Redesigning Guided Reading Instruction: Achieving Equity through Homogeneity

Tiffany Taira Young

University of Nebraska-Lincoln, tif.t.young@gmail.com

Follow this and additional works at: https://digitalcommons.unl.edu/teachlearnstudent

Part of the Curriculum and Instruction Commons, Early Childhood Education Commons, and the Elementary Education Commons

Young, Tiffany Taira, "Redesigning Guided Reading Instruction: Achieving Equity through Homogeneity" (2019). Theses, Student Research, and Creative Activity: Department of Teaching, Learning and Teacher Education. 99.
https://digitalcommons.unl.edu/teachlearnstudent/99
REDESIGNING GUIDED READING INSTRUCTION:
ACHIEVING EQUITY THROUGH HETEROGENEITY

by

Tiffany T. Young

A DISSERTATION

Presented to the Faculty of
The Graduate College at the University of Nebraska
In Partial Fulfillment of Requirements
For the Degree of Doctor of Philosophy

Major: Educational Studies

Under the Supervision of Dr. Guy Trainin

Lincoln, Nebraska

May 2019
Elementary school students are often placed into groups with peers of similar reading ability in a practice called within-class ability grouping for guided reading instruction. Through this practice, students are differentially exposed to reading skills, strategies, and texts that are presumed to match their current level of ability. This widespread practice is particularly problematic given that (1) current notions of matching early readers to texts for reading instruction are based on traditional instructional practice rather than empirical evidence, (2) poor, minority students are overrepresented in the lowest ranked groups, (3) students in higher ranked groups make greater academic gains than those in lower ranked groups, and (4) teacher perceptions of students’ abilities are often inaccurate. Conversely, several studies have shown that when students are presented with texts of increased difficulty and given appropriate instructional support, they are able to make accelerated reading progress.

The purpose of this design-based research study was to develop innovative classroom practices and theoretical insights on the use of heterogeneous grouping for guided reading instruction to increase the reading achievement of all students. Qualitative data, in the form of fieldnotes, semi-structured interviews, and documents, were collected. Data analysis included structural and process coding to result in the explication of five design principles to assist in the application of this design in other contexts. In
addition, student progress was monitored using comprehensive reading assessments. On average, students made the equivalent of one year’s worth of literacy growth in two and a half months of design implementation. This dissertation is concluded with specific attention to the technical, normative, and political aspects inherent in the dissemination and sustainability of the proposed design.
DEDICATION

To Beckett and Baby Young -

so that you know anything is possible
ACKNOWLEDGEMENTS

Did you ever see the way the clouds love a mountain? They circle all around it; sometimes you can’t even see the mountain for the clouds. But you know what? You go up top and what do you see? His head. The clouds never cover the head. His head pokes through, because the clouds let him; they don’t wrap him up. They let him keep his head up high, free, with nothing to hide him or bind him. – Toni Morrison

I would like to express my sincere and deepest gratitude to the multitude of individuals who contributed to both my professional and personal development throughout my doctoral journey. First, my advisor and mentor, Dr. Guy Trainin, you offered clear, consistent, and expert advice throughout every step of my journey. I cannot imagine a more open-minded, supportive advisor. Second, Dr. Wayne Babchuk, you not only served as a member of my committee but as a co-instructor, co-presenter, and co-author during my time at the university. Through it all you became not only a mentor, but a role model for my future work with students. Dr. Stephanie Wessels and Dr. Elaine Chan, the other two outstanding members of my committee, you have both offered unwavering support and encouragement, as well as contributed to my interest in early literacy and equity, which serve as the foundation to my scholarly work. I would also like to thank Dr. Loukia Sarroub. You believed in my potential early on in my graduate studies and continued to serve as a mentor in my work long after your official mentorship had ended.

I am indebted to my husband and life partner for spending longs nights at Old Chicago with me while I wrote, for experiencing every high and low alongside me, and for giving me the time and space to do the work I love. Austin, we made it! My parents, Greg and Lori Teichmeier, you have been walking beside me through the long journey of my education and have taught me that wisdom isn’t measured in degrees, but in the work
we do for others. You two have truly earned the highest distinction here. My sister, Nikki Janssen, you have long been my role model and constant supporter. You told me years ago that one day I would write a book about education. Perhaps this is that book, or perhaps there are more books to be written. I am thankful for my lifelong friends Emily Kottwitz and Brittany Russell for late night study snacks and words of encouragement. I also want to thank my grandmother Bonnie Burianek, my principals Paula Baker and Rhonda True, as well as Dr. Ted Hamann for believing in me before I believed in myself.

Finally, Mrs. Poppy, thank you for welcoming me into the sacred space of your classroom and agreeing to go on this complicated and sometimes uncomfortable journey with me. You are a shining example of the power of a strong and incredibly dedicated teacher.

To those mentioned here by name, and to the many, many more unnamed who have played a role in supporting me in this work, I thank you all dearly for being the clouds that surrounded me during this thirteen-year-long journey.
CONTENTS

CHAPTER I

INTRODUCTION ..............................................................................................................1
The History of Ability Grouping ..................................................................................2
Types of Ability Grouping ..........................................................................................5
Conflicting Ideas .........................................................................................................11
Ability Grouping in the Primary Reading Classroom ..............................................15
Exploration of the Research Problem .......................................................................19

CHAPTER II

LITERATURE REVIEW ...............................................................................................22
Ability Grouping and Achievement .........................................................................24
Ability Grouping and Opportunity ............................................................................33
Ability Grouping and the Self .....................................................................................39
Text-Reader Match ....................................................................................................43
   Operational Text Matching ......................................................................................44
   Developmental Text Matching ................................................................................46
   Dispositional Text Matching ...................................................................................48
Ability Grouping and Text Difficulty .........................................................................49
Increasing Text Difficulty .........................................................................................51
Alternatives to Ability Grouping in Elementary Reading Instruction ......................56
Theoretical Framework: Oakes’s Dimensions of Change ........................................62

CHAPTER III

METHODOLOGY ........................................................................................................67
<table>
<thead>
<tr>
<th>Chapter/Journey</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose Statement</td>
<td>67</td>
</tr>
<tr>
<td>Research Questions</td>
<td>68</td>
</tr>
<tr>
<td>Exploration of Design-Based Research</td>
<td>68</td>
</tr>
<tr>
<td>Rationale for a Qualitative Design-Based Research Approach</td>
<td>71</td>
</tr>
<tr>
<td>Philosophical Paradigm</td>
<td>73</td>
</tr>
<tr>
<td>Researcher Positionality</td>
<td>74</td>
</tr>
<tr>
<td>Research Design</td>
<td>77</td>
</tr>
<tr>
<td>Research Site</td>
<td>77</td>
</tr>
<tr>
<td>Research Participants</td>
<td>78</td>
</tr>
<tr>
<td>Timeline</td>
<td>79</td>
</tr>
<tr>
<td>Initial Design Framework</td>
<td>81</td>
</tr>
<tr>
<td>Data Collection</td>
<td>83</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>86</td>
</tr>
<tr>
<td>Research Ethics</td>
<td>88</td>
</tr>
<tr>
<td>Validity, Reliability, and Generalizability</td>
<td>89</td>
</tr>
<tr>
<td>CHAPTER IV</td>
<td></td>
</tr>
<tr>
<td>FINDINGS</td>
<td>92</td>
</tr>
<tr>
<td>Students’ Literacy Progress</td>
<td>94</td>
</tr>
<tr>
<td>Casey</td>
<td>94</td>
</tr>
<tr>
<td>Zane</td>
<td>96</td>
</tr>
<tr>
<td>Marcus</td>
<td>98</td>
</tr>
<tr>
<td>Jill</td>
<td>100</td>
</tr>
<tr>
<td>Hattie</td>
<td>102</td>
</tr>
</tbody>
</table>
CHAPTER V
DISCUSSION ...........................................................................134

Design Principles .....................................................................136

Design Principle I: Comprehensive Literacy Program ...............136
Design Principle II: Research-based Guided Reading Instruction ...144
Design Principle III: Access to Grade Level Literacy Content ..........150
Design Principle IV: Synchronous, Explicit, Systematic Instruction .....154
Design Principle V: Give the Design Time to Flourish .................158

Summary of Design Principles ..................................................161

Theorizing the Why? .................................................................163

Implications ..............................................................................171

Limitations and Future Research ..............................................174
CHAPTER VI

CONCLUSION..................................................................................................................177

Technical Dimension.................................................................................................177

Normative Dimension.................................................................................................179

Political Dimension.....................................................................................................181

Final Words..................................................................................................................184

REFERENCES..............................................................................................................187

APPENDICES..............................................................................................................197

Appendix A: Institutional Review Board Approval Letter............................................198

Appendix B: Oral Recruitment Form for Teachers.........................................................199

Appendix C: Interview Protocols..................................................................................201

Appendix D: Approved Teacher Informed Consent Form..............................................202

Appendix E: Approved Parent/Guardian Informed Consent Form...............................204

Appendix F: Approved Verbal Assent for Student Participants.....................................206
LIST OF TABLES AND FIGURES

Tables

Table 1.1 Types of Ability Grouping Practices ......................................................... 6

Table 4.1: Text Level Correlation Chart ................................................................. 94

Table 4.2: Summary of Students’ Progress ......................................................... 131

Figures

Figure 5.1: Reading Strategy Card ................................................................. 157

Figure 5.2: The Heterogeneous Guided Reading Groups Model ......................... 166
CHAPTER I
INTRODUCTION

Educational equality is an idea that has fallen from favor (Oakes, 1985/2005, p. xv). Written in the 1980s, this statement echoes loudly into the present. As a potential solution to the increasing ethnic, racial, linguistic, and economic diversity of American classrooms, ability grouping for early reading instruction has one again become a common practice in classrooms across America (Loveless, 2013). For decades, debates over heterogeneous versus homogeneous grouping of students has proliferated. In systems of homogeneous grouping, also referred to as tracking in its strictest sense, students are placed with peers of similar ability to receive differentiated instruction. High achievers receive accelerated instruction on rigorous content; low achievers receive basic instruction at remedial levels. In heterogeneous grouping, students of all abilities are grouped together and strategies of differentiation are utilized to ensure that students can both access and excel at grade level content.

Drawing on the field of Sociology, the effects of group composition, whether homogeneous or heterogeneous, influence not only the experiences of students, but also their academic achievement (Dreeben & Barr, 1988). Grouping practices therefore should be given significant consideration in discussions of students’ social, psychology, and cognitive development. In fact, by law, we have the obligation to closely examine these practices. According to the landmark case, Brown v. Board of Education (1954), “Where a state has undertaken to provide a benefit to the people, such as public education, the benefits must be provided on equal terms to all people unless the state can demonstrate a compelling reason for doing otherwise.” This assertion requires us to define the word
equal and begs us to answer the question of whether or not differential access to instruction via ability grouping practices is considered to be equal? And, if not, whether there is a reason compelling enough to continue the practice?

**The History of Ability Grouping**

These questions are not so easily answered and, as a result, ability grouping remains a “perennial theme” in education studies (Loveless, 2013, p. 3). For over a century, researchers have studied the effects of the practice of ability grouping on American students (Steenbergen-Hu, Makel, & Olszewski-Kubilius, 2016). Its nascent forms are linked to W.T. Harris’s plan in 1867 in which high ability students were accelerated through elementary school. The Santa Barbara Concentric Plan was another early model. In this plan, students were sorted into a triumvirate of classes separating high, average, and low ability students (Kulik & Kulik, 1982). This grouping structure was further reinforced by the introduction of Terman’s concept of the IQ test in 1916 (Oakes, 2017). Assigning a scientific test score to students’ abilities marked the beginning of the seemingly objective task of ranking students and placing them into groups based on their endemic characteristics (Oakes, 2017). By 1919, the city of Detroit introduced an ability-grouped class assignment plan in which students were sorted into the top 20%, middle 60%, and bottom 20%, though no major forms of curricular adjustments were made and all students learned the same content (Courtis, as cited in Kulik & Kulik, 1992). These practices were bolstered in the 1950s as the nation focused on individual excellence (Lou, Abrami, Spence, Poulsen, Chambers, & d’Apollonia, 1996). Curricular adjustments by level soon became the norm as technological
competition with Russia provided the impetus for increasing the academic potential of the highest achievers (Kulik & Kulik, 1987).

However, by the 1960s, in line with the sentiments of the Civil Rights Movement, concerns of educational inequality dominated discussions of grouping. Whether or not students in separate groups could be given equal opportunities to learn was vehemently debated, and reviews of the research on ability grouping proliferated (Braddock & Slavin, 1992; Kulik & Kulik, 1982, 1984, 1987, 1992; Lou et al., 1996; Slavin, 1987, 1990, 1993). As a result, efforts to end ability grouping in schools across the country were prevalent and effective (e.g. Cunningham, Defee, & Hall, 1991, 1998; Oakes, 2005), and drastic changes in grouping structures followed. By the mid-1980s to 1990s ability grouping practices were largely eschewed (Steenbergen-Hu et al., 2016). While in 1961, 80% of elementary schools practiced ability grouping for reading, by the mid-1990s only 27% of elementary teachers reported using this practice (Loveless, 2013). According to Loveless (2013), “Grouping students by ability, no matter how well it is done, will inevitably separate students by characteristics that are correlated statistically with measures of ability, including race, ethnicity, native language, and class” (p.15). In fact, ability grouping is considered by some to be a form of second-generation segregation (Kalogrides & Loeb, 2013). For these reasons, in the 1990s the National Governors Association, the American Civil Liberties Union, the Children’s Defense Fund, and the NAACP Legal Defense Fund outwardly denounced ability grouping (Oakes, 1992).

As a proxy for the amount of attention ability grouping has received over the years, Loveless (2013) analyzed the number of times “ability grouping” was mentioned in Education Week, a popular journal that reports on K-12 education, from 1983 to 2012.
At the peak of the detracking movement, between 1990 and 1993, reports of ability grouping reached an all-time high with up to 20 mentions per year. These numbers gradually decreased and since the 2000s there have been no more than four yearly mentions. While these numbers appear to suggest that the debate was settled in the 1990s with the rejection of ability grouping, an interesting phenomenon has occurred. As discussions of ability grouping dissipated, the practice slowly and quietly began once again galvanizing traction in classrooms across the country. In 1996, 28% of fourth grade students were grouped by ability for reading instruction, but by 2009 this number reached 71% (National Assessment of Educational Progress, as cited in Loveless, 2013). As evidenced by these percentages, not only is ability grouping reaching historic numbers, but it is simply becoming the status quo. Incredulously, it is now hardly even a point of discussion.

There is some speculation as to why this is occurring. Concern for the education of highly able students has resulted in the push back against the detracking efforts of previous decades due to the presumption that the learning of high achievers will be unfairly limited by the slower pace of their peers (Oakes, 2005). Furthermore, increases in Federal accountability measures have led to a scramble to find practices that will leave no child behind (Loveless, 2013). In addition, the Common Core State Standards have called for an increase in rigor to prepare students to become college and career ready (Hiebert & Mesmer, 2013). Teachers are left to bridge the gap between past and present expectations and ability grouping may appear for some to ameliorate the problem.

Ironically, while No Child Left Behind (NCLB) demanded the disaggregation of achievement data by race and class, it has also contributed the disaggregation of students
by race and class in our nation’s classrooms. As our poor and minority students continue
do less well on tests of academic achievement than their middle class and white
counterparts, these students are more likely to be placed in the lowest groups and the
lowest tracks in our schools in an attempt to provide them with instruction at their so-
called level (Darling-Hammond, 2007). Ability grouping practices are now being
cemented into policy in some parts of the country and new forms of ability grouping have
erupted in others (Oakes, 2008). It appears that collective amnesia regarding the inherent
corns of ability grouping has fallen over the field of education.

Types of Ability Grouping

There are many terms used to describe the variety of practices used to group
students for academic instruction based on their ability. While there is some overlap in
the ways in which these terms are used, the following definitions will be adopted and
used for the purpose of clarity. Table 1.1 provides a summary of these practices, which
are further explored in this section.

- **Programmatic tracking**, the most rigid of these practices, refers to “broad,
  programmatic divisions that separate students for all academic subjects”
  (Gamoran, 1992, p. 11). Students are placed into either academic, general,
or vocational tracks. These track placements are highly stable and students
from different tracks have little interaction with one another (Oakes,
2005). Once students are placed into a given track, their curriculum is
ddictated for the remainder of their high school career. Rising concerns of
inequity beginning in the 1960s and 1970s (Kulik & Kulik, 1987) led to
its near disappearance in the United States (Loveless, 2013).
Table 1.1 Types of Ability Grouping Practices

<table>
<thead>
<tr>
<th>Ability Grouping Type</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programmatic Tracking</td>
<td>Students are separated by ability for all academic subjects into highly stable academic, general, or vocational tracks.</td>
</tr>
<tr>
<td>Course-by-course tracking</td>
<td>Students are grouped by ability in some subject areas (e.g. math or English) and leveling for each course is independent of the others.</td>
</tr>
<tr>
<td>Ability-grouped class assignments</td>
<td>Homerooms in elementary school are grouping homogeneous based on ability within the grade level (e.g., high class, average class, remedial class).</td>
</tr>
<tr>
<td>Multilevel grouping</td>
<td>Homerooms in elementary schools are grouped heterogeneously, but students are regrouped by ability levels for select subjects such as math and reading.</td>
</tr>
<tr>
<td>Between-class grouping</td>
<td>Students are grouped heterogeneously for most of the day, but are regrouped across grade levels for instruction in some subjects such as math and reading (e.g. the Joplin plan).</td>
</tr>
<tr>
<td>Grouping for selected subjects</td>
<td>Students are placed in small, homogeneous groups within the larger heterogeneous class for part of the day. While the teacher works with students in one group, the other students complete independent tasks.</td>
</tr>
<tr>
<td>Cross-grade grouping</td>
<td>Students are placed in small, homogeneous groups within the larger heterogeneous class for part of the day. While the teacher works with students in one group, the other students complete independent tasks.</td>
</tr>
<tr>
<td>Nongraded plans</td>
<td>Students are placed in small, homogeneous groups within the larger heterogeneous class for part of the day. While the teacher works with students in one group, the other students complete independent tasks.</td>
</tr>
<tr>
<td>Within-class ability grouping</td>
<td>Students are placed in small, homogeneous groups within the larger heterogeneous class for part of the day. While the teacher works with students in one group, the other students complete independent tasks.</td>
</tr>
<tr>
<td>Intraclass grouping</td>
<td>Students are placed in small, homogeneous groups within the larger heterogeneous class for part of the day. While the teacher works with students in one group, the other students complete independent tasks.</td>
</tr>
</tbody>
</table>

- Course-by-course tracking, a legacy of programmatic tracking, continues to endure in high schools and middle schools. In this grouping structure, classes in a given subject are hierarchically leveled and the curriculum offered in each course varies in difficulty (Chmielewski, 2014). For example, in English, students may take an “honors” course, general English, or a reading course. Similarly, in math, students may take algebra, pre-algebra, or basic math. In contrast to broad programmatic tracking, students may be in the highest level of one subject, and the lowest level in another based on their personal strengths and weaknesses.
In addition, students typically have some choice in which tracked course they would like to take, though they are often highly influenced by suggestions of academic advisors as well as their own past performances (Oakes, 2005). In schools using course-by-course tracking, some subjects such as math and English may be grouped by ability, but others such as physical education and those in the social sciences may be heterogeneously grouped. Students of various abilities have the potential to take courses with one another throughout a typical school day, but ability grouping some courses can lead to de facto ability grouping of others due to course scheduling constraints. In most comprehensive high schools and some middle schools, course-by-course tracking is now the norm.

- *Ability-grouped class assignments* or *multilevel grouping* is a practice that responds to the tradition of assigning students to homerooms in elementary schools. This makes grouping structures in elementary schools unique from those in middle and high schools. In ability-grouped class assignments (Slavin, 1987), or multilevel grouping, student heterogeneity is reduced by grouping all students across a given grade level into homeroom classes based on ability (Kulik & Kulik, 1992). For example, there may be a designated high class, average class, and remedial class into which students are sorted at the beginning of the school year. Students remain in these classes for the entire day and instruction in all subject areas is based upon the group’s presumed generalized ability. This system,
resembling programmatic tracking, may be particularly problematic for students who are high achievers in one subject area, but are struggling in another.

- **Between-class grouping**\(^1\) was created to counteract this concern, which allows students to be grouped by ability for only some subjects, similar to course-by-course tracking (Kulik & Kulik, 1987). This practice is also sometimes referred to as *grouping for selected subjects* (Slavin, 1987). Using this grouping structure, homeroom classrooms are grouped heterogeneously, which allows students the opportunity to interact with peers at various levels. During select subjects, such as math and reading, students are regrouped across classes in the same grade level to receive instruction at an appropriate level of difficulty.

- **Cross-grade grouping** (also known as *nongraded plans*) is another grouping structure utilized in elementary schools. In cross-grade grouping, students are grouped heterogeneously most of the day, but students from several grade levels are regrouped together for instruction in some subjects. The Joplin plan, introduced by Cecil Floyd in Joplin, Missouri is the most well-known grouping structure of this type (Kulik & Kulik, 1992). Using this structure, for example, high achieving fourth graders

---

\(^1\) The terms *between-class grouping* and *multilevel grouping* are sometimes used interchangeably although they mean different things per the definitions explored above. This can lead to confusion when comparing the results of research studies. Regardless of the term used by the original author of the studies reviewed, this paper will use the definitions presented in this section to avoid confusion. For this reason, the terms used in the present text do not always match the terms used in the original study.
may be grouped with average fifth graders and below average sixth graders. Group assignments are flexible and students are switched to different groups as their reading performance warrants. It is argued that when heterogeneity of the whole classroom is reduced in cross-grade grouping plans, the teacher is able to spend more time on direct instruction at the level assumed to be appropriate for all students (Slavin, 1987). In contrast to other practices in which little curricular adjustment is made based on group level, such as ability-grouped class assignments, students in cross-grade grouping structures are given access to different curricula designed to meet their current level of academic performance.

- **Within-class ability grouping or intraclass grouping** (Lou et al, 1996), an alternative to between-class ability grouping, is a practice in which students are placed in small, homogeneous groups within the larger heterogeneous class for part of the day (McCoach, O’Connell, & Levitt, 2006). Within-class ability grouping allows a single classroom teacher to provide differentiated instructional content to small groups of students while the rest of the class works on independent learning tasks (Loveless, 2013). This form of grouping is the most flexible of the ability grouping practices since teachers are independently able to make changes in group composition. It is commonly used for both reading and math instruction and is sometimes used in science and the social sciences as well (Kulik & Kulik, 1992). For example, based upon reading test scores, a teacher might divide her students into four groups. Learning objectives and student texts
are uniquely planned for each group. While one group may be learning how to decode phonetically regular words, another may be focusing on reading with fluency. Similarly, one group might be reading informational text to support content knowledge, while another is reading a narrative to understanding cause and effect. Theoretically, when students are grouped for instruction within the classroom, instructional content can be specifically designed to match the academic levels of each group (Kulik & Kulik, 1992).

- Other grouping structures exist throughout the K-12 continuum that are beyond the scope of this study. In some schools, a school-within-a-school structure also serves to separate students by ability. The International Baccalaureate program is available in some high school across the country. Through this program, students can gain credit commensurate with qualifications of universities around the globe. Students in the International Baccalaureate program take all, or a majority, of their coursework together and rarely interact with others in the school who are not in the program. Special classes for high achievers and special classes for low achievers are also common within schools (Slavin, 1987). These types of grouping systems often require students to have an identification label, such as gifted or special education in order to be grouped in these classes. Additionally, students who are new to learning the English language are also sometimes grouped separately for instruction in English language classes. These classes enable instruction to focus primarily on
the skills needed to be able to speak, listen, read, and write in English. In each of these grouping structures, the remainder of the students continue to be grouped heterogeneously for instruction. While the above grouping systems also warrant further research and analysis, classes geared specifically towards certain groups of students such as those identified as gifted, special education, and English language learners are beyond the purview of this study.

**Conflicting Ideas**

While wide variation in the structure of ability grouping exists, several common arguments undergird these practices. The primary argument of those in favor of ability grouping is that it is the most effective way for teachers to meet the academic needs of a diverse group of students (Braddock & Slavin, 1992; Kulik & Kulik, 1982; Slavin 1987, 1990, 1993). Hypothetically, teachers are able to provide enrichment for high achievers and remedial instruction for low achievers (Lou et al, 1996). By nature of this separation, it is presumed that the growth of high ability students will not be stunted by slower learners and that slower learners will not feel overwhelmed by instruction that moves too fast or is too rigorous. All students, therefore, are said to benefit from this grouping structure.

Vygotsky’s (1978) well-known theory of the zone of proximal development, which suggests the provision of teacher scaffolding at students’ unique stages of learning, may initially seem to reify the argument for homogeneous grouping. However, the logic in this argument assumes that students in ability grouped classes are placed in groups that are truly homogeneous and that those placements are accurately made. Research
indicates, however, that a wide range of abilities exists in each group (Hallinan, 1994; MacIntyre & Ireson, 2002; Pallas, Entwisle, Alexander, & Stluka, 1994; Rosenbaum, 1980). Furthermore, several studies have found that teachers’ perceptions of students’ abilities are often inaccurate (Ready & Wright, 2011; Tach & Farkas, 2006). These empirical studies call into question the basic premise undergirding the argument for homogeneity.

Advocates of ability grouping also suggest that placing lower ability students in groups with students of a higher ability will lead to feelings of inferiority and superiority as students become increasingly aware of the cognitive differences that exist between them (Oakes, 2005). Using this line of reasoning, separating students by ability will have greater positive effects on the social and psychological development of students than heterogeneous practices. To better understand these effects, research studies and reviews have analyzed the interaction between grouping practices and psychological development. Results are conflicting. While Lou and colleague’s meta-analysis (1996) reported that students’ self-concepts are higher in grouped classes, Braddock and Slavin (1992) found that when disaggregated by ability, low students in low groups had lower self-esteem than their non-grouped but also low-achieving counterparts. Furthermore, in several reviews, Kulik and Kulik (1982, 1984, 1992) found only trivial effects of ability grouping on students’ self-concepts.

Opponents of ability grouping practices counter these arguments by insisting that not only is ability grouping ineffective, but it is inequitable, antithetical to democratic ideals, and has pernicious effects on students (Gamoran, 1992). According to Oakes (2005), hundreds of empirical studies lead to one conclusion regarding the effects of
ability grouping on student achievement: “*no group of students has been found to benefit consistently from being in a homogeneous group*” (p. 7, emphasis in the original). While some studies suggest that high achievers benefit from ability grouping, others suggest that greater benefits result from heterogeneous grouping. Even meta-analyses provide conflicting accounts. For example, the reviews done by Kulik & Kulik (1982, 1984, 1987, 1992) generally found small positive effects of ability grouping practices on achievement, while those done by Slavin (1987, 1990, 1993; Braddock & Slavin, 1992) generally found a near zero or slightly negative effect. Accounts are likely conflicting due to the multiple dimensions of ability grouping practices, including grouping type, grade level, age of student, subject area, and class size. Ultimately, however, arguments for the overall effectiveness of ability grouping gain little traction from the inconsistent findings in these studies.

Second, opponents of ability grouping suggest that such practices result in an inequitable *opportunity to learn* for students. This phrase, first coined by Carroll in 1963 (as cited in Schmidt, Burroughs, Zoido, and Houang, 2015), has been used to describe the range of opportunities students from various group placements receive as a direct result of such placement. Research in this area provides results of great clarity: students placed in the lowest ranked groups are met with a variety of confounding factors that systematically limit their opportunity to learn (Allington, 1983; Chorzempa & Graham, 2006; Condron, 2008; Eder, 1981; Gamoran, 1992; Hallinan, 1994; Lleras & Rangel, 2009; MacIntyre & Ireson, 2002; Oakes, 1992, 2005; Rowan & Miracle, 1983). As students’ opportunities to learn are reduced, their academic achievement is directly affected. This argument against ability grouping can be summed up by the *Matthew*
*Effect:* in ability grouping structures, the academically rich get richer while the academically poor get poorer.

The third argument against ability grouping logically follows the last. Certain students, by nature of their demographic characteristics, are more likely to be placed into the lowest groups resulting in de facto segregation by race and class (Hallinan, 1994). Since poor and minority students consistently achieve at lower levels than their white, middle class counterparts, these students are more likely to be placed into the lowest groups and therefore have the fewest opportunities to learn (Braddock & Slavin, 1992). To further compound the problem, there is evidence to suggest that placement of these students into lower groups is more likely even after controlling for previous achievement (Tach & Farkas, 2006). Given that ability groups are relatively stable over time (Dreeben & Barr, 1988), a feedback loop is created in which low students, particularly poor and minority youth, are placed into the lowest groups. Once placed, they receive the fewest opportunities to learn. The less they learn, the more stable their placement in the lowest group becomes and the cycle repeats itself over the course of their public education from kindergarten through high school.

Ultimately, the arguments for and against ability grouping are reflective of the differences in personal beliefs as to whether public education should serve convergent or divergent goals (Deunk, Smale-Jacobse, de Boer, Doolaard, & Bosker, 2018). Those aligned with a divergent theoretical point of view believe that schools ought to devote equal time and resources to all students to help them reach their highest potential. They are likely to view ability grouping as an appropriate instructional practice in meeting that goal. They find the widening of the achievement gap to be a natural result of schooling.
Alternatively, those with a convergent viewpoint believe that true equity is only achieved when struggling students are provided with the additional time and resources needed to meet equally high standards of proficiency. Ideologically, they contend, it is the purview of the school to systematically close the achievement gap. This conflict between convergent and divergent goals is also evident in the conflation of policy and practice. The political sphere in America has promulgated the convergent goal of leaving no child behind; ironically, in an effort to do so, the practice of divergent instruction has become the norm (Loveless, 2013; Oakes, 2008).

**Ability Grouping in the Primary Reading Classroom**

The practice of within-class grouping by students’ reading ability begins as early as the first weeks of kindergarten and continues throughout the elementary years (Fountas & Pinnell, 1996, 2017; Richardson, 2009, 2016). According to the Early Childhood Longitudinal Study-Kindergarten cohort (ECLS-K) dataset, 70% of first grade teachers grouped students by ability for reading instruction (as cited by Lleras & Rangel, 2009). While all children are typically exposed to the same grade level content during certain aspects of balanced literacy programs, such as shared reading and shared writing, during guided reading students are placed in presumed-to-be homogeneous ability groups. In these groups, instruction is varied according to students’ ability levels demonstrated on formal and informal literacy assessments.

According to Fountas and Pinnell (1996), the goal of guided reading is “to help children learn how to use independent reading strategies successfully” (p. 2) with the close guidance of the teacher. To provide this guidance, primary teachers typically divide their students in three to four groups with four to six students in each group. Groups meet
for approximately fifteen to thirty minutes most days of the week while other students
work on independent literacy activities. The groups are ranked from high to low and
instruction is planned accordingly. Teachers provide each group with texts at their
respective levels of difficulty and pair them with instruction needed to access the text

What occurs in these reading groups often contributes to the feedback loop
explained above. Students placed in the highest reading groups are exposed to higher
level content, while those in lower groups are more likely to receive instruction in basic
skills (Allington, 1983; Chorzempa & Graham, 2006; Gamoran, 1992). The provision of
differentiated content instruction is based on the argument that students in lower groups
are not ready for more challenging content. However, without the critical exposure to
higher level content, students will certainly never catch up to their higher grouped
counterparts. Furthermore, not only do higher groups have fewer behavioral disruptions
(Eder, 1981), but simply being assigned to a higher ability group has been shown to have
positive effects on a child’s individual behavior (Tach & Farkas, 2006). Even if groups of
varying levels are given equal instructional time, those in lower groups spend less of that
time engaged in learning opportunities (Allington, 1983). Given these considerations, it is
unsurprising that the achievement gap between higher-achieving and lower-achieving
students increases over time through the use of ability grouping (Lleras & Rangel, 2009).

Learning to read is a complex, multidimensional process (e.g., Chall, 1983; Ehri,
1995, 1999), and appropriate text selection for readers at various stages of reading
development has been given much attention over the years (e.g., Betts, 1946; Fountas &
for example, designate 26 unique levels of text difficulty, ranging from A to Z. These levels break down the broad phases and stages of reading development (Chall, 1983; Ehri, 1995) into minute specification of skills. Each level of text corresponds with the introduction of unique skills and strategies outlined for instructional purposes. Due to the incremental nature of this system, student placement into groups is highly controlled. It is suggested that students are placed in groups with peers who vary from them by no more than one reading level (Richardson, 2009, 2016). When students demonstrate independent levels of accuracy and comprehension at a given text level, they are advanced to the next level (Fountas & Pinnell, 1996, 2017). However, the higher group often remains an unreachable target as students in that group are also advanced to the next level of difficulty. This system of leveling continues until students reach the end of the continuum sometime near the end of elementary school or middle school (Fountas & Pinnell, 1996).

Determining when students are ready to progress to a higher level of text difficulty has been guided by tradition rather than research (Brown, 2009). Betts (1946) was the first to introduce the idea of independent (basal), instructional, and frustration levels for students. He suggested that the accuracy and comprehension percentages with which a student is able to read a text should be used to determine whether the text would be most beneficial for independent reading, reading instruction, or not at all. This criterion has been widely proliferated in the field of education (Fisher, Frey, & Lapp, 2012; Fountas & Pinnell, 1996; Mesmer & Cumming, 2009) and over the decades has remained largely unquestioned. However, these percentages are not based upon empirical evidence (Shanahan, 2011). In fact, while there has been ample theoretical discussion on
the importance of matching readers to appropriate texts, there is strikingly little evidence on the level of difficulty that is considered to be an appropriate match (Compton, Appleton, & Hosp, 2004; Hiebert & Mesmer, 2013; Kuhn & Stahl, 2003; Morgan, Wilcox, & Eldredge, 2000).

This problematizes the widely used practice of ability grouping for guided reading instruction, professed to be beneficial for the very reason that it effectively matches readers to an appropriate text level. The lack of research on traditional designations of text difficulty calls into question not only whether or not students are correctly placed into their reading groups, but also whether or not students are receiving early reading instruction that best fits their instructional needs. In fact, recent research suggests that students can benefit from reading text at more difficult than expected levels given appropriate instructional supports (e.g., Cunningham et al. 1991, 1998; Kuhn, Schwanenflugel, Morris, Morrow, Woo, Meisinger, Sevcik, Bradley, & Stahl, 2006; Morgan et al., 2000; Stahl & Heubach, 2005). In these studies, ability grouping structures for reading instruction were replaced with highly successful alternative approaches. Results indicated that students who would have likely been relegated to the lowest reading groups were indeed able to successfully read grade level material. Furthermore, there was no indication in these studies that the achievement of high performing students was hindered by the removal of ability grouping structures. Yet, despite these promising alternatives, ability grouping practices remain the status quo in elementary reading instruction.
**Exploration of the Research Problem**

Although the Common Core State Standards call for an increase in text difficulty beginning in elementary school (CCSS Initiative, 2010), the percentage of fourth grade students reading above a basic level has not shown statistically significant changes in over a decade (National Center for Education Statistics, 2017). On the National Assessment of Educational Progress in 2017, 32% of fourth graders failed to meet even basic levels of reading proficiency (National Center for Education Statistics, 2017). Perhaps uncoincidentally, this percentage is almost identical to the percentage of students likely placed in the lowest reading group in their classroom. Ability grouping practices, by nature of limiting students’ exposure to high level texts, arguably exacerbates the problem of reading deficient youth in our nation.

This concern and others have echoed for over a century as the merits and flaws of ability grouping have been debated. Arguments of effectiveness have been pitted against those of inequity, the results of one meta-analysis against another, the beliefs of one teacher against another. Through it all, Slavin (1987, 1990, 1993) suggests that the “burden of proof” ought to lie with those in favor of ability grouping. Since ability grouping proponents rely primarily on arguments of effectiveness, and the opponents on issues of equity, the latter argument remains plausible regardless of whether or not research finds ability grouping to be effective.

Research has long attempted to provide this much sought-after proof, but as will be explicated in the review of the literature, **consistent** proof has never been found. This lack of consensus in the research is reflected in the ebb and flow of ability grouping in classrooms across the nation. Teachers, along with administrators and district personnel,
are required to make consequential decisions about grouping that are often based more on curriculum publishers’ decisions and time allocations than empirical evidence (Park & Datnow, 2017). Opportunities to learn are limited by archaic traditions. Students enter the feedback loop created and reinforced by ability grouping within weeks of kindergarten entrance and spiral, often unproductively, until they are dismissed into the workforce. Oakes (2005) aptly concludes, “Those at the bottom of the social and economic ladder climb up through twelve years of ‘the great equalizer,’ Horace Mann’s famous description of public schools, and end up still on the bottom rung” (p.4).

While research on alternatives to ability grouping practices has dotted the professional landscape over the past 30 years, this body of research has lacked the detail necessary for teachers to implement these new practices in their own classrooms. Research must attend to the challenges that real teachers face in real classroom settings if ability grouping practices are to be replaced by alternatives that are both more effective and more equitable for all students (Tach & Farkas, 2006). Thick descriptions of the systematic moves that teachers make in meeting the diverse literacy needs of all students can serve as a realistic model for practitioners. This sentiment has been echoed throughout the years by various researchers on ability grouping (e.g., Hamilton & O’Hara, 2011; Hong, Corter, Hong, & Pelletier, 2012; Oakes 2017; Slavin, 1987).

The research study presented in this document responds to this much heeded call. The purpose of this design-based research study is to collaboratively and iteratively design an approach to primary reading instruction that meets the needs of a diverse group of students without the use of ability grouping. This study builds off of a two-year action research study that replaced homogeneous grouping with heterogeneous grouping...
practices. Many aspects central to research-based guided reading instruction remained intact, but all students were given the instruction necessary to successfully read grade level text or higher. The primary question under further investigation is: How can a teaching strategy focusing on heterogeneous grouping practices for guided reading support the reading achievement of students of all levels of reading ability? The literature on ability grouping and text difficulty will serve as a foundation for designing instructional practices that draw on the successful components of ability grouping structures while abandoning the ineffective cognitive, social, and psychological ramifications.
CHAPTER 2
LITERATURE REVIEW

The literature on ability grouping is vast. It began just over a century ago, in 1916, when a man by the name of Guy Whipple was the first to seriously study homogeneous grouping (Kulik & Kulik, 1982). Since that time, research in the area has proliferated. There are studies documenting ability grouping in nearly every subject, from biology (e.g., Saleh, Lazonder, & de Jong, 2005) to mathematics (e.g., Leonard, 2001), of every grade level from kindergarten (e.g., Condron, 2008) to higher education (e.g., Alba & Lavin, 1981), of every type from programmatic tracking (e.g. Oakes, 2005) to within-class ability grouping (e.g., Ready & Wright, 2011), and in nations around the world from Kuwait (e.g., Saleh et al., 2005) to Scotland (e.g., Hamilton & O’Hara, 2011). In addition, various research methods have been employed to better understand this practice, including surveys (e.g., Chorzempa & Graham, 2006), propensity score matching (e.g. Condron, 2008), interviews (e.g., Boaler, William, & Brown, 2000), observations (e.g., Rist, 1970), hierarchical linear modeling (e.g., Tach & Farkas, 2006) meta-analyses (e.g., Kulik & Kulik, 1984), best evidence syntheses (e.g., Slavin, 1993), and many more.

Given the extent of the literature on ability grouping and the purpose of this research study, the review in this document narrows its scope and focuses primarily on grouping practices in early elementary school. According to Fountas and Pinnell (1996), “The primary years are our chance to alter the trajectory of failure, make a difference in students’ lives, and in turn make a difference in our educational system” (p. 193). Students enter elementary school with a wide range of previous experiences, opportunities, and knowledge. Many kindergarten students are not able to identify the
letters of their own name, while others have already developed a wide range of early literacy skills including the ability to read and write many words. Given these disparities, some students already have a greater likelihood of failure and need instruction that can circumvent the circumstances. When students are grouped by ability, the opposite occurs. Research indicates that there is a positive relationship between the frequency and duration of ability grouping and the widening of the achievement gap (Lleras & Rangel, 2009). When this grouping begins in elementary school, it furthers the case for the need for ability grouping in middle and high schools to respond to disparate abilities of students. Based on this logic, grouping students by ability in elementary school arguably impacts the findings of studies on ability grouping for students in middle and high school as well.

To further narrow the literature, when such research is available, this review will focus primarily on grouping practices within early literacy instruction. The decision to focus on literacy instruction in this research is twofold. First, learning how to read and write is arguably the most important goal of instruction in the early years. This is not only reflected by common sentiment, but also in the number of instructional minutes devoted to literacy instruction in the primary grades. As prerequisite skills for later academic tasks, future success in school rests heavily on these foundational skills. Furthermore, literacy is a prime subject for study due to the prevalence of ability grouping for instruction in this area. Recent accounts have indicated that upwards of two-thirds or more of elementary students are sorted for reading instruction (Chorzempa & Graham, 2006; Loveless, 2013). Due to the interaction between its significance and its prevalence,
ability grouping practices for elementary students in literacy instruction warrant further study.

**Ability Grouping and Achievement**

The effects of ability grouping practices on academic achievement have been the primary topic of inquiry in ability grouping debates over the years. Several large-scale meta-analyses have been conducted to sort through the wealth of studies in this area. Findings are unique to each analysis, therefore close attention to methods is necessary to understand the results. Kulik and Kulik (1984) conducted a seminal meta-analysis on the effects of between-class ability grouping and ability grouped class assignments versus no ability grouping in elementary schools. Studies of other forms of ability grouping, such as within-class ability grouping, were excluded. Inclusion criteria consisted of experimental studies of grouping practices in any subject. After reviewing 28 studies, Kulik and Kulik (1984) found the overall benefits of these practices to be small but significant with an effect size of +.19. However, of the studies reviewed, less than half of them had results that reached levels of statistical significance. Furthermore, nearly a quarter of those reaching significance found significant negative effects for ability grouping on achievement. Ultimately, the small, positive overall effect found in this study must be considered in light of the variation that existed among the individual studies analyzed.

Using a new approach of combining features of meta-analysis and narrative review, Slavin (1987) conducted separate best-evidence syntheses of between-class ability grouping and within-class ability grouping versus no grouping. Studies of elementary students in grades first through sixth were included in the analysis, as well as studies of any subject area. Overall, Slavin (1987) found an effect size of .00 for ability
grouped class assignments and stated, “unequivocally the research evidence refutes the assertion that ability-grouped class assignments can increase student achievement in elementary school” (p.307). Slavin argued that differences in selection criteria, such as including studies focusing on programs for the gifted (e.g., Kulik & Kulik, 1984), resulted in the disparities between his analysis and those previously conducted. In addition, Slavin (1984) noted that Kulik and Kulik (1984) conflated the effects of ability-grouped class assignments and between-class grouping practices. When these effects were separated, Slavin found the effect-size of ability grouped class assignments to be zero. There were only seven studies focusing on the effects of regrouping for reading and mathematics versus no grouping and results were found to be inconclusive. Slavin (1987) also found a medium effect of +.45 for the cross-grade grouping assignment in reading, known as the Joplin Plan. Of the cases that reported results disaggregated by students’ ability levels, none found that one group benefited at the expense of the others. Finally, Slavin (1987) found a +.34 effect for within-class ability grouping, however, the paucity of reviewable research in reading makes this effect attributable to math only. Slavin (1987) suggests that due to the prevalence of ability grouping for reading instruction in elementary schools, studies of randomized assignments to ungrouped control groups are difficult to conduct.

While Kulik and Kulik (1984) and Slavin (1987) reported only the effects of multiple types of grouping compared to no grouping, Lou and colleagues (1996) conducted a much-needed meta-analysis of the effects of within-class homogeneous grouping versus within-class heterogeneous grouping to isolate the homogeneity variable from that of the grouping variable. However, this meta-analysis reported on studies of
students from elementary to post-secondary schools. Additionally, all subject areas were included. A slight, but significant, effect size of +.12 was found in favor of homogeneous grouping, but the range of studies examined varied from -1.75 to +1.12. Upon further analysis, the authors concluded that both subject area and initial ability level served as moderators of this practice with high ability students and those grouped for reading benefiting most. However, when interpreting these results for early reading instruction, caution is warranted as only four of these studies examined within-class ability grouping in reading with no indication of the ages of students. In addition, collaborative learning groups in which students serve as group leaders were included, which contrasts with the common teacher-led within-class ability groups in early reading instruction.

When the controversy of ability grouping decreased in the late 1990s, so too did research in the area. At least for the time being, the debate seemed to be settled in some respects and ongoing in others; ability grouped class assignments were deemed ineffective, while within-class ability grouping was found in need of further study. As the primary grouping structure for early reading instruction, attention to this area is of critical importance. Fortunately, additional inquiry into within-class ability grouping was assisted years later with the access to the Early Childhood Longitudinal Study-Kindergarten (ECLS-K) database, which is part of the Early Childhood Longitudinal Study conducted by the Institute of Education Sciences and the National Center for Education Statistics. The dataset includes information on a cohort of students from their entry into kindergarten in 1998 until eighth grade (Deunk et al., 2018), and has been analyzed by multiple researchers over the past several decades.
Tach and Farkas (2006) used this data to estimate both the determinants and consequences of placing kindergarten and first grade students in within-class ability groups for reading. They found the overall effect of ability grouping versus nongrouping in first grade reading achievement to be negative, and the overall effect in kindergarten to be positive. When analyzing the results based upon student ability levels, Tach and Farkas (2006) found the overall positive effect of within-class ability grouping for kindergarteners to be a result of the relative gains of high-ability students in comparison to the relative losses of low-ability students. In perhaps the most consequential finding, Tach and Farkas (2006) report that regardless of prior performance and behavior, kindergarteners’ placements into a higher ability group resulted in increased student performance. Therefore, if two students of the same ability were placed in separate groups, the student placed in the higher group would have higher achievement than the student placed in the lower group simply as a result of such placement.

Using the same dataset, Condron (2008) used propensity-score matching to compare students placed in within-class reading ability groups to similar ability peers in nongrouped classrooms. For example, students placed in the high reading group were compared to those students in the nongrouped class who would have been placed in the high reading group had this practice been used in their classrooms. This method accounts for the possibility that high ability students might learn more as a function of other unmeasured variables that contribute to learning rather than grouping itself. Using this improved method, data revealed that students placed in low-ranked groups still learned significantly less than their nongrouped peers who would have been placed in a low-ranked group. In contrast, ability group placement had an ambiguous effect on reading
gains for middle-ranked students, and a slight positive effect on high-ranked students. Condron (2008) ultimately concluded, “Ascriptive inequalities in educational achievement may stem primarily from non-school factors (Downey, von Hippel, & Broh, 2004), but curriculum differentiation is probably one of the most salient school factors contributing to such gaps” (p.386, emphasis in the original).

All of the studies reviewed thus far support the differential effects hypothesis of ability grouping (Lleras & Rangel, 2009). This hypothesis states that students who are placed in lower groups for reading instruction learn substantially less than their higher grouped counterparts. Critical consumption of the research on ability grouping practices demands consideration of the differential effects. The reporting of average effect sizes of a given type of ability grouping often masks the fact that ability grouping is detrimental to low achieving readers. For example, as mentioned above, when studies of separate classes for the gifted students were removed from the Kuliks’ (1984) study, the results no longer showed a positive effect for ability-grouped class assignment. Given the long-standing effectiveness versus equity debate, the differential effects of ability grouping suggest that within-class ability grouping is neither effective nor equitable for our lowest achieving students.

To examine ability grouping practices at a broader level, Nomi (2010) used the propensity-score method on the ECLS-K data to study how school level characteristics influence the practices of first grade ability grouping in reading. Results indicated that schools using within-class grouping practices are more likely to be public, high-minority, low-socioeconomic status, and low-performing schools with heterogeneous ability compositions. Though ability grouping was more prevalent in these schools, analysis
found that ability grouping had either no or negative effects on student achievement, especially for low ability students. This analysis once again supports the differential effects hypothesis. Alternatively, ability grouping led to increased achievement for all ability levels within schools that were least likely to use grouping practices, particularly private schools with regulated admission processes. Surprisingly, in this type of school setting, low ability students were found to benefit most from ability grouping practices. Nomi (2010) explains that in these schools, there are fewer low ability students. As a result, low ability students are more likely to be placed in groups with peers of higher levels.

This contradicts McCoach and colleagues’ (2006) study of the same data set that found schools in which teachers reported greater uses of ability grouping had higher average gains in measures of kindergarten reading. Although there is a correlation, there are perhaps other confounding variables involved. Particularly problematic is the fact that data from students who were identified as learning disabled, emerging multilinguals, and those who moved schools (perhaps an indicator of highly mobile students) were excluded from analysis. According to studies supporting the differential effects hypothesis (e.g., Condron, 2008; Tach & Farkas, 2006), excluding these students who are more likely to be considered low ability may skew the data in favor of the positive effects on high ability students. In addition, the ECLS-K data was collected at a time when kindergarten classrooms were just beginning to shift from a developmental focus to an academic focus (Russell, 2011). While some kindergarten teachers focused their instruction on personal and social development, others focused instruction on developing academic skills such as reading and writing. Kindergarten students in an academically-focused classroom are
more likely to be sorted into groups based upon their academic reading performance than those in developmentally-focused classrooms. The former kindergarteners are also more likely to receive explicit reading instruction. Given this context, the supposed correlation between ability grouping and kindergarten reading achievement might actually be a correlation between academically-focused kindergartens and reading achievement as a result of such a focus.

Several other studies also used the ECLS-K dataset to explore the comparative effects of ability grouping versus nongrouping in early reading instruction. Hong and colleagues (2012) reported that while high-ability students showed no overall benefits of ability grouping, low and average ability students showed mixed results depending on the time and intensity of grouping. These students benefited from homogeneous grouping practices if ample instructional time was provided. Conversely, homogeneous grouping practices were detrimental, particularly to low-ability students, if overall reading instruction time was limited. Lleras and Rangel (2009) contributed to knowledge in this area by finding that when students are grouped by ability from first to third grade the achievement gap increases. Furthermore, they found that this gap decreases and even disappears in some instances within the context of non-grouped classes. In combination, the set of studies based on the ECLS-K dataset provides empirical evidence that ability grouping for reading instruction in elementary school can yield different results depending on extant circumstances, including initial ability of the student, amount of instructional time provided, and school level factors. These studies, however, are conclusive in one area: ability grouping negatively impacts learning for those already at a disadvantage.
Following these studies, Steenbergen-Hu et al. (2016) conducted a second-order meta-analysis of one hundred years of research on ability grouping, spanning 1922 to 1994. Thirteen meta-analyses, including those done by Kulik and Kulik (1984), Slavin (1987), and Lou and colleagues (1996) were analyzed. A majority of these studies were conducted between the 1960s and 1980s during the peak of the equity debate. Studies were inclusive of kindergarten through twelfth grade and included multiple subject areas. Steenbergen-Hu et al. (2016) found that the following ability grouping practices benefit students at least to a small degree: within-class grouping, cross-grade grouping, and special grouping for the gifted. On the other hand, ability grouped class assignments were found to have negligible effects on student achievement. Furthermore, overall results rejected the differential effects hypothesis and found that students from all initial ability levels benefited equally from grouping practices.

While Steenbergen-Hu and colleagues’ research review provides the widest support of ability grouping practices, its results should be considered carefully. First, all of the meta-analyses reviewed are at least a quarter of a century old or older. Since that time, new methodological techniques such as propensity score matching, regression discontinuity designs, and multilevel modeling were introduced which may produce more accurate estimates of the effects of ability grouping (Steenbergen-Hu et al. 2016). Several of the studies explored above, such as Condron (2008) and Nomi (2010), used these methods and found unfavorable results particularly for low ability students. Furthermore, this second-order meta-analysis had a much wider scope than early reading instruction and therefore caution is warranted in its application to the primary area of focus for this study.
Two other meta-analyses have been recently conducted. Puzio and Colby’s (2010) study is unique in its particular attention to within-class ability grouping for reading instruction. However, this study focused on second through tenth grades and did not include studies on the earliest stages of reading development. In addition, many of these studies included other interventions such as collaborative learning, which are not components of early guided reading group instruction (Fountas & Pinnell, 1996, 2017). Overall, Puzio and Colby (2010) found a positive weighted effect size of +.22 in their meta-analysis of within-class ability grouping on reading. However, in a more recent meta-analysis and systematic review, Deunk and colleagues (2018) examined the differentiation practices of language and math instruction in primary education from 1995 to 2012 They found no effect of ability grouped class assignments, between-class grouping, and within-class homogeneous grouping on the achievement of high and average ability students. They did, however, find that these grouping structures had a small negative effect on low-ability students. Considering the contradictory results of these analyses, it is important to note that each analysis focused on different subject areas, grade levels, and time periods.

Overall, the decades of research on ability grouping practices for early reading achievement have produced conflicting results. Meta-analyses have found ability grouping practices to be both effective and ineffective, but ultimately there are no meta-analyses that focus solely on within-class ability grouping for early reading instruction. Furthermore, although several individual studies indicate that ability grouping is detrimental particularly for low ability students (e.g. Condron, 2008; Lleras & Rangel, 2009; McCoach et al., 2006; Nomi, 2010; Tach & Farkas, 2006), there is an over-reliance
of these studies on the ECLS-K dataset that is now two decades old. Analysis of the newly available data from the ECLS-K dataset as well as other data sources is warranted. In the interim, ability grouping practices for early reading instruction have failed to pass Slavin’s (1987, 1990, 1993) “burden of proof” test. Chary educators and administrators would be wise to consider further evidence about other effects of ability grouping before making instructional decisions.

Of importance to educators and researchers alike is not only the academic outcomes of ability grouping, but also the social and emotional effects of this practice on students. According to Rosenbaum (1980), “[Ability] grouping systems are more than educational practices; they are also social entities which set in motion social processes which have social effects” (p. 391, emphasis in the original). Examination of these social processes demands that we not only know what happens, but also how it happens to better understand the overall effects of ability grouping.

**Ability Grouping and Opportunity**

The concept of students’ *opportunity to learn* has been a central topic of inquiry for researchers studying ability grouping (e.g., Allington, 1983; Condron, 2008; Chorzempa & Graham, 2006; Gamoran, 1992; Hallinan, 1994; Lleras & Rangel, 2009; MacIntyre & Ireson, 2002; Oakes, 1992, 2005; Rowan & Miracle, 1983, Schmidt et al., 2015). This broad and overarching phrase encompasses all of the ways in which the interactions within the classroom environment, in this case those related to ability grouping, serve to either hinder or promote learning. In a seminal article, Rowan and Miracle (1983) reviewed two hypotheses to explain the ways in which different opportunities may lead to the disparate learning outcomes explored in the previous
section. Both the differential instruction and differential peer hypotheses have found support in the research on ability grouping.

The differential instruction hypothesis suggests that when students are grouped based on their presumed abilities, teachers provide better instruction for students in higher ranked groups. In Chorzempa and Graham’s (2006) study of first through third grade students, students in lower groups were less likely to be asked critical comprehension questions, had fewer opportunities to select their own reading materials, and spent more time overall on non-instructional activities. Similarly, Hallinan (1994) reported that both the quantity and quality of instruction increased in higher groups as they worked at a much faster pace. Several studies also reported that children in higher groups were given more meaningful, less skill-based instruction and assignments (MacIntyre & Ireson, 2002; Oakes, 1992, 2005). Such differences in opportunities to learn based on grouping structures are not unique to the United States, but are perhaps an endemic characteristic of ability grouping itself (Schmidt et al., 2015). Subscribing to this hypothesis in his research review, Allington (1983) argues that “good and poor readers differ in their reading ability as much because of differences in instruction as variation in individual learning styles or aptitudes” (p. 548).

However, other studies suggest that differences in instruction between ranked groups is exactly what is needed in order for ability grouping practices to demonstrate positive effects. In their review of ability grouping for math and reading instruction in kindergarten through sixth grade, Deunk and colleagues (2018) found that the greatest degree of curricular adjustments produced the largest effects on learning. This reaffirmed the Kuliks’ (1992) prior conclusion that adaptations in course content was the key factor
of success in ability grouping structures. Furthermore, Rowan and Miracle (1983) found that while differential instruction was detrimental to low ability fourth grade students in ability grouped class assignments, it was beneficial for them in within-class grouping structures due to a faster pace of instruction and more direct teacher interaction. Given the conflicting findings regarding the effects of differential instruction, descriptive research in this area would help to elucidate the ways in which differential instruction interacts with other variables to influence students’ opportunities to learn.

The differential peer hypothesis proposes that ability grouping stratifies peer contexts and that peer contexts influence achievement. This hypothesis was drawn from research on high school tracking, which suggested that ability grouping practices may lead to differences in peer allocation, peer values, and peer activities (Rowan and Miracle, 1983). Studying these interactions in elementary school, Eder (1981) found that first grade students in lower groups exhibited more inattentive behavior. This inattentive behavior required more teacher management, which led to an increase in disruptions during reading instruction. Ultimately, these factors contributed to lower reading achievement. Furthermore, Tach and Farkas (2006) found that placement in a higher ability group actually had positive effects on student behavior. In combination, these studies indicate that students who need the best reading instruction are often placed in groups not only with other low readers who need additional support and teacher time, but also those struggling with behavior causing disruptions in the learning environment (Allington, 1983; Gamoran, 1992).

Regardless of which facets of opportunity are analyzed, research supports the conclusion that the opportunity for students to learn in the lowest ranked groups is less
than the opportunity to learn in the highest. This is particularly problematic given that research clearly documents that ability grouping practices consistently result in the overrepresentation of poor and minority students in the lowest ranked groups (e.g., Braddock & Slavin, 1992; Chmielewski, 2014, Condron, 2008; Hallinan, 1994; Kalogrides & Loeb, 2013; Oakes, 1985/2005, 1992, 1994; Tach & Farkas, 2006). In fact, this stratification is so prevalent, as mentioned above, it is sometimes referred to as a “second-generation segregation” practice (Kalogrides & Loeb, 2013; Oakes, 1992).

From a sociological standpoint, schools have the capacity to either reduce, reproduce, or exacerbate preexisting social realities (Condron, 2008). Given the explicit egalitarian goals of the United States as well as the historical and current patterns of racial and economic stratification, public education arguably must function in ways to reduce such stratification. However, when ability grouping is used, cultural norms are reproduced and social stratification is exacerbated. According to Oakes (2008), “Stratification of educational opportunities is driven by deep cultural beliefs about native abilities, cultural deficits, and by the match between schooling and social inequality” (p. 710).

In perhaps the most well-known study of grouping, Rist (1970) conducted a microethnography of a group of Black children in an “urban ghetto school.” Through extensive observation and teacher interviews over the course of the students’ kindergarten year and half of their second-grade year, he uncovered egregious findings. Students were sorted into groups within the classroom on the eighth day of kindergarten, and those same groupings persisted through their second-grade school year. Not only were the groups static, but placements were based on middle-class norms of appropriate
language, dress, and overall tidiness rather than on any sort of objective academic assessment. The teacher ridiculed and ignored students in the lower groups and labeled them as “slow learners.” The “fast learners” adopted the teachers’ patterns of behavior towards the “slow learners.” Eventually both low and high learners adopted this ascribed identity and their achievement levels began to match those expected of them by the teacher.

Rist’s (1970) seminal study contributed not only to the research on ability grouping, but also to the theory of the self-fulfilling prophecy. According to Merton (1948), “The self-fulfilling prophecy is, in the beginning, a false definition of the situation evoking a new behavior which makes the originally false conception come true” (p. 195, emphasis in the original). In Rist’s (1970) study, the false definition of students as slow learners in kindergarten ultimately became true by their second-grade year.

Ability grouping in today’s classrooms creates the same risk. According to several studies, students are often assigned to ability groups based on factors other than their reading achievement, such as behavior, relative age, maturity, social class status, and teacher judgment (Dreeben & Barr, 1988; Eder, 1981; Pallas et al., 1994; Slavin, 1990; Tach & Farkas, 2006). For example, Ready and Wright (2011) found that teachers perceived Black, Hispanic, poor, and male students to have lower literacy abilities than their peers even though only half of these disparities in perception were linked to actual between-group differences. Based on this research, students of certain demographics have a greater chance of being placed in a low-ranked reading group in which they don’t belong. Given the effects of the self-fulfilling prophecy, this false definition of low ability may indeed become true over time through the systematic differences in teacher
and peer behavior. As Merton aptly writes (1948), “the specious validity of the self-fulfilling prophecy perpetuates a reign of error” (p. 195).

Unfortunately, this reign of error is unlikely to have term limits. While Rist’s example might be extreme by today’s standards (Hallinan, 1994), research documents that ability groups continue to be relatively stable over time (Dreeben & Barr, 1988; MacIntyre & Ireson, 2002; Oakes, 1992; Pallas et al., 1994). Although leading proponents of ability grouping in elementary reading instruction encourage grouping that is dynamic and flexible to meet the changing ability levels of students (Fountas & Pinnell, 1996, 2017; Pinnell & Fountas, 2007; Richardson, 2009, 2016), research on the self-fulfilling prophecy suggests that significant changes in group composition are unlikely to occur.

When students are placed into groups, the mechanisms by which the self-fulfilling prophecy occurs can be better understood by examining students’ opportunities to learn in ranked groups. These societal factors congeal to form a vicious and self-sustaining system of inequity. Poor and minority students, whether accurately or inaccurately, are more likely to be placed in the lowest ranked groups. Once placed, it is probable that they will have limited opportunities to learn due to the slow pace of instruction, low expectations, and frequent behavioral disruptions. Given these circumstances, it is plausible that they will make limited academic growth. Consequently, even if they are fortunate enough to get an astute teacher who can accurately perceive their abilities during their subsequent elementary years, the self-fulfilling prophecy likely already had ample time to ensure that the students’ abilities correctly match their current low placement.
Ability Grouping and the Self

Both proponents and opponents of ability grouping agree that such practices can also directly impact a child’s psychological development. Those in favor of ability grouping contend that when students are placed with peers of similar ability they are less likely to develop feelings of inferiority based on their current level of ability. This is because they will no longer constantly compare themselves and their work to that of students of higher abilities. Furthermore, all students will have the opportunity to be a leader in their respective groups. Given these factors, some believe that ability grouping ultimately helps the psychological development of students in all groups (see Oakes, 2005 for a more elaborate discussion). However, opponents argue that separation by ability leads to further stigmatization as students are acutely aware of the differing status levels of groups (Braddock & Slavin, 1992). Students may feel stuck in their current group placement resulting in decreased motivation and an overall negative self-concept. According to Rosenbaum (1980), “The problematic issue is whether students evaluate themselves by comparing themselves with others in their own ability group or whether they evaluate themselves by identifying with their ability group and its position in the class as a whole” (p.371).

Several seminal research reviews have focused on the effects of ability grouping on the self-concept of students. In a narrative review of research, Rosenbaum (1980) reported that although there were mixed results, a majority of the studies found that ability grouping practices hurt the self-evaluation of both low and average ability students. In addition, Kulik and Kulik (1984) conducted a meta-analysis of nine studies that analyzed the effects of ability grouped class assignments on the self-concept of
elementary students in a variety of subjects. Four of the studies found a trivial or small
positive effect on self-concept, while five of the studies found a trivial or negative effect
on self-concept. Overall, the studies reduced self-concept by .06 standard deviations.

Kulik and Kulik’s (1982) study of secondary students as well as their study of elementary
through secondary students (Kulik and Kulik, 1992) also found near zero effects on self-
estee.

Years later, Lou and colleagues (1996) conducted a rudimentary meta-analysis on
the effects of within-class ability grouping versus no grouping on students’ self-concept
in elementary through post-secondary school. They found that students in grouped classes
had significantly more positive attitudes towards the subject area in which they were
grouped (d+=+.18). They also had significantly higher general self-concepts than those in
the nongrouped classes (d+=+.16). Several other items were analyzed, including attitudes
toward instructional approach and academic self-concept, but no other significant results
were found. While Lou and colleagues (1996) also analyzed the achievement effects of
homogeneous within-class grouping versus heterogeneous within-class grouping, the
same comparison was not made for studies of self-concept. Therefore, although two
significant, positive effects were found, it is unknown whether such differences should be
attributed to ability grouping or simply to grouping in general which allows for greater
student-teacher interaction and more individualized support. Furthermore, although the
two effects were found to be statistically significant, their practical significance is
considered to be small.

Since the 1990s, there has been little new research on the effects of ability
grouping on students’ self-concepts. It seems that research reviews sufficiently settled the
debate: ability grouping has negligible effects on students’ overall self-concept. However, these results should still be regarded with caution. Changes in self-concept may be much more difficult to detect and measure than those of an academic nature. For example, Lou and colleagues’ (1996) criteria for inclusion explicated that grouping practices only had to be in place for more than one day. Few would argue that changes in a child's self-concept would occur within merely a few days of ability grouping. Even studies conducted over a full academic year may not be able to truly account for the impacts of ability grouping on a child who has been in the lowest or highest ranked group for thirteen years of formal schooling. In addition, measuring an abstract concept such as a child’s self-concept is much more challenging than measuring the specific skill growth of a child. Not only does the data collection method need to be reliable, but it also needs to be a valid representation of the concept. Between the measurement difficulties and the length of time needed to document changes in self-concept, further research in this area is needed.

Three general areas of research on ability grouping have been reviewed, including the ways in which ability grouping affects academic achievement, learning opportunities, and self-perceptions. Given the numerous and sometimes conflicted findings throughout the century, it is unsurprising that ability grouping has remained a consistent area of interest over time. However, the compilation of research in this review points toward a few insights on the implications of ability grouping on elementary school students. First, in terms of academic achievement in literacy, the cohort of studies using the ECLS-K dataset suggest that ability grouping has positive effects on high ability students to the detriment of low ability students. Second, research indicates that students’ group
placements influence their overall opportunity to learn, with those in the lowest groups receiving the fewest opportunities. This is particularly troubling given that poor and minority students are most likely to end up in the bottom groups regardless of initial academic ability. Finally, research has not found consistent evidence that ability grouping either positively or negatively affects students’ self-concept development. In combination, empirical evidence suggests that the practice of ability grouping for early reading instruction needs replacement. Braddock and Slavin (1992) succinctly write, “Ability grouping must end because it is ineffective, harmful to many students, and damaging to interracial relations and democratic society. Effective and practical alternatives exist” (abstract).

While research on the effects of tracking and detracking has nearly brought broad programmatic tracking to an end (Loveless, 2013), the elementary school classroom still clings tightly to its tradition of sorting students, particularly for reading instruction. Some claim that within-class ability grouping is itself an alternative to the more stringent practices of tracking (Boaler et al., 2000), but research shows that even within-class ability grouping fails to pass Slavin’s (1987, 1990, 1993) “burden of proof” test as an effective practice that outweighs its inherent inequalities. Research has also shown that there are effective alternatives to within-class ability grouping for elementary reading instruction (e.g., Cunningham et al., 1991, 1998; Kuhn et al., 2006; Morgan et al., 2000; Stahl & Heubach, 2005), but unfortunately these methods have yet to replace the stronghold ability grouping has on reading instruction (Lleras & Rangel, 2009; Loveless, 2013). This may be a result of the long-held beliefs regarding the importance of systematically matching readers to an appropriate level of text difficulty.
Text-Reader Match

From rudimentary sentences such as “Lin can sit. Lin can sit on Sid.” to the classic stories of Dick and Jane, the field of education is steeped in tradition when it comes to matching texts to developing readers. According to Brown (2009), “With regard to the optimum oral reading accuracy and rate benchmarks for pacing students’ progression in leveled text, the field has been and continues to be, driven by tradition rather than research findings” (p. 195). Brown is referring to the widely held benchmarks set forth by Betts in his 1946 text Foundations of Reading Instruction. This text, created for preservice teachers to use in their reading methods courses, sets stringent expectations for reading accuracy and comprehension. Though these benchmarks were neither based upon nor supported by empirical evidence (Fisher et al., 2012; Shanahan, 2011), they continue to remain the most salient in the field and are often used to inform ability grouped placements. An examination of text difficulty, text-reader match, and current reading instructional practices provides compelling evidence to support a break from years of uninformed tradition.

According to Billman, Hilden, and Halladay (2009), “Research has indicated that it is important wherever and whenever possible to provide students with texts that are appropriately matched to their reading abilities” (p. 219). Indeed, much research has been done on the ways in which texts of varying levels of difficulty and skill can support students’ reading development (e.g., Kuhn, 2005; Morgan et al., 2000; O’Connor, Bell, Harty, Larkin, Sackor, & Zigmond, 2002; Schwanenflugel, Kuhn, Morris, Morrow, Meisinger, Woo, Quick, & Sevcik, 2009). Mesmer and Cumming (2009) outline three dimensions of text difficulty to consider when appropriately matching readers to text:
operational, developmental, and dispositional. While operational matching requires attention to text features, the developmental and dispositional dimensions require special attention to the reader. These three dimensions will be used to frame a discussion on the ways in which teachers can define the level of text difficulty with the reader in mind.

**Operational Text Matching**

Operational text matching refers to the level of text at which a student is able to read at designated percentages of accuracy and comprehension. As explained above, the field of literacy education continues to be dominated by Bett’s (1946) criteria. Given these standards, students reading a text with accuracy rates of 99% percent or higher and comprehension levels of 90% or higher are said to be able to read the text at a *basal* or *independent* level. These texts would be considered an appropriate choice for the student to select during independent reading times in the classroom or at home. When accuracy rates fall between 90% and 95%, and comprehension scores fall between 75% and 90%, the texts are considered to be at the reader’s *instructional* level. Reading of these texts is most effective when guided by the support of a more skilled reader, and are most often chosen for instructional use in within-class ability-grouped lessons. Finally, those texts that a student reads with accuracy rates below 90% and without the ability to comprehend are considered to be at a student’s *frustration* level and are therefore said to be inappropriate for reading even given teacher support. These texts, however, may be used for read aloud purposes in which an adult reads the text to the child. Other researchers and leaders in early reading instruction continue to promulgate guidelines similar to these percentages (e.g., Fountas & Pinnell, 1996, 2017; Richardson, 2009, 2016).
Several leveling systems have been created to assist primary reading teachers in selecting texts that account for the operational dimension. Fountas and Pinnell (1996, 2012, 2017) created a holistic text leveling system that spans kindergarten through eighth grade texts. Per their suggestions, kindergarteners begin reading texts designated as a level A and progress throughout their elementary and middle school careers until they reach a level Z. Using this system, students are typically grouped for instruction with similarly leveled peers. These 26 levels are individually defined for instructional purposes complete with outlines of reading skills and strategies students need to be successful at each level (Pinnell & Fountas, 2007; Richardson, 2009, 2016). This leveling system requires fairly rigid adherence, suggesting that students should not be grouped with peers who are more than one level apart (Richardson, 2009, 2016). In addition, Pearson Education, Inc. (2011) has established the Developmental Reading Assessment, which is also widely used to assess readers on designed levels of increasingly difficult texts. This system uses numbers rather than alphabet letters to denote difficulty, but text progression is comparable to that developed by Fountas and Pinnell (1996).

To assist in the process of text selection, several readability formulas have also attempted to quantify the variety of text factors that contribute to text difficulty. These formulas of text difficulty focus on both word factors and syntax factors, such as word frequency and mean sentence length (Hiebert & Mesmer, 2013). While readability formulas can be helpful, they fail to consider important reader variables that can only be addressed by qualitative measures, such as the necessary background knowledge needed to access the content or students’ motivation and interests (Fisher et al., 2012).
addition, readability formulas, such as the Lexile levels referenced in the Common Core State Standards, are inaccurate when used on texts at the kindergarten and first grade levels because the nature of these texts often lead to skewed results (Fisher et al., 2012). Words used in these texts may be challenging, but highly supported by the picture. For example, in an emergent text about a birthday party, the author may simply include the following words individually along with a corresponding picture that greatly assists the reader in solving the word: balloon, candle, present, surprise, party. If these same words were used within sentences without pictures, they would be much more difficult for children to read. Although not perfect, readability formulas offer teachers a quick indication of the level of text difficulty to guide them in appropriate text selection.

**Developmental Text Matching**

Readability formulas and text leveling systems account for only half of the text-reader relationship, and attention must also be given to the developmental characteristics of the reader (Hiebert & Mesmer, 2013). To better understand the developmental needs of students, Ehri’s (1995, 1999) phase theory of learning to read words by sight is instrumental. According to this theory, development begins at the pre-alphabetic phase. During this phase, children do not use alphabetic knowledge to read words, but rather use visual clues, such as the shape and color of a stop sign, to identify and memorize words. In the partial alphabetic stage, children begin to learn the letter names and sounds and can use this information to predict words using beginning or final consonant sounds (Ehri, 1995). For example, if a child sees the word dog next to a picture of a dog, they may use the initial /d/ sound to predict that the word is dog rather than puppy. As students progress in their phonological awareness and ability to decode words, they enter into the
full alphabetic phase. Although children at this stage still have difficult reading multisyllabic words, their sight word knowledge increases as the relationship between the graphemes and corresponding phonemes solidifies. Finally, in the consolidated alphabetic phase, readers begin to consolidate letters into multi-letter units such as syllables, rimes, and morphemes (Ehri, 1995).

Using this knowledge, teachers can select from a variety of texts to meet the developmental needs of students. Cunningham, Spadorcia, Erickson, Koppenhaver, Strum, and Yoder (2005) suggest that defining a text’s purpose is more important than defining its difficulty level. They outline four categories of texts, including vocabulary-controlled, decodable, predictable, and multiple-criteria, which serve varying purposes. Predictable texts often provide repetitive phrases with one new word per page that is paired with an illustration to serve as a meaning cue (Brown, 2009). These texts are supportive of students in the pre-alphabetic phase because they allow students to simply memorize the patterns presented in the texts as they begin to learn that written words hold meaning. Predictable texts can also support students in the partial alphabetic phase who are just beginning to understand the relationships between initial graphemes and their corresponding phonemes in words. Decodable texts, on the other hand, are books that are composed of phonetically regular words designed to include target letter/sound relationships taught during phonics instruction (Cunningham et al., 2005). These texts can help students enter the full alphabetic phase of development by encouraging text directionality and spelling patterns (Jager Adams, 2009). As a third text type, vocabulary-controlled texts systematically introduce and repeat targeted words so students can learn to develop automatic word recognition of core words (Mesmer & Cumming, 2009). Such
repetition of words may also increase automaticity in recognizing multi-letter word units, which is necessary in order for students to enter into the consolidated alphabetic phase. Finally, multiple-criteria text combines decodable texts, vocabulary control, and predictable elements to support a wide range of skill development (Cunningham et al., 2005). In addition to these four categories, high-low texts provide students with reading material on high-interest topics with lower levels of reading demands to support students who may not have the skills necessary to read grade level texts (Mesmer & Cumming, 2009). In developmental-text matching, without specific leveling systems, teachers can support text-reader matching by selecting texts from various categories to assist students in their development of certain reading skills.

**Dispositional Text Matching**

Dispositional factors, including reader engagement and motivation, also need consideration when selecting appropriate texts for students to read (Mesmer & Cumming, 2009). In a review of studies on reading motivation, Morgan and Fuchs (2007) found support for the hypothesis that reading skill and reading motivation have a bidirectional relationship. That is, higher student reading achievement is predictive of reading motivation, and, alternatively, reading motivation is predictive of reading achievement. Given this research, finding texts that are engaging and motivating to students, particularly those of high interest topics, may lead to an overall increase in reading achievement. Background knowledge, previous experiences, task persistence, and the ability to connect to prior knowledge are just a few of the reader characteristics a teacher can assess prior to text selection to increase motivation (Billman et al., 2009). Conversely, when readers are continually given texts that are uninteresting, over time
they may be less motivated to read and ultimately their reading achievement may decrease. Certain categories of texts, including decodable and vocabulary-controlled texts, have been criticized for their forced and unnatural language (Jager Adams, 2009), resulting in a shift in the late 1980s towards higher quality, less controlled texts for early readers (Hiebert & Martin, 2009). While there is once again renewed interest in controlled texts (Cunningham et al., 2005), teachers have a challenging task in balancing the idiosyncratic operational, developmental, and dispositional dimensions of text difficulty.

**Ability Grouping and Text Difficulty**

Elementary teachers often use the operational dimension when assigning students to within-class ability groups for guided reading instruction. In most classrooms using this grouping structure, the classroom teacher will assess all students using a running record or other assessment instrument to determine the level of text difficulty at which each child is able to read with a designated level of accuracy and comprehension (often those suggested by Betts, 1946). After assessing all students, the teacher then forms three to four groups with students of similar ability. If, say, four to six students could read a level B text with approximately 95% accuracy, they would likely be placed together in one ability group. The teacher might also have a group reading at a level A, another at a level D, and the final group reading at a level F. Often times, there are a couple of students that do not fit perfectly in any group. Richardson (2009, 2016) suggests that students are not placed in a group that differs more than one instructional level from them. If no such group exists, it is recommended that the teacher instructs this child individually. Based on this operational component, each group would receive guided
reading instruction with a book at their designated level. Skills and strategies necessary to read texts at that level would be taught until students demonstrate proficiency. Once demonstrated, the students would move on to the next text level.

Advocates of homogeneous ability groups for guided reading instruction (e.g., Fountas & Pinnell, 1996, 2017; Richardson, 2009, 2016) highly encourage these groups to be flexible, allowing students to move to higher or lower groups as needed. However, based on the leveled nature of instruction and the explicit skills scheduled to be taught at each level, instructional practices serve to make these groups highly stable. For example, students receiving instruction at level B are unlikely to learn the skills and strategies needed to move up into the level D group when they have not been given instruction at the interim level C. Meanwhile, this level D group is a moving target; the closer a level B student gets to reading a level D text, the farther that group has progressed. Unless the teacher accelerates the instruction in the lower group, students are likely to either stay in their current group or move down if they are unable to keep pace with their peers.

This systematic control over text level exposure may partly contribute to the differential effects hypothesis which suggests that students who are in lower groups for reading instruction learn substantially less than their counterparts placed into higher groups (Lleras & Rangel, 2009). Simply stated, students cannot be expected to learn what they are not taught. In the lowest groups, students are often not exposed to grade level material, but are still expected to meet proficiency levels by the end of the year. Instead, a self-perpetuating cycle ensues. A student is assigned to a low-level group. He is matched with low-level text and corresponding low-level instruction. Formative assessments of that student’s progress reveal that he continues to be below grade level, ensuring that he
will stay in the low-level group and continue to be denied grade level reading content. Ultimately, in this system of within-class ability grouping for guided reading instruction, student progress is limited by current ability and current ability is limited by instruction.

On the other hand, if the developmental and dispositional factors of text selection are also considered when grouping students for reading instruction, an avenue may exist for students to exit this cycle. These components allow for a much broader interpretation of students’ current levels of ability and similarly a broader range of texts considered to be appropriate for instruction. Considering Ehri’s (1995, 1999) four phases of word recognition and Chall’s (1983) six stages of reading development, breaking down the process of reading into 26 identifiable levels to which readers are strictly assigned may be too constricting. The nature of reading development is both fluid and complex. Research on early literacy development reminds us that exposure to text and literacy practices are instrumental to development (Chall, 1983). Students who have more early exposure are better readers. Given this research, breaking the process down into such controlled, hierarchical steps may unnecessarily limit students’ learning. Furthermore, when these limitations are placed as early as kindergarten, they are compounded year after year until ability grouping becomes increasingly necessary to provide instruction for students who are no longer in the same developmental phase as their peers.

**Increasing Text Difficulty**

Guided by this same logic, several studies have attempted to provide a better understanding of the interactions between students and texts at varying levels of difficulty. Since current operational guidelines are based on tradition rather than empirical evidence (Shanahan, 2011), these studies are imperative in understanding the
nature of an appropriate text-reader match. Considering within-class ability grouping for guiding reading relies heavily on this construct, pedagogical implications of these studies are particularly important.

In a study of word learning in three first-grade classrooms, Johnston (2000) found that predictable texts, which are often given to our lowest readers, provide too much support and thereby limit the number of words that students learn. When the same text was printed on sentence strips without the illustrations, students exhibited greater word learning. Furthermore, when those words were printed in isolation in a word bank, students learned the most. In this study, there was an inverse relationship between word learning and the structure of the task. While students may initially appear more successful in the structured task of reading predictable texts, ultimately students benefit in the long-term by creating their own structure, such as attending to visual cues, which can be later used when the predictably structured text is removed.

Similarly, in a study of second grade readers of all levels of achievement, Benjamin and Schwanenflugel (2010) studied the effects of increased text difficulty on students’ prosody, or expressive reading. While the authors hypothesized that increases in text difficulty would result in decreases in prosody, they found the opposite to be true. The young readers in this study used prosody, especially purposeful pausing, as a scaffold to read and understand more challenging text. Given this prosodic support, students who might have initially failed at reading a higher-level text learned to independently structure the task in a way that allowed them to be successful. The authors ultimately concluded that children may benefit from the oral reading of text that is considered to be too difficult based on traditional notions of text difficulty.
Morgan, Wilcox, and Eldridge (2000) studied 51 struggling second grade readers as they participated in an instructional intervention using a dyad reading approach. In this approach, first introduced by Eldredge in 1988, struggling readers (assisted readers) were partnered with a more successful peer (lead reader) for daily reading practice. During this fifteen-minute intervention, the lead reader read grade level texts aloud with fluency while the assisted reader read aloud in unison using the cues from the lead reader on words as needed. For the pretest-posttest control-group design of this study, students were randomly assigned to one of three dyad reading treatments. The group in the first treatment read text at their instructional level, the second read text two grades above their current reading level, and the third treatment read text four grades above their reading level. Students participated in 95 intervention sessions lasting 15 minutes each. At the end of the study, each student was assessed using the Burns/Roe Informal Reading Inventory and a running record of a popular trade book. Results indicated that all groups exhibited improvement with dyad reading, but the group that read texts two grade levels above their reading level gained the most. Those that read text at their instructional level made the least growth. This finding is demonstrative of the potentially negative effects of placing low-performing students in the lowest ability group and providing them with text at their supposed instructional level. The benefits of increased difficulty, however, tapered off somewhere between two grade levels and four grade levels above students’ instructional level. This seems to indicate that there may a threshold for appropriate levels of text difficulty that future research can explore.

Kapur’s (2008) theory of productive failure provides insight into what may be occurring as students interact with challenging texts. In this theory, students benefit from
tasks in which limited structure, defined as scaffolding, expert help, tools, and resources, is provided. Ill-structured tasks, Kapur proposes, allow students to develop structures which will help them in subsequent tasks of a similar nature. Indeed, in Kapur’s study (2008), students initially presented with ill-structured tasks performed better on ensuing ill-structured and well-structured tasks than those who were provided with well-structured tasks from the onset. While this study was conducted on eleventh grade science students, the theory of productive failure may provide insight into matching students with more challenging text that traditionally assumed. When students are presented with challenging texts they must develop structures and strategies to support their reading.

As described above in Benjamin and Schwanenflugel’s (2010) study of reading prosody, when students were given challenging text they were able to use prosodic elements to create their own structures in an ill-structured task given their ability. Similarly, the students in Johnston’s (2000) study benefited most when the structures of illustrations and patterns were removed, thereby allowing students to create their own structures that they could utilize on future texts. Furthermore, the second graders in the study of dyad reading (Morgan et al., 2000) performed better when reading above grade level texts because they enabled the students to create structures, such as reading strategies, which they later used on the independent testing materials. Conversely, those students reading four grade levels above may have been less effective at creating structures that bridged such a large gap.

The idea of a threshold of difficulty is further supported by a study of third through fifth grade students with a learning disability by O’Connor and colleagues
(2002). In this study, students who read texts at their reading level during one-to-one tutoring sessions performed better on oral reading fluency than those who read text at their grade level. Students reading at their instructional level may have performed better because it more closely matched the materials used for assessment, but there is also a possibility that the nature of their learning disability limited their ability to create sufficient structures to use on subsequent tasks. Instruction on scaffold creation may need to be more explicit for some students.

Finally, O’Connor, Swanson, and Geraghty (2010) conducted an experiment investigating text difficulty and its interactions with improving the reading rate of poor readers in second through fourth grades. Students met individually with an adult listener for 15 minutes, three days a week, for a 20-week period. Some students read texts that would be considered to be at their independent reading level, while others read texts they could only read with 80% to 90% accuracy. In both treatments, adults were trained to simply give students words that they were unable to read without offering any instruction on the use of reading strategies such as decoding. At the end of the intervention, there were no significant differences in reading growth between the two treatments. However, the lack of instruction provided by the adult listener coupled with the lack of time for students to independently solve unknown words likely limited their ability to create structures that could support their future reading attempts. Conversely, as evidenced above, when students are given the opportunity develop word solving strategies via effective reading instruction on complex texts, reading achievement is positively affected
Alternatives to Ability Grouping in Elementary Reading Instruction

Based on the evidence above, as well as the evidence that follows, perhaps less attention should be directed towards matching text to reader, and more focus ought to be placed on matching instruction to the interactions between text and reader. According to Shanahan (2011), a renowned researcher in literacy development, “Instructional level theory posits that the text difficulty level relative to the student reading level is the important factor in learning. But that ignores the guidance, support, and scaffolding provided by the teacher” (web log). In practice, instructional level theory often results in the poorest instruction for the lowest readers who have the least opportunity to learn when they are placed in the lowest reading groups (e.g., Allington, 1983; Condron, 2008; Chorzempa & Graham, 2006). The studies below evince that when teachers provide adequate instruction for all learners, the daily practice of grouping students by ability is not only unnecessary, but also less effective than other approaches.

As a researcher, curriculum coordinator, and first-grade teacher respectively, Cunningham, Hall, and Defee (1991) developed an approach to reading instruction that met the needs of a diverse group of students without resorting to the practice of ability grouping. The study was conducted in Defee’s classroom located in a large elementary school in the southeastern part of the United States. Reading instruction was purposefully designed to combine all four major approaches to reading instruction, including the whole group basal approach with gradually increasing levels of text difficulty, the phonics approach which focused on letter/sound relationships, the literature approach which exposed children to a wide range of quality literature, and the writing approach in which students learned to read by reading their own writing and that of their classmates.
They titled this approach the Four Blocks and all students were instructed using grade level materials. Students’ progress was observed daily and instruction was carefully planned and adapted as necessary, which included an additional intervention for one student who made little progress. At the end of the year, results on a reading test indicated that the Four Blocks approach was effective for children of all levels, especially those who would have been placed in the lowest reading group. Furthermore, eight years later, the authors reviewed data on the same group of students and found that as the elementary years progressed more and more students were able to meet grade level expectations (Cunningham et al., 1998).

Stahl and Heubach (2005) also developed their own program of reading instruction, titled Fluency Oriented Reading Instruction (FORI). Designed to increase automatic word recognition and reading fluency, the program included repeated readings, partner reading, choice reading, and reading at home. These practices increased the amount of time students spent reading both at school and at home. In their study of second grade readers, the significant amount of scaffolding provided from these practices enabled all children to read the grade level basal reading text and select independent reading materials that were more difficult than suggested by traditional leveled reading programs. FORI instruction served to reduce the achievement gap by setting high expectations for all students and providing the scaffolded instruction needed for success. As a result, in the first year of implementation students made an average gain of 1.88 years and 1.77 years in the second. These results ultimately led Stahl and Heubach (2005) to reject traditional notions of text difficulty levels, such as those promulgated by Betts (1946) and instead “argue that the instructional reading level for a given child is inversely
related to the degree of support given to the reader. That is, the more support given, the lower the accuracy level needed for a child to benefit from instruction” (p.55).

Kuhn and colleagues (2006) further explored this premise in an experimental study examining the effects of FORI in comparison to a wide-reading approach on fluency gains for second graders. While the two programs are similar, FORI focuses instruction on a single story repeated over the course of the week; in the wide-reading approach students are introduced to three different texts each week. In both treatments, students read grade level texts rather than texts at their presumed ability level. The researchers examined whether repeated readings or wide exposure to texts was more beneficial to fluency achievement. They found that both programs were more effective at increasing fluency and comprehension skills than the control group, but the benefits of wide reading emerged earlier. In relation to providing instruction to match text difficulty, it is particularly important to note that in this study the lowest readers were provided with an additional intervention focusing on their phonological development. This is an exemplar of the responsiveness of teachers in increasing instructional supports to ensure that all readers are successful when working with difficult texts. Other studies also documented the positive effects of wide reading (Kuhn, 2005; Schwanenflugel et al., 2009) and FORI (Schwanenflugel et al., 2009) on reading fluency and comprehension.

Reis, McCoach, Little, Muller, and Kaniskan (2011) corroborated this finding in their experimental study of the Schoolwide Enrichment Model-Reading (SEM-R) of second through fifth grade students. In this approach, all students were exposed to high quality literature and discussions via read alouds, along with individual conferencing to support the independent reading of texts that were approximately one to two grade levels
above their current independent level. Extension activities, including buddy reading, discussion groups, and projects followed. In this program, in which reading was individualized rather than group-based, students using the SEM-R approach had the same or higher scores on measures of reading fluency and comprehension than the control group depending on the school context. This suggests difficult texts combined with individualized supports can yield positive results for students.

All five alternative approaches to reading instruction described above (dyad reading, Four Blocks, FORI, wide reading, and SEM-R) eschewed traditional notions of text difficulty that systematically limit students’ access to grade level texts and instruction. As a result, all five approaches resulted in positive results on reading achievement. Even results that showed no differences from the control group (e.g., Reis et al., 2011) are arguably positive considering the inequitable practice of ability grouping was replaced with an approach with equally effective results. In combination, these studies provide a strong argument against the practice of grouping students by their ability and providing them with low-level texts that fail to result in sufficient reading growth.

These effective alternatives to ability grouping share several important features. First, all of the approaches expose all students to grade level instruction and content. When considering the research on students’ opportunities to learn, this instruction is likely critical to student attainment of grade level proficiency in reading. Students cannot be expected to learn what they are not taught. This is not to say that materials far above students’ reading level should be presented to them in the proverbial sink or swim fashion. Instead, literacy instruction must also include the second thing all alternative
approaches explored above have in common: scaffolded instruction. Echo reading, partner reading, choral reading, and reading at home are just a few of the ways students of all levels were supported in their attempt to access grade level text. Furthermore, additional targeted instruction in the skills students needed to bridge the gap between their current level and that of grade level proficiency was provided in several of these studies. This individualized instruction is not the core of the program, but rather an additional scaffold to accelerate students’ progress. For example, after six weeks of limited progress, Cunningham, Hall, and Defee (1991) met with one student daily for an additional ten minutes to provide the support needed for success, and later they formed a small group of students that met flexibly to support individual needs. In the study on wide reading, Kuhn and colleagues (2006) formed a similar intervention group for the six lowest readers. However, these groups did not replace instruction with grade level text, but rather supplemented the instruction given to all students. In this way, high expectations and a pace sufficient to master grade level material remained.

As evidenced above, instead of matching texts to readers and readers to reading groups, teachers can match instruction to the interaction that occurs between readers and texts of various difficulty levels. The research explored here provides evidence that calls into question traditional notions of text-reader matching and provides insight into the types of practices that guarantee access to grade level reading material for all students. According to Stahl & Heubach (2005), such access can begin a cycle of success.

Since our struggling readers had more exposure to the materials, through additional readings at home and through some additional work in class, they were able to read materials of much greater than expected difficulty. In turn, the
reading of more difficult material aided their growth as readers, allowing them to read second-grade material with more ease. (p. 56)

Such exposure to grade level texts is systematically denied our lowest readers in classrooms in which ability grouping is the norm. Easy text selection that offers the lowest readers the fewest chances to scaffold and structure their learning is just one of the many ways in which students are denied the opportunity to learn in the lowest groups (e.g., Allington, 1983; Chorzempa & Graham, 2006; Condron 2008; Gamoran, 1992; Oakes, 2005) Referring once again to the differential effects hypothesis (Lleras & Rangel, 2009) students placed in the lowest reading group learn substantially less than higher grouped students as well as their nongrouped counterparts (Condron, 2008; Lleras & Rangel, 2009).

Proponents of ability grouping rest their argument primarily on its effectiveness to efficiently meet the instructional needs of a diverse group of learners (Braddock & Slavin, 1992; Gamoran, 1992; Slavin, 1987, 1990, 1993). If, however, students can successfully read texts outside the levels demarcated by traditional notions of text difficulty, as evinced by the research presented here from multiple classrooms of various grade levels across the United States, the argument for ability grouping quickly erodes. That is not to say that practices of individualized differentiation are implicated in the process. On the contrary, responsiveness to the diverse needs of individual learners is heightened as instruction is finely tuned to meet the needs of each learner. Fisher, Frey, and Lapp (2012) argue, “The text difficulty level is not the real issue. Instruction is. Teachers can scaffold and support students, which will determine their amount of learning and literacy independence” (p. 7). In several studies explored within this section,
additional instruction was systematically provided to readers who demonstrated the need for extra assistance (e.g., Cunningham et al., 1991; Stahl & Heubach, 2005). Such instruction is likely to be even more responsive to individual needs than that in ability grouped systems in which the true range of abilities is much wider than expected but masked by group designations (Pallas et al., 1994).

**Theoretical Framework: Oakes’s Dimensions of Change**

Given the current prominence of ability grouping practices for literacy instruction in early elementary schools, simply attending to the nature of reading development and text complexity are unlikely to result in systemic changes in instructional practice. For decades, alternative approaches to ability grouping structures have proven themselves equally if not more effective than the current inequitable practice (e.g., Kuhn et al., 2006; Stahl & Heubach, 2005, etc.), yet none have engendered salient and long-lasting changes in elementary reading instruction. To better understand the complexity of educational reform, Oakes (1992) describes three dimensions of change that must accompany reform in ability grouping practices: the technical, the normative, and the political. These three dimensions will be used in this study to guide and evaluate the design process in ways that attend to the multiple criteria needed for change.

The *technical* dimension, which includes changes to curriculum, instruction, and assessment, has been the most widely addressed dimension in efforts to replace ability grouping practices (Oakes, 1992). Studies on various approaches such as FORI, Four Blocks, wide reading, dyad reading, and SEM-R successfully addressed the technical aspects of change. These studies explored instructional practices that lead to increases in students’ reading achievement and served as exemplars of equitable alternatives to ability
grouping. However, these interventions ultimately fell short of describing the contexts surrounding the interventions. Has the context been restructured in a way that encourages these changes or is there pressure on the classroom teachers to return to business as usual? Do school or district structures support these changes or do they make it easier for teachers to return to traditional practices? Thick descriptions of contextual features provide information on how these technical dimensions function in realistic settings and offer information regarding both the strengths and challenges of the new approach (Design-Based Research Collective, 2003). Understanding these components is essential to the spread and sustainability of educational reform in this area.

The collaboration between Cunningham, Hall, and Defee (1991, 1998) in the Four Blocks approach was insightful in effectively creating long term, although not far-reaching, change. In this study, the collegial relationship between a researcher, curriculum specialist, and classroom teacher resulted in sustained educational reform nearly a decade after the original study ended. Together this trio understood and accounted for not only the technical dimension, but also the normative and political dimensions as will be explicated below. Oakes (2017) explains, “Technical solutions are surely needed, but they are not enough. Inequity and exclusion reflect and are sustained by cultural norms and power relations” (p. 96).

The *normative* dimension of change refers to the traditions and norms that sustain educational practices. In terms of ability grouping in the United States, over 200 years of slavery, discrimination, and segregation in education serve as the backdrop for current norms that favor practices designed to provide differential access to opportunity based on supposed merit. This false pretense of meritocracy is so deeply ingrained in the national
norms that schools are accustomed to inequity. These norms must be explicitly confronted, questioned, and revised. This requires attention to mindsets, beliefs, and attitudes of teachers, administrators, and even students. Oakes (1992) writes that this process,

asks people to challenge their entrenched views of such matters as human capacities, individual and group differences, how schools and classrooms should be organized, and, ultimately, whether sorting students to prepare them for a differentiated work force with unequal economic rewards is what schools should do. (p. 19)

Educational innovations must also fit within the structural norms of the classroom. Designing an intervention that asks teachers to make grand systemic changes that require unavailable time and resources are likely to end as quickly as the researcher leaves the site. Instead, efforts must be made to balance the ideal with the realistic. Cunningham and colleagues (1998) found,

Teachers who are widely criticized for not being willing to try anything new will change when the innovation has lots of familiar elements, is doable within the time frame and materials they currently have, and results in observably better readers and writers. (p. 662)

In other words, when innovations fit within the current norms, long-term adoption is more likely. In the case of ability grouping, some norms such as inequity must be adapted to fit the intervention; on the other hand, the intervention must be adapted to fit other norms such as time and resource availability.
Finally, the political dimension accounts for the new relations that must develop in the school. Ability grouping practices are utilized by schools, but they are sustained by the interests of educational stakeholders, such as parents, students, and policymakers. Often, the politics of compared advantage result in parents advocating for their children to be placed in the highest group knowing that this placement will widen future opportunities. This often occurs through the intergenerational transmission of advantage as those parents who themselves were advantaged know how to access similar resources for their own children (Oakes, 2005). Thus, changes in ability grouping practices require the redistribution of power (Oakes, 1992). Systemic policies and structures must be confronted and people at multiple levels of decision-making must be on board. For example, even if teachers are informed of the research on ability grouping and subscribe to a discourse of equity, higher levels of power such as administrators and curriculum programs may thwart teachers’ autonomy in determining grouping practices (Park & Datnow, 2017). To increase the likelihood of forming a coalition that advocates for all students, detracking efforts must ensure that the new innovations “will create educational opportunities that are at least as rich and rigorous as those previously enjoyed by high students” (Oakes, 1992, p. 19).

As Oakes (2008) writes, “absent a concerted effort to establish new norms about whether and how all students can learn to high standards, high-stakes incentives seem to have driven well-intentioned people to intensify the old patterns of differentiation and inequity” (p. 707). As a result, decades after the debate of ability grouping seemed to end, we are once again faced with increasing numbers of students whose instruction is limited by their current level of literacy achievement, which is often simply a proxy for the
amount and quality of early literacy experiences they have encountered. Before any sweeping changes in ability grouping practices in elementary reading instruction are likely to occur, research is needed that attends to the normative and political aspects of change that must accompany the technical (Oakes, 1992). According to the Thomas Theorem, “If men define situations as real, they are real in their consequences (as cited in Merton, 1948, p. 193). It is time to stop defining students as low readers and begin defining the instructional approaches that will better serve them.
CHAPTER III

METHODODOLOGY

Purpose Statement

Elementary school students are often placed into groups with peers of similar reading ability in a practice called within-class ability grouping (Loveless, 2013). Through this practice, students are differentially exposed to reading skills, strategies, and texts that are presumed to match their current levels of ability. This widespread practice is particularly problematic given that (1) current notions of matching early readers to the appropriate levels of text difficulty for reading instruction are based on tradition rather than empirical evidence (Brown, 2009), (2) poor and minority students are overrepresented in the lowest ranked groups (Braddock & Slavin, 1992; Condron, 2008) (3) students in higher ranked groups make greater academic gains than lower ranked groups (Tach & Farkas, 2006), and (4) teacher perceptions of students ability are often inaccurate (Ready & Wright, 2011). Conversely, several studies have shown that when students are presented with texts of increased difficulty and given adequate instructional support, they are able to make accelerated reading progress (e.g., Benjamin & Schwanenflugel, 2010; Johnston, 2000; Kuhn et al., 2006; Morgan et al., 2000; Stahl & Heubach, 2005). The purpose of this design-based research study was to collaboratively and iteratively design an approach to primary reading instruction with a kindergarten teacher leading to the development of practical and theoretical insights on alternative grouping strategies and the necessary instructional supports to increase the reading achievement of all students.
Research Questions

The research questions at the heart of this study are based off of Bakker’s (2018) suggestions for targeted research questions in design-based research projects. Therefore, the central question of this study is: (RQ1) How can a learning and teaching strategy focusing on heterogeneous grouping practices for guided reading support the reading achievement of students of all levels of reading ability? Inquiry into this question will result in the exploration of the following subquestions:

- RQ2: What is an appropriate goal for the reading achievement for kindergarten students?
- RQ3: What is a teaching-learning design that would help students to achieve this goal?
- RQ4: How well was this design implemented?
- RQ5: How well did students meet the specific reading goal?
- RQ6: What would an improved design look like?

Exploration of Design-Based Research

Nearly thirty years ago, Brown (1992) and Collins (1992) are credited with the emergence of design-based research (Reimann, 2011). This relatively new and evolving research paradigm was designed to attend to the problem of the gap between educational research and instructional practice (Dede, 2004). Since that time, it has gone by many names, including design research, formative research, design experiments, and developmental research (McKinney & Reeves, 2012). Modeled after design sciences, such as aeronautics and artificial intelligence as noted by Brown (1992), “design experiments have both a pragmatic bent –‘engineering’ particular forms of learning -- and
a theoretical orientation -- developing domain specific theories by systematically studying those forms of learning and the means of supporting them” (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003, p. 9). Design-based research is not a new methodology per se, rather it relies on qualitative and quantitative data collection and analysis techniques to design educational innovations (McKenney & Reeves, 2012).

Anderson and Shattuck (2012) outline eight characteristics of design-based research. (1) It often uses both quantitative and qualitative data, as described above, to tell the whole story of what occurred throughout the research. (2) It occurs within the context of a real classroom or other educational context, rather than a laboratory setting. (3) The focus of the research is on designing and testing an educational intervention in direct collaboration with practitioners. (5) The design team documents design principles, which reflect the necessary conditions for success of the intervention. (6) These principles are a result of multiple iterations of the design which are designed to directly impact educational practice. (8) It is inherently different than action research in that it seeks not only practical implications, but also the creation of theoretical insights. According to the Design-Based Research Collective (2003), “Design-based research, by grounding itself in the needs, constraints, and interactions of local practice, can provide a lens for understanding how theoretical claims about teaching and learning can be transformed into effective learning in educational settings” (p. 8). This marriage of theory and practice allows researchers to investigate not only what works, but also how, when, and why it works as well (Cobb et al., 2003).

Design-based research is a particularly strong fit for this study given the research questions as well as the identified gap in the literature. Throughout the years, researchers
have called for research on grouping practices that is both naturalistic and accountable for the lived experiences of teachers and students (e.g., Hamilton & O’Hara, 2011; Hong et al., 2012; Oakes, 2017). Several authors have also encouraged research that documents what is feasible in real classroom settings, the challenges that real teachers face, the consequences of detracking, and the support that is needed to implement alternative grouping practices (e.g., Nomi, 2010; Oakes, 1992; Puzio & Colby, 2010; Tach & Farkas, 2006). Furthermore, Slavin (1987) suggests that not only is more research needed on within-class ability grouping in reading in the primary grades, but also that descriptive research studies ought to develop theoretical insight to encompass the research findings. Although over 40 years have passed, a literature-base focusing on the interaction between theory and practice in early literacy grouping structures continues to be absent. This study sought to provide a foundation for that work upon which future research can build.

As a central tenet of design-based research, special attention was given to the real concerns, such as increases in diversity, that real teachers face in real classroom settings. Alternatives to ability grouping practices must be more than idealistic; they must be realistic in terms of students’ needs, teachers’ time, and availability of resources. Design-based research allows for on-the-ground instructional design and innovations that “result in greater understanding of a learning ecology-- a complex, interacting system involving multiple elements of different types and levels” (Cobb et al., 2003, p. 9, emphasis in the original). As suggested by Reimann (2011), employing design-based research will capitalize on the knowledge and expertise of classroom teachers to find an alternative that is both feasible and sustainable. Furthermore, design-based research demands an iterative
design process through which interventions are constantly designed, enacted, analyzed, and redesigned to respond to the realities of the classroom (Zheng, 2015).

As the multiple real-life variables in these studies are documented and analyzed rather than controlled and removed, the relationships between and among theory, design, and practice can be studied (Design-Based Research Collaborative, 2003). Attending to Oakes’s (1992) Dimensions of Change framework, design-based research allowed the focus to be placed not only on the technical aspects of the innovation, but also on the norms and politics surrounding them. Attention to all three dimensions was necessary for the potential spread and implementation of the design in broader educational contexts. According to Anderson and Shattuck (2012) design-based research “seeks to increase the impact, transfer, and translation of education research into improved practice” (p. 16). This was the ultimate goal the research study presented here.

Rationale for a Qualitative Design-Based Research Approach

While design-based research can include both quantitative and qualitative approaches to data collection, this study relied primarily on qualitative research methods. Qualitative inquiry best fit the proposed study for several reasons. First, qualitative designs uniquely allowed for the contextualization of the data that was collected (Creswell & Poth, 2018), which is not only a core characteristic of design-based research, but also is particularly important in understanding the political and normative dimensions of change (Oakes, 1992). Second, qualitative research enabled the “thick descriptions” (Geertz, 1973) necessary to share the instructional design process with other educators interested in adopting the innovation. To highlight this importance, Cobb and colleagues (2003) write in reference to the instructional design, “‘What works’ is underpinned by a
concern for ‘how, when, and why’ it works, and by a detailed specification of what, exactly, ‘it is’” (p. 13). The description of the “what” of the design allows for adoption of the design by practitioners, while the how, when, and why descriptions were instrumental in theoretical development. Finally, qualitative research allowed for the fluidity of boundaries between researcher and participants as well as flexibility in intervention designs. It allowed for change to be documented, rather than controlled, and was responsive to and representative of real-life in the classroom. As is the case in many design-based research studies, quantitative data was also be included in this study, but was limited to descriptive statistics of the students’ literacy progress on benchmark assessments.

Several critiques of design-based research were heeded throughout the study. Dede (2004) admonishes that most design-based research is “under-conceptualized and over-methodologized.” In other words, there is often too much data collected which results in perfunctory analysis yielding common sense results without theoretical contribution. Given the qualitative nature of my research, an abundance of data was collected via participant observations, interviews, and data collection. Throughout the study, careful attention was placed on the analysis process to delve deeply into the theoretical rather than remain at surface level conclusions. The goal throughout the entirety of the study was to develop a sharable theory regarding early literacy grouping structures.

Furthermore, Brown (1992) cautions against what she terms the Bartlett effect, named after a researcher who selectively chose amidst the data corpus to prove his theory. According to this effect, when a researcher has a large data corpus, especially
common in qualitative research, the researcher may be more likely to select that data which validates the researcher’s claim while ignoring the rest. It is therefore, of great importance, that the researcher examines her beliefs, biases, theories, and fundamental assumptions throughout the course of the study and remains open to alternatives. It is with this in mind, that I began the process of unpacking my own beliefs and experiences.

**Philosophical Paradigm**

According to Bogdan and Biklen (2007), critical social science work “is critical of social organization that privileges some at the expense of others” (p.22). Through the lens of this philosophical paradigm, ability grouping is viewed as a systematically unjust practice. As explored in the research above, ability grouping privileges high achievers at the expense of low achievers, white students at the expense of minority students, and middle-class students at the expense of poor students. Furthermore, critical theory is shaped by the ontological assumption of “historical realism” meaning that reality exists, but it is influenced and reinforced over time through the constructs of gender, race, class, and culture among others (Lincoln, Lynham, & Guba, 2018). This is reflective of the political and normative dimensions of change (Oakes, 1992). Creswell and Poth (2018) explains that epistemologically this reality is not only understood through the study of “social structures, freedom and oppression, power and control” (p.36), but methodologically those social structures can, and should, be changed through social science research and subsequent actions. Design-based research serves as a platform of both action in attending to the present, and insight in attending to the future.
Researcher Positionality

I closely align with the epistemological and ontological view of critical social science explored above. As an elementary teacher for nearly a decade, my work has been driven by creating social change in my classroom. I have witnessed the multiple realities that are constructed for students in my classroom based on historical inequities that exist as a consequence of race, gender, culture, and especially social class. I am particularly interested in the ways in which social class and its subsequent realities influence school achievement, especially in the area of early literacy development.

After working at a low-income elementary school for several years and being reminded of our failing status according to No Child Left Behind, I made a purposeful transition to work at a middle-class school to better understand the ways in which social class interacts with literacy achievement. I began to see firsthand the differences in early literacy knowledge with which children entered school. This was not at all surprising, as I was already well-aware of the literature documenting the early literacy gaps that exist by class (National Early Literacy Panel, 2008). However, throughout my tenure there, I slowly became mindful of an unsettling reality. It was not so much the past literacy experiences (or lack thereof) that resulted in below grade level reading, but rather the ways in which the norms in school limited the opportunities for students to advance their literacy knowledge.

I witnessed the prolific use of within-class ability grouping for reading instruction in both schooling contexts; I even subscribed to this practice. It was the way I was trained in my preservice literacy methods course and it was reinforced by the expectations of the district in which I was serving. Many of my students entered my classroom already
considered to be low-ability readers. They were grouped into a low-achieving cohort in my classroom for guided reading instruction and, uncoincidentally, they left my room as low-achieving readers, year after year after year.

When I transferred to a new school, I was initially unaware that it was a community of transitioning demographics. Though historically a middle-class school, changes in residential patterns and school programming, such as the introduction of an early childhood program, resulted in a much more economically and academically diverse community. As a result, my highest reading group read texts at levels far above those in my lowest groups. Based on the logic of ability grouping, I taught those with the most literacy experiences the highest level of instruction, and those with the fewest literacy experiences remedial levels of instruction. After teaching kindergarten in this manner for a few years and seeing the results, I slowly began to reason that there was no way the students in my lowest groups would ever reach grade level proficiency, let alone catch up to their more experienced peers; I was not providing them with the instruction they needed to get there. I remained a proverbial cog in the machine that promulgated the Matthew effect which effectively helped the literacy rich get richer and ensuring the literacy poor get poorer.

I decided to create an alternative to this practice. During the last two years of my tenure, I began grouping students heterogeneously for guided reading instruction. All students received instruction on the reading skills and strategies necessary to reach grade level proficiency. In addition, I disregarded the traditional notions of text difficulty and provided all of my students with access to grade level text, and the scaffolding necessary to read them successfully. To say I had promising results is an understatement. In two
years, all of my students met grade level expectations for reading fluency and nearly all
for reading comprehension. This number includes special education students, emerging
multilinguals, and students identified with language processing concerns. Furthermore, in
counter to those claiming heterogeneous grouping is detrimental to high ability students,
my highest achievers also performed at higher levels than my highest students in the
previous years. When my instruction matched the interactions between reader and text,
all students made significant progress.

As a researcher, I am committed to expanding from the practical aspects of this
work to the theoretical. Consistent with the tenets of qualitative research, I do not claim
to be an unbiased scientific instrument, nor do I think that is possible for any researcher. I
know that my philosophical beliefs as well as my past experiences remained a part of me
throughout the study. Erickson (1984) writes of this phenomenon, “It was I who was
there doing the fieldwork, not somebody else. My fundamental assumptions and
prejudices are part of me. I cannot leave them home when I enter a site” (p. 60, emphasis
in the original). It was with this in mind, that I attempted to minimize the effects of my
fundamental assumptions, biases, and interpretations by writing memos after each field
visit to capture and analyze my own thoughts throughout the research process. I reflected
on the ways in which my identity and experiences as a teacher might have resulted in bias
favoring the design, and continued to remain open to new emerging ideas and insights
presented in the data and throughout the analysis process.
Research Design

Research Site

Consistent with qualitative research methodology (Creswell & Poth, 2018; Merriam & Tisdell, 2016) and due to the importance of the context in design-based research, I used criterion sampling to select a school with a reputation for welcoming innovative practices. Due to the entrenched nature of ability grouping, this criterion was essential in being accepted into a school community in order to do research that asked a teacher to try instructional approaches rejecting the status quo. Without such building level support, I would not have been able to conduct my study. In addition, I quickly realized how hard it was to gain teacher participants as I pitched my research idea to many who thought it was a “great idea” but simply couldn’t “imagine how to make it work without grouping students by ability levels.” I ended up selecting a fantastic school, although its students were admittedly less diverse and of a higher social class than I had originally imagined for my study.

The site selected, referred to throughout this study as King Elementary School, serves preschool through fifth grade students in a Midwestern City in the United States. It is a parochial school and has developed a reputation in the community as a good school, offering competition to the local public school for students whose parents are “school shopping.” King elementary is located in a middle-class neighborhood and costs several thousand dollars to attend each year. There are scholarships to attend this school, but certainly financial stability plays a role in determining the student population. In addition to support from the administration, given the school’s parochial status, teachers at this school have much more instructional autonomy than their public school counterparts.
Research Participants

Criterion sampling was also used for participant selection. According to Dede (2004), it may be difficult to find practitioners that are willing to participate in design-based research due to its iterative and loosely-defined nature. I found this to be very true in my recruitment efforts. Not only did the study require a long term stay in the participants’ classroom, but it also required the participant to be a true collaborator in the design and implementation process (Reimann, 2011). My criteria for participation was that the teacher had at least five years of teaching experience and was willing to be a partner in designing, evaluating, and redesigning the instructional innovation as necessary to best meet the needs of his or her students. Since kindergarten is most often the first grade in which ability grouping is utilized, the additional criterion of being a kindergarten teacher was added. After recruiting one kindergarten teacher based on the selection criteria, all students her classroom were also invited to participate in the study.

The teacher participant selected was a highly qualified, veteran teacher known throughout this study as Mrs. Poppy. With experience teaching in a variety of schools and grade levels and an advanced degree in early literacy development, she was a prime candidate for partnership in this study. She had thirteen students in her classroom, and all thirteen students, given the consent of their parents and/or guardians, agreed to participate. This low number of students was unusual for Mrs. Poppy, as she had nearly double that number the previous year. Another kindergarten teacher was hired right before school started due to the rising number of kindergarten students. The student diversity in terms of race, ethnicity, and social class was limited, but representative of the surrounding community.
Timeline

McKenney and Reeves (2012) explicate three core phases that are undertaken during a design-based research study. These phases can occur multiple times throughout a given study. While they do not always occur in linear fashion, they will be explained as such for purposes of clarity. The process begins with *analysis* and *exploration*, which includes initial orientation to the research problem, a focused literature review, field-based investigation of the problem, site visits, professional meetings, networking, and goal development. This dual process of analysis and exploration is followed by *design* and *construction*. During this phase, design solutions are explored, mapped, and constructed through the creation of prototypes. Finally, during the *evaluation* and *reflection* phase, guided questions are framed, methods and strategies are determined, data is collected and analyzed, and findings are reported.

According to this framework, a design-based research study unfolds over several years through multiple iterations of these six core processes. To better understand the overall design process, McKenney and Reeves (2012) further differentiate between micro-, meso-, and macro-cycles of design. A micro-cycle of design consists of just one core phase, such as *design* and *construction*. The work in one micro-cycle determines which core processes will be addressed next. For example, a researcher’s work during the micro-cycle processes *analysis* and *exploration* will inform his or her work in the subsequent *design* and *construction* micro-cycle. However, this micro-cycle may lead to new information or questions that forces the researcher back into another micro-cycle of *analysis* and *exploration* rather than onto the *evaluation* and *reflection* phase. A meso-cycle contains more than one core phase, but less than a complete project. A meso-cycle
may even include all three core phases as long as evaluation and reflection lead to further analysis and exploration. Finally, a macro-cycle is a design-based research study in its entirety. A macro-cycle may contain several micro- and meso-cycles. Due to the multiple iterations of design in this type of research, knowing when the study is truly complete can be a challenge for design researchers (Anderson & Shattuck, 2012).

Using McKenney and Reeves’s (2012) structure, the macro-cycle of this design-based research began in the Fall of 2016. Throughout the course of two academic school years, I enacted the instructional design process to find an alternative to ability grouping practices in my own classroom. Several iterations have evolved during this time as I fine-tuned the design. Focused mostly on practical implications rather than on the theoretical insights, the first two years of this project might best be described as action research rather than design-based research (Anderson & Shattuck, 2012). The study presented in this document is an extension of the design used in my own classroom and could therefore be considered the third meso-cycle in the overall design process.

While the initial design was created and tested within the context of my own classroom, the aim of this portion of the project was to refine the design and construction of the design principles as well as to analyze and evaluate its success in another classroom under the instruction of another teacher. Data collection for this study took place from November 2018 until mid-March 2019, or two of the four academic quarters of the school year. The first phase of data collection, November through December (quarter 2), focused on gaining access to the setting, building rapport with participants, (Creswell, 2018) and familiarizing myself with the curriculum and classroom routines. The second phase of data collection, January through mid-March (quarter 3), focused on
designing, implementing, and redesigning an alternative to ability grouping with the classroom teacher.

**Initial Design Framework**

Based on two years of action research and review of the literature on ability grouping, the following design principles served as a springboard for this study. Each design principle is written in the format suggested for design-based research principles suggested Van den Akker (2013 as cited in Bakker, 2018).

*If you want to design guided reading instruction in which heterogeneous grouping is used to support the literacy achievement of all students, you are advised to:*

- **Distribute students equitably across groups, paying special attention to reading achievement, native language, and behavior** because this increases the opportunities to learn for all students.

- **Pull students who are reading significantly higher (at least one grade level) than all other students into their own group for instruction.** An emerging hypothesis is that students who are reading well beyond grade level content may need to continue to be pulled into their own group. This is supported by research on the gifted (e.g., Kulik & Kulik, 1987).

- **Devote almost the entire fifteen minutes of guided reading instruction to in-text reading** because students need to spend significantly more time practicing reading skills and strategies in more challenging texts.
• **Select text that is at grade level or higher** because students will not be able to read grade level text if they have no exposure to it. Students need time to practice reading texts at the level they are expected to master.

• **Provide the instruction and practice necessary for students to be successful at grade level text or higher** because students cannot be left to unproductively struggle while reading a text. Students may need a richer book introduction, oral practice using the language of the text, and additional teacher guidance while reading. They may benefit from participating in two guided reading groups per day to allow for sufficient practice of the text.

• **Carefully plan guided reading instruction based upon the needs of all students in the group** because all students need to benefit from guided reading instruction. For example, if the highest achieving students in the group are ready to learn digraphs and the lowest achieving students still need more practice reading short a words, then during the word work portion of guided reading teach all students digraphs but focus on reading and writing words that also have the short a sound (e.g., catch, that, sham). In this way, all students will receive instruction and practice on necessary skills while still being exposed to grade level content.

• **Synchronize instruction across contexts** so students, particularly those who find reading to be challenging, have a cohesive program of instruction. As much as possible, all of the teachers working with the students, (e.g., classroom teacher, English language teacher, special education teacher, reading intervention teachers, paraeducators, etc.) should align instruction so that students are hearing the same instruction and the same language across contexts.
• Provide explicit instruction on reading strategies because reading is a complex process that can be broken down with scaffolds to support students’ learning. When a student gets to an unknown word while reading, rather than providing students with a random assortment of strategies to try, teach an explicit problem-solving sequence, such as: (1) Put the first letter in your mouth. (2) Think what would make sense. (3) Blend the letters. (4) Reread from the beginning of the sentence. (5) Read ahead to the end of the sentence. When students are presented with multiple strategies, but no instruction on when to use them, strategy instruction becomes haphazard.

Data collection

During phase one of data collection from November through December 2018 (quarter two), I conducted participant observations (Spradley, 1979/2016) of literacy instruction one time per week for one and a half hours at various times of the day. Observations focused on the teacher’s instructional practices, the literacy curriculum, and the students’ literacy development. Jottings of all observations were recorded in a digital notebook during the observation or immediately afterwards as suggested by Emerson, Fretz, and Shaw (2011). The jottings were transformed into field notes with the intent to create “thick descriptions” (Geertz, 1973) of the classroom culture including the technical, normative, and political dimensions of design implementation (Oakes, 1992). My jottings and fieldnotes purposefully included the content, structure, and organization of the literacy instruction as well as the actions of both the teacher and students related to literacy instruction and development. The purpose of this phase was to gather the
information necessary to be a knowledgeable contributor to the design process in this particular classroom with these particular students.

During phase two from January to mid-March 2019 (quarter three), I maintained the role of a co-instructional designer, an observer, and an active participant in the classroom. I purposefully assisted in routine classroom happenings, such as passing out supplies, in an effort to become a routine part of the classroom to reduce the influence of my presence on the data. Additionally, my role as researcher also included modeling instructional practices (Reimann, 2011) similar to the support provided to classroom teachers by a literacy coach. I was in the classroom two times per week for up to three hours per day, including all components of literacy instruction. While the instructional design only directly focused on guided reading time, I purposefully observed the cumulative literacy experiences and instruction that served as the foundation to guided reading instruction. Observations served as the primary method of data collection in accounting for how the design functioned in authentic settings, including both the successes and failures of implementation (Design-Based Research Collective, 2003).

During phase two, Mrs. Poppy and I also had weekly collaborative planning sessions. These meetings typically occurred during her planning time when the students were at specials and lasted for approximately 20 to 30 minutes. At these meetings, we discussed the instructional design components already in practice, revised them as necessary to meet the needs of all students, and collaboratively created new ideas for implementation. We also routinely talked about the current instructional groups to ensure their heterogeneity as well as problem solved at the individual student level for those not making sufficient progress. While our meetings directly attended to the practical
components of guided reading instruction without the use of ability grouping, they also contributed to theoretical development as well.

Interviews were used to provide insight into what could not be observed (Merriam & Tisdell, 2016) including Mrs. Poppy’s perceptions and beliefs. I conducted a pre- and post- semi-structured interview with Mrs. Poppy focusing on her evolving beliefs regarding grouping practices and the nature of ability. The post-interview also attended to Mrs. Poppy’s perceptions of the progress and development of the instructional design. The interviews were audio recorded and transcribed verbatim for analysis. In addition, frequent informal interviews, more closely resembling conversations, were used to gather insights into the observed instructional decisions and her perceptions of students’ literacy progress.

Student work samples and literacy assessments administered by the classroom teacher were collected throughout the study as evidence of students’ literacy growth as a reflection of the instructional innovation. Mrs. Poppy assessed all students at the end of quarters two and three using Rigby PM Benchmark (2000) and/or MONDO Bookshop levels (2008). During the administration of both assessments, students were asked to read a short book, similar to those read during guided reading groups. Each book was written as a designated level of difficulty to monitor students’ progress over time. In order to pass a level, students were expected to read with a given level of accuracy and comprehension deemed proficient by the assessment (similar to those promulgated by Betts at the independent level). In addition, Mrs. Poppy administered assessments on letter name and sound identification, phonological awareness, concepts of print, and sight word knowledge. Since this study was designed to be attentive to the current assessment
standards of the classrooms, no additional assessments were given beyond those already administered by the classroom teacher to assess students for this study.

**Data analysis**

Data analysis focused on two goals: the identification of design principles for practical implementation and the development of theoretical understanding to contribute to the literature on alternatives to ability grouping practices. When identifying the design principles, Brown (1992) suggests paying special attention to the question “what are the absolutely essential features that must be in place to cause change under conditions that one can reasonably hope to exist in normal school settings?” (p. 173). The explication of these features and the inferences made enable others to both understand and critique the analysis process (Cobb et al., 2003). To differentiate design-based research from action research (Anderson & Shattuck, 2012) analysis must progress beyond the *what* of the design and into the *why*. According to Cobb and colleagues (2003) “A design theory explains why designs work and suggests how they may be adapted to new circumstances” (p.9). This theory will be “relatively humble” (Cobb et al., 2003) in that it will result in the understanding of learning within the specific ecological space of a few classrooms (McKenney & Reeves, 2012).

Analysis of both aspects began at the onset of the study to avoid the threat of over methodologizing and under conceptualizing the data (Dede, 2004). Furthermore Brown (1992) warns of the tendency to romanticize the research findings based on a few promising observations. Early attention to data analysis focused energy on what was occurring with special attention to data which appeared to be conflicting. Per Emerson et al.’s suggestions (2011), I wrote memos after each field visit that assisted in attending to
emerging patterns for design principles and the development of theory while still in the field as will be explicated below.

Throughout the data collection period, I used MAXQDA 18.1 software to organize, code, and analyze the data. Structural coding was used to code the data, including my analytic memos, based on the research questions. According to Saldaña, (2016), structural coding is a first cycle coding method which allows the data corpus to be coded by the research questions. It served as a preparatory organization phase for more detailed subsequent coding and analysis. I also coded the data at the individual student level so I could monitor the literacy growth of each student throughout the course of the study. Structural coding of each observation, collaborative meeting, and interview was also done immediately following data collection so that the analysis process could further inform the design. Through this process, I was able to begin to see how the instructional design was being implementing, evaluate existing practices, and develop new design ideas.

In addition, I conducted a second round of coding using the method of process coding (Saldaña, 2016) to analyze the specific actions of the teachers and students. Creating gerund-based codes assisted in determining the design procedures or the what of the instructional design. Process coding was particularly useful in explicitly attending to Mrs. Poppy’s systematic instructional moves decided upon during the design meetings, as well as those unconscious moves that provided insight into the normative and political classroom environment in which the design was implemented. In addition, process coding can “clue the researcher to a sequence of process in action” (Saldaña, 2016, p. 112) showing how change occurs over time. In this study, process codes, such as sharing
ideas with teammate enabled me to document a shift not only in the students’ literacy development, but also in the teacher’s perspective of the instructional design.

Several rounds of code mapping (Saldaña, 2016) followed to categorize and organize my data into design principles for further analysis. Initially all codes were listed randomly and carefully read through to ensure they captured the heart of the corresponding data. During the second iteration of code mapping, I sorted the codes into categories by comparing codes against one another. Several categories essential to the implementation and evaluation of the design were identified and considered. Eventually these categories were condensed to themes that were incorporated into the design principles. Throughout this process my hypotheses were developed abductively as suggested by Agar (1996). Patterns begin to emerge through the processes of close reading of fieldnotes and transcriptions, careful coding of data, and developing categories and themes guided by the research questions (Emerson et al., 2011).

**Research Ethics**

This study was approved by the University of Nebraska’s Institutional Review Board (see Appendix A). The adult participant in this study signed an informed consent form notifying her of the potential costs and benefits of participation. The parents and/or guardians of all student participants also signed an informed consent form for their child to participate in the study. Additionally, all student participants gave their verbal assent prior to participating in the study. Confidentiality and anonymity were granted to all participants.

According to Lichtman (2013), “Ethical behavior represents a set of moral principles, rules, or standards governing a person or a profession. We understand that to
be ethical is to ‘do good and avoid evil.’ (p. 51). Yet, in social science research, even a straightforward definition is plagued by the context specific, daily decisions a researcher must make (Neuman, 2011). Tracy (2013) delineates the differences between procedural ethics, situational ethics, and relational ethics (as cited in Merriam & Tisdell, 2016). While procedural ethics are those prescribed by IRBs (such as do no harm), situational ethics and relational ethics require the researcher to reason and be reflective of concerns that arise throughout the research process. Throughout this study, I have made daily choices in an effort to maintain the highest ethical standards in all areas.

Attention to situational and relational ethics was particularly important given that the literacy development of students was affected by the research study. While the literature review and previous meso-cycles of the design pointed to promising results for the students in this study, a different teacher, different students, and a different school context could interact in different ways with the research design. Fortunately, the direct collaboration with the practitioner as well as the ability to redesign the innovation throughout the study, enabled the research project to be accountable to both the teacher and student participants. If students’ progress appeared halted or the teacher felt the workload was overbearing, we were immediately able to alter the design.

Validity, Reliability, and Generalizability

The terms reliability, validity, and generalizable have been contested in terms of their use in qualitative research (Creswell & Poth, 2018; Merriam & Tisdell, 2016). While some researchers choose to use other terminology that they deem more appropriate to qualitative research (such as Lincoln and Guba’s (1985) use of the terms credibility, transferability, dependability, and confirmability), I believe the use of the more common
terms of validity, reliability, and generalizability will provide clarity for the larger audience even though the application of these terms in qualitative research is pragmatically different.

Creswell and Poth (2018) recommend that researchers utilize at least two validation strategies, or ways to assess “the accuracy of the findings as best described by the researcher and the participants” (p. 255). For this study, I engaged in five of the nine possible validation strategies suggested by the authors. (1) I engaged in researcher reflexivity by disclosing my biases and personal experiences with the instructional design within this study. Furthermore, before beginning data collection as well as throughout the duration of the study, I wrote memos to reduce the risk of such bias entering the data collection and analysis process. (2) I maintained rich collaboration with Mrs. Poppy as her expertise was essential to a design that fit within the context of her classroom. (3) Thick descriptions of the design and implementation are provided in the findings and discussion so the reader can assess whether the research claims are warranted from the data. (4) Triangulation of multiple data sources including observations, interviews, and documents was used to corroborate evidence to shed light on the design process. (5) I maintained prolonged engagement in the field (half of an academic school year) to build rapport with the participants and learn the culture of the classroom so as to reduce the influence of my presence on the data.

Obtaining reliability, or “the extent to which the research findings can be replicated” (Babchuk, Guetterman, & Garret, 2017, p. 4), is inherently different in qualitative research. Since human behavior is dynamic, replication is unlikely to lead to the same results as is expected in quantitative studies (Merriam & Tisdell, 2016). Instead,
in terms of reliability, qualitative researchers ask “whether the results are consistent with the data collected” (Merriam & Tisdell, 2016, p. 251). In order to ensure that the results were dependable, I kept an audit trail to demonstrate the rigorous methods of data collection and analysis employed throughout the study. In addition, much description and verbatim talk of the participants is provided in this document in order for readers to gauge that the results are consistent with the data. Furthermore, I have engaged in the process of research reflexivity in this document so readers can gauge whether my potential biases have influenced in the findings.

In terms of generalizability, also referred to as external validity or transferability (Babchuk et al., 2017), Lincoln & Guba (1985) write, “the burden of proof lies less with the original investigator than with the person seeking to make an application elsewhere. The original inquirer cannot know the sites to which transferability might be sought, but the appliers can and do” (p. 298). Therefore, in this study, it is my responsibility to provide both data and clarity of my interpretations in a way that enables readers the ability to transfer or extrapolate on my work. The provision of “thick descriptions” (Geertz, 1973) are evident throughout this document to allow for theoretical generalization in other contexts.
CHAPTER IV

FINDINGS

The instructional design of heterogeneous grouping for guided reading instruction was implemented between the months of January through mid-March. Each week, Mrs. Poppy and I met to discuss and evaluate the design as implemented as well as to determine the next steps for instruction. Since ability grouping practices are so entrenched in early reading instruction, the process of grouping students heterogeneously was relatively slow. Each week, I gently nudged Mrs. Poppy in the direction of the intended design as well as incorporated her own ideas in the process. Throughout the two and a half months of design implementation, Mrs. Poppy met with students approximately two to three times per week for small group guided reading instruction. During this instruction, students practiced sight words, independently read short texts with teacher prompting, answered comprehension questions, and participated in word work.

In the sections that follow, each individual student who participated in this study will be introduced, followed by a presentation of their literacy development over the course of their kindergarten year (Richardson, 2009), including their interaction with the instructional design. Although it is recommended that students are placed within (homogeneous) guided reading groups within the first six weeks of kindergarten, Mrs. Poppy did not begin grouping students for guided reading until the beginning of the third quarter (January) and then did so heterogeneously. Therefore, the results of the literacy progress of the thirteen students described below demonstrate the progress students can
make when all students are exposed to grade level literacy content in both whole group and small group contexts.

At the beginning of the study, Mrs. Poppy and I set a reading goal for her students based on the leveling systems that have been suggested by the MONDO Bookshop levels (2008) and Rigby PM Benchmark Assessment (2000), which were used as the assessment instruments at King Elementary. The MONDO Bookshop levels (2008) correlate with the Fountas and Pinnell levels (1999, 2001) and use a lettered system in which A is the easiest level and texts get progressively more difficult towards the end of the alphabet. The Rigby PM Benchmark levels, on the other hand, begin with a level one and progress in difficulty as the numbers increase. For the purposes of this study, the Rigby PM Benchmark levels will be reported as their correlate MONDO Bookshop/ Fountas and Pinnell letter designation to ease in the understanding of the findings.

*Table 4.1* provides an adaptation of the St. Paul School District’s (2009) text correlation chart for the MONDO Bookshop levels, Fountas and Pinnell reading levels, and the Rigby PM Benchmark levels. King Elementary was not part of the St. Paul School District, however, this was the chart used by Mrs. Poppy to guide her reading instruction and assessment. It provides not only the text level correlations between the three systems, but it also provides the grade level equivalent for each text level.

Mrs. Poppy stated that typically one or two students in her classroom every year are unable to pass a level C text, which is the benchmark goal for the end of kindergarten. Since this study ended at the end of quarter three (mid-March) rather than at the end of the year, we expected that all students would be able to pass at least a level B, or middle kindergarten level. Mrs. Poppy used both the MONDO Bookshop levels (2008) as well as
the Rigby Benchmark Assessment (2000) to assess her students. While she preferred the Rigby system, the early childhood team in her building required the use of the MONDO levels, so she switched between the two while assessing students.

*Table 4.1: Text Level Correlation Chart*

<table>
<thead>
<tr>
<th>Grade Level Equivalent</th>
<th>Fountas &amp; Pinnell</th>
<th>MONDO Bookshop Levels</th>
<th>Rigby PM/PM Plus Benchmark Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginning Kindergarten</td>
<td>A</td>
<td>A</td>
<td>Starters 1</td>
</tr>
<tr>
<td>Middle Kindergarten</td>
<td>B</td>
<td>B</td>
<td>Starters 2</td>
</tr>
<tr>
<td>End Kindergarten</td>
<td>C</td>
<td>C</td>
<td>3-4</td>
</tr>
<tr>
<td>Beginning First Grade</td>
<td>D</td>
<td>D</td>
<td>5-6</td>
</tr>
<tr>
<td>Middle First Grade</td>
<td>E</td>
<td>E</td>
<td>7-8</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>F</td>
<td>9-10</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>G</td>
<td>11-12</td>
</tr>
<tr>
<td>End First Grade</td>
<td>H</td>
<td>H</td>
<td>13-14</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>I</td>
<td>15-16</td>
</tr>
</tbody>
</table>

Adapted from St. Paul Public School District (2009)

**Students’ Literacy Progress**

**Casey**

When I first arrived in the classroom in November, before I got to know any of the students as readers or writers, I took the time to look through their daily independent writing journals. Most students at that time were writing approximately five to seven words in one sentence each day. Although I was quite surprised at how similar the
students were in terms of their writing skills, I picked out Casey as the “highest” writer in the classroom. He seemed to have the best grasp of grapheme-phoneme correspondences and consistently wrote a readable story with a strong idea. After looking through the November writing samples, Mrs. Poppy got out the August writing journals for me to take a look at. I was immediately surprised to find out that I would have ranked Casey as one of the lowest writers in the class at the beginning of the year.

*I asked Mrs. Poppy to tell me about this student as a writer while the students transitioned. She told me that he had grown a lot as a writer over the school year and she would consider him to be about in the middle of the group in terms of his relative writing ability.* (FN 11-30-18)

This discrepancy between Mrs. Poppy’s assessment of Casey’s current writing ability and my own observation was particularly interesting to me in terms of understanding how teacher perceptions are tied to the long-term knowledge of students. While as an outside source I rated this student as the highest in the class, Mrs. Poppy’s perception of this child was much different. Since this student started the year as one of the lowest in the class, perhaps she was unable to fully see the significant amount of progress he made. Casey was also a very confident writer and rarely asked adults or peers in the classroom how to spell an unknown word.

Although Casey was a fairly quiet child, he was always eager to engage in class discussions during interactive read alouds. Almost every time he was called on for both comprehension questions and questions about how to spell words (sight words and others), he was able to answer correctly. It was clear that he quickly absorbed instruction and was an active learner in the classroom. Casey’s ability to write consonant-vowel-
consonant words in isolation was stronger than most of his peers. Perhaps, it was due to his phonetic spelling that I picked him out as the highest writer early in the year. He consistently was able to represent all short vowel sounds, and his errors, although there were few, usually occurred with the final consonant.

Based on letter name and sound identification, phonological awareness, and concept of print assessments conducted throughout the first quarter of kindergarten, Casey’s scores placed him about in the middle of the class. At the end of quarter two, before guided reading groups started, Casey passed a level A text. Although his reading skills still placed him in the middle of the class range, he was slowly becoming relatively stronger. This score put him on track to pass a level B by the end of the third quarter and meet the goal of a level C by the end of the year. After only a little more than two months of heterogeneous guided reading instruction in which Casey was given the opportunity to read texts at a higher than suggested level of difficulty and read alongside peers at higher levels, he strongly passed a level D text, which is considered to be a beginning first grade level text.

**Zane**

Throughout my time at King Elementary, Zane was the life of the class. He was by far the loudest, most boisterous child in the room, but certainly still followed the classroom expectations. He was excited and wanted to share his ideas, always and to whomever would listen. He was also intensely curious and constantly wanted to know what I was doing while on my iPad or computer in the classroom. He demanded the attention of the teacher often and commanded the attention of his peers. Sometimes this got in the way of his literacy work.
[During writing time] Zane and Logan spent most of the time discussing where they should meet up for fun night. I prompted them to get started several times. Finally, Zane wrote, “I am meting (meeting) Logan at fun night.” (FN 1-25-19)

He also struggled to stay on task during partner reading opportunities and literacy centers, but somehow always seemed to be learning even when he tried not to do so.

Zane was part of this group and was hoarding all of the letters. He wasn’t “working” but would pass letters to other students as they needed them [to make words]. He often found the letter before they requested it so he was solving the words more quickly and accurately even though he wasn’t working on his own card. (FN 1-16-19)

Even when Zane didn’t appear to be paying attention, he often muttered the answers to Mrs. Poppy’s questions under his breath when others were slow to respond or didn’t know the answer.

Although I would consider Zane to be a confident student in almost all respects, he was surprisingly unsure of his abilities as a reader and writer. He asked me to spell words for him during writing time even though I always encouraged him to stretch them out and he was able to do so quite well. He was also able to read things in the classroom that other peers could not, such as the morning message and daily question. Yet, when asked to read independently on the rug, I saw Zane sit down with a book and loudly proclaim, “I’m bad at reading lots of books!” He repeated something similar again a few moments later as if he was talking to himself. (FN 1-9-19). As guided reading groups got up and running during quarter three, I noticed that he was quickly becoming one of the strongest readers in his group, making relatively few errors while reading. He was able to
read most consonant-vowel-consonant words automatically, without first blending the sounds as his peers did and was developing a rich sight word vocabulary.

At the end of quarter two before the study began, Zane passed a level A text. While this put him on track to meet grade level expectations, I was surprised based on my anecdotal evidence from November and December that he did not score higher. In combination with his other test scores, he ranked in the bottom third of the class. After being placed in a heterogeneous guided reading group, Zane quickly flourished while listening to the prompts Mrs. Poppy gave his peers and reading texts that would traditionally be considered at his frustration level. Beginning in mid-February, Zane was reading so well that Mrs. Poppy and I agreed he should start working in the group with two others who were reading mid-first grade level texts. At the end of quarter three, Zane impressed us all by passing a Level G text. Challenging text paired with individualized instruction enabled him to quickly rise to the very top of his class in terms of relative literacy achievement.

**Marcus**

Marcus was a quirky and silly kindergartener. While he was much less social than Zane, he always seemed to be up to something. It seemed as if he was always trying to get someone to laugh at him, but didn’t necessarily know how to go about it. For example, one day he incessantly poked everyone as he walked by them, while grinning widely. Though he clearly expected a laugh from his peers, his friends responded with looks of irritation. Another time, Marcus continued to put his foot on Mrs. Poppy’s big book and laughed and laughed each time he did it. He continued to do so even after a few reminders to keep his feet on the floor so Mrs. Poppy asked him to sit in the hallway for a
few minutes to calm down. For the duration of the study, Marcus was the only student I observed being asked to go to sit in the hallway due to behavioral disruptions.

Marcus was a very confident reader and writer and was one of the strongest in the class. In January and February, his writing skills skyrocketed and he began filling his entire page with words. For example, he wrote,

“My stuft (stuffed) anumuls (animals) are so hugee (huge) and they hug me to (too) and the uthr (other) wuns (ones) are his freins (friends) and they are cyut (cute) ok.” Then he proudly said “I could write for like two hours all over the table!” (FN 2-4-19).

After the students finished their work, Mrs. Poppy would “teacher write” underneath words that were spelled incorrectly. Sometimes she used it as an instructional moment to teach students a specific word or sound. When Marcus was finished with his writing, he would raise his hand and eagerly ask me if he spelled everything right. I always praised him for stretching out the words he wrote as I wrote the conventional spelling underneath to follow Mrs. Poppy’s procedures.

At the beginning of the study, Mrs. Poppy agreed to group students heterogeneously with the exception of Marcus and another student (Rebecca) who were reading significantly above their peers. I agreed this would be appropriate and necessary given his current reading level. This suggestion was also in line with the design principles crafted after the first two mesocycles of this work in my own classroom, as well as research on grouping students together who are significantly advanced. At the end of quarter two before guided reading groups started, Marcus passed a level E text. With the exception of Rebecca, the next highest reading level in the class was a level B which was
considered one full grade below Marcus’s current reading score. At the end of the third quarter, Marcus passed a level G text, which is considered to be a mid to late first grade level. While his fluency, including reading rate and accuracy, was very strong, his comprehension level held him back from passing further levels. This aligned with my observations of Marcus while answering questions during other components of literacy instruction.

Jill

Jill was a good friend to all students in her classroom. She was skilled at talking with peers in a mature fashion for her age, often organizing games and play time. She had the ability to think beyond herself and seemed to be the “mother hen” of the group. Before going out to recess, she would often help her friends with their outdoor wear, and even brought several pairs of mittens one day so she could lend them to friends who did not have any. She was always kind, polite, and never seemed to be involved in the squabbles that often happened between other girls in the class. She most often had a quiet presence, but assumed a leadership role when necessary.

At the beginning of the study, Mrs. Poppy told me that Jill had made significant progress up to that point. Although she started the year as one of the lowest students in the room in terms of literacy achievement, she was getting stronger and stronger. In December, I documented that her reading growth continued to be impressive.

*The wreath had the words “hope, Jesus, peace, joy,” and “love” on it. I asked Jill what the words said. She pointed to each candle and read each word with no problem. (These are pretty challenging words and despite the fact that Mrs.*
Poppy said she was one of the lowest students at the beginning of the year, she was able to read these words with no trouble.) (FN 12-4-18)

In January, Jill’s confidence began to grow. She asked me to read a Beauty and the Beast book to her during independent reading time, a text she often asked me to read. I told her I would like her to read to me instead and that she could read me the pictures. She looked at me and immediately said, “I can read some of the words!” (FN 1-9-19). We then shared the responsibility of reading this challenging text together.

When reading, Jill relied primarily on meaning to solve unknown words. Decoding words, particularly those with short e were difficult for her to read without a picture or context of a story. However, her continuous text reading was strong enough that at the end of January Mrs. Poppy told me that she was considering moving Jill into the first-grade texts with Marcus and Rebecca. I agreed, but after just a few days, Mrs. Poppy became worried about Jill’s level of frustration in the group despite the fact that she also told me Jill was doing quite well with the advanced level texts.

To continue to encourage heterogeneous groups [even within a significantly above level group], I suggested Zane and Leah could be pulled with Jill, Marcus, and Rebecca and that perhaps this would make Jill feel more confident. I suggested that perhaps it is not the text level that is frustrating her, but rather that her peers can read it faster so she doesn’t feel like she is doing a good job. I said mixing up the group might give her the confidence boost she needs. (FN 2-6-19)

A few days later, Mrs. Poppy made this adjustment to the reading groups and Jill continued to flourish (along with the others in the group).
At the end of second quarter, Jill passed a level B text which is considered to be a mid-kindergarten text, placing her on track to meet grade level requirements by the end of the year. Surprisingly, at the end of quarter three, just a little over two months later, Jill passed a level E text placing her reading score at a mid-first grade level. While she needed continued support to strengthen her visual cueing system, Jill made tremendous growth over the course of the year especially considering where she started the year. Jill’s progress, moving from one of the lowest in the class to one of the highest in terms of relative literacy achievement, demonstrates the power of heterogeneous grouping by exposing all students to advanced level content even when it may seem beyond their current achievement level.

**Hattie**

Hattie always seemed to be caught up in some sort of girl drama, sometimes declaring she needed a “break” from some of her friends. She also liked to pretend she couldn’t do things, such as put on her own shoes, to get extra attention from the adults in the room. Her mom volunteered in the classroom nearly every morning to assist students during independent writing time and most often this did not interfere with Hattie’s ability to work, although sometimes her mom would help her solve some of the issues that arose with peers. Otherwise, Hattie was fairly quiet and reserved and often followed the rules and expectations of the classroom.

At the beginning of the study, Hattie received speech services in the area of articulation from the speech-language pathologist who visited from the local public-school district. She was dismissed from the program during the middle of the study as it was determined she no longer needed this support. Perhaps due to this boost in her
articulation, throughout the course of the study, Hattie’s confidence in writing grew rapidly. While in December she sought my help to write down words and seemed confused when I asked her to sound them out, by mid-January I noted she confidently engaged in the writing process without asking for any help. In February, she still showed evidence of difficulty decoding and encoding short vowel words, particularly those with short e. However, after some small group explicit phonics instruction she made significant gains as noted by both my anecdotal notes and Mrs. Poppy’s observations.

As a reader, Hattie also demonstrated increased levels of confidence throughout the study. When asked to read a word during whole group, she would decode or search for meaning quietly in her head until coming up with an answer. Most often, her answers were correct, but sometimes she needed additional prompting from an adult to help her arrive at an answer.

*I listened to Hattie read first. She did not know the word “here,” which is a popcorn [sight] word. I helped her find the word “he” in it and then to blend the /r/. She got it and was able to read the rest of the sentence. She got to the word “wheels” and did not know it. I helped her to use the strategy card to get her lips ready and to check the picture. (FN 1-25-19)*

Hattie was able to quickly learn from this responsive, individualized support while reading. She began to take increased ownership of the problem-solving process.

At the beginning of the year, Hattie likely would have been placed in the lowest reading group based on initial literacy assessments such as letter name and sound identification, concepts of print, and phonological awareness. Yet without this grouping structure, by the end of quarter two Hattie passed a level A text and scored high enough
on other assessments that she would have been placed in the middle reading group. By
the end of the quarter three, Hattie’s increased confidence and hard work enabled her to
pass a level D text, which is considered to be a beginning of first grade level.

Molly

Molly was selectively mute when she started her kindergarten year at King
Elementary School. Although her family said she talked at home, Mrs. Poppy said Molly
was silent at school, which was also the case during preschool. Molly also worked with
the speech language pathologist from the local public-school district to support her
language development. By November she was talking but only with a whisper voice. This
continued throughout the course of the study, although Mrs. Poppy started prompting her
to use a louder whisper voice to which she accommodated. Despite only using a whisper
voice to communicate, Molly developed several friendships in the classroom with her
peers. Mrs. Poppy felt that pressure from peers to talk while playing together was part of
the reason she began using a whisper voice.

Being selectively mute also did not hold Molly back as a reader or writer. For
example, she evidenced higher level thinking in the following excerpt: Molly wrote “I am
going to fun nie (night).” I asked her how she knew how to write “going” and she said “I
know ‘go’ and I know ‘-ing!’” (FN 1-25-19). Then she continued to write an additional
detail to her story. By mid-February, her reading was also beginning to take off. During
independent reading time, I grabbed a book she had been reading during her guided
reading group and asked her to read it to me.

Molly read a level C book called “The Little Snowman.” She did a fantastic job,
making only a few errors throughout the entire book. She read “snowman” for
“snowballs,” but corrected it with prompting. She read “the” for “a,” and on several pages said “snowman” instead of “snowman’s.” Near the middle of the book, I showed her the “-s” at the end of the word and she was able to read the rest correctly. She also read “snow” for “big” (which made sense), and then self-corrected with independence. (FN 2-4-19)

As with many of her classmates, Molly most often used meaning cues to read unknown words, relying much less on visual cues. As can be seen in the excerpt above, however, she was beginning to notice these visual cues and use them to self-correct (e.g. snow for big.) Decoding consonant-vowel-consonant words without a picture continued to be difficult for Molly, particularly when reading short o and short e words.

Molly’s end of quarter one literacy assessment scores, including phonological awareness, letter name and sound identification, and concepts of print placed her about in the middle of her class. By the end of quarter two, Molly maintained her relative class ranking, passing a level A text and was on track to pass end of kindergarten expectations. In mid-February, Molly was reading so strongly in her reading group that Mrs. Poppy decided to give her the benchmark assessment early. She passed the level B text and the level C text, making about a half a year’s worth of literacy progress in just one month. Mrs. Poppy said she, couldn’t believe how well her students were doing and that she thinks it is because she is giving them more challenging books than she would have in the past since they are grouped heterogeneously (FN 2-11-19). One month later, at the end of quarter three, Molly demonstrated her continued accelerated progress. She passed a level E text, which is a mid-first grade level.
Bailey

Bailey was a tiny and quiet little girl who most often had a smile on her face. She was kind, polite, and very easily distracted. Due to this, she often did not get her work finished and had to complete it before she could play at the end of the day, which didn’t really seem to bother her. Often times when Mrs. Poppy asked her a question, it took several prompts and rephrasing of the question before Bailey would answer. Sometimes this was because she simply did not know the answer.

_Bailey was shown the word “or” and couldn’t read it. Mrs. Poppy gave her several prompts saying, “it is a little word in other words like ‘morning’.” Then she pointed to the word “for” behind her and showed Bailey how “or” was hiding in that word. This did not help because Bailey did not know the word “for” either. Eventually Mrs. Poppy simply told Bailey the word._ (FN 1-16-19)

Other times it seemed that it took longer for Bailey to answer because she was not paying attention.

_Bailey was asked which punctuation mark to use after “Good morning kindergarten” She said “i?” and proceeded to half lay/roll on the carpet. Mrs. Poppy said “i?” “What punctuation mark?” Then Bailey said “a question mark?” Then, “a stop sign?” And, finally, she was able to say “an exclamation mark.” (FN 2-8-19)

This lack of attention often meant that Bailey missed out on important opportunities to practice her reading skills throughout the day. For example, when Mrs. Poppy asked the girls in the classroom to read the poem, _Bailey just looked around_ (FN 1-25-19). These occasions were noted frequently throughout the duration of the study.
Despite Bailey’s difficulty in attending to tasks, she made a surprising amount of progress in her literacy development. At the end of quarter one, Bailey would have been placed in the lowest reading group in her classroom due to her first quarter test scores. At the end of quarter two, she passed a level A text, but still likely would have remained in the lowest reading group. In early February, Bailey continued to have difficulty reading consonant-vowel-consonant words, specifically short o, short u, and short e words. After one month of explicit phonics instruction, she became much stronger in this skill. At the end of quarter three, her benchmark test scores showed tremendous progress. She passed a level D text, or beginning of first grade level.

**Brittany**

Brittany was a tall and confident girl, often seen bouncing around the classroom. She was also most often the other half of the girl drama, alongside Hattie. When I arrived in the classroom at the beginning of the study, and perused through the students’ writing journals, I initially thought Brittany fell in the bottom third of her class in terms of literacy development based on her phonetic representation of words. Over the course of the next couple of months, Brittany’s writing skills began to improve. At the end of January, she independently wrote, *I am going to fun Night Ies (it’s) going to be so fun.* (FN 1-25-19). As evidenced in this writing, Brittany was developing a repertoire of sight words, as well as a growing knowledge of grapheme-phoneme connections. She also used environmental print to write certain words such as “fun night.”

Brittany was a confident reader who was willing to try new and difficult tasks. By early February, she was able to read and write most consonant-vowel-consonant words correctly with the exception of short e and sometimes short u words. By the end of
February, after explicit instruction in this area, she was much more proficient and rarely misspelled or misread these words. When reading continuous texts, Brittany was quite proficient at using meaning and structural cues, as well as her growing knowledge of visual cues. She also demonstrated developing comprehension skills and was eager to engage in discussions about the texts.

As with her classmates, heterogeneous grouping for guided reading instruction supported Brittany’s impressive reading growth. At the end of quarter two, Brittany passed a level B text. She had strong phonological awareness skills and knew all of her letter sounds, but was still working on a few letter names. At the end of quarter three, she successfully passed a level E text. In just a little over two months of instruction, Brittany made nearly a year’s worth of reading progress. A miscue analysis of her end of quarter three benchmarks tests demonstrated that she most often used meaning and structure to solve unknown words, but only sometimes self-corrected based on visual cues. For example, she read *cat* instead of *kitten* all four times it appeared in the book and did not self-correct even though *cat* is a word she could certainly decode. This pattern of miscues matched what I observed of Brittany while reading in the classroom. She decoded with accuracy when prompted to do so or when reading a word in isolation, but often ignored this strategy when she felt that she could rely on meaning.

**Rebecca**

Rebecca was the oldest and most outspoken student in the class. She was most often engaged in the classroom learning activities and always had questions and ideas to share. Rebecca loved to dance and could often be seen flopping around the rug playing with peers. Most academic tasks came fairly easy to Rebecca and she had already
mastered many foundational academic skills prior to kindergarten. As a result, she was rarely met with an academic challenge, particularly in the beginning of kindergarten. When a challenge was presented or something did not go quite as she had planned, Rebecca became easily frustrated. During these moments of frustration, she lacked the coping skills to be able to problem solve and often ended up in tears until a teacher was able to assist her. Rebecca was also a perfectionist and wanted her drawings and her letters to be exactly the way she envisioned them, often leading her to erase and re-attempt independent work several times.

Rebecca’s intense desire to do things exactly right meant that she rarely took risks in learning. In fact, when I assessed the students’ independent writing journals prior to knowing anything about the students, I placed her in the lowest third of the class. In contrast, her reading assessments placed her significantly above her peers (with the exception of Marcus). It seemed that less structured tasks were the most difficult for her. For example, although Rebecca had strong language skills and was thrilled to be the fox in the class play of The Gingerbread Man, she had the most difficult time of all the students when it was time to practice her oral retelling.

Rebecca, who has been outspoken the whole time, froze when it was her turn to be the fox. Mrs. Poppy prompted several times and then eventually gave her the words she was supposed to say. (FN 12-13-8).

Rebecca often asked me how to spell words for her and became frustrated when I told her to try and stretch them out. She was quite skilled at knowing where to find environmental print in the classroom and would often leave her desk to go find a word after I declined to spell it for her. As a result, she often wrote her words with one hundred percent accuracy,
such as “I am playing with my friends and I love to play with my friends” (FN 1-25-19). As you can see, while all the words are spelled correctly, Rebecca did not try to include any unknown words that would require her to use phonetic spelling.

When she did try, Rebecca had some difficulty systematically encoding words. In early February, she wrote, “I love the days for the werckqxs (weeks) bckus (because) my berchthay (birthday) is in January” (FN 2-8-19). As noted in this writing sample, Rebecca attempted to represent unknown phonemes with multiple possible graphemic representations. For example, to make the /k/ sound in week, Rebecca included a c, k, q, and x as possible spellings, all of which are logical given that the graphemes -ks at the end make the /x/ sound. In addition, unsure which digraph to use in birthday, she included both th and ch. More explicit instruction in phonics during the study helped support her continued development in representing each phoneme with one grapheme or grapheme pair, but she continued to struggle with this skill since she did not often practice writing words phonetically unless asked to do so.

At the beginning of the study, Rebecca was the other student that Mrs. Poppy felt needed to be grouped with Marcus because they were already reading first grade level texts. Yet, despite her ability to read continuous text, Rebecca struggled to read unknown consonant-vowel-consonant words in isolation. Either she knew them automatically and could just say them (ex: map, pod, fin, hug, bet FN 2-8-19), or she struggled to decode them. Near the end of quarter three, I played a game with a small group of students, including Rebecca and two other students who scored at much lower reading levels. The students had to quickly decode words and determine if they were real or nonsense words. If they could make a real word, they had to grab the letter chip that completed the word
before others snatched the chip. Rebecca struggled significantly with the task and was able to make only a few words. She often created real words using meaning that were visually similar but inaccurate. For example, she insisted *pit* made the word *pet*.

Meanwhile the student beside her was racking up the words. Rebecca lamented, “*How is she so good at this?*” (FN 3-4-19).

Similarly, when Rebecca encountered a word in a text that she did not know and could not use meaning to solve, she had few other strategies at her disposal to help her. Mrs. Poppy said when this occurred during guided reading group, *she would simply skip over the word* (FN 3-4-19). In fact, when Mrs. Poppy was explicitly teaching Rebecca’s reading group the sequence of steps to solve an unknown word, Rebecca openly questioned it. *Rebecca asked why they had to use all [the strategies], and Mrs. Poppy said it makes them a better reader.* (FN 1-25-19). Such strategy usage likely seemed to be a waste of time to Rebecca who rarely had to rely on explicit strategy use to solve an unknown word. Pushing her to work in harder texts in which she would encounter more unknown words was difficult given her current level of coping skills when met with challenges.

At the end of quarter one, Rebecca scored very highly on all literacy skills that could be learned by rote memorization, such as letter names, letter sounds, and concepts of print. However, skills such as phonological awareness, which required active problem-solving skills were comparatively much more challenging for Rebecca. She scored in the lowest third of her class in phonological awareness skills, although she was the highest in all other areas. At the end of quarter two, she passed a level D text, placing her at a beginning first grade level. At the end of quarter three, she passed a level G text, making
about a half of a year’s worth of progress in just one quarter. While this was not as much progress as some of her peers, Mrs. Poppy and I were pleased with her progress given her level of effort and ability to cope with the unknown.

**Leah**

Leah was a happy and rule following kindergartener. She often raised her hand to participate, most often having the right answer. She also liked to make people laugh and sometimes would pretend to know less than she really did. When Mrs. Poppy asked Leah to spell the word *friends* during their shared writing time,

*Leah spelled friends “fir” and then laughed and laughed, clearly knowing this was not correct. Then she zipped through the word saying “friende” and then laughed again before saying the “-s.”* (FN 1-25-19).

While she was developing a strong sight word vocabulary, such as the word *friends* which was spelled aloud almost every day in the morning message, Leah also had difficulty decoding and encoding phonetically regular words as did many of her classmates. At the beginning of the study, I placed her among the lowest third of the writers in her class based on her independent writing skills, which often lacked grapheme-phoneme correspondence for medial vowels.

Reading short e words in isolation continued to be challenging for Leah even after several weeks of explicit phonics instruction in February. Most often she read and represented short e words with short a, which makes sense given the proximity of the places of articulation of these two sounds in the mouth. The excerpt below shows how heavily Leah relied on meaning while reading. The students were given a worksheet
which had six pictures. The students were expected to cut out the six words at the bottom of the page and glue them next to the matching pictures.

_I asked Leah to read the words to me after she had already cut and pasted them [on her worksheet]. She looked right away at the pictures so I covered them up to see what she could do without them. She said “sud” for “sub” and “train” for “ten.”_ (FN 2-25-19).

Leah insisted that the word _ten_ was in fact _train_ even after I prompted her to reread the word. Although these words begin and end with the same sounds, Leah did not seem to notice that the middles of these words did not visually match. Eventually I asked Leah to say each sound and then blend them together with me to read the word. After we arrived at the word _ten_, I uncovered the picture and showed her how it matched.

At the end of quarter one, Leah’s literacy assessment scores likely would have placed her in the middle reading group if grouped based on ability. She had learned all of her letter names and most of her letter sounds. She was developing her phonological awareness skills and had mastered her concepts of print. At the end of quarter two, she passed a level B text, placing her on track to meet grade level expectations of a level C at the end of kindergarten. By the end of quarter three, Leah surpassed this expectation and passed a level E text (although barely meeting the requirements for accuracy). While she still struggled to decode phonetically regular words, her reliance on meaning cues was strong enough to propel her through several levels.

**Vivian**

Vivian was a student who loved to seek out adult connections. She always greeted me by name when I arrived at King Elementary school and asked me when I would be
coming back when it was time for me to leave. She also often offered compliments about my jewelry or attire. Vivian made friends with many students in the classroom as well. She was kind and polite to everyone around her. She followed the classroom expectations and clearly loved to be at school. On the carpet, her spot was right in front of the white board and her voice could always be heard reading, answering questions, or suggesting ideas. Of all the students in the classroom, Vivian was perhaps the most genuinely engaged in the learning opportunities and seemed to gain real pleasure from it.

Interestingly, despite her attentiveness and desire to learn, reading and writing were challenging for Vivian. In fact, after Mrs. Poppy administered Vivian’s second quarter assessments, she told me she was most concerned about Vivian’s progress (along with Logan’s) (FN 1-9-19). I also noticed her difficulty in literacy skills early on during my observations.

*I noticed Vivian’s note [from her mom] on the table. I asked her to read what it said. The note read “Be the star I know you are” and was decorated with stars in multiple colors. I asked Vivian to read it to me. She read “be” and “the” with no problem, but stopped at “star” and looked at me. I asked her to put the first letter in her mouth. She said “ssssssss” and I asked her what would make sense. She was unable to tell me. I prompted again and reread “be the” and pointed to the star while saying /s/. She was still unable to tell me. I prompted again similarly and she excitedly said “star!” (FN 12-13-18).*

The excerpt above is representative of Vivian as a reader. It often took several prompts of increasing levels of support to assist her in figuring out an unknown word, however, she was always willing and eager to keep trying until she got it.
Similarly, when students were introduced to the new strategy card to help them problem solve to figure out unknown words, *Vivian was engaged but had a hard time applying the strategies in the text* (FN 1-13-19). This continued over time. *Mrs. Poppy passed out the strategy card and asked students what to do when they get to a word they don’t know. Vivian proudly shouted “lips!” Mrs. Poppy asked what to do next and Vivian said “picture!”* (FN 1-16-19). Vivian’s voice was the first and loudest voice in the group. Still, when given the book that day, she didn’t understand exactly how to use these strategies in the context of continuous text.

*I worked with Vivian and she struggled [to read the words]. When she got to a word she didn’t know, she relied mostly on meaning. She did not use the “get your lips ready” strategy and I prompted her to use this. Then she used visual only and did not choose a word that made sense (e.g. she read “up” for “under”).* (FN 1-16-19).

Over time, Vivian did get slightly better at using the strategy card, but it took a considerable amount of practice.

In February, as I began to work with the students on explicit phonics instruction, (this will further be explored in the articulation of the design principles in chapter five) I quickly noticed that Vivian had a very difficult time articulating the vowel sounds. While her oral language skills were proficient and she could correctly articulate these vowel sounds within words, when asked to make the short vowel sounds in isolation, or while segmenting words, she stuck her tongue out of her mouth while attempting to make all five of the short vowel sounds. As a result, decoding words was very challenging for her. In February, I asked her to read a list of phonetically regular consonant-vowel-consonant
words, including, map, pod, fin, hug, and bet. She was unable to read any of them with independence. For example, she read make for map which shows that she did not make use of phonemic blending. She did, however, get “her lips ready” for the initial sound and thought of a word that might make sense (a known word rather than a nonsense word) since there was no meaningful context to support her reading (FN 2-8-19).

Encoding words was equally challenging for Vivian, often leaving out vowel sounds and sometimes final sounds as well. During independent writing time in mid-February,

*I noticed Vivian attempted to write popcorn [sight] words, but almost all other words had only the beginning sound. For example, she wrote “s” for “scratches,” “h” for “hand,” “c” for “coming,” and “f” for “from.” I worked with her on a couple of the words and then told her I know she can add more sounds by herself to the other words so I left her to do this work independently.

She was able to write some additional letters, but also included some that did not belong. (FN 2-11-19).*

Even when using sound boxes as a scaffold to write all three sounds in a consonant-vowel-consonant word, Vivian still struggled hearing and recording the medial vowel sounds.

*When Vivian was writing [in the sound boxes], she often left out the middle vowel sound. One time, she even wrote the last sound in the last box, but did not write any other sounds. She had an especially difficult time hearing the difference between the short e and a sounds. (FN 2-11-19).*
Short i sounds also continued to be problematic for Vivian, although she did show some growth after receiving additional explicit phonics instruction.

Despite having difficulty with blending words, Vivian was able to use meaning and structure to help her successfully read continuous texts. During guided reading groups, Vivian read a level E text with few errors as evidenced below in the transcription of Vivian’s oral reading of the text (words read incorrectly are followed by the correct word in parentheses).

Meow, I like this said the h- (listened to the prompt given to her neighbor) hungry little kitten. Go away said a big cat. Meow, I like this said the hungry little cat(kitten). Ruff (grr, grr, grr) go away said a big dog. Go away little kitten, ruff (grr, grr, grr). (FN 3-6-19).

As evidenced in this text, almost all of Vivian’s errors made sense and were structurally correct, however, at the point of difficulty, Vivian made little use of visual cues.

At the end of first quarter, Vivian’s scores on early literacy assessments would have likely placed her in the lowest reading group if grouped on the basis of ability. She was similar to many of her peers in terms of her knowledge of letter names and sound identification, but scored much lower (along with Kelley) in her phonological awareness.

At the end of second quarter, when Mrs. Poppy noted she was concerned about her progress, Vivian passed a level A text. This score still put her on track to potentially meet grade level expectations by the end of the year, although she would have to pass a new text level each consecutive quarter. At the end of quarter three, even after missing nearly a week and a half of school, Vivian returned and passed a level E text, meaning that in just one quarter, Vivian progressed from a beginning of kindergarten level to a middle of
first grade level. This would be impressive growth for any student, but is even more impressive for a student who would have likely been identified as a struggling reader early in the study.

Kelley

Kelley was a tiny, sweet little girl who got along well with the others in her classroom. She was a hard worker when at school, although Mrs. Poppy said that when work was sent home, such as rereading a guided reading book, it did not get done. Some days Kelley seemed a bit “off” as Mrs. Poppy would put it, meaning that she seemed confused at ordinary tasks or skills she was asked to do. For example, one day Kelley drew me a beautiful picture of a rainbow-colored heart that clearly took a lot of time and effort to make. On the top, she wrote To: Miss Tiffany From: Kelley. While I was working in the hallway, I heard Mrs. Poppy tell Kelley she could go give Miss Tiffany the drawing. Kelley walked into the hallway and right past me. After she walked a little further, she stopped and looked around, seeming unsure of what to do. I called her over, thinking maybe she didn’t see me, but when I did she didn’t seem to recognize me. She told me she was looking for Miss Tiffany and I said that was me. She looked confused, but gave me the letter and went back to class.

At the onset of the study, I grouped Kelley’s writing in the bottom third of the class. As I became better acquainted with the students and their literacy development, I realized that while Kelley’s literacy skills were indeed comparatively lower than many of her peers, she remained a confident and enthusiastic writer. In fact, most mornings, Kelley was the very first one done with her independent writing. She rarely asked me how to spell a word and would instead sit down, stretch out each word, and record the
sounds that she heard along the way. In January she wrote, *I ws pl on the slad* (*I was playing on the sled*) (FN 1-13-19). In this example, it is clear that Kelley needed additional support to represent vowel sounds in words, as she either incorrectly represented them or did not represent them at all in most of the words in this sample. However, just a few weeks later, Kelley wrote, *I am hafnd a sep ovr wif my Nan* (*I am having a sleepover with my Nana*) (FN 2-4-19). Here we see that Kelley was able to include at least three sounds for each word, including medial vowels, and correctly spell several sight words.

In terms of reading work, in the beginning of December, Mrs. Poppy told me that Kelley was also probably her lowest reader. This was reaffirmed by my observations of her in class reading.

*I noticed that Kelley frequently started the page [of the pattern book] by reading “I am” rather than “Look at” [which were the words on the page]. When I prompted her to point to each word while she was reading she was able to read the words correctly. She got stuck a few times at the unknown word on the page. All of her errors were meaning related (e.g. Kelley read “drawing” instead of “writing”).* (FN 12-4-19)

As with many of her classmates, Kelley developed an over reliance on meaning with little attention to the visual cues of words. In addition, while Kelley often attempted to participate in the shared readings of the poem or morning message, the class often read too fast for Kelley to keep up.

When asked to read phonetically regular consonant-vowel-consonant words in isolation, Kelley had even more difficulty. At the beginning of February, she was asked
to read *map, pod, fin, hug, and bet*, but she was unable to read any of them with independence. While she could identify almost all of the sounds at an individual level, when she blended them together, she arrived at the wrong word (for example, reading *pond* for *pod*) (FN 2-8-19). It seemed that she, too, was trying to make meaning of these words while reading. When attempting to decode a word, Kelley often said each letter sound several times in a choppy manner, making it difficult for her to come up with the word.

*When Kelley got to the word boat she said /b/ /b/ /b/ /o/ /o/ /o/ /a/ /a/ /a/ /t/ /t/ in quick succession. She skipped very quickly to step three [on the strategy card] except her blending was ineffective.* (FN 1-25-19)

In addition, Kelly had a difficult time articulating the short e sound, often moving her mouth in awkward ways and producing a sound that more closely resembled the short a sound. However, after explicit instruction in this skill, focusing on smooth blending of sounds, Kelley soon became one of the most accurate decoders in her classroom. In early March, we played a phonetic blending game in which Kelley beat everyone in the group, including Rebecca. *By the end of the game, Kelley’s confidence skyrocketed. She left in all smiles* (FN 3-4-19). Mrs. Poppy and I celebrated that perhaps for the first time, Kelley felt truly proud and competent as a reader in front of her peers.

During quarter one, Kelley’s literacy assessments placed her at the very bottom of her class. She had acquired just 33 out of 52 letter names, and 12 out of 33 sounds. Her phonological awareness skills were also just developing, as was her knowledge regarding the concepts of print. After quarter two testing, Mrs. Poppy shared with me that she was very concerned about Kelley’s reading growth (FN 1-9-19). Kelley was the only student
in her classroom who was not able to pass at least a level A text, although she did make considerable progress in letter names, sounds, and phonological awareness. After two and a half months of exposure to higher level texts that required active problem solving along with some explicit phonics instruction, we were shocked by Kelley’s literacy progress. At the end of quarter three, Kelley passed a level D text, making an entire year’s worth of progress in just one quarter and catching up with several peers in her classroom.

**Logan**

Logan was one of the most reserved students in the classroom. He answered when called upon, but rarely volunteered to participate. He seemed shy in all respects, including his socialization with peers, although he had a budding friendship with Zane, which seemed to push him out of his comfort zone a little bit. This also seemed to get him side-tracked and sometimes Zane told him answers to his work, curtailing his problem-solving efforts. During independent writing time, Logan and Zane sat next to each other and below is an excerpt of a typical morning.

*Logan had a difficult time getting started [on his writing]. We talked about his sentence and it was really long. I encouraged him to make it a little bit shorter so that he could remember it. He chose “I like being at school” because he was gone several days after being sick. Logan correctly wrote “I” but stopped. I reminded him that “like” is a popcorn word and he looked for it. Mrs. Poppy prompted Logan to look in a new spot for the word [it had moved to the cupboards] and it took him awhile to find it. He finally found it but was slow to get started writing it. Zane, having heard this whole conversation, started spelling “like” for Logan to write. Logan still did not write. Zane became impatient and said it louder.*
reminded him to let Logan do his own work. Eventually Logan wrote the word
and I moved on to work with another student. (FN 1-13-19)

As evidenced above, Logan often needed additional processing time when reading or
writing and other classmates often jumped in to help. Mrs. Poppy always reaffirmed in
front of the class that Logan (and others) could do the work on their own and gave them
as much wait time as needed.

After looking through the students’ writing journals at the onset of the study, Mrs.
Poppy shared with me that Logan was the lowest writer in the class. Yet given her
perception of his relative writing abilities, as I perused through his writing journals from
August to November, it was clear that Mrs. Poppy did not limit his exposure to higher
level skills.

She said he could not even write his name when he started kindergarten…[but]
even this student wrote a word with “-ing” on October 18 and in November he
wrote the chunks “oo,” “th,” and “or” correctly as well. (FN 11-30-18).

In fact, as I observed throughout the literacy block, there was no indication given by Mrs.
Poppy that Logan was struggling in his literacy development. She asked him questions
with the same level of difficulty as the rest of the class and he completed all of his work
with support (if necessary) rather than modification.

By February, Logan was still working on reading and writing phonetically regular
consonant-vowel-consonant words as evidenced in my observations. Logan had a
difficult time representing vowel sounds, either skipping them entirely (writing “hg” for
“hug”) or misrepresenting them (writing “pag” for “peg”) (FN 2-20-19). When asked to
read map, pod, fin, hug, and bet in isolation, he was unable to read any of them
independently (FN 2-8-19). For example, when attempting to read *map*, he said “ap,” then “am,” then “meep” and clearly did not have a well-developed system of decoding to help him read unknown words. After some explicit instruction on blending, he made some progress in this work by the end of February, however, his lack of consistency in letter/sound correspondence continued to result in many errors. For example, he was able to read *map* and *fin* accurately, but read *bod* for *pod*, *huj* for *hug*, and *dit* for *bat* (FN 2-27-19).

At the beginning of the year, Logan’s literacy assessments matched my own observations and Mrs. Poppy’s perceptions. Logan scored at the bottom of his class, similarly to Kelley. His phonological awareness skills were emerging and he knew about half of his letter names and sounds at the end of quarter two, Mrs. Poppy was thrilled that he passed a level A text, and that he was able to retain his reading progress after the long break. Throughout quarters one and two, his progress remained slow by steady.

During quarter three, Logan’s confidence as a reader and writer noticeably grew as he was presented with texts of increased complexity and offered the instructional support necessary to successfully read them.

*Then I pulled Logan and asked him to read another level C book called “Wake Up, Dad.” He also did a great job, he needed help with the name “James,” and read “says” instead of said several times, but clearly maintained meaning throughout the book. He read “sleeping” for “asleep” and when I asked him if it looked right he self-corrected. I praised him as he was reading at a few tricky parts and he was beaming.* (FN 2-4-19)
He transitioned from being one of the last students to get started on his writing work, to being one of the first.

_I was surprised to see that Logan was the first one at his table to get started on writing...Logan was working hard stretching out the sounds in the words. He wrote, “I had a slp ov at miy gam hs” (I had a sleepover at my grandma’s house.) As can be seen, he is very rarely including vowel sounds, however his confidence has increased. He did raise his hand and ask me for help on the word grandma, but I encouraged him to stretch it as he had the other words. (FN 2-25-19)_

At the end of quarter three, Logan surpassed the end of year goal to read a level C text by accurately reading, self-correcting, and comprehending a level D text. Considering level D text is a beginning of first grade text, Logan successfully progressed from what many would have considered to be a “struggling reader” to be a student who was exceeding literacy expectations for his grade level.

Mrs. Poppy’s Progress

Mrs. Poppy was a veteran teacher with over two decades worth of teaching experience. She taught a variety of grade levels in both the public and private sectors in rural and urban areas of the Midwest. In addition to her ample teaching experience, she took graduate courses in early literacy development and had her reading specialist endorsement as well. Mrs. Poppy has also worked for some time with a local university to support the development of elementary writing teachers throughout the state. Given Mrs. Poppy’s background, she was a prime participant for the study as a design collaborator.
From the onset of the study, Mrs. Poppy agreed to participate with the stipulation that I be as transparent as possible with her throughout the process. She wanted me to tell her if I noticed anything I thought she should change, and frequently asked my perspective on literacy instructional practices. I reminded her often that my role was not that of an evaluator, but as co-designer of the research study. I frequently reassured her that she was doing fantastic work with her students and that they were fortunate to have a teacher so dedicated to their success. She often apologized to me during initial visits when the schedule was off track or they had to skip something. I reminded her that this study focuses on the real-life context of teaching and that flexibility was an important component of elementary instruction.

Overtime, however, I did begin to worry about the number of times that guided reading instruction was skipped or limited. Holidays, school programs, testing, and catching up on other content were among the several reasons that Mrs. Poppy decided not to teach groups. It seemed that if something in the schedule had to be cut, guided reading was the first thing to go. After noting this several times, I considered talking with Mrs. Poppy about trying to fit in guided reading instruction more often. However, as students began to show progress, Mrs. Poppy began meeting more frequently with the groups without my suggestion. She spoke with me about how Wednesdays are problematic for scheduling because of library and chapel time, but that she was thinking about cutting out whole group reading on those days instead because she saw the importance of guided reading instruction. This was a big change for Mrs. Poppy considering in her pre-interview she talked about possibly foregoing guided reading instruction completely because she felt as if she could meet the needs of all of her students during whole group
instruction (INT 12-10-19). During Mrs. Poppy’s post-interview, she reflected on this further stating,

The thing that I’ve learned the most is I’ve always waited to start small groups…
you know I haven’t started them right away. I’ve usually waited until October,
and I’m not going to do that next year. I’m going to pick it up right from the beginning. (INT 3-19-19)

After several decades of teaching guided reading instruction, it was not until she started grouping students heterogeneously that she started to see the immense value in small group instruction.

From the beginning of the study until the end, Mrs. Poppy expressed that she felt grouping students by ability level was ineffective for early reading instruction. However, even though she was invested in heterogeneous grouping, she had a difficult time applying this in practice due to the entrenched nature of ability grouping in early reading instruction. This is the approach promulgated by the leading texts on guided reading instruction on which Mrs. Poppy based her instruction. It was also the method promulgated by the professional development in at least some of the school districts in which she worked. As many of the teachers I attempted to recruit for this study, Mrs. Poppy had a difficult time seeing true heterogeneous grouping in practice.

In January, when it was time to begin grouping her students for guided reading instruction, I noticed Mrs. Poppy had some misconceptions regarding what I considered to be heterogeneous grouping. The following situation occurred in early January just as she started meeting regularly with her self-identified heterogeneous guided reading groups.
She showed me another basket, [this group would be] working on level C. I was thrilled and said this was so good even for those students who only passed level A because it gives them great exposure to problem solving even if they can’t finish the book. Then she pulled out another basket and showed me her last group. I was surprised because I thought the rest of the students (with the exception of Rebecca and Marcus) were grouped heterogeneously. This was not the case. The third and final group was set to read level A and B books. Mrs. Poppy took me over to the center board and showed me the groups. There were five students in the “higher” group and six in the “lower group.” There were three students whom she felt “could go either way” but she wanted to leave them in the “lower” group as models for the three others. She said as long as she continues to push them in the group it shouldn’t hurt them. (FN 1-13-19)

Mrs. Poppy clearly felt the way she had grouped students was aligned with my thinking of heterogeneous grouping. By pulling the three “higher” students into the “lower” group, she felt that her “lower” readers would benefit from the opportunity of working with higher level learners. In reality, nearly all of the students who passed a text level A at the end of quarter two were placed together in a group, and those that passed a level B, along with some of the stronger level A students, were placed in another group and given a higher-level text to read. In addition, Mrs. Poppy gave the lower group a lower level text with the thought that she would be able to provide a sufficient amount of challenge to those students who might actually need a higher-level text.

Mrs. Poppy’s initial interpretation of the design was exactly opposite of its original intention. Rather than giving higher students lower level books in heterogeneous
groups, lower level students should be given challenging books at the higher level of their peers supported by scaffolding and instruction. In Mrs. Poppy’s initial version of heterogeneous grouping, the learning of the higher students would likely be negatively impacted as is often stated as a concern of heterogeneous grouping (e.g. Oakes, 1985/2005). In the intended version, the higher students receive instruction very similar to what they would have received in homogeneous grouping, but the lower students benefit from increased access and opportunity to read grade level materials.

In the next few visits, I slowly challenged her to reconsider her groups and her thinking on how to best support all students’ literacy development. Eventually,

I asked her if she would consider switching up groups so that the lowest students were split between the groups. I complimented her on the rich book introduction that she did with the students and explained how this could provide the scaffolding needed for the lower students in the group. I was pleasantly surprised when she enthusiastically said “sure!” She went right over to the center board and began thinking of how to switch groups. (FN 1-25-19)

This gradual shift in practice was reflective of the embedded nature of the norm to group students by ability. It took time and attention to deconstruct traditional methods of instructional grouping.

Referring back to the theoretical framework of Oakes’s (1992) dimensions of change, we began the study by attending to the technical dimensions. Mrs. Poppy agreed to group her students heterogeneously during guided reading instruction, but in practice attending only to the technical dimension resulted in misconceptions and misapplications of the design. The students were still clearly grouped homogeneously, but Mrs. Poppy
felt she had made real changes in practice. The norm of grouping students by ability for instruction was so ingrained in her way of thinking that even this subtle shift, moving some higher students to the lower group felt like a big change. It took several weeks of graceful navigation to alter this way of thinking.

In fact, I was met with a similar situation while trying to recruit participants for my research. The first district I approached denied my study on the basis that they do not group students by ability for guided reading instruction. Although I was privy to district trainings that promulgated and demanded fidelity to that very structure, grouping students by ability was so wholly integral to the system that they couldn’t “see” it and identify it as such. Claims of flexible and dynamic grouping replaced the language but not the practice of ability grouping, which remained the steadfast approach to instruction. Attending to the definition of ability grouping was essential to influencing the normative dimension of change in this study, bringing light to an otherwise subliminal system of inequity.

**Summary of Class Progress**

As documented in each of the student cases explored above, all students who participated in heterogeneous guided reading instruction in this study made considerable literacy progress during just one quarter of the intervention. Each student far surpassed the end of year goal of reading a level C text by the end of quarter three. Table 4.2 provides an overview of the text level that each student passed at the end of quarter two as well as the text level they passed after the intervention at the end of quarter three. Each student’s overall reading growth in terms of grade level equivalency after just one quarter is also reported. It should be noted that these numbers are rough estimates given then that
neither MONDO Bookshop levels (2008) nor Rigby PM Benchmark levels (2000) provide quarterly grade level expectations.

As can be noted, all students benefited from the intervention, including students who were considered to be below grade level, on grade level, and above grade level in their literacy ability. In fact, twelve out of the thirteen students in the study made over a quarter of a year’s worth of growth in just one quarter, with the final student making exactly one quarter’s worth of growth. This demonstrates that not only is grouping students heterogeneously as effective as the status quo, but it is, in fact, more effective than the instruction that occurs in a majority of classrooms around the country.

Furthermore, while Lleras and Rangel (2009) document that homogeneous grouping serves to widen the achievement gap, this study evidences that heterogeneous grouping actually has the potential to close it. While all students made at least expected progress or greater, averaging as a class about one grade level of progress, the lowest achieving readers made the greatest gains. Kelley, Logan, and Vivian, the three students in this study who needed to most scaffolding and responsive instruction to read grade level text made approximately 1.25 years’ worth of reading progress, while the highest students at the onset of the study, Rebecca, Marcus and Jill, averaged approximately .5 years’ worth of reading progress. With the continued use of heterogeneous grouping for guided reading instruction over time, the achievement gap may indeed effectively close.

At the conclusion of the study, I asked Mrs. Poppy how her students this year were doing in comparison to the reading levels of students in a typical year at the end of quarter three. She stated,
Table 4.2: Summary of Students’ Progress

<table>
<thead>
<tr>
<th>Student</th>
<th>End of Quarter 2</th>
<th>End of Quarter 3</th>
<th>Growth in Grade Level Equivalency from Q2-Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marcus</td>
<td>E Middle First</td>
<td>G Middle First</td>
<td>.25</td>
</tr>
<tr>
<td>Rebecca</td>
<td>D Beginning First</td>
<td>G Middle First</td>
<td>.5</td>
</tr>
<tr>
<td>Leah</td>
<td>B Middle K</td>
<td>E Middle First</td>
<td>1</td>
</tr>
<tr>
<td>Jill</td>
<td>B Middle K</td>
<td>E Middle First</td>
<td>1</td>
</tr>
<tr>
<td>Brittany</td>
<td>B Middle K</td>
<td>E Middle First</td>
<td>1</td>
</tr>
<tr>
<td>Casey</td>
<td>A Beginning K</td>
<td>D Beginning First</td>
<td>1</td>
</tr>
<tr>
<td>Hattie</td>
<td>A Beginning K</td>
<td>D Beginning First</td>
<td>1</td>
</tr>
<tr>
<td>Molly</td>
<td>A Beginning K</td>
<td>E Middle First</td>
<td>1.5</td>
</tr>
<tr>
<td>Zane</td>
<td>A Beginning K</td>
<td>G Middle First</td>
<td>1.5</td>
</tr>
<tr>
<td>Bailey</td>
<td>A Beginning K</td>
<td>D Beginning First</td>
<td>1</td>
</tr>
<tr>
<td>Vivian</td>
<td>A Beginning K</td>
<td>E Middle First</td>
<td>1.5</td>
</tr>
<tr>
<td>Logan</td>
<td>A Beginning K</td>
<td>D Beginning First</td>
<td>1</td>
</tr>
<tr>
<td>Kelley</td>
<td>Pre-A Pre-K</td>
<td>D Beginning First</td>
<td>1.25</td>
</tr>
</tbody>
</table>
Pretty similar except for the lower end... not pulling them out and keeping them in the smaller groups has allowed them to achieve more than I imagined. You know...normally at this point, I panic about how they are doing so I pull them out for separate instruction and now it’s like, “Oh my gosh! Why did I do this [pull them out for separate instruction]?” Because I wanted them to be able to meet the assessments at the end of the year and now they are way above. (INT 3-19-19)

As described here, the students in this study did exceedingly well. The findings may even seem a bit incredulous. After all, it is rare that all students respond so well to an instructional intervention, particularly those who have demonstrated over time that reading presents particular challenges for them, as with students such as Kelley, Logan, and Vivian. What was it exactly that occurred in this classroom to produce such tremendous results? How can a teacher utilize heterogeneous guided reading groups in his or her own classroom and achieve similar results? What challenges can one expect to encounter when disrupting the status quo of deeply entrenched tradition when trying a new approach to meet the literacy needs of all students?

The following chapter provides a discussion which attends to all of the questions above in order for practical application of theoretical insights presented in the form of instructional design principles. The design principles explicated in chapter five are identified as essential characteristics for the success of heterogeneous guided reading groups and ultimately the success of students’ reading growth. These principles highlight that guided reading instruction does not occur within a vacuum, but rather within a myriad of literacy instructional practices, all of which contribute to or hinder the literacy achievement of students. The design principles, therefore, necessarily cover aspects of the
entire literacy curriculum. It is now time to attend to the heart of this matter: *how* and *why* was this approach to literacy instruction so successful?
CHAPTER V
DISCUSSION

“Most education research describes or evaluates education as it currently is. Some education research analyzes education as it was. Design research, however, is about education as it could be or even as it should be” (Bakker, 2018, p.3, emphasis in the original). The type of heterogeneous guided reading group learning documented in study is unique to this project. It cannot simply be observed in a naturalistic setting because it is not done in practice. Mrs. Poppy, a veteran teacher, and myself, a teacher and researcher, set out first to provide a proof of principle (Bakker, 2018) that students can be grouped heterogeneously for guided reading instruction and have equal, if not better, levels of literacy achievement. Given Slavin’s burden of proof argument (1987, 1990, 1993), ability grouping should be eschewed if this is possible simply because heterogeneous grouping is a more equitable practice. Second, if such alternative instruction is possible, our goal was to answer the question how (Bakker, 2018) turning the findings into a warranted claim rather than simply a narrative account. I believe we have achieved both aims.

In this section, I will present the five design principles that emerged from the data, followed by examples and evidence from the study. I will also explore how well each of these principles was implemented in the research site as well as our (sometimes complicated) journey towards implementation. This section will provide the rich descriptions of an alternative to ability grouped literacy instruction in a real classroom context that was called for from decades of researchers (e.g., Nomi, 2010; Oakes, 1992; Puzio & Colby, 2010; Tach & Farkas, 2006; Slavin, 1987). It will specifically
demonstrate the ways in which a teacher can support a diverse group of students without resorting to grouping students by their ability, which has shown to have deleterious, particularly effects for our lowest readers (e.g. Chorzempa & Graham, 2006; Condron, 2008; Tach & Farkas, 2006). Each principle will follow the format for design principles suggested by Van den Akker (2013, as cited in Bakker, 2018) followed by a description of the essential features necessary for student success when placed in heterogeneous guided reading groups.

Many of these features described in the design principles below are also arguably necessary for a student to succeed in homogeneous reading groups, although I argue the definition of success is different in heterogeneous versus homogeneous contexts. In homogeneous grouping structures, it is quite normal to have students who are reading below grade level. The classroom teacher responds to this disparity in achievement by providing the higher readers with higher levels of text and the lower readers with lower level of texts. If all students make progress, regardless of whether or not they reach grade level norms, this is considered to be successful instruction. However, in systems of heterogeneous grouping, teachers must offer the kind of responsive and effective instruction throughout the literacy block that enables all children to read grade level text (or higher) in order for the grouping system to work. This heightens the importance of quality instruction in all aspects of the literacy curriculum so that students are prepared to read texts of increased complexity during guided reading. Thus, in order for students to be successful in the context of heterogeneous guided reading groups, the design principles must account for all of the ways in which literacy instruction can support the success of all students when reading grade level texts. This is, as one might imagine,
quite challenging work.

**Design Principles**

**Design Principle I: Comprehensive Literacy Program**

*If you want to design guided reading instruction in which heterogeneous grouping is used to support the literacy achievement of all students, you are advised to do so within a comprehensive and balanced literacy program.*

Mrs. Poppy was a firm believer that in order to be a reader, students have to love literacy. She provided her class with a variety of enjoyable literacy activities each day that gradually released the responsibility of the reading and writing work to students. A typical day of literacy instruction will be described below to demonstrate the comprehensive nature of the literacy curriculum to which students in this classroom were exposed. Furthermore, it will highlight the critical components of literacy instruction that needed further attention in this classroom in order to support the work of heterogeneous guided reading work.

Upon arrival to school each day, students checked in for lunch and *then read the question of the day requiring a yes or no answer. Examples included, “Are you going to fun night?” (FN 1-25-19) and “Did the groundhog see his shadow?” (FN 2-4-19). The questions often had several sight words, known as popcorn words in this classroom, as well as some words students would need to use meaning and visual cues to solve. Then students went to their desks to write independently in their notebooks. In November, students were expected to write one sentence each day along with a picture, but, in January, Mrs. Poppy began expecting two sentences. As students finished, Mrs. Poppy*
held individual writing conferences with students to support their writing growth while “teacher writing” the correct spellings underneath their words.

Later in the day, after math, students met on the rug to do some phonological awareness activities that were projected onto a screen. These activities included identifying the number of syllables/sounds in a word, rhyming, phoneme deletion, and counting the number of words in a sentence (FN 1-13-19; 2-4-19). After this five-minute activity, students sat on the rug to help complete the morning message as a shared writing experience. Mrs. Poppy wrote these messages out in advance each day. The students had to use meaning to figure out the missing words and then help Mrs. Poppy conventionally spell them. For example, on February 25th the message looked as follows:

_G__ morning_

_D___ y ____ l ___ the snow_ We h___ a busy d____ _ We will d___ something new in ma___ _

_L____,

Mrs. P___

Mrs. Poppy called on individual students to fill in the blanks using meaning. When finished, the message read, “Good morning! Do you like the snow? We have a busy day. We will do something new in math. Love, Mrs. Poppy” (FN 2-25-19).

This was typically followed by an interactive read aloud in which Mrs. Poppy read a book from a self-created unit on a given topic (e.g. snow and making pizza).

Below is an excerpt of one such read aloud.

Mrs. Poppy read a book called “Hold the Anchovies” to prepare the students for the field trip they will be taking in two weeks to a pizza shop. She told them they would be thinking about the order in which the pizza was made and that this was
called sequencing. They sang a song about the author and illustrator (in this case a photographer) to help students learn about these roles. She asked them a few questions while reading, such as “why are there tomatoes?” and “do you know what anchovies are?” After reading the story, Mrs. Poppy called students up to the pocket chart to sequence the pizza ingredient cards. When students struggled, she showed them how they could look at the book to help them remember. (FN 11-14-18)

As demonstrated above, Mrs. Poppy often used this time to teach and practice comprehension skills and strategies.

Following the read aloud, Mrs. Poppy and the students worked on a poem in a pocket chart. One new poem was introduced each week and the class spent time each day choral reading the poem or even adding their own new words into the poem. The students really enjoyed this shared reading activity, although it was common for several students to sit quietly rather than practice reading with the class. Below is an example of a typical shared reading experience.

*Mrs. Poppy flipped the chart paper to reveal a poem.*

The Turkey

The turkey is a funny bird.

His head goes wobble, wobble.

But all he knows is just one word.

Gobble, Gobble, Gobble!

Students read the poem chorally with the teacher. Then Mrs. Poppy handed out pieces of paper to the students. Each paper had one of the words of the poem on
it. Immediately several students shouted out their words, including sight words and other words such as “turkey.” As Mrs. Poppy read and pointed to the poem on the chart paper, students stood up if she pointed to their word and they placed their word in the pocket chart. Students had to pay attention to capital letters, lower case letters, and punctuation marks (identified as stop signs, breath marks, and excited marks). She also talked to students about when to “swoop down” to the next line. (FN 11-14-18)

Conventions of print, phonological awareness, and sight word work were all part of this shared reading experience. At the end of the week, students were given a paper copy of the poem to cut out and put into their poetry notebooks.

Some days the students would engage in letter work for a brief period of time. Mrs. Poppy would share a big book about a given letter and then play a song that went along with it. She encouraged the students to sing along, but it usually went too fast for many of the students to follow. Below is an excerpt of a typical lesson on letters.

They focused on the letter Q. Mrs. Poppy reminded them that another letter always goes with Q. They looked through the book “Quilla Queen” and found all of the letter u’s that followed. Then Mrs. Poppy played a song that matched the words in the text. She did not reference the text while reading, but the students sang along. Some students did not read/sing at all. Then Mrs. Poppy talked with the students about how to write the letter Qq and they did a page from their Handwriting Without Tears book. (FN 1-13-19)
This work was often very short and not referenced during any other part of the day. During independent writing, students often wrote the letters that were taught with the incorrect formation, but this did not seem to be a major point of focus for Mrs. Poppy.

The students also had many opportunities throughout the school day to read independently, sometimes from their personal book boxes containing small books they created, and other times from the classroom library. When finished with the independent work, students were told to go to the rug and read while Mrs. Poppy continued to support those working at desks. Since she was busy and there was little active monitoring of the rug, most of the time students talked with one another and played around. Other times, Mrs. Poppy would specifically ask the students to get their book boxes while she walked around the room and checked in with students. In general, students were much more productive during this format.

There were some aspects of the literacy instruction that could be slightly improved upon, such as increased print referencing, increased expectations during independent reading, and decreased focus on conventional spelling during writer’s workshop. These items are minimal and, as in any classroom and at any time, there will always be minor aspects of instruction we can tweak to bolster the effectiveness of instruction. Not all of these minor things need to be addressed in order for heterogeneous guided reading groups to be successful. However, if a major component of a balanced literacy program is absent or limited, it will likely have a negative influence on students’ ability to access increasingly complex texts.

After analyzing a wealth of research regarding how children learn to read, the National Reading Panel (2000), including administrators, teachers, and researchers
identified five areas of instruction that are essential to early reading success. Four of the five were evident in Mrs. Poppy’s literacy curriculum, including phonemic awareness, fluency, vocabulary, and comprehension. However, after spending some time in the classroom, I could see that explicit phonics instruction, the fifth identified element, was given little instructional attention. For example, in mid-November I wrote,

*I have yet to see systematic phonics instruction. Is this in another part of the day I have not observed? Have students not yet started this work? There is a lot of focus of correct spelling (rather than phonetic spelling) of words.* (FN 11-16-18)

While Mrs. Poppy did introduce letter names and sounds, and even talked about word chunks, such as inflectional ending and digraphs, there was limited direct instruction in how to decode or encode unknown words. Sometimes Mrs. Poppy told students to “stretch” a word, but the students seemed to be unsure of how to do this effectively as it was not often modeled or practiced in a shared setting.

During our first interview, I probed into Mrs. Poppy’s thinking and understanding of phonics instruction.

*I asked her what she thought about phonics instruction versus whole language. She seemed to reticently (as if she knew she shouldn’t) say that she falls more on the whole language side of the continuum. She said that she has been doing some [phonics instruction] because she feels like she has to. She said that the other kindergarten teacher believes that phonics should be taught…and that they tried to keep their instruction aligned to make parents happy. She said she had the kids listen to some phonics videos for about five minutes a day which go over some*
phonemic awareness activities. She said she thinks phonics instruction is really boring for kids. (12-10-18)

The five-minute phonics video to which she was referring actually focused on phonological awareness and did not delve into the grapheme/phoneme connections targeted in phonics instruction.

When I listened to students read, and as evidenced throughout many of the individual student discussions, I quickly realized that students had a difficult time reading unknown words using visual cues. Mrs. Poppy’s rich read alouds and shared reading experiences resulted in students developing a strong ability to read for meaning, but when a meaningful context was not presented, even some of highest readers struggled. For example, when presented with words in isolation, Zane read hang for the word hug and Jill read pit for pod and but for bet (FN 2-8-19). To help support students’ growth, I asked Mrs. Poppy if I could start serving as a volunteer might in her classroom by having students meet with me for one of the literacy rotations to practice automaticity of short vowel sounds as well as reading and writing consonant-vowel-consonant words. She graciously agreed, and I met with students approximately two times per week for about 15 minutes during the month of February.

I hesitated to serve in this role in the classroom because I wanted to document a design that was feasible and sustainable with just the classroom teacher. However, I knew the students needed this instruction to be successful in heterogeneous groups. I purposefully made sure that the work I was doing could be easily replicated by a volunteer with little training, since Mrs. Poppy had access to several volunteers and we talked about having a volunteer take my place upon conclusion of the study (INT 3-19-
19). Each time I met with students, we practiced flashcards of the alphabet sounds (specifically short vowels), reading consonant-vowel-consonant words using continuous blending, and writing consonant-vowel-consonant words using sound boxes and white boards.

After a few weeks, Mrs. Poppy really started noticing a difference in their reading and writing skills and commended me on the work I was doing with students. On subsequent visits to the classroom, I noticed increased attention to phonics instruction during whole class instruction (e.g. FN 2-25, 2-27, 3-4). She began explicitly showing students how to articulate the short vowel sounds, using gestures that highlighted the ways in which the mouth moves to make each sound, after she observed this instruction in my small group. She also started building CVC words on the pocket chart for students to read and manipulate. In addition, during guided reading groups, students started working on the encoding process as part of the word work portion of the lesson. In our final interview, Mrs. Poppy spoke about her belief in the importance of working with letters and sounds and hoped to include more of this instruction in the future.

Prior to conducting this mesocycle of the research, I wrongly assumed that all five components of early literacy instruction identified by the National Reading Panel (2000) nearly two decades ago, including phonemic awareness, explicit phonics instruction, comprehension, vocabulary, and fluency could be found with consistency in most primary classrooms. Since that is not necessarily the case, a comprehensive literacy program is an essential characteristic of this design. Adding this missing aspect, in combination to increased levels of text difficulty in which to apply this newly developed skill, may account for why the growth in literacy skills was so drastic in just one quarter.
In order for students to have the instructional support necessary to be able to successfully read texts of increased difficulty, they need to develop a strong and efficient cueing system that enables them to use meaning, structure, and visual cues to solve unknown words (e.g. Fountas & Pinnell, 2009). A majority of the students in Mrs. Poppy’s class were adept at using meaning and structural cues supported by the rich work of read alouds and shared reading, however, they lacked the ability to consistently use visual cues to solve unknown words. While students in homogeneous ability groups also benefit from a well-developed cueing system, the need becomes less apparent as students who have a less developed cueing system end up in the bottom group and receive easier texts on their “level” which requires fewer demands on their cueing system. However, in heterogeneous groups, the pressure quickly mounts when trying to support several students at once who have limited ability to decode unknown, context independent words. Thus, a rich and balanced literacy program that teaches readers a wealth of word solving skills is an essential feature of the design.

**Design Principle II: Research-based Guided Reading Instruction**

*If you want to design guided reading instruction in which heterogeneous grouping is used to support the literacy achievement of all students, you are advised to adhere to the research-based essential characteristics of guided reading.*

While many kindergarten classrooms begin grouping students for guided reading instruction within the first six weeks of school (e.g. Richardson, 2009, 2016), Mrs. Poppy had considerable instructional autonomy and felt that such a direct focus on reading was not needed until the second semester of kindergarten. In fact, early in the study she even admitted to me that she wondered if whole group instruction was more effective in
meeting her students’ needs because it allowed her more instructional time than if they were split up into groups.

Although the topic of guided reading is well-researched, there are perhaps no two scholars who have contributed more to guided reading than Irene C. Fountas and Gay Su Pinnell. Publishing their seminal text *Guided Reading: Good First Teaching for all Children* (1996) and an updated version with nearly all new content titled *Guided Reading: Responsive Teaching Across the Grades* (2017) along with multiple other texts and articles, they have provided a framework for teachers around the world on how to provide purposeful, responsive, individualized reading instruction. Their immense contribution to guided reading instruction informed the core of what occurred during guided reading in this study. Mrs. Poppy has used the original version as a seminal text throughout her career and purchased their second text during this study to learn more. Fountas and Pinnell (2017) identify eight characteristics as essential to guided reading instruction. All eight, albeit with some redefining, were utilized in this study and are considered to be essential to the overall design.

First, students were (1) placed into small groups for instruction, with four to six students per group. Throughout this study, students were regrouped once a month on average, as suggested by Richardson (2009), although we did so with the purpose to maintain (rather than to avoid) the heterogeneity of the groups. Not only does the National Reading Panel (2000) identify small group instruction as beneficial to the development of early literacy skills, such as phonological awareness, but it also fits nicely within the current structure of small group literacy instruction, which Cuban (1993) points out as helpful in creating lasting change.
Mrs. Poppy was also skilled at providing (2) a rich book introduction, as is evidenced below.

From 12:49 to 12:53 Mrs. Poppy led the students on a picture walk through “Ben’s Red Car” which is a Rigby level 2. Mrs. Poppy asked the students to find “Ben,” “red,” and “car” in the title. Then they went page by page as she asked. What is he drawing? What color is it? What did he add? Doors and then what on them? Oh look, he added windows. What is up front? What’s that called? (steering wheel). What’s that brown thing called? It’s a roof-rack. What did he put on the back? A trailer and what? (a boat). (FN 1-25-19).

As noted in the except, Mrs. Poppy helped students practice reading key words, planted potentially new vocabulary words, and effectively built students’ background knowledge about the story. This allowed (3) all students in the group to read the same text. In this particular study, that meant all students in all groups were able to access grade level texts, including those who might have otherwise been hindered by the language demands of this particular book. A word such as “roof-rack” is neither considered phonetically regular for an emergent reader, nor is it a word that is likely to be in their vocabulary. By introducing this word prior to independent reading time, Mrs. Poppy increased access to the text for a wide range of readers.

To build comprehension skills and strategies, Fountas and Pinnell (2017) include (4) a text discussion as another essential component of guided reading. Mrs. Poppy was particularly skilled at this and, in fact, spent perhaps too much time on this element given the limited time in each reading group. Her questions often moved from text-based to the inferential.
Mrs. Poppy began asking comprehension questions such as “Why didn’t he want to play?” “Why do you not want to play?” “How would you pick someone to play with?” (FN 1-25-19).

This focus on the meaning of texts is particularly important when grouping students heterogeneously. Though the text complexity is increased and students may be devoting significant amounts of cognitive energy to solve unknown words, it is important to keep the overall focus of reading on comprehension and engagement with the text.

Focusing on the grapheme-phoneme connection during (5) word work portions of guided reading helps students learn the skills needed to tackle unknown words using the visual cueing system. Although Mrs. Poppy did not include explicit phonics instruction during whole class literacy activities (until February), she did include some word work activities in her guided reading lessons. Sometimes, however, the instruction was too easy or could have more efficiently been done during whole group instruction. For example, sometimes the students would simply draw a game piece with a letter on it and say its name, individually practicing only a handful of letters each day. As the study progressed, the word work portion was more robust and included reading or writing words.

Perhaps the most consequential feature of guided reading instruction for this particular study is (6) purposeful instruction. In order to meet the needs of learners of a variety of levels, Mrs. Poppy had to be acutely aware of and responsive to the reading growth of each student. Differentiating for individualized instruction within each group rather than between the groups is the cornerstone of heterogeneous grouping. Below is an
excerpt of an observation of Kelley reading a level B book shortly after she was unable to pass a level A.

Kelley needed the most support to read this book, but also was very eager to read. She was definitely not frustrated as the traditional leveling system might suggest, especially given scaffolded support. (FN 1-25-19)

In fact, I never observed a student who showed obvious signs of frustration while reading challenging texts during guided reading group (although Mrs. Poppy did mention at one point that she felt Jill was getting frustrated as explored above). I actually began to notice quite the opposite. Allow me to refer again to an excerpt of my reflection after listening to Logan read a level C book, which would have been presumed to be far too difficult for him at the time.

He did a great job. He needed help with the name “James,” and read “says” instead of “said” several times, but clearly maintained meaning throughout the book. He read “sleeping” for “asleep” and when I asked him if it looked right he self-corrected. I praised him as he was reading during a few tricky parts and he was beaming. (FN 2-4-19)

This praise undoubtedly felt genuine to Logan as he knew that he was doing really challenging reading work. He also saw that he was reading, actually reading, the same text everyone else in the group was reading, too. Overtime, his confidence grew tremendously.

The last two essential characteristics of guided reading group that Fountas and Pinnell (2017) suggest are also in-line with the design principles of this study. However, this design requires a much broader interpretation of these final two characteristics than
originally intended. Fountas and Pinnell (2017) suggest that teachers (7) group students together who can read at similar literacy levels and (8) make appropriate text selections. While Richardson (2009, 2016) suggests that students ought to be grouped together only if they are no more than one text level apart, this design shows that “similar” can refer a much wider range, allowing most students of the same grade level to be grouped together. In addition, a much wider range of texts would also be considered appropriate than those traditionally provided to students. Shanahan reminds us that it is not the text level or the student’s ability that is important, but rather the teacher’s ability to provide the guidance, support, and scaffolding needed for students to successfully access the text (web log, 2011).

That said, the three mesocycles of this design support the fact that there are some students who are so significantly advanced in their reading development that true heterogeneous grouping is ineffective in supporting their reading development. This is supported by research on ability grouping that focuses on gifted students (e.g. Kulik & Kulik, 1984, 1987, 1992; Steenbergen-Hu et al., 2016). In this study, Rebecca and Marcus were reading an entire grade level higher than their next closest peer at the onset of the study. Eventually Jill, Leah, and Zane made accelerated progress and were also grouped with Rebecca and Marcus. However, it should be noted that with the inclusion of the final three students, this group also became heterogeneous as these five students passed benchmark assessments at more than the suggested one-level apart, which would traditionally mean they could not be grouped together for guided reading instruction (Richardson, 2009, 2016).
**Design Principle III: Access to Grade Level Literacy Content**

*If you want to design guided reading instruction in which heterogeneous grouping is used to support the literacy achievement of all students, you are advised to expose all students to grade level literacy content (or above) and expect that they can be successful.*

From the beginning of the study, it was obvious that Mrs. Poppy’s teaching philosophy was rooted in the fact that all of her students could and should be successful in their literacy endeavors. Thus, in a sense, it would be reasonable to state that this intervention work began as soon as the students started their kindergarten year, which may also help account for the drastic increase in scores observed between quarters two and three. Merton’s (1948) explanation of the self-fulfilling prophecy is enlightening of the importance of this design principle.

Men respond not only to the objective features of a situation, but also, and at times primarily, to the meaning this situation has for them. And once they have assigned some meaning to the situation, their consequent behavior and some of the consequences of that behavior are determined by the ascribed meaning. (p. 194)

While the self-fulfilling prophecy can be applied to both teacher and students, let’s take a moment to explore this through a student’s perspective. As children are developing their own reading self-concept, they may ascribe to the identity imposed by others. For example, in this study, Logan may have realized that the objective feature of the situation was that reading was really challenging for him, but when Mrs. Poppy exposed him to the same content and instruction as his peers and provided him with the instruction necessary for him to feel success, Logan walked away with the understanding that he can indeed
learn the content. This meaning, ascribed by the situation, potentially altered his subsequent behavior (e.g. putting forth more effort, spending more time reading, increasing risk-taking) to meet those expectations. This self-fulfilling prophecy creates a self-extending cycle (opposite in many ways from the feedback loop explored in chapter two) in which the teacher’s expectations and beliefs alter the child’s behavior in ways that result in reading successes which then reinforce the teacher’s initial beliefs.

During my initial observations, I attempted to see if I could find the “lowest” or “highest” students in the classroom, simply by watching the questions Mrs. Poppy asked of them and the ways in which she interacted with them during the many components of literacy instruction. After several weeks, I was still unable to make these determinations. If I could not tell the difference between the proverbial literacy haves and have nots, then I imagine students were unaware of this as well. Mrs. Poppy made her thinking behind this practice explicit during an interview, *I think they need, I don’t know how to say this, but I really want them to hear everything that we’re doing. Um, because I’m amazed how much they pick up* (FN 12-10-18).

Mrs. Poppy was specifically referring to the introduction of digraphs, vowel pairs, diphthongs, and inflectional endings during whole group instruction, even before all students had mastered the single letter grapheme-phoneme correspondences. For example,

*When talking about the word “today,” Mrs. Poppy drew their attention to the letters “ay” and reminded them that these two letters work together to say the long a sound. Then she had them use their knowledge of “ay” to read “stay”, “may” “say,” and “pay”* (FN 2-27-19).
While this was the type of high-level instruction Rebecca and Marcus needed, all students benefited from this exposure, as evidenced in Logan’s early use of oo, th, and -ing in his independent writing as described in chapter four (FN 11-30-18). Although Mrs. Poppy did not expect that all students would master this higher-level content with the first exposure, over the course of the year students slowly picked up on these skills after hearing them many times.

In addition to believing that all students should be exposed to the higher-level content, Mrs. Poppy established a classroom culture in which all students were expected to answer higher level questions and take risks in learning. In fact, during one of my first observations, I wrote the students in the classroom do not seem to be afraid of getting an answer wrong. Students excitedly answer questions and respond happily when they need to seek help (FN 11-16-18). As a result, Mrs. Poppy often did not wait for hands to be raised, but rather called on various students throughout the day. If they were unable to answer a question with independence, support was provided.

I have already noticed that when a student is asked to answer a question, they are either prompted by the teacher or asked to phone a friend [consult with a peer] until they are able to get the answer correct. Simply calling on another student in the classroom is not an option in this space. (FN 11-16-18)

In Mrs. Poppy’s classroom, the original student called upon was always the student who provided the answer to the class. I wrote about this over and over again as it clearly supported the work we asked students to do during heterogeneous grouping. She doesn’t let the student give up when they don’t get it [the answer]. She gives them time to process and come up with the answer with guidance if needed (FN 1-13-19). Their
increased levels of confidence and stamina when met with difficulties was evident when reading. Below is an excerpt demonstrating what this looked like in the classroom.

Vivian was asked what the word was [that started with t and was missing from the morning message] but she didn’t know. Mrs. Poppy said, “what are we doing tomorrow?” Another student wanted to say the answer, but Mrs. Poppy said, “no don’t phone her.” Eventually Vivian said “field trip!” And Mrs. Poppy reread “Are you excited for our tr__” and Vivian said “trip!” When asked how to spell it, she started stretching the word and saying /ch/ /ch/. Mrs. Poppy corrected her saying /trrr/. Then Vivian said “p”! Mrs. Poppy told her she was missing a sound and helped her stretch it. Finally, she was able to say “i” and “p.” (FN 2-13-19)

Mrs. Poppy clearly expected that Vivian could be successful in solving the word, which Vivian undoubtedly felt and then behaved accordingly. Each step she got closer to the final answer, the more excited she became as she genuinely solved a challenging task and was met with success.

Vivian was perhaps the slowest student to show progress in this study. However, with continued exposure to higher level content paired with intensive instruction based on the belief that she could achieve grade level proficiency, Vivian’s progress was incredible. Of all the students in her classroom, Mrs. Poppy was most surprised by Vivian’s growth saying, Vivian just three weeks ago really still didn’t have a lot of interest and, all of a sudden, it’s just like WOAH! It’s clicking and it’s huge just how fast…even her writing has improved over night! (INT 3-19-19). Ultimately, Mrs. Poppy “believed that working and pushing beyond where [students] feel like they can” (INT 3-19-19) was instrumental in their success and Mrs. Poppy certainly acted on this belief
time and time again with Vivian until Vivian joined the ranks of her peers who were exceeding the grade level reading expectations.

I argue that what occurred in this classroom, and what is essential to the design, is that students’ level of frustration is considered malleable. Traditional levels of text difficulty suggest if a student reads a text with less than 90% accuracy that the text is deemed to be in the child’s frustration level (e.g. Betts, 1946; Richardson, 2009). This study documents that through direct instruction, scaffolding, and believing that all students can and will be successful, teachers can directly impact the level at which a text becomes too frustrating to allow for productive learning. When a climate of productive learning is created, as evidenced in this study, traditional notions of text difficulty are revealed as archaic and artificial boundaries that do nothing more than limit students’ opportunities to deeply problem solve and engage with literacy learning.

Design Principle IV: Synchronous, Explicit, and Systematic Instruction

If you want to design guided reading instruction in which heterogeneous grouping is used to support the literacy achievement of all students, you are advised to provide students with synchronous, explicit, and systematic strategy instruction to solve unknown words.

Learning how to read is a complex process (e.g. Chall, 1983, Ehri, 1995). It involves mastering a variety of skills, such as learning how text functions, mastering the alphabetic principle, and so much more. While learning this complex process, students are taught a variety of reading skills and strategies to help them read an unknown word. For example, Mrs. Poppy taught her students about characters such as “Lips the Fish” and “Chunky Monkey” to help students remember ways in which they can solve an unknown word. The image of “Lips the Fish” reminded students that they need to get their lips
ready for the first sound in the word, while “Chunky Monkey” reminded them to look for chunks or parts of words they knew to help them solve.

However, during the first two mesocycles of this research I noticed that many students, particularly those struggling to learn the complex process of reading, did not know which of the many strategies they learned to use when they approached an unknown word. They haphazardly selected from the repertoire of strategies until they either solved the word or gave up trying. When asking students to read more complex text during heterogeneous guided reading, students will encounter even more unknown words than in the traditional leveling system. As a result, it is paramount that students develop a strong and efficient system for solving unknown words. Arguably, as mentioned above, this is ideal for all students regardless of how they are grouped for instruction, but in heterogeneous systems it changes from the ideal to the essential in order to support the learning of all students.

Throughout the multiple mesocycles of this project, Figure 5.1 has evolved as a tool to teach and systematically support the development of an effective problem-solving process. The strategy card makes explicit the “in your head” processing system that proficient readers use without even noticing. The steps are purposefully organized to encourage efficient problem solving, by having readers do the least amount of work necessary to solve a word, which assists in both ready fluency and reading comprehension.

Using the strategy card, students are taught that when they get to an unknown word in a text, the first step is to (1) “get their lips ready” by putting the first sound in their mouth. This is the most efficient of the problem-solving steps and often students can
arrive at the word without any further problem solving. Take for example the following sentence: *I saw the cow go in the b____.* Many readers can use the context of the sentence along with the initial /b/ sound to come up with the word *barn*. They would not need to spend additional time “sounding out” the word as students are often told to do. If, however, they were unable to solve this word using the first step, they would move on to strategy (2) “look at the picture and think what makes sense.” Keeping the /b/ sound on their lips, they would look at the picture, which in this case would show a big red building with white lines across the big doors in the middle of an open field. By getting their lips ready, looking at the picture and thinking about what makes sense they are likely to arrive at the word *barn*. If that didn’t work, and only if that didn’t work, they could move on to strategy (3)” find a chunk you know and blend.” This step, closely resembling the “sound it out” strategy, requires that students visually process the entire word. In this instance, we might notice the chunk *ar*. Using that chunk, we could read /b/ /är/ /n/ to read the word *barn*. In this instance it would be unlikely that a student would not successfully arrive at the word *barn* using the first three steps, however, if they were still struggling they would (4) “Reread” to put the context of the story back in their head and then (5) “Read ahead” to see if more text would provide them with additional clues that they could use to solve.

I introduced a slightly different version of this strategy card to Mrs. Poppy at the beginning of the study and we adjusted it together to meet the language of instruction she had already been using in the classroom. For example, for the first step I previously had “Get your mouth ready,” but we changed this language to match up with the language of
“Lips the Fish” which she had already introduced to her students. Our collaboration also led to more explicit images and an enlarged version for whole class use as well.

*Figure 5.1 Reading Strategy Card*

<table>
<thead>
<tr>
<th>cat</th>
<th>cat</th>
<th>cat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Get your lips ready.</td>
<td>Look at the picture and think “What makes sense?”</td>
<td>Find a chunk you know and blend.</td>
</tr>
<tr>
<td>I can see the cat.</td>
<td>Reread.</td>
<td>Read ahead.</td>
</tr>
<tr>
<td>The cat is brown.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This reading strategy card has been shared with many teachers across the state in an editable document so that it can be adjusted to match the language and visuals of instruction used in their schools. What is important here is not the exact language or images we chose, but rather the systematic and explicit nature of teaching students how to solve an unknown word. Students no longer randomly guess or use inefficient problem-solving methods such as “sounding out a word” when it is not necessary. Not only does this explicit strategy instruction lead to increases in fluency, but it also contributes to reading comprehension. When students know exactly how and when to use each step of the strategy card through explicit teaching and teacher (as well as peer) modeling, they can begin to take ownership of the problem-solving process.

At the beginning of the formal intervention phase of the study, when I was still working with Mrs. Poppy to truly confront the norm of grouping students by ability, I asked her if we could introduce the strategy card to students as a “first step” in the process. She agreed and our work began. We spent time explicitly teaching it to students,
one step at a time, while covering the rest of the steps with sticky notes (which unintentionally peaked their interest in learning subsequent strategies).

*Then Mrs. Poppy got out the strategy card and reviewed with the students to get their lips ready and look at the picture to see what makes sense. Vivian and Kelley eagerly shouted out these strategies to Mrs. Poppy. She passed out a strategy card to each child.* (FN 1-25-19)

Over time, Mrs. Poppy introduced additional steps to the strategy card until students were familiar with them all.

In previous mesocycles of the design, collaboration with support staff, such as special education teachers, intervention teachers, and speech/language pathologists was an essential part of the design to support synchronous learning across texts. In this mesocycle, only two students (Logan and Hattie) were pulled out of the classroom for articulation services, so this strategy card and other elements of the design were not shared with additional support staff. However, based on the success of this portion of the design within the previous two mesocycles, it continues to remain part of this design.

**Design Principle V: Give the Design Time to Flourish**

*If you want to design guided reading instruction in which heterogeneous grouping is used to support the literacy achievement of all students, you are advised to give the design time to flourish before abandoning it.*

It has become a common sentiment through the language and use of “independent,” “instructional,” and “frustration” levels (e.g. Betts, 1946; Richardson, 2009, etc.) that frustration has no place in learning how to read. As explored above in design principle three, frustration can be considered a malleable characteristic and
redefined as productive struggle, given the right supports. However, this requires explicit attention and deconstruction to the normative claims that (1) students have a similarly defined threshold of tolerance for frustration and (2) once students cross this threshold they can no longer productively learn. As Oakes’s (1992) suggests, if normative dimensions, such as the place of frustration or productive struggle in learning how to read, are not addressed alongside the technical, then the intervention will likely be unsustainable.

I was worried that the design might come to an untimely end when Mrs. Poppy first began truly grouping students heterogeneously for guided reading instruction. As could be expected, the students needed additional support and encouragement to read texts at a much higher level of difficulty than they had been experiencing.

*Overall, Kelley read with many errors, but did not seem to get frustrated at all.*

*She got stuck several times on the word here, saying “had” and “have” and usually correcting herself. When she got to the word “big” she said /b/ /i/ /i/ /g/ “bug.” Kelley finished the book, but I was surprised how difficult it was given the number of times she had practiced.* (FN 2-8-19)

As I listened to students read those first few weeks, my fieldnotes were inundated by the word “support.” It was becoming a theme of guided reading groups.

*Then I moved on to work with Vivian. She needed a lot of support to read the book. She read “it tires” instead of “it has” clearly using meaning, but not visual cues.* (FN 1-25-19)
The experiences above felt qualitatively different than when listening to students read texts that provided them with only little challenge as is typically common during guided reading groups.

While the benefits of providing students with texts of increased complexity have been well-documented (e.g. Benjamin & Schwanenflugel, 2010; Johnston, 2000; Morgan et al., 2000), the classroom teacher’s job becomes much more difficult when presenting students with this additional challenge. There is immediately an increased responsibility to teach and provide scaffolding to meet the needs of the students who encounter many more problem-solving situations. Shanahan writes in reference to the traditional leveling system for guided reading, “The problem with guided reading and similar schemes is that they are focused on helping kids to learn with minimal amounts of teaching (something Pinnell and Fountas have stated explicitly in at least some editions of their textbooks)” (web log, 2011). Ultimately, when the normative format for guided reading requires little teacher support as noted by Shanahan, meeting the literacy needs of students in heterogeneous groups can at first feel quite uncomfortable and overwhelming.

Throughout this transition period, I connected frequently with Mrs. Poppy to see how she was feeling about the grouping structure. I wondered how she would respond to the real productive struggle in word solving that I had not witnessed prior to that point. Overall, she felt very positive, and I was pleasantly surprised by her continued enthusiasm despite the increase in challenge for both the students and herself. The only point at which she began to feel uncomfortable with the level of a students’ frustration was when working with Jill as described in chapter four. However, after suggesting that perhaps the homogeneity of the group (since Marcus and Rebecca both assessed at a
much higher level than Jill), rather than the text level was causing this frustration, Mrs. Poppy added Leah and Zane to the group which alleviated any signs of frustration she observed from Jill.

After getting over these initial growing pains, it didn’t take long for the results of heterogeneous grouping for guided reading instruction to begin to show in students’ literacy progress. In early February, Mrs. Poppy started to assess many of her students using the Rigby PM Benchmark system because she couldn’t believe how much progress they were making. She was always eager to show me the results.

*Mrs. Poppy showed me a copy of Molly’s running record. In just one month she was able to pass a level B and a level C text. She said she couldn’t believe how well her students were doing.* (2-11-19)

Although getting to the point of literacy success was challenging, the results of students’ literacy progress reinforced the importance of the work we had done together.

**Summary of Design Principles**

According to Bakker (2018), design principles are guidelines, advice, or heuristics that are used in subsequent iterations of the design and adapted, revised, or dropped after further exploration. They are considered essential features for successful implementation of the design in other contexts (Brown, 1992). After the first two mesocycles of the design, eight design principles emerged. During this mesocycle, several of those design elements were subsumed under larger categories and combined with additional design elements for more clear and accessible guidelines for teachers to implement the design. Other design principles, such as, “Devote almost the entire fifteen minutes of guided reading instruction to in-text reading” were dropped as elements of the design. While in
the first two mesocycles, this principle seemed to be essential, in this iteration Mrs. Poppy thoroughly attended to all aspects of guided reading suggested by Fountas and Pinnell (2017) and her students still demonstrated tremendous literacy growth. Thus, while I think providing students more time with continuous text reading is certainly important, data does not support it as an essential feature of this design.

Below are the five design principles that arose from the data collected in three mesocycles of grouping students heterogeneously for guided reading instruction. They are considered to be “the absolutely essential features that must be in place to cause change under conditions that one can reasonably hope to exist in normal school settings” (Brown, 1992, p. 173). This is the heart of the study in terms of practical application.

If you want to design guided reading instruction in which heterogeneous grouping is used to support the literacy achievement of all students, you are advised to:

1.) do so within a comprehensive and balanced literacy program.

2.) adhere to the research-based essential characteristics of guided reading.

3.) expose all students to grade level literacy content (or above) and expect that they can be successful.

4.) provide students with synchronous, explicit, and systematic strategy instruction to solve unknown words.

5.) give the design time to flourish before abandoning it.

As can be noted, several of these principles are already well-executed in many classrooms as part of research-based literacy instruction. Making the transition from homogeneous to heterogeneous grouping in some classrooms will therefore be easier than others. For some teachers, there are only two to three additional design principles that
need to be integrated into their work in order for heterogeneous grouping to be successful. For others, it will require direct attention to all aspects of the literacy curriculum from guided reading to shared reading and beyond.

**Theorizing the Why?**

Analyzing the substantial gains that students made in just one quarter of the instructional design, demands the question *why*? Anderson and Shattuck (2012) suggest that attention to the *why* question is essential in taking this work beyond that of action research and into the scholarly arena. Particularly, attention to the theoretical insights can lead others to identify how this work can be adapted to new circumstances (Cobb et al., 2003). In this case, the design theory can not only provide insight to the malleable nature of ability, appropriate text-reader match, and grouping structures, but also the ways in which other educational interventions can be sustainable and feasible in real-life classrooms. In addition, with aims to support the explicit, but not always realized, egalitarian goals of public education, this theoretical work can serve as a bridge between the inequitable present and a more equitable future.

While the central aspect of the design in this study is grouping students heterogeneously for guided reading instruction to increase literacy achievement for all students, the dataset suggests it is not heterogeneous grouping in and of itself that is the panacea of the design. Literacy development is a complex, multifaceted process that occurs over time (Ehri, 1995, 1999; Chall, 1983). The design principles outlined in this study reflect the instruction necessary in the entire literacy block in order for heterogeneous grouping to be successful. Thus, unlike other prepackaged literacy interventions that attend only to the timeframe and instruction that is core to the
intervention itself, this instructional design and underlying theory recognize that heterogeneous grouping of students for guided reading instruction does not occur within a vacuum.

Let’s take a look, for example, at the following scenario. If Mrs. Poppy simply placed students in heterogeneous grouping, but acted during whole group instruction in ways that undermined the expectation that all students could read grade level texts, how might this influence students’ overall reading achievement? Imagine Logan being asked only easy questions, or allowing others to answer for him while he was processing an answer. Consider Kelley never being called on to read the morning message because Mrs. Poppy thought it would be too difficult and cause her unwarranted frustration. Based on the self-fulfilling prophecy (Merton, 1948), it is reasonable to expect that reading scores at the end of quarter three, particularly for the lowest achieving readers, might have looked much differently.

Perhaps a second example will address this idea of the embedded nature of heterogeneous grouping on a more concrete level. Although there were several strong components of a comprehensive literacy program in place within Mrs. Poppy’s classroom, explicit phonics instruction was notably limited. It was clear that the students, even those with the highest levels of literacy achievement, could not decode phonetically regular words with consistency. In order to read texts of increased complexity, the students needed to be adept problem solvers at the word level with a fully developed cueing system in which they could use meaning, structural, and visual cues to read unknown words. Explicit instruction in how to systematically decode and encode words was necessary in supporting their reading of continuous, complex texts in guided reading
work. Without this instruction and the subsequent strengthening of students’ problem-solving abilities, the combined needs of all students’ in a heterogeneous group may have exceeded the support one teacher was able to provide.

*Figure 5.2* illustrates the Heterogeneous Guided Reading Groups model created from this research. This model provides understanding of *why* this design worked. It is complex, multifaceted, and covers all components of literacy instruction. For an intervention to be successful, this type of holistic overview demonstrating how it fits in, is supported by, and contributes to existing programming. It offers a more realistic view of the implementation process leading to a sustainable design. As described above, heterogeneous grouping is embedded within quality, research-based guided reading instruction which is further embedded in a comprehensive literacy program. However, all of these must be placed within a context in which all students are exposed to grade level content throughout the day with a teacher that expects and believes all students can achieve at grade level or beyond. Synchronous, explicit, and systematic teacher instruction serve as the significant inputs in the design spurring the effectiveness of the heterogeneous guided reading instruction. Time is considered necessary in unpacking the norms of grouping students by ability and restructuring guided reading as a space of intensive instruction. Finally, the benefits of heterogeneous guided reading emerge: access to grade level texts, positive literacy identity construction, peer modeling, and curricular integration. These outputs ultimately lead to increased reading achievement for students of all levels.
This design champions the professionalism of teachers by proposing that their instruction, rather than a student's current level of literacy achievement, is the defining factor in determining not only the level of books children can access but ultimately their level of reading achievement. This major input of teacher’s synchronous, explicit, and systematic instruction into the design moves heterogeneous grouping from possible to
probable by empowering teachers to empower students, while still recognizing the hard work and multiple dimensions that are necessary to ensure success. Mrs. Poppy spoke to the increased demand on teachers saying,

*making sure I know the text well enough before I present to the kids has made a huge difference. With those lower [books] you get used to it and you’re like, “Oh yeah, yeah,” but now where it’s more challenging...I have to make sure that I have set them up for success.* (INT 3-19-19)

By adding the dimension of time, this design also recognizes that this work requires patience and dedicated attention to the practical application of these components before change may become apparent. It ensures that teachers get the opportunity to see the benefits of their hard work.

Within the very heart of this design is heterogeneous grouping for guided reading instruction. What is it exactly about this grouping structure that contributes to the overall success of the design? To begin with, as mentioned throughout this study, heterogeneous grouping puts grade level texts in the hands of students. When we have students consistently reading below level texts in guided reading groups, we cannot expect that they will pass a grade level text at the end of the year when they have never had experience reading text of that difficulty level. The provision of complex texts (always combined with appropriate scaffolding) increases the opportunities students have to problem-solve and engage in the deep, cognitive work of reading (e.g. Johnston, 2000). Imagine a child reading a simple pattern book with only one opportunity per page to really engage in the problem-solving experience (likely the last word in the sentence). This child will likely skim through even those words which they should problem solve.
because it is been reinforced throughout the book that they do not need to “read” the text. However, imagine presenting the same child with a more complex text that requires them to slow down and attend to each word. Every word in this text presents an opportunity to cross check their cuing systems to make sure that each word makes sense and looks right because it is not part of a predetermined, memorized pattern. Not only will this child have more opportunities to problem solve, but he or she is also likely to feel genuinely proud of the effort expended to successfully read those pages.

Furthermore, by grouping students heterogeneously, students avoid the ascribed identity of “struggling reader.” In homogeneously grouped classrooms, this identity work can become so pervasive that teachers often refer to the students in their lowest groups by names such as “lowbies” or “low babies.” Not only do these labels reflect the teacher’s view of their literate identities, but they also infantilize these students making it seem even more unlikely that they will be able to catch up to their grade level peers. Such terms often connote pity for the child and stability of their ability rather than the belief in their agency as a teacher to ensure that this child is never deserving of such a label in the first place.

Another key feature of heterogeneous grouping likely to contribute to the success of all students is the heterogeneous peer modeling that occurs while placed in heterogeneous groups. While Fountas and Pinnell (2017) suggest that students read in a whisper voice during guided reading instruction, *strong reading*, or loud and confident reading, allows students to literally hear themselves and others problem solving unknown words. This supports the decoding process by allowing them to hear the sounds as they are blended together to assist the recognition process, and it also enables them to hear
peers model this process that they can later try themselves. After encouraging Mrs. Poppy to have her students in heterogeneous groups engage in strong reading, she reflected, “having them read above a whisper read has made a huge difference because they are hearing each other and going. “Wait I can do this too!” (INT 3-19-19). She also told the following story about Bailey.

*I think they learn so much from their peers...because I see Bailey waiting for Brittany to get to that word today. You know what I mean? Because she knew they were on the same page. She was listening in to her read text to reassure herself that, “Yep! That’s what it was!” It was just cool! (INT 3-19-19)*

While students listening in to the other students in traditional system of guided reading might have been frowned upon, this design model suggests that students can and should use one another as models in the reading process.

Strong reading also provides the teacher with valuable information regarding each child’s active participation in the reading process. When students are whisper reading, the teacher only becomes privy to one child’s use of problem solving, or lack thereof, at a time. Meanwhile, other students in the group may be skipping unknown words, breaking down into frustration after attempting a word for extended periods of time without success, or completely checking out and daydreaming while the teacher is none the wiser. Heterogeneous grouping of students combined with loud problem solving enables teachers to provide responsive instruction to the students, students to provide models of problem solving to one another, and the space to take on new problem-solving strategies in a scaffolded context.
So far, most of the theory behind why the heterogeneous grouping aspect of this design works revolves around how it supports students who would have been placed in the lowest ranked reading group. However, in all three mesocycles of the design, even the highest readers made greater than anticipated reading gains (with the exception of one student, Marcus, who made exactly the gains anticipated). Much of the criticism about heterogeneous grouping focuses on how it will be detrimental to high achieving students. Why was that not the case in this study? In traditional systems of homogeneous grouping for guided reading instruction, the higher-level reading content would be reserved for the fifteen to thirty-minute guided reading group of the highest learners. This higher-level content could not be integrated into other components of the literacy curriculum because only a few students were privy to that knowledge. When groups are heterogeneous and all students are exposed to this higher-level content within their guided reading groups, it can be discussed, reviewed, and practiced during all components of literacy instruction. In this way, higher level students’ literacy instructional needs are no longer only supported during their short reading group, but throughout the entire day, allowing this material to be learned at a faster and more appropriate pace.

Heterogeneous grouping of students, then, is embedded within a comprehensive literacy program in which all students are exposed to grade level content with the expectation that they can succeed. Furthermore, it is surrounded by the essential components of guided reading instruction suggested by expert researchers in the area, primarily Fountas and Pinnell (2017). The work proposed in this design is not suggesting that teachers reinvent the proverbial wheel, rather it encourages them to build off of well-established, well-known research for effective literacy teaching while slightly tweaking
two of the essential characteristics of guided reading (appropriate text selection and grouping students of similar ability) in ways that offer students access to grade level reading content materials. As is the goal in design-based research (Cobb et al., 2003) this study contributed to both the practical application and theoretical insights on the nature of grouping students heterogeneously for guided reading instruction.

**Implications**

For over one hundred years, researchers have studied the effects of grouping students homogeneously by ability for instruction (Steenbergen-Hu et al., 2016). While many have warned of the negative effects of homogeneous grouping practices (e.g. Condron, 2008; Oakes, 1985/2005; Rist, 1970; Slavin, 1987; Tach & Farkas, 2006), there has been a lack of descriptive research documenting both feasible and sustainable replacement practices (Hamilton & O’Hara, 2011; Hong et al., 2012; Oakes, 2017).

This study not only provides evidence of “proof of principle” (Bakker, 2018, p. 101) that heterogeneous grouping for guided reading instruction is possible, but it also answers the “how question” (Bakker, 2018) assisting in practical application of the design. Through the delineation of student progress over time, as well as the “thick descriptions” (Geertz, 1973) of the design principles, practitioners are given the opportunity to “see” the influences of the design within the context of this classroom. This allows for “transferability” (Lincoln and Guba, 1985), or the ability for teachers to transfer or extrapolate upon the design principles in their classrooms as they see fit. In addition, in an age of scripted, teacher-proof, curriculum and intervention programs, this instructional design values the professionalism of teachers who use their knowledge and
skills to provide increased amounts of responsive and scaffolded instruction to meet the literacy needs of their students.

Furthermore, school administrators and district personnel who are eager to support real, research-based change in students’ literacy outcomes would be wise to use this study as a springboard for the professional development of teachers. After close reading and understanding of the design in application to alternative contexts, a cohort of teacher mentors could serve in much the same capacity as I did in this study. Their responsibilities would include encouraging other teachers to slowly combat the norms of ability grouping, supporting their process of instructional change, and collaborating with teachers on innovative instructional ideas. This model is low in cost and high in reward. No additional materials, instructional time, or “experts” are needed to implement this highly successful design, rather it builds off the comprehensive literacy program, “hallowed niche” of guided reading, and professionalism of teachers and staff already available in the school or district.

Given federal and state level policies to support the reading achievement of all students, such as Every Student Succeeds Act, policy makers should consider the enormous potential of heterogeneous guided reading groups to increase reading achievement. While within-class ability grouping disproportionately disadvantages poor and minority students, a group often targeted in legislative policies, three mesocycles of this design demonstrate how heterogeneous grouping for guided reading instruction benefits all students and actually serves to decrease the achievement gap. By offering increased exposure to grade level reading materials in addition to supportive and
responsive classroom instruction, teachers can potentially alter a child’s schooling trajectory from one of struggle to one of success.

This study also has significant implications for early literacy research. First, it offers a new educational innovation prime for further study. While the three mesocycles of this design-based research study focused on the benefits of heterogeneous grouping for kindergarten learners, exploration into other grade levels is warranted. Can the effects of year(s) worth of ability grouping be counteracted by heterogeneous grouping or do the effects become so crystalized and students’ abilities so disparate that heterogeneous grouping is no longer possible? Additionally, this study suggests a theoretical lens that can be utilized to better understand the effects of this educational innovation and others.

It is unique in its holistic, comprehensive nature, acknowledging that heterogeneous grouping (or any other educational innovation) does not exist in isolation. Rather for a design to be both feasible and sustainable, a bird’s eye view is important in understanding the mechanisms by which learning occurs.

This study not only explores the ways in which grouping students homogeneously by their ability is inequitable, but it also provides a design pointing towards a more equitable future. According to Condron (2008), curriculum differentiation, such as that utilized in homogeneous grouping for guided reading instruction, is one of the most salient factors contributing to the achievement gap. It is time that educational practices live up to their egalitarian goals. For teachers, administrators, policymakers, and researchers, the design principles offer an exciting new view of what literacy education can and should look like. For students, particularly those who have been historically underserved and underrepresented, it offers so much more; it offers the opportunity to be
exposed to grade level content, the opportunity to receive instruction and scaffolding needed to access that content, and the opportunity to avoid the label of a “struggling reader.” For these students, it means literacy success is finally within reach.

**Limitations and Future Research**

As noted by Brown (1992), design-based research trades experimental control for “richness and reality” (p. 152). While lacking experimental control can certainly be considered a limitation, the surge of the detracking movement followed by a return to the status quo of ability grouping suggests that teachers either did not find a viable and sustainable alternative to ability grouping, or that available alternatives did not sufficiently fit into a “hallowed classroom niche” (Brown, 1992). The richness and reality offered in this study provides researchers, teachers, administrators, and policymakers with insight into not only a possible alternative to the inequitable practice of homogeneous grouping for guided reading, but also descriptive insight that allows for practical application of the design. That said, there are several limitations in this study as well as avenues that are prime for further study.

This study documented the process of redesign and implementation of an alternative to ability grouping practices within one classroom in one school. While this sample size allowed for the thick descriptions of the research context and the literacy growth of all student participants, inviting a team of teachers in the design process may have provided additional insight into the ways in which teachers grapple with the technical, normative, and political dimensions when implementing this educational design. Working with a team was considered at the onset of this study development, but I was unable to find an entire team of teachers willing to participate, which may in itself
speak to the entrenched nature of ability grouping for guided reading. Many teachers were interested in the idea behind the design, but simply stated that they couldn’t “see” how heterogeneous grouping would actually work in the classroom setting. Being able to show future teachers both the findings and the design principles of this study may help them better envision how this practice might look within the context of their own classroom setting and therefore allow for deeper investigation into the design.

Another limitation in this study is that the class of students under investigation was both less diverse and smaller than many of the classrooms in which teachers might hope to implement this design. There were no students receiving special education services or English language instruction, and only three students received speech/language support in the area of articulation. Just one of the thirteen students in this group would be considered a racial/ethnic minority. In addition, all students were able to attend a parochial school, albeit perhaps some with the assistance of a scholarship, in a middle-income neighborhood. As described above, research documents that poor and minority students are overrepresented in the lowest groups in ability-grouped classrooms regardless of their actual reading achievement (Condron, 2008). Thus, those students standing to benefit the most from this practice were underrepresented in this study. While the first two mesocycles of this research had greater heterogeneity in terms of social class and racial/ethnic makeup and had similarly promising results, further research of this design should be conducted in schools serving higher percentages of poor and minority students.

Literacy skills are arguably among the most important skills for a person to develop. Being able to read and write opens doors to content area learning, engagement
with storied worlds, expansion of vocabulary, and much more. The students in this story, particularly those who were considered to be struggling readers, made tremendous gains within the one quarter of design implementation. Following this group of students, or others, into first grade and beyond may provide additional insight into the ways in which alternatives to ability grouping affect students’ overall schooling trajectories. Furthermore, the study of heterogeneous grouping in later grades would be instrumental in understanding if the effects of ability grouping, such as the increased achievement gap over time (Lleras & Rangel, 2009) can be reduced through this design.
CHAPTER VI

CONCLUSION

Jeannette Oakes, perhaps the most well-known champion of detracking efforts, acknowledged the hard work of educational reform in this area (1992). Ability grouping is an educational practice that is steeped in tradition and although there has been a barrage of research encouraging its demise. Still it remains, not just in a few classrooms here and there behind closed doors, but in a majority of classrooms across the country. As discussed in chapter two, Oakes (1992) articulated three dimensions that must be attended to in order to produce lasting educational change: the technical, the normative, and the political. As I close this study on an alternative to within-class ability grouping for guided reading instruction, the proposed design will be measured against Oakes’s dimensions of change framework. As mentioned above, this study demonstrated not only proof of what is possible, but also suggested a clear path for implementation in other classroom contexts. Still, does this design have what it takes to withstand the crushing weight of tradition? Or, is it destined to remain merely a good idea without the depth and breadth needed for sustainability?

The Technical Dimension

The technical dimension is the dimension most commonly attended to in educational reform. It focuses on the what of the design, the curriculum or materials that must be implemented and utilized to achieve the desired results. Over time, research studies have attended to this dimension (e.g. Cunningham et al., 1991, 1998) with some success, but ultimately ability grouping has held steadfast. In the present study and in
response to calls from researchers in the hopes of more widespread change, the technical
dimension has not only been introduced in this study, but fully fleshed out through the
identification and description of the design principles. These design principles have been
created, adapted, and revised based off of three mesocycles of grouping students
heterogeneously for guided reading instruction. Data regarding what works, what doesn’t
work, and what else could be tried has been analyzed and refined to result in five clearly
defined and articulated principles that can be further adapted in new classrooms. Through
this work, classroom teachers can envision what it takes to make heterogeneous grouping
for guided reading successful in their own classroom context and slowly take ownership
of the design process. Even more exciting, teachers do not any need additional materials,
money, or resources to implement this design. This work can begin in their very own
classrooms tomorrow.

As described above, not only has the what of this design been clearly articulated
and described, but it has been shown to have truly tremendous results. Three mesocycles
of the design have demonstrated that all students, of all initial achievement levels, benefit
from this design. It has supported the work of those students who are hurt the most from
systems of homogeneous grouping (e.g. Condron, 2008, Lleras & Rangel, 2009, Tach &
Farkas, 2006), while avoiding the potential concerns of heterogeneous grouping, such as
hindering the learning of our highest students. According to Oakes (1992), the new
technologies that are suggested must “create educational opportunities that are at least as
rich and rigorous as those previously enjoyed by the high track students” (p. 19). In the
present study, the design has demonstrated exactly that.
Normative Dimension

The normative dimension focuses on the ways in which the design attends to the current traditions and norms that sustain educational practices. This dimension is particularly difficult to address given that inequity in educational opportunity has been a covert cornerstone of American education, masked by the explicitly expressed egalitarian goals of public education. Grouping students separately based on presumed ability, which often leads to de facto segregation based on race and social class, has become so prominent in primary classrooms that even well-intentioned teachers and district officials have a difficult time “seeing” this practice and naming it as such. As mentioned above, when recruiting for this study, a district official was adamant that ability grouping for guided reading did not exist in the district even though it was in fact part of the official curriculum. Furthermore, Mrs. Poppy, who already believed in the value and importance of heterogeneous grouping and regrouped her students to attempt this ideal, ultimately still ended up with a high, middle, and low reading group.

Attending to this norm means unpacking our expectations and beliefs about students’ ability. It means valuing the place of frustration redefined as productive struggle, and even sometimes as productive failure (Kapur, 2008) Admittedly, when students were finally grouped heterogeneously, it was initially uncomfortable for me as a teacher and later as a researcher in Mrs. Poppy’s classroom. The norms of reading groups have become such that students are presented with limited amounts of challenge, but this design suggests that the presentation of challenge means the presentation of new opportunities: the opportunity to access grade level text, the opportunity to problem solve, and ultimately the opportunity to achieve. Getting to see the power of this newly
proposed norm for reading groups means that teachers must give the design time to prove itself thereby deconstructing the current norm. This is exactly why time was written into the design principles as an essential feature of the design. Without this component, the norms of inequitable opportunity and limited amounts of challenge will likely result in a preemptive end to the design.

This design also purposefully attends to the normative structure of the classroom during literacy instruction. According to Cuban (1993), if an educational innovation is to be long-lasting, it helps if it fits nicely within the sacred traditions of the classroom. Following this idea, Brown (1992) writes,

Reading group has a long-standing place in the school day -- teachers are used to arranging their classes to accommodate reading groups; all (all!) that is needed is to redefine the activities that take place in a socially sanctioned niche. (p. 172)

This statement speaks to both the superficial ease and the realistic challenge of replacing ability groups with an alternative practice. By retaining the time-honored tradition of small group reading instruction, this design fits into Brown’s “socially sanctioned niche.” Not only are teachers comfortable with this practice, but research has shown that small group instruction is beneficial to students (National Early Literacy Panel, 2008; National Reading Panel, 2000). In heterogeneous grouping for reading instruction the small teacher to student ratio and purposeful instruction remain the same as those in homogeneous grouping structures, but what changes are the limitations placed on students based upon their presumed ability.
Political Dimension

In this study, the technical and normative dimensions have been thoroughly attended to through the design construction and design principles. The final dimension which must be attended to is the political. As a public entity, schooling is inherently political, and as a result it has been the subject of fierce debates and critiques over time. In his seminal work *Schoolteacher*, Dan Lortie (1979) describes the *apprenticeship of observation* through which teachers feel as if they know how to teach based on their twelve (or more) year apprenticeship as a student. The knowledge gained during this apprenticeship is hard to overcome during only a few years of teacher training focusing on new teaching methods. Not only does this influence the normative dimension, as we are more likely to teach that which we have seen and experienced, but it also influences the political sector as well. Teachers, administrators and policymakers vie for their vision of education often informed by their own experiences.

In the technical section of this chapter, the ease of implementation of this design was highlighted, stating that teachers could begin this work in their very own classrooms tomorrow. However, the implementation process may be more complex from a political standpoint. Unfortunately, ideas about teaching and learning such as heterogeneous grouping for reading instruction can be thwarted at many levels. Often district or building level decisions regarding instruction are made to ensure that all children receive uniform instruction. This was indeed the case for Mrs. Poppy in her previous school district. All teachers in the primary grades were instructed to group students for guided reading instruction using the traditional notions of text difficulty and teachers had little autonomy to decide otherwise. However, teachers in some buildings, such as those in King
Elementary school, have considerably more freedom in instructional decision-making.

When I talked with the principal of King Elementary School about conducting my research in Mrs. Poppy’s classroom, he warmly said, “all types of learning are welcome here” (FN 11-14-19) and gave us permission to do whatever we felt was best for the kids.

Although the teachers in this school had considerable autonomy compared to many in the public sector, the political maneuvering amongst the staff was still quite apparent. Mrs. Poppy often felt pressured by the early childhood team to teach in ways that did not align with her training as a reading specialist or the research she read on early literacy development. Some of these pressures had the potential to directly impact the heterogeneity of her guided reading instruction.

*Mrs. Poppy said that the other [kindergarten] teacher wants them to switch their two highest and two lowest students for reading groups [to make grouping more homogeneous by ability]. Mrs. Poppy said she was hesitant about having her two lowest students pulled into a low group. She said she thinks they need the positive role model of their higher-level peers.* (FN 1-13-19)

In this situation, Mrs. Poppy approached me asking for advice. She knew what she felt was best for her students, but was unsure about how to navigate the conversation with her colleague. Ultimately, she was able to retain her own students for guided reading instruction and continue to group them heterogeneously, but this was not always the case when such discussions arose. For example, decisions by the early childhood team meant that Mrs. Poppy had to use an assessment system she did not feel was valid and teach a list of sight words she felt were not appropriate for kindergarten students.
However, Politics within the school can also lead to positive changes as well. When Mrs. Poppy’s students started to show progress in heterogeneous guided reading groups, she began to share this information with her colleague. It is exactly this type of informal conversations that can begin to change the political landscape in a school.

*Mrs. Poppy and her teammate started talking about their students and the levels they were reading at. Mrs. Poppy said that her highest group will be reading Fs and that all of her other students will be reading Cs and Ds. (It seemed she was happy to share with her colleague that her students were doing so well.) The other teacher talked about using the guided reading format and Mrs. Poppy said she wasn’t really using the format [of grouping students homogeneously], but was instead giving the students what they need.* (FN 2-11-19)

While I did not see any changes in the teaching practices of the other kindergarten teacher during the study, there is potential for this design to spread through their partnership.

Finally, politics can also work at a parent level as well. Even those who are not teachers experience the *apprenticeship of observation* and often have ideas regarding the type of schooling they expect for their children. As a result, parents often advocate for students to be placed in the highest groups, knowing that such grouping structures exist and that higher placement will result in increased future opportunities. To combat this, I am reminded of Oakes’s words (1992) when she suggests that the alternative to ability grouping practices must offer instruction that is at least as effective and rigorous as that which is offered to the highest groups in ability grouped classroom. In this study, Mrs. Poppy did not have any concerns from parents regarding the changes in the grouping
structure, and attributed this to her relationship and communication with parents, stating, “I think they trust me to do what’s best for kids” (INT 12-10-19). Developing this trust with parents, colleagues, and administrators is essential in this type of grassroots change starting at the teacher level.

Regardless of the level at which this design is first introduced, it is likely to be met with at least some opposition at either the philosophical or practical level. After all, Loveless (2013) reminds us that ability grouping has been a “perennial theme” of education for over a hundred years (Steenbergen-Hu et al., 2016). There will be some who see the tremendous results of this study who will ultimately feel that school is meant to serve divergent rather than convergent goals (Deunk et al., 2018) and that we still should be providing inherently different instruction for different groups of students. There will be others who think the results are fantastic and even desire a more equitable approach to education who just can’t imagine how it would work in their own classroom even given the clearly explicated design principles. After all, reading about it and seeing it in action are two entirely different experiences, particularly when confronting more than twelve years of experience in an alternative model. Yet, there will be some who read about this design, believe in its potential, give it a try, see the results firsthand, and never go back. This firsthand evidence of effectiveness, combined with developing relationships of trust with all stakeholders, is perhaps the very best way to navigate and ultimately overcome the political dimension.

**Final Words**

Educational interventions are only as powerful as they are sustainable and feasible. No matter how effective the programming, if the technical, normative, and
political dimensions are not attended to, the design will likely come to an untimely end. Through the comprehensive, context-focused nature of design-based research, this study has attempted to provide a holistic and realistic account of the implementation process, thereby attending to all three dimensions. It started with the ideas of one teacher who was discontent with the feedback cycle witnessed year after year and resulted in a fully-fledged design that answers the call of decades of research with better than anticipated results. It’s a design drawing off the professionalism of one teacher to spark the professionalism of all teacher implementers.

When asked what others design elements she would like to try, it was clear Mrs. Poppy had already been thinking about the future of this work. She stated,

*You know when I did small groups it was more word work and it was more word sorts kinds of things in the small groups that I did, but I’m going to start right away with the guided reading. Now it’s just taking the time to make sure I’ve got that all ready this summer. But I think it will be huge and I think I’m guessing we’ll see even more growth. Don’t you think?* (INT 3-19-19)

She also mentioned there had been a “flip in her attitude.” At the beginning of the study, she told me that she felt she could meet students’ needs in whole group and when there were changes in schedule, guided reading instruction seemed to be the first thing to be cut from the schedule. In fact, during twelve of my visits, guided reading groups were either shortened or skipped due to a variety of reasons from watching a movie to catching up on other work. I asked Mrs. Poppy to elaborate on what she meant by this attitude flip. She said,
I think your study...the amount of time actually spent with kids in text, and I’m not talking in their just right basket, I’m talking teacher instruction in text needs to be a bigger part of what I was doing. Because I was doing...I guess I read so much [research] that I was picking pieces here and there and not doing enough of anything. (INT 3-19-19)

She proceeded to talk about ways in which she might change up her scheduling to cut back on whole class reading rather than guided reading on days when they had extra activities. It was evident throughout this conversation that not only did she see the power of the design, but that she also saw the importance of her role as the teacher in providing access to literacy opportunities for students. Although Mrs. Poppy’s commitment to heterogeneous grouping began as a research study, she will continue this work as a commitment to her students. At the end of our time together she reflected on the importance of this work.

One little boy who didn’t recognize the letters in his name at the beginning of the year is now reading beyond what we would expect at the end of kindergarten.

For me, that just says it all.

(INT 3-19-19)
REFERENCES


APPENDICES
APPENDIX A:
UNL INSTITUTIONAL REVIEW BOARD APPROVAL LETTER

Official Approval Letter for IRB project #18572 - New Project Form
November 9, 2018

Tiffany Young
Teaching, Learning and Teacher Education
1644 Sawyer Street

Guy Trainin
Teaching, Learning and Teacher Education
HENZ 11B, UNL, 665080355

IRB Number: 20181118572EX
Project ID: 18572
Project Title: Instructional Grouping for Early Literacy Instruction

Dear Tiffany:

This letter is to officially notify you of the certification of exemption of your project for the Protection of Human Subjects. Your proposal is in compliance with this institution's Federal Wide Assurance 00002258 and the DHHS Regulations for the Protection of Human Subjects (45 CFR 46) and has been classified as exempt. Exempt categories are listed within HRPP Policy #4.001: Exempt Research available at: http://research.unl.edu/researchcompliance/policies-procedures/.

- Date of Final Exemption: 11/9/2018
- Review conducted using exempt category 1 at 45 CFR 46.101
- Funding (Grant congruency, OSP Project/Form ID and Funding Sponsor Award Number, if applicable): N/A

We wish to remind you that the principal investigator is responsible for reporting to this Board any of the following events within 48 hours of the event:
- Any serious event (including on-site and off-site adverse events, injuries, side effects, deaths, or other problems) which in the opinion of the local investigator was unanticipated, involved risk to subjects or others, and was possibly related to the research procedures;
- Any serious accidental or unintentional change to the IRB-approved protocol that involves risk or has the potential to recur;
- Any publication in the literature, safety monitoring report, interim result or other finding that indicates an unexpected change to the risk/benefit ratio of the research;
- Any breach in confidentiality or compromise in data privacy related to the subject or others; or
- Any complaint of a subject that indicates an unanticipated risk or that cannot be resolved by the research staff.

This project should be conducted in full accordance with all applicable sections of the IRB Guidelines and you should notify the IRB immediately of any proposed changes that may affect the exempt status of your research project. You should report any unanticipated problems involving risks to the participants or others to the Board.

If you have any questions, please contact the IRB office at 402-472-6965.

Sincerely,

[Signature]

Jenn Klein
for the IRB
APPENDIX B:
ORAL RECRUITMENT FORM FOR ELEMENTARY TEACHERS
Title: Instructional Grouping for Early Literacy Development
Recruitment Script
IRB#: 18572

Hello, Mr./Ms. _________?
My name is Tiffany Young. I am a doctoral student in the Teacher’s College at the University of Nebraska-Lincoln. I have been a kindergarten and first grade teacher for the past nine years.

I am working on a research study regarding the instructional grouping practices for literacy instruction in early elementary school. May I explain a bit more about the project to see if your school might be interested in participating in this study?

As you know, elementary school students are often placed into groups with peers of similar reading ability for guided reading instruction. Through this practice, students are differentially exposed to reading skills, strategies, and texts that are presumed to match their current level of ability. Research suggests this practice is problematic given that (1) current notions of matching early readers to the appropriate levels of text difficulty for reading instruction are based on tradition rather than empirical evidence (Brown, 2009), (2) poor, minority students are overrepresented in the lowest ranked groups (Braddock & Slavin, 1992; Condon, 2008) (3) net of prior performance students in higher ranked groups make greater academic gains than lower ranked groups (Tach & Farkas, 2006) (4) teacher perceptions of students ability are often inaccurate (Ready & Wright, 2011). Conversely, several studies have shown that when students are presented with texts of increased difficulty and given adequate instructional support, they are able to make accelerated reading progress (e.g. Kuhn, Schwanenflugel, Morris, Morrow, Woo, Meisinger, Sevcik, Bradley, and Stahl, 2006; Stahl & Heubach, 2005, Reis, McCoach, Little, Muller, & Kaniskan, 2011).

The purpose of my study is to design innovative classroom practices with primary educators that lead to the development of theoretical insights on alternative grouping strategies and the necessary instructional supports to increase reading achievement of all students.

Beginning the fourth week of the 2018-2019 school year, I would observe literacy instruction in your classroom. After initial observations, I would meet with your team weekly to collaboratively design heterogeneous grouping practices for guided reading instruction and other instructional supports to meet the needs of students of all ability levels in your classroom. Observations would continue throughout the school year and together we would change the instructional design as needed to best meet the needs of all students in the classroom. You would also participate in three interviews focusing on your beliefs about literacy instruction and grouping structures. Your students’ literacy
development will be observed and their progress on the Developmental Reading Assessment will be collected each quarter.

Are you interested in participating in this study? If you have any questions, you can read me either by phone or email. tiffanytairayoung@gmail.com 402-429-0964
# APPENDIX C: INTERVIEW PROTOCOLS

<table>
<thead>
<tr>
<th>Interview Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interview 1:</strong> Initial Interview</td>
</tr>
<tr>
<td>1. Tell me about yourself as teacher.</td>
</tr>
<tr>
<td>2. Think about the ways in which you, your building, and/or your district groups students for literacy instruction. How are these groups determined? What student characteristics are considered?</td>
</tr>
<tr>
<td>3. What are your thoughts on grouping students by ability for literacy in elementary school?</td>
</tr>
<tr>
<td>4. Tell me about the nature of “ability.” What does this word mean to you as a teacher?</td>
</tr>
<tr>
<td><strong>Interview 2:</strong> Post Interview</td>
</tr>
<tr>
<td>1. How has this quarter reaffirmed/complicated/problematized your beliefs about grouping students by ability?</td>
</tr>
<tr>
<td>2. Tell me about the progress your students made in literacy development this quarter. What surprised you? What were some frustrations?</td>
</tr>
<tr>
<td>3. Tell me about the nature of “ability.” What does this word mean to you as a teacher?</td>
</tr>
<tr>
<td>4. Beyond the instructional practices designed during this study, what further considerations are needed when thinking about heterogeneous grouping students for literacy instruction?</td>
</tr>
</tbody>
</table>
APPENDIX D:
APPROVED TEACHER INFORMED CONSENT FORM

Title: Instructional Grouping for Early Literacy Development

Purpose:
The purpose of this study is to design innovative classroom practices with a primary educator that leads to the development of theoretical insights on alternative grouping strategies for guided reading instruction and the necessary instructional supports to increase reading achievement of all students. You are invited to participate in this study because you are a primary level educator involved in the literacy instruction of students.

Procedures:
Beginning in November 2018 through May 2019, I will observe literacy instruction in your classroom. During the observations I will take notes about instructional grouping practices, literacy instruction, and students’ literacy development. From November through December, observations will occur once per week for up to 1.5 hours. From January through May observations will occur two times per week for up to three hours. You will also be asked to participate in three audio-recorded interviews regarding your beliefs about grouping practices and the nature of ability, which will last no longer than 45 minutes each. In addition, from January through May you will be asked to participate in weekly 20-30 minute collaborative planning sessions during which we will focus on grouping students for instruction and creating instructional supports to meet their literacy needs.

Benefits:
As a participant in this study, you will have access to weekly collaborative sessions with the lead researcher to plan literacy instruction as well as instructional supports created by the researcher. Results will benefit society by adding to the existing literature on early literacy instructional practices. It may be particularly beneficial to students who have been historically underserved as it seeks to provide an alternative to within-class ability grouping, which disproportionately disadvantages poor and minority students.

Risks and/or Discomforts:
There are no known risks or discomforts associated with this research.

Confidentiality:
Any information obtained during this study that could identify you will be kept strictly confidential. The data will be stored in the investigator’s office and on a password-protected computer within a password-protected folder and will only be seen by the investigators during the study and for up to ten years after the study is complete. The information obtained in this study may be published in scientific journals or presented at scientific meetings, but the findings will either be reported using pseudonyms for all names and locations or in aggregate form.

Opportunity to Ask Questions:
You may ask any questions concerning this research and have those questions answered before agreeing to participate in or during the study. You may contact the investigator(s) at the phone numbers below. Please contact the University of Nebraska-Lincoln Institutional Review Board at (402) 472-6965 to voice concerns about the research or if you have any questions about your rights as a research participant.

118 Henniker Hall / P.O. Box 880355 / Lincoln, NE 68588-0355 / (402) 472-2231 / FAX (402) 472-2837
Freedom to Withdraw:
Participation in this study is voluntary. You can decide not to participate or withdraw at any time without harming your relationship with the researchers or the University of Nebraska-Lincoln, or in any other way receive a penalty or loss of benefits to which you are otherwise entitled.

Consent, Right to Receive a Copy: You are voluntarily making a decision whether or not to participate in this research study. Your signature certifies that you have decided to participate having read and understood the information presented. You will be given a copy of this consent form to keep.

The University of Nebraska-Lincoln wants to know about your research experience. This 14 question, multiple-choice survey is anonymous. This survey should be completed after your participation in this research. Please complete this optional online survey at: http://bit.ly/UNLresearchfeedback.

Signature of Participant:

________________________________________  __________________________
Signature of Research Participant          Date

Name and Phone number of Investigator(s)
Tiffany Young, M.A., Principal Investigator (402) 429-0964
Guy Trainin, Ph.D., Investigator (402) 472-2231
APPENDIX E:
APPROVED PARENT/GUARDIAN INFORMED CONSENT FORM

Hello parents/guardians.

By signing this form, you agree to have your student participate in the study, “Instructional Grouping for Early Literacy Instruction” conducted by Tiffany Young, a doctoral student in the College of Education and Human Sciences at the University of Nebraska-Lincoln. She is exploring how guided reading instruction can best meet the needs of all students in close partnership with your child’s classroom teacher.

By agreeing to allow your child to participate in this study, you understand that Ms. Young will observe the literacy instruction in your child’s classroom and may take notes on your child’s literacy development as it is influenced by instruction throughout his or her kindergarten school year. Ms. Young may also collect samples of your child’s literacy work, such as papers, drawings, and tests given by the classroom teacher. Your child will not be asked to participate in any additional activities not already done in the classroom for the purpose of this study.

Ms. Young will analyze the data collected in the classroom observations and work samples. She will also write and speak about this research to scholarly audiences. She will always use a pseudonym for your child’s name both in the data she collects and when writing or speaking about the research. She will store all data in secure place in her office and on a password-protected computer within a password-protected folder. Only Ms. Young and her advisor, Dr. Guy Trainin, will have access to the data. Reporting of the data will be used for research or teaching purposes and may include unidentified examples of your student’s work or notes taken during observations.

Participation in this study is voluntary. You can refuse to allow your child to participate or withdraw your child’s participation at any time without adversely affecting your relationship with Ms. Young, the university or your child’s school. Your decision will not result in any loss of benefits to which you are otherwise entitled. If you have any questions about your rights as research participants that have not been answered by Ms. Young, you may contact the University of Nebraska-Lincoln Institutional Review Board at 402-472-6965.

Permission can be granted on the back of this form. Thank you for your consideration.

Tiffany Young, M.A., Principal Investigator (402) 429-0964
Guy Trainin, Ph.D., Investigator (402) 472-2231

118 Henzlik Hall / P.O. Box 880355 / Lincoln, NE 68588-0355 / (402) 472-2231 / FAX (402) 472-2837
Parents/Guardians, please note your consent by checking (✔) the following:

1. Yes ___ We agree to have our student observed during literacy instruction.
2. Yes ___ We agree to have our student’s work samples collected as data.
3. Yes ___ We agree to have our students’ literacy assessments collected as data.

Printed name: ________________________________

Parent signature: ________________________________

Name of student: ________________________________

Date: ________________________________
APPENDIX F:
APPROVED VERBAL ASSENT FOR STUDENT PARTICIPANTS

The lead researcher will read the following statement to each student participant prior to data collection:

My name is Mrs. Young. I will be in your school to learn about teaching kids how to read. I was wondering if it would be okay with you if I watched you sometimes during reading time this year to see how much you grow as a reader. I’d also like to look at some of the notes your teacher writes when she listens to you read (DRA assessment form – students do not always know that this is a reading test). Would that okay with you?