and marking) were associated with meaningful sequential patterns at Lag 1 in the homogeneous below average ability group. Parallel to the heterogenous groups and the homogeneous average group, findings indicated a statistically significant transitional probability between teachers’ prompting and students’ elaborated explanations (52.9%, p< .001). In Figure 4-6, students in the homogeneous below average ability group were discussing *Lost City: The Discovery of Machu Picchu* (Lewin, 2013). In turn 7, the teacher prompts a student to expand on why they speculated that a box the main character carried would contain gold coins. In his response, the student elaborates to say that people keep rare [reason 1] or valuable stuff [reason 2] in boxes.

![Discourse example from the fourth-grade homogenous below average ability group of prompting eliciting an elaborated explanation.](image)
Figure 4-7: Discourse example from the fourth-grade homogenous below average ability group of an elaborated explanation eliciting marking.
In contrast to the sequential nature of students responding to teachers’ prompting with elaborated explanations, the remaining two statistically significant sequential patterns involved teachers’ responses to students’ elaborated explanations, suggesting that teacher scaffolding moves and student elaborated explanations may share a bidirectional relationship. Elaborated explanations increased the likelihood that teachers would employ both marking and challenging in the homogeneous below average group. There was a 4.8% probability (p < .001) that the teacher would mark an elaborated explanation and a 3.2% probability (p = .015) that the teacher would challenge an elaborated explanation. See Figure 4-7 for a sample of a teacher marking an elaborated explanation during the homogeneous below average group’s discussions.

Despite the fact that challenging was not associated with any statistically significant transitional probabilities in the other fourth-grade groups, the student’s elaborated explanations did elicit challenges from the teachers during the homogeneous below average group’s discussions. In the following discourse example students were discussing Horse Heroes: True Stories of Amazing Horses (Petty, 2013), an anthology of stories about horse that had an impact on history. Specifically, students were talking about why characters in one of the stories may have decided to hide some of the perils associated with the pony express. In turn 6, a student proposes that hiding the perils was a good idea [claim] because then people would be less worried [reason] and provides evidence from the story an event that would have made characters worry [evidence]. The teacher then challenges the student’s reasoning in the adjacent turn by saying that it was not a good idea because people would not know how to keep themselves safe if they did not know what had happened (see Figure 4-8).
In contrast to the three heterogeneous groups and the homogeneous average group, there was no evidence of student’s elaborated explanations snowballing in the homogeneous below average group’s discussions. A possible explanation for the lack of snowballing is that students in the homogeneous below average group engaged in more parallel talk than their peers in other groups. Parallel talk is characteristic of talk episodes where students do not engage with each other’s ideas or build on what has been said but respond to questions in-turn and individually instead. There is little collaboration between group members during parallel talk. A lack of collaboration may also explain why there was a meaningful association between teachers’ challenges and elaborated explanations in the homogeneous below average group. In other groups, where students were engaged with each other’s ideas, students took on the challenging role as seen in Figure 4-5. It may be then, that the teachers were trying to model how to engage with each other’s ideas and challenge each other, albeit not explicitly as defined in Wei et al. (2018).

Figure 4-8: Discourse example from the fourth-grade homogenous below average ability group of an elaborated explanation eliciting challenging.
Influence of teachers’ scaffolding moves across group composition

My third research question pertained to the extent to which the strength of the sequential relationships between teachers’ scaffolding moves and students’ productive talk differ as a function of group composition. The transitional probabilities reported in the previous section represent the conditional probability of a target event (e.g., elaborated explanation) occurring given the occurrence of a criterion event (e.g., teacher scaffolding move) within grouping type. Transitional probabilities should not, however, be used to index the relative strength of sequential patterns across discussion groups because transition probabilities are influenced by the frequency at which meaningful utterances occur. Even if the strength of the association between a criterion event and a target event stays the same, the transitional probability would increase if the frequency of meaningful utterances doubled. As such, transitional probabilities provide a biased representation of the differences in sequential relationships across discussion groups unless the base rate of the meaningful utterances are the same.

The current recommended practice, therefore, is to provide and compare measures of association that are not influenced by the number of data points in a sequence such as Yule’s Q (Bakeman & Gottman, 1997). Yule’s Q is an index of the strength of an association between pairs of meaningful utterances with values that range from -1 to 1 indicating the strength and direction of the association. The interpretation of Yule’s Q is analogous to the Pearson correlation coefficient; positive values indicate that the criterion event increases the likelihood of the target event, negative values indicate that the criterion event decreases the likelihood of the target event, and the absolute value of the coefficient indicates the strength of the relationship with zero indicating no association between the criterion event and the target event.

As shown in Table 4–6, the sequential patterns between teachers’ scaffolding moves and students’ productive talk differed in magnitude and direction across groups. Prompting was the
only scaffolding move associated with a statistically significant transitional probability in all three
group types that met the assumptions for the analysis. Furthermore, the association between
prompting and elaborated explanations was strong and positive across groups suggesting that
prompting increases the likelihood of productive talk regardless of group composition. Although
the association was strong across the three groups types in the analysis, it was strongest in
homogeneous below average group (see Table 4-6).

Table 4-6: Yule’s Q associated with sequential patterns across group composition in fourth grade

<table>
<thead>
<tr>
<th>Sequential Pattern</th>
<th>Homogenous Groups</th>
<th></th>
<th>Pooled Heterogenous Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>Below Average</td>
<td></td>
</tr>
<tr>
<td>Teacher Initiated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prompting x Elaborated Explanation</td>
<td>.602*</td>
<td>.915*</td>
<td>.557*</td>
</tr>
<tr>
<td>Student Initiated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elaborated Explanations x Marking</td>
<td>.556*</td>
<td>.863*</td>
<td>-.087</td>
</tr>
<tr>
<td>Elaborated Explanation x Challenging</td>
<td>-1.00</td>
<td>.712*</td>
<td>-1.00</td>
</tr>
<tr>
<td>Elaborated Explanation x Elaborated Explanation</td>
<td>.484*</td>
<td>.212</td>
<td>.312*</td>
</tr>
</tbody>
</table>

Note. * indicates statistically significant transitional probability.

My analysis of the sequential patterns during small-group discussions also indicated three
unexpected relationships where students’ productive talk served as the criterion event and
teachers’ scaffolding moves served as the target event. These sequential patterns were not,
however, associated with statistically significant transitional probabilities across all group
compositions. First, elaborated explanations increased the likelihood of marking in both
homogeneous groups, but not the heterogeneous groups. Results indicated that like prompting,
the association between elaborated explanations and marking was strongest in the homogeneous
below-average group. In contrast, the association between elaborated explanations and marking
was close to zero and negative in the heterogeneous groups, however, this sequential pattern was
not statistically significant.
Findings also indicated a strong, positive sequential pattern between elaborated explanations and challenging in the homogeneous below-average group that was not present in the other two group compositions. In fact, there was a very strong, negative association between students’ productive talk and challenging in the homogeneous average and heterogeneous groups indicating that the teachers never challenged a student’s elaborated explanation at Lag 1 in these groups over the course of the 54 discussions included in the analysis. Given that transitional probabilities are affected by the base rate of events as well as the frequency of the events cooccurring, the lack of statistical significance may be due to the lack of transitional frequency between these codes in both groups. It does seem meaningful, however, that a strong, positive association in one group is completely absent from others.

The final sequential pattern represented in the data was the snowballing of students’ productive talk which was present in the homogeneous average and heterogeneous groups. There was a moderate positive association between elaborated explanations in both the homogeneous average and heterogeneous groups. While it was not associated with a statistically significant transitional probability, there was a small positive relationship between elaborated explanations in the homogeneous below-average group as well.

Fifth-Grade Results

Descriptive statistics

A total of 5,206 meaningful utterances were analyzed across 54 discussions in fifth grade. Table 4-7 displays the frequency of meaningful utterances teachers generated over the course of the intervention by category of teacher move and participation and the frequency of meaningful utterances students generated over the course of the intervention by elaborated explanations and
participation. Similar to fourth grade, fifth-grade students generated the majority of meaningful utterances and the majority of talk generated by teachers (70.2%) represented a teacher move from Wei et al. (2018). Teacher talk, however, accounted for a smaller percentage (12.9%) of turns in fifth grade than in fourth grade (21.2%). The fifth-grade teachers also generated less than half the number of teacher moves that the fourth-grade teachers generated even though the overall frequency of meaningful utterances across grades was moderately similar.

The fifth-grade teachers employed more management moves than scaffolding moves overall. There appeared to be some variance in the frequency by which teachers employed scaffolding moves across groups (see Table 4-7). As displayed in Table 4-7, teachers’ scaffolding moves accounted for approximately the same proportion of meaningful utterance across the three homogeneous groups but varied across the three heterogeneous groups. There appeared to be less variance across group composition, however, in the frequency at which teacher employed management moves. Teachers’ management moves accounted for the smallest proportion of meaningful utterances in the homogeneous above-average group and one of the heterogeneous groups and the greatest number of meaningful utterances in another heterogeneous group, though the proportion was similar to the homogeneous average group (see Table 4-7).

A different frequency pattern emerged within scaffolding moves in fifth grade then fourth grade. Table 4-8 displays the frequency at which the fifth-grade teachers employ individual scaffolding moves by discussion group. While prompting, marking, and challenging were the most frequently employed moves in fourth grade, the fifth-grade teachers employed marking most often, followed by prompting, modeling, challenging, and summarizing respectively (see Table 4-8). Interestingly, the three most frequently used scaffolding moves used in each grade account for a similar percentage of teacher moves employed over the course of the intervention. Prompting, marking, and challenging accounted for 22.1% of teacher moves employed in fourth
grade while prompting, marking, and modeling accounted for 22.6% of teacher moves employed the fifth grade.

Fifth-grade students produced 647 elaborated explanations. The homogeneous-above average group produced more elaborated explanations than any other group while the homogenous below average group produced the fewest (see Table 4-7). It should be noted however, that the homogenous below-average group produced a similar overall number of elaborated explanations to the remaining four discussion groups irrespective of group composition. Moreover, this is a different pattern than fourth-grade where the homogeneous average group produced the highest frequency of elaborated explanations and the homogenous above average and homogenous below average produced the same number. Furthermore, elaborated explanations account for a higher percentage of student talk in fifth-grade (i.e., approximately 14%) than in fourth-grade (i.e., approximately 9%). In the following sections I present the findings from the fifth-grade data for my three research questions.
Table 4-7: Frequency of type of turns by group composition in fifth grade.

<table>
<thead>
<tr>
<th>Type of Turn</th>
<th>Homogeneous Ability Groups</th>
<th>Heterogeneous Ability Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Above Average</td>
<td>Average</td>
</tr>
<tr>
<td></td>
<td>Turns</td>
<td>% of GT</td>
</tr>
<tr>
<td><strong>Teacher</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffolding Moves</td>
<td>19</td>
<td>2.2</td>
</tr>
<tr>
<td>Management Moves</td>
<td>51</td>
<td>5.9</td>
</tr>
<tr>
<td>Participation</td>
<td>45</td>
<td>5.2</td>
</tr>
<tr>
<td>Total</td>
<td>115</td>
<td>13.3</td>
</tr>
<tr>
<td><strong>Student</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elaborated Explanations</td>
<td>124</td>
<td>14.4</td>
</tr>
<tr>
<td>Participation</td>
<td>623</td>
<td>72.3</td>
</tr>
<tr>
<td>Total</td>
<td>747</td>
<td>86.7</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td>862</td>
<td></td>
</tr>
</tbody>
</table>

Table 4-8: Frequency of scaffolding moves by group composition in fifth grade.

<table>
<thead>
<tr>
<th>Type of Scaffolding Move</th>
<th>Homogeneous Ability Groups</th>
<th>Heterogeneous Ability Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Above Average</td>
<td>Average</td>
</tr>
<tr>
<td></td>
<td>Turns</td>
<td>% of GT</td>
</tr>
<tr>
<td><strong>Challenging</strong></td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td><strong>Marking</strong></td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td><strong>Modeling</strong></td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td><strong>Prompting</strong></td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td><strong>Summarizing</strong></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>19</td>
<td>20</td>
</tr>
</tbody>
</table>
Relationship between teacher moves and group composition

To fully address my first research question pertaining to the relationship between teachers’ facilitation practices and group composition, I mirrored the analysis I performed of the fourth-grade data in fifth-grade. I conducted a chi-square test for independence between teacher move category and degree of heterogeneity in group composition. All expected cell frequencies were greater than five. There was not a statistically significant association between teacher move category and degree of heterogeneity in group composition in fifth grade, $X^2 (1) = .435, p = .526$ (see Table 4-9). Results suggest that the observed frequency of scaffolding and management moves employed by fifth-grade teachers was not different than expected when considering the homogeneous groups and heterogenous groups overall. This finding is different than fourth grade where teachers employed fewer management moves than expected with heterogenous groups and more scaffolding moves than expected with homogeneous groups.

Table 4-9: Crosstabulation of teacher move category and group composition in fifth grade.

<table>
<thead>
<tr>
<th>Teacher Move Category</th>
<th>Group Composition</th>
<th>$X^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Homogeneous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffolding Moves</td>
<td>54</td>
<td></td>
<td>.435</td>
</tr>
<tr>
<td></td>
<td>(-0.7)</td>
<td></td>
<td>.526</td>
</tr>
<tr>
<td>Management Moves</td>
<td>173</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogenous</td>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.7)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A chi-square test of independence was also conducted to between category of teacher move and category of homogenous group composition in fifth grade. All expected cell frequencies were greater than five. Results indicated no statistically significant association between the frequency of scaffolding or management moves were employed and type of homogeneous group composition, $X^2 (2) = .474, p = .789$ (see Table 4-10). This finding mirrors
the fourth grade where teachers observed use of scaffolding and management moves across the three homogeneous groups was not different than expected.

Table 4-10: Crosstabulation of teacher move category and ability level in fifth grade.

<table>
<thead>
<tr>
<th>Teacher Move Category</th>
<th>Ability Level</th>
<th>$\chi^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Above Average</td>
<td>Average</td>
<td>Below Average</td>
</tr>
<tr>
<td>Scaffolding Moves</td>
<td>19</td>
<td>20</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>(0.2)</td>
<td>(-0.7)</td>
<td>(-0.5)</td>
</tr>
<tr>
<td>Management Moves</td>
<td>51</td>
<td>64</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>(-0.2)</td>
<td>(0.7)</td>
<td>(0.5)</td>
</tr>
</tbody>
</table>

Influence of teachers’ scaffolding moves on students’ productive talk

I also investigated the turn-by-turn effect of teachers’ facilitation practices on students’ productive talk in fifth grade. The fifth-grade analysis was limited to marking, modeling, and prompting because they were the three more frequently employed scaffolding moves in fifth grade. Despite narrowing the analysis to the most frequently used scaffolding moves, the decreased base rate of scaffolding moves in fifth grade created instability in the sequential patterns. Lag-sequential analysis requires at least five data points of each meaningful utterance to ensure stability in the sequential pattern and accurate interpretation of transitional probabilities (Bakeman & Gottman, 1997). This requirement was met when data was pooled across the heterogenous groups but was not met in the homogeneous above-average, average, or below-average groups. Therefore, I limited the analysis to the scaffolding moves that exceeded five instances when the data was pooled within group for the three homogeneous discussion groups. As a result, marking was the only scaffolding move represented in the sequential data from the homogeneous above-average discussion group and the homogeneous average group; marking and prompting were represented in the sequential data from the homogeneous below-average group.
Assumptions

The sequential data from fifth grade was also assessed for dependence, stationarity, and homogeneity. As displayed in Table 4-11, the pooled heterogeneous groups and the homogeneous average group failed to meet dependence. This suggests that like the fourth-grade homogenous above-average group there was no statistically dependent sequential structure in the turn-by-turn meaningful utterances generated during discussion. Thus, even though the data are theoretically dependent by nature, the statistical significance of individual sequential relationships should not be examined (O’Conner, 1999). For the homogeneous above-average and the homogeneous below-average groups, the likelihood ratio chi-square test for dependence was significant (see Table 4-11).

Table 4-11: Likelihood ratio tests for nonstationarity, independence, and heterogeneity in fifth grade

<table>
<thead>
<tr>
<th></th>
<th>Nonstationarity</th>
<th>Independence</th>
<th>Heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LRx</td>
<td>df</td>
<td>p</td>
</tr>
<tr>
<td><strong>Homogenous Groups</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above Average</td>
<td>72.16</td>
<td>48</td>
<td>.014</td>
</tr>
<tr>
<td>Average</td>
<td>46.31</td>
<td>48</td>
<td>.542</td>
</tr>
<tr>
<td>Below Average</td>
<td>85.13</td>
<td>96</td>
<td>.779</td>
</tr>
<tr>
<td><strong>Heterogenous Groups</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>63.25</td>
<td>160</td>
<td>1.00</td>
</tr>
<tr>
<td>Group 2</td>
<td>64.65</td>
<td>160</td>
<td>1.00</td>
</tr>
<tr>
<td>Group 3</td>
<td>136.22</td>
<td>160</td>
<td>1.00</td>
</tr>
</tbody>
</table>

I assessed both groups for stationarity. No statistically significant results were found for the homogeneous below-average group, $X^2 (96) = 85.1324$, $p = .7786$, suggesting that the sequential patterns were stable across time and that data from their nine discussions could be pooled together (see Table 4-11). The homogenous above-average group, however, violated the stationarity assumption, $X^2 (48) = 72.1591$, $p = .0136$. Thus, the data from their nine discussions could not be pooled. Unfortunately, being unable to pool the data from their nine discussions
created two additional problems for interpretation. Individually the nine discussions do not exceed the threshold of number of turns needed for the analysis. Furthermore, there was not five instances of marking within one discussion. As such, the analysis was limited to the homogeneous below-average group in fifth-grade.

**Fifth-grade homogeneous below average ability group**

Four types of turns were represented in the homogeneous below-average ability group’s sequential data (i.e., elaborated explanations, marking, prompting, and a neutral participation code). Modeling was dropped from the analyses due to an inadequate frequency of instances across the nine discussions. Two statistically significant sequential patterns concerning teachers’ scaffolding moves and students’ productive talk emerged. There was a 76% chance that a student would respond to a prompt with an elaborated explanation \(p < .001\) in the adjacent turn and a 4% chance that the teacher would mark an elaborated explanation \(p = .005\). Results also indicated that the relationship between prompting and elaborated explanations as well as the relationship between elaborated explanations and marking to be strong, positive associations \(Q= .886, Q= .659\), respectively). See Figure 4-9 and Figure 4-10 for a representation of prompting, elaborated explanations, and marking in fifth grade.
Figure 4-9: Discourse example from the fifth-grade homogenous below-average ability group of prompting eliciting an elaborated explanation.

Figure 4-10: Discourse example from the fifth-grade homogenous below-average ability group of an elaborated explanation eliciting marking.
Chapter 5

DISCUSSION

Small-group, text-based discussions can enhance students’ reading comprehension. The effectiveness of discussion as an instructional tool, however, is dependent on the quality of students’ talk during discussion. Further, there is some evidence that environmental factors such as group composition and teachers’ discourse moves can mediate the types of social interactions students engage in during discussion and thus, subsequently influence student’s comprehension outcomes. As such, there is a need to investigate the relationship between these two factors as well as their influence on the quality of students’ talk during small-group discussions.

The purpose of this secondary analysis was to extend our understanding of the sequential nature of discourse by examining the relationship between group composition (i.e., homogeneous versus heterogeneous) and teachers’ facilitation practices (i.e., discourse moves) and their influence on student talk within the context of a discussion-based intervention known to promote high-level comprehension (i.e., Quality Talk). Within the context of the primary student, we divided students into homogeneous and heterogeneous reading ability groups at baseline based on oral reading fluency. After which, teachers began implementing Quality Talk (QT) in their classroom. Implementation included presenting mini-lessons addressing specific discourse indicators, facilitating weekly, small-group discussions, and participating in professional development sessions aimed at supporting the teachers as they learned to use teacher discourse moves to promote productive talk. This study drew on 108 discussion observations taking place over the course of a school year which we later coded for individual teacher discourse moves and student’s oral arguments. Our coding procedures allowed us to analyze the teachers
overall use of the individual teacher moves across group composition as well as the nature of the sequential relationships between them.

A major finding of this study was the absence of sequential dependence in half of the discussion groups (i.e., both homogeneous above average groups as well as the fifth-grade heterogeneous groups and homogeneous average group). The null hypothesis of no dependence can be understood as the conditional probability of the target event (i.e., elaborated explanations) given the criterion event (i.e., teacher move) is equal to the unconditional probability of the target event (Allison & Liker, 1982). That is, in this context, dependence is a comparison between the proportion of times elaborated explanations and specific teacher moves co-occurred compared to the proportion of times they occurred individually. The absence of dependence in the sequence of the aforementioned groups, therefore, suggests that students in these groups were just as likely to produce elaborated explanations spontaneously as they were in response to any initiating event. Furthermore, teachers were also just as likely to employ teacher moves in response to elaborated explanations as they were to employ them independently of those events.

As such, the lack of dependence could indicate that these groups required less explicit teacher modeling and scaffolding to engage in productive talk while other groups require more direct facilitation (e.g., prompting). Students in the homogeneous above average groups as well as fifth-grade may produce more produce talk independent of what the teacher does than other groups. As part of two preliminary usability studies of Quality Talk, Reninger and Wilkinson (2010) and Wilkinson, Soter, Murphy, and Li (2008), found that implicit teacher modeling and scaffolding were not sufficient independent of direct instruction for students to acquire the fundamental discourse skills required for productive discussions. Instead, students needed to be explicitly taught through the discourse mini-lessons how to ask meaningful questions and engage in oral argumentation. It may be that such mini-lessons are enough to ensure that above average
students engage in those fundamental discourse elements while other groups still require additional, explicit teacher modeling and scaffolding during discussion.

Importantly, many of the fifth-grade students had participated in a year of Quality Talk prior to this iteration of the intervention and thus, may have already appropriated the normative expectations of productive discussions which led to the need for less explicit teacher facilitation. In other words, due to the direct instruction and scaffolding they received the previous year, they no longer required such direct facilitation. This is not to say that the teachers’ discourse moves didn’t matter, but that the previous teachers’ discourse moves served as a successful scaffold. It is also possible, therefore, that grouping type moderates the pace at which students adopt fundamental discourse skills. Both interpretations may seem to contradict those of Jadalla et al. (2011) and Lin et al. (2015) who demonstrated dependence within the sequential patterns of their data. Both studies, however, took place in the context of fourth-grade classrooms with students who were both new to Collaborative Reasoning and were grouped heterogeneously. Furthermore, Collaborative Reasoning does not explicitly teach students to engage in productive talk.

It is also possible, however, one or more of the groups did not reach dependence due to a lack of power. While all cells exceeded the minimum cell values for the analysis, the Quality Talk model espouses student-directed, teacher-facilitated discussions. As such, the teachers were encouraged to release control of the discussions to their students and necessarily make far fewer utterances. Jadallah et al. (2011) found, however, that the teacher can have a substantial impact on students’ use of evidence with limited discourse moves. More research is therefore necessary to determine the extent to which group composition impacts the pace at which student’s appropriate discursive behaviors as well as how group composition effects the type of facilitation needed.

The analysis also showed that prompting was the most effective move at increasing student’s use of argumentation overall. This finding is consistent with Lin et al. (2015) who
found that direct requests increased the likelihood of productive talk in the following turn above and beyond over moves. Interestingly, however, the power of prompting was not consistent across all grouping type. Specifically, prompting was the most effective in the homogeneous below average group indicating that students in these groups were less likely than others to produce arguments absent of a direct request.

Moreover, there was evidence that while students’ arguments snowballed in other groups, they did not in the homogeneous below average groups. The absence of the snowball phenomenon is contradictory to previous findings which show that students’ productive talk serves as a catalyst for their peers to do the same (Anderson et al., 2001). This finding provides further evidence that group composition may moderate the nature of the teacher moves needed to elicit argumentation. It may also explain why others (e.g., Salah et al., 2005) have found that homogeneous below average groups produce fewer individual arguments, on average, then other groups. These findings indicate that researchers and teachers who are facilitating small-group discussions comprised of student’s with below average reading fluency may want to employ prompting more regularly. However, more research should be done to determine if there are ways to initiate the snowball phenomenon or if this pattern is consistent in additional domains and contexts.

The findings provided evidence that group composition may influence the extent to which the relationship between teachers’ facilitation practices and students’ use of argumentation is bidirectional. There were only meaningful patterns concerning how the teacher responded to students in the homogeneous average and below average groups. Marking was associated with students’ arguments in the homogeneous average and below average groups, but not the heterogeneous groups, while challenging was associated to students’ arguments in the homogeneous below average group in fourth grade. Hogan, Nastasi, and Pressley (2010) found that teachers tended to be more engaged with groups that they perceived to be struggling to co-
construct knowledge. It may be that the teachers perceived these groups as needing more marking overall. Murphy, Greene, et al. (2017) also showed that students in the homogeneous below average group were less likely to challenge each other which is a key component of argumentation. As such, the teachers may have challenged students’ arguments more often to substitute for the students not doing it themselves.

Instances of marking and challenging did not, however, have a positive effect on student’s immediate use of oral arguments in any group. Lin et al. (2015) showed that specific praise (i.e., praise for cognitive and social strategy use as well as personal traits) can have both immediate and delayed effects on student’s relational thinking. While it did not have an immediate effect, praise could have had a delayed effect on student’s oral arguments in some or all group compositions. Jadalla et al. (2011) also found that challenging increased the likelihood of students use of evidence in the immediate turn. Both contradictory findings could be due to how we operationalized our outcome variables. For example, teachers may have praised *pieces* of student’s arguments (e.g., I like your use of evidence) instead of their *full* argument (e.g., that was a really nice elaborated explanation) and thus, students were more likely to respond with a piece of evidence but not, necessarily, a full argument.
References


https://doi.org/10.1016/j.cedpsych.2017.09.003

http://dx.doi.org/10.1037/a0015576


## Appendices

### Appendix A List of Coded Discourse Elements

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Description</th>
<th>Example Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scaffolding Moves</td>
<td>Challenging</td>
<td>The teacher challenges a student to consider justification of the response and an alternative point of view.</td>
<td>“Does that prove she was nice to them though?”</td>
</tr>
<tr>
<td>Marking</td>
<td></td>
<td>The teacher draws attention to or reinforces specific aspects of a student’s discourse by explicitly pointing it out.</td>
<td>“That was great -- bringing up another text that we’ve all read about.”</td>
</tr>
<tr>
<td>Modeling</td>
<td></td>
<td>The teacher exhibits an aspect of discourse that she/he would like students to make by explicitly stating what she/he is going to do.</td>
<td>“So, let me ask a follow up question.”</td>
</tr>
<tr>
<td>Prompting</td>
<td></td>
<td>The teacher helps a student construct an elaborate response. For example, sometimes the teacher may ask for reasons and evidence from the students.</td>
<td>“So why do we think that? Can we think of any evidence from the text?”</td>
</tr>
<tr>
<td>Summarizing</td>
<td></td>
<td>The teacher overviews a part of the discussion thus helping build coherence for students.</td>
<td>“So I think we’re -- we’ve come to the conclusion that this story took place, you know, in the past, not in the present, and that it takes place in the South, but in the United States.”</td>
</tr>
<tr>
<td>Type</td>
<td>Code</td>
<td>Description</td>
<td>Example Quote</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Management Moves</td>
<td>Checking</td>
<td>The teacher tries to make sure that every student has a basic literal understanding of the text.</td>
<td>“So, now did the people on the Pony Express, did they go the whole route?”</td>
</tr>
<tr>
<td>Clarifying</td>
<td></td>
<td>The teacher prompts a student to provide a clearer response by asking a question that sometimes includes a teacher’s refined version of the student’s response.</td>
<td>“So you think that -- you’re talking about when this took place?”</td>
</tr>
<tr>
<td>Instructing</td>
<td></td>
<td>The teacher gives explicit instruction on background knowledge, content of the text, and discussion-related skills or rules.</td>
<td>“Remember, we’re talking to each other and not -- like, we’re not just talking to me, OK?”</td>
</tr>
<tr>
<td>Procedural</td>
<td></td>
<td>The teacher manages the flow and directs the focus of the discussion.</td>
<td>“Let’s move onto a new topic, because we’re getting away from the text, OK?”</td>
</tr>
<tr>
<td>Reading</td>
<td></td>
<td>The teacher reads the text aloud to the students as a read-aloud activity or as a reference to the text for information during the discussion.</td>
<td>“I’d like to ask you a question about, on page 444, it says, ‘After Paul and Babe settled on the river, they -- and spent the day rolling the logs down the river, and he was so tired, he said …’”</td>
</tr>
<tr>
<td>Productive Talk</td>
<td>Elaborated Explanation</td>
<td>Statement comprised of a claim (e.g., position, opinion, or belief) that is based on at least two independent, conjunctive, or causally connected forms of support (i.e., reasons or evidence)</td>
<td>“I would feel fortunate (claim) because my family was given the opportunity to go out west and start this new life (reason) and you would not think of the chores as boring; you would think of them as fun because of the space you had to do them in (reason).”</td>
</tr>
</tbody>
</table>

Source. – Adapted from Wei, Murphy, and Firetto (2018). Adapted with permission.
Appendix B Human Subjects Research Approval

APPROVAL OF SUBMISSION

Date: August 21, 2014
From: Jodi Mathieu, IRB Analyst
To: Pricilla Murphy

<table>
<thead>
<tr>
<th>Type of Submission:</th>
<th>Modification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title of Study:</td>
<td>QUALITY TALK: DEVELOPING STUDENTS' DISCOURSE TO PROMOTE CRITICAL-ANALYTIC THINKING, EPISTEMIC COGNITION, AND HIGH-LEVEL COMPREHENSION</td>
</tr>
<tr>
<td>Principal Investigator:</td>
<td>Pricilla Murphy</td>
</tr>
<tr>
<td>Study ID:</td>
<td>PRAMS00044096</td>
</tr>
<tr>
<td>Submission ID:</td>
<td>MOD00001314</td>
</tr>
<tr>
<td>Funding:</td>
<td>United States Department of Education</td>
</tr>
<tr>
<td>IND, IDE, or HDE:</td>
<td>Not Applicable</td>
</tr>
<tr>
<td>Documents Reviewed:</td>
<td>*JES_QT_protocol_8.20.14.pdf (0.01), Category: IRB Protocol</td>
</tr>
<tr>
<td>Review Level:</td>
<td>Expedited</td>
</tr>
</tbody>
</table>

On 8/21/2014, the IRB approved the above-referenced Modification. This approval is effective through 8/6/2015 inclusive. You must submit a continuing review form with all required explanations for this study at least 45 days before the study’s approval end date. You can submit a continuing review by navigating to the active study and clicking ‘Create Modification / CR’.

If continuing review approval is not granted before 8/6/2015, approval of this study expires on that date. To document consent, use the consent documents that were approved and stamped by the IRB. Go to the Documents tab to download them.

In conducting this study, you are required to follow the requirements listed in the Investigator Manual (HRP-103), which can be found by navigating to the IRB Library within CATS IRB (http://irb.psu.edu). These requirements include, but are not limited to:
- Documenting consent
- Requesting modification(s)
- Requesting continuing review
- Closing a study
- Reporting new information about a study
- Registering an applicable clinical trial
- Maintaining research records

This correspondence should be maintained with your records.
VITA

Rachel M. V. Croninger

EDUCATION

B.A. 2011 Calvin College Psychology

SELECTED SCHOLARLY PUBLICATIONS


SELECTED CONFERENCE PRESENTATIONS

