

WITHIN-CLASS GROUPING DURING LITERACY INSTRUCTION:
A LOOK AT EQUITY

by

Jessica St.Louis

A Project Submitted to the Faculty of

The Evergreen State College

In Partial Fulfillment of the Requirements

for the degree

Master in Teaching

2010

This Project for the Master in Teaching Degree

by

Jessica St.Louis

has been approved for

The Evergreen State College

by

Jon Davies, Ed.D., Member of the Faculty

June 2010

ACKNOWLEDGEMENTS

Completing this paper has been a significant part of a very long journey. I wish to thank everyone who has helped me along the way. Jon Davies, Ed.D., Member of Faculty at Evergreen, has provided me invaluable help and advice along the way. His comments and advice concerning student growth have greatly impacted my thinking. Quotes from our conversations will remain in my thoughts, and writings for years to come. I'd also like to thank Sonja Wiedenhaupt, Ph.D., Member of Evergreen Faculty, for the many ways in which she pushed me, and stretched my mind. To Grace Huerta, Ph.D., Member of Evergreen Faculty, for her multi-cultural awakening, and encouragement for us to be "de-centered." To Terry Ford, Ph.D., Member of Faculty at Evergreen, for her detailed and insightful comments, striving toward excellence at teaching and reaching my students. I'd also like to extend a warm hug and appreciation to my cohort, class of 2010, who stood beside me in this journey, you have demonstrated to me the ideal professional community of love and respect for each other, a level of insightful communication I will always strive to achieve with both my students, and my colleagues. And most of all, my family, for their patience and support as I invested and sacrificed to complete this program to its fullest extent. To all of my attending faculty, my cohort members, and my family, your continued commitment to my goals have helped me achieve my ambitions in teaching and learning, thank you.

ABSTRACT

In this literature review I will be asking what grouping strategies and methods are most effective at creating equity in learning literacy across ability levels. The use of homogenous groups in education gained popularity in the nineteenth century. More recently teachers and administration have begun turning toward heterogeneous grouping methods. The research on this subject will be divided into five sections; group sizes, small group dynamics, the effect of grouping on students' needs, and teacher's practices. This review will be keeping an eye toward producing equity in learning, not equality, when considering results on grouping methods. Findings showed that small-groups create more favorable learning experiences, and that benefits exist in a heterogeneous group over a homogenous group.

TABLE OF CONTENTS

TITLE PAGE.....	i
APPROVAL PAGE	ii
ACKNOWLEDGEMENTS.....	iii
ABSTRACT.....	iv
CHAPTER 1: INTRODUCTION	1
Introduction	1
Rationale.....	1
Definitions.....	7
Summary	10
CHAPTER 2: HISTORICAL BACKGROUND.....	12
Introduction	12
Historical Background.....	12
Grouping Methods Reviewed	15
Summary	23
CHAPTER 3: A CRITICAL REVIEW OF THE LITERATURE	24
Introduction	24
Group Sizes	24
Small Group Dynamics	44
The Effect of Grouping on Student Needs	64
Teacher’s Practices: Instruction and Organization	96
Student Perceptions	124
CHAPTER 4	141
Introduction	141
Summary of the Findings	142
Implications for Teaching	148
Suggestions for Further Research	150
Conclusion.....	153
REFERENCES	157

CHAPTER 1: INTRODUCTION

Introduction

This literature review will be looking at the use of ability grouping in the classroom to teach literacy. The research and reviews include a focus on the implications of various grouping methods to create equity, rather than equality. More specifically, the research I will be asking what grouping strategies and methods are most effective at creating equity in learning literacy across ability levels. To do this I will be reviewing and discussing the research across five themes.

Each theme is reviewed in a section included in chapter three. The first section will be looking at the differences in implementing varying group sizes, including the use of small groups or whole class. In the second section small group dynamics, and how they affect student learning, will be reviewed. The third section includes students' needs, and the use of scaffolding and differentiation in lesson plans. The fourth section will contain what teachers perceptions and practices in the classroom, and the final, fifth section will cover what students think.

Rationale

The importance of reading in our society cannot be disputed. It is required as part of higher level thinking skills for almost, if not every, line of work. While the act of reading comes almost second nature to those of us who know how, getting there is quite complicated. Though we may remember reading books as a child, or participating in a school lesson involving text, in general we don't remember learning how to read, or recall a day where we suddenly became a reader. According to The National Council of Teachers of English (NCTE), this never happens, "there is no fixed point at which we suddenly become readers" (1999). Learning to read, or

becoming literate is a gradual process that happens over the course of many years, and never really ends. Even as an adult we encounter new words, or an unfamiliar voice, perhaps in places like legal documents or research analysis, and as we decipher it, gaining meaning from the text, we increase our reading ability, furthering our development as a literate individual. This process begins when we are very young, according to the NCTE, starting with our very first “interactions with print” (1999, ¶ 2).

Reading, the act of reading, in definition is hard to pinpoint. The NCTE (2004) defines reading as:

A complex, purposeful, social and cognitive process in which readers simultaneously use their knowledge of spoken and written language, their knowledge of the topic of the text, and their knowledge of the culture to construct meaning. Reading is not a technical skill acquired one and for all in the primary grades, but rather a developmental process (¶ 2).

Frank Smith (1985) found that the definition of reading was so vast, that it was not worth making (p. 93). Any implied definition of reading, Smith says, would only be an argument of semantics (1985, p. 93). However, the NCTE and Smith agree that text alone is not enough to create meaning. The NCTE says reading comprehension and interpretation happens when we bring together our understanding of the spoken language, our own prior knowledge, experiences with the world, and the extensive skills and strategies we develop over time (1999). Smith feels text in and of itself does not have an inherent meaning, but rather, we, as readers, bring meaning to it (1985, p. 48). Because there is so much more that goes into reading then can be summed up in a simple definition, the methods of instruction

used in schools must take into account all that readers draw upon to gain the fullest understanding of text.

Literacy is the foundation which a student's academia stands on. Their ability to read becomes the true "cornerstone of a student's success," marking not only a score in reading and writing, but "impacting a person's total school experience more than any other ability" (Harris, Harrison, 1988, p. 3). As students progress through school, and advance on to college, or other professional tracks, their level of reading ability is often connected with their level of success or advancement. Students are tested and graded on their reading ability right up into college. In fact, a student whose reading ability isn't high enough won't be allowed many privileges of society, including admission into universities, trade schools, the military, and a wide range of jobs.

The importance of reading is acknowledged beyond the academic and professional level. The NCTE's position is that reading is necessary in order to "better understand ourselves, others, and the world around us" (1999, ¶ 1). Under this belief reading could make us better citizens, more caring individuals, and have an increased sense of responsibility for the world at large. Along with the ability to think critically, this understanding is one of the most important aspects of growing into adults in our modern world. Literacy is also intricately tied up with entertainment, communication, and self esteem (Harris, Harrison, 1988, p. 3). An adult or older child who can't read may even suffer social consequences, being looked down upon by peers. Literacy is of high importance in our society, with expectations continually growing higher.

One of the biggest reasons our demands for literate individuals continues to grow is due to the increasing levels of technological advances, including computer literacy. Harris and Harrison (1988) back this saying, "higher and higher levels of literacy will be required in the

future” (p. 3). As higher demands are laid on our adults and society to reach higher literacy ability levels, these expectations trickle down to our children and students. While we have always taught literacy as a beginning, or even primary, skill in schools, the demands placed on children to become literate at a very young age are ever increasing alongside the demands of society. For students who struggle with reading, this demands may be even more so threatening on their futures. Vaughn, Linan-Thompson, Kouzekanani, Bryant, Dickson, and Blozis, S., (2003) quote Juel, (1988) and Morris, Shaw, & Perney, (1990) saying;

As the literacy demands of our society continue to increase, the gap between successful and struggling readers broadens. Because students who do not learn to read in the first or second grades are likely to struggle with reading throughout their lives, effective reading interventions for students early in their educational careers are critical. (p. 301)

As such, we must take our children’s ability to read very seriously, beginning every students’ career in school with literacy skills that allow them to excel to the best of their own ability. We cannot underestimate the importance by which reading happens, and as an extension of that, how it is taught in schools. As an elementary school teacher, a parent, and a contributing member of society this issue is of great importance to me. I want to be the best teacher I can, making sure that I do all I can to help my students grow into successful adults. As such I must help them become the best readers they can. This issue is of equal importance to teachers and school administration everywhere.

Currently literacy education is one of the primary focuses upon bringing students into public schools. The first years of school, up to third grade, are commonly called the “learning to read years,” followed by “reading to learn” in higher grades (L. Karamatic, Reading Specialist,

Horizons Elementary, personal communication, 2009). As educators, we see a wide range of ability levels in students at every grade level, developing from their personal breadth of prior knowledge and learning experiences in the content area of literacy. In kindergarten students may come in experienced at reading simple word books, while other students may not even recognize the letters, or know where to start on a page of text. As students progress through their development in reading, teachers struggle with the best methods of teaching students. One of these considerations is the use and method of grouping in the classroom.

Grouping is a convoluted subject, with multiple variables and applications. Grouping is a common instructional strategy used in classrooms to teach literacy. Typically grouping involves breaking students into groups of 3 to 5 students, with a common assignment. The dynamics of what happens in the group, the outcomes of the group, and why students are placed together in a group are all disputed and diverse.

Used as an instructional tool, and not a method to lead into tracking, grouping strategies in the classroom can take on many faces. As an ability-based method, it can be heterogeneous, where mixed ability students are grouped together, or homogenous where same-ability students are placed together. Another prime consideration to take into account is size, and whether to use small groups at all, or stay with whole class lessons. Teachers may use whole-class instruction, small groups of three or more students, or pairs. Teachers must also consider the dynamics within a group, and a student's perceptions. Of all of these, ability grouping is discussed the most. Should the more capable students be placed with other more capable students, and less capable with less capable in homogenous groups? Or, should groups be designed to cross ability levels, so more capable readers are reading alongside less capable peers?

This discussion of grouping methods is an important one. Research in teaching and learning theories create support and evidence for varying aspects of grouping. For example, to achieve real learning students need to be challenged, without being over challenged. Zull (2002) explains this through a biological process. He explains that all learning turns into prior knowledge, and gets biologically stored in neuronal pathways in the brain (Zull, 2002, p. 98). As such, when new material is introduced it must be built upon existing neuronal pathways, or the existing knowledge of the student (Zull, 2002, p. 102). When a student gets introduced to material that is completely foreign to them, and they have no neuronal pathways to build upon, they may encounter a “brink wall,” or a lack of a pathway by which they can gain understanding of the material (Zull, 2002, p. 101). So the introduction of a concept or a set of skills, that a student is not yet ready for, will not be understood. Scaffolding and differentiation need to happen so all students receive their optimal reading and learning needs. From high achieving students, to the learning disabled. The use of the Zone of Proximal Development (ZPD), developed by Vygotsky, is also worth considering when looking at the use of group work. Wells (2000) defines Vygotsky’s ZPD as the edge of knowing where a student, or individual, is “able to achieve more with assistance than he or she can manage alone” (p. 57). This zone is an important aspect of learning, and available with the application of group work in a classroom.

When looking at the social implications of ability grouping there is important implications in the research also. Self-esteem and feeling capable as a learner is an important aspect of learning, including developing as a reader. Our brain, containing our knowledge, our memory, the real ‘heart’ of where learning takes place, is also at the heart of our emotions. Therefore, the emotions of the student must be recognized, and taken into account when

teaching. James Zull (2002) studied the biological origins of learning in the brain, and found that emotions are a key aspect, “we could argue that the entire brain is an organ of emotion, and that emotion, reason, and memory are all linked together” (p. 65). It is as impossible to separate emotion from learning as it is to separate the trees from the forest. One without the other would result in the death of both.

Emotion occurs primarily in the amygdala, a section of the brain wrapped inside the cortex, where thought primarily occurs. Between them is a complex array of connections that communicate back and forth. Thought communicates to emotion, and emotion communicates back to thought. Within these signals there are outgoing and incoming pathways. The outgoing connections of the amygdala travel mostly to the areas of the brain “involved in the memory part of reflection (temporal cortex) and the creative and judgment part of abstraction (frontal and anterior cingulated)” (Zull, 2002, p. 74). When looking at the direction of these pathways there are more outgoing passages traveling from the amygdala to the cortex, then visa versa, creating a biological setup where our “emotions influence our thinking more than our thinking influences our emotion (Zull, 2002, p. 74).” Because of this biological set up, emotions cannot be separated from learning, and are another important learning element. As the student continues to encounter new experiences through internal and external learning elements, the emotions are first to respond. This means that a student who is made to feel depressed from their placement in a group, is also biologically depressed from learning

Definitions

Terms used for group sizes include; *whole class*, where the entire class is taught the same lesson, followed by individual seat work where all students have the same expectations, *small*

groups, where students are broken into groups of 3 or more students for lessons, and *pairs*, where students work in groups of 2.

Ability grouping, or ability-based groups, is synonymous with homogeneous groups. In its most broadest sense refers to the process of placing students in a lesson or cooperative group project based off their pre-assessed, or pre-perceived ability in the concepts and content being taught. Huerta defines it as “the practice of sorting elementary or secondary students on the basis of their skills into specific groups within a class” (2009, p. 90). McCoach et. al. (2006) defined ability grouping as “the process of teaching students in groups that are stratified by achievement, skill, or ability levels” (p. 339). Ability grouping can happen either inside a single class, across classes of the same grade, or even across classes of multiple grade levels. Teachers may use it to group students with similarly assessed students to create same-level groups, or they may group high-achieving students with low-achieving to create cross-ability groups. In a classroom where homogenous groups are being implemented a teacher will typically have 3 (or more) groups in the class; a low, middle, and high achieving group, with student placement based primarily off ability level. Homogeneous grouping is commonly referred to as ability grouping. For purposes of this review I am only looking at within-class ability grouping, as it is used for instructional purposes, not across grade, or across classrooms.

Heterogeneous grouping refers to small groups where students of different ability levels are placed together into one group. This is also known as *cross-ability grouping*. This contrasts with *homogeneous grouping*, or *same-ability* grouping. A further discussion of this is provided in chapter 2.

Tracking is a form of ability grouping involving grouping students, based off previous assessments, into generally permanent courses or “tracks” of classes. These classes are rated as high or low achieving, and students are placed in a set of courses based off difficulty and specific academic curriculum. Huerta (2009) defines tracking as sorting “secondary students into specific academic curricula, such as vocational, college preparatory, and general education programs” (p. 90). McCoach et. al. (2006) refers to Renzulli & Reis (1991) when defining tracking as “the general and usually permanent assignment of students to classes taught at a certain level” (p. 339), and also acknowledges that tracking generally happens at the secondary levels. Though the beginnings of tracking may take place in grade school, particularly upper grade school, where students’ ability levels are determined, not tracking, nor the effects of tracking, will be looked at in this paper. My focus here will lie solely on with-in class grouping strategies, as they apply to instructional strategies in the classroom. Any research containing across-class grouping studies, or grouping leading into tracking has been excluded.

Literacy instruction generally refers to a combination of reading, writing, and oral language surrounding reading and writing. Writing involves planning, drafting, editing, and publishing creative or nonfiction works. Reading typically includes gaining meaning from, summarizing, and orally repeating written text. Paterson et. al. (2003) defined literacy as “reading, writing, and oral language” (p. 180). I will be looking at literacy as a whole, while reading can happen independently of writing, they are learned simultaneously, and studies focusing on either or both of these activities have been utilized. While I widely acknowledge that learning to read and literacy as a whole never ends in a person’s life, I have also focused my research on the first years of primary school where students are learning to read more heavily.

During these first few years teachers and students spend a larger portion of their time on reading activities.

Equity in literacy education refers to each student gaining what they need to create individual growth in their own learning, while *equality* refers to providing the same amount of resources, time, and level of lesson difficulty for all students, regardless of individual needs for growth. In this review I will be keeping an eye toward equity, and not equality, with a goal of every student achieving to their highest potential. To account for this aspect I have sought varying studies that looked at high achieving, low achieving, or learning disabled students, and included them all for review. The combined results of the varying studies will be discussed in the third and fourth chapter. With my final conclusions and reviews of every study included, I will be keeping an eye toward all students' success in each grouping method.

Summary

There are many considerations for teachers to weigh out when designing a literacy lesson, and a lot of research to wade through to help sway those decisions. In this literature review I will be asking what grouping methods should be used in the classroom to create the highest level of learning for all students, creating a focus on equity and not equality? As such, I will be sorting through many of the possible decisions to be made in grouping for literacy lessons.

In the following chapter, historical background, I will cover the historical background of using various grouping methods to teach literacy. I will also review the contemporary controversies, and various proposed implications of, common grouping methods, including

whole class instruction, small-group instruction, and homogeneous versus heterogeneous grouping, as they relate to creating equity over equality.

Chapter three will include a critical review of the literature, and be divided into five sections. In the first section I will be looking at the differences in implementing various group sizes, including the use of small groups or whole class. In the second section I will be reviewing small group dynamics, and how they affect student learning. In the third section I will be looking at students' needs, including the use of scaffolding and differentiation in lesson plans. This includes differences in learning for all ability levels; learning disabled, high achieving, and low achieving. Through this section, and the course of this paper I will be keeping an eye toward producing equity in learning, not equality. Or rather, I will be examining grouping situations that favor all students, and not a select group. To finish the paper off I will end with a section on what teachers are doing, and a final, fifth section on what students think.

CHAPTER 2: HISTORICAL BACKGROUND

Introduction

This chapter is divided into three sections, the first of which is this introduction, the second includes the historical background, and the third covers the implications of grouping methods as they pertain to creating equity in learning for all students. These grouping methods include whole-class instruction, small group instruction, and heterogeneous versus homogeneous groups. Each of these methods is discussed in detail, providing detailed descriptions of their use, how they pertain to learning theories, and what their implications are in creating equity in the classroom.

Historical Background

Grouping, as an instruction strategy in the classroom, can be traced back to the beginnings of education. Schumm, Moody, and Vaughn (2000, p. 477) cited the work of Barr, & Dreeban (1991), and Kulik, & Kulik (1984) saying, the use of homogenous, within-in class ability grouping is recorded back as far as the turn of the century, and has been the most popular and common form of grouping until recently. Things have changed as additional grouping options have been offered to teachers and school administration, including across-grade, across-class, and heterogeneous considerations. A trend toward heterogeneous grouping in particular is taking hold (Schumm et al., 2000, p. 477).

Research on the use of grouping in the classroom began in the 1950's, with few studies existing at that time (Lou et al., 1996, p. 425). The brunt of the research, particularly through the 1950's and 60's covered the use of ability grouping in reading and mathematics. In the 1970's

and 80's research in grouping turned to the use of cooperative learning, defined by Abrami et al. (1995), and cited by Lou et al. (1996) as small group instruction utilizing both "positive interdependence and individual accountability" (p. 426). Positive interdependence involves each member contributing to the learning of others. This includes acknowledging each person's contribution to the group's work by the teacher, other group members, or other peers. Positive interdependence contrasts with individual accountability where each member of the group is held responsible for their own work or role in the group. In this arrangement each member would be congratulated or graded separately based off their own production, whether the end product was a group project or each member produced their own work. Both of these methods are used in classroom grouping arrangements, though the research comparing and contrasting them is only beginning. What has been researched is the more common forms of grouping, particularly heterogeneous and homogeneous grouping strategies.

Homogeneous ability groups as an instructional tool in literacy lessons is a common activity. McCoach et. al. (2006) quotes Loveless (1998) by saying that heterogeneous "ability grouping for reading instruction appears nearly universal, especially in the early grades" (p. 339). Teachers commonly place near-leveled students into matching reading groups for literacy activities. One of the most common uses of homogeneous ability grouping is the implementation of the three-tiered, homogeneous ability groups. The primary reason this strategy is so popular is due to the required use of basal reading programs in school districts.

Reading programs, generally decided on and purchased by school districts or school administration, include teacher guidance books with lesson plans and sets of books or stories for the students to read and do lesson plans around. Each reading program comes with a prescribed

set of limitations based off their design. When looking at the use of grouping, these limitations may include limitations in variance of the reading materials, a prescribed set of groups, or a prescribed applications of groups. Because of the implementation of these reading programs by school administration the use of grouping strategies, whether ability based or not, may not even be up to the teacher.

Reading programs assigned by the school or district administration come with varying degrees of freedom in implementation in the classroom. In some cases teachers have little or no freedom of choice about either the materials the students read, nor how grouping for literacy instruction will be used in their own classroom. Moody and Vaughn (1997) quote a general education 3rd grade teacher; “it was pretty much mandated by the administration that we were going to change to HBJ [Harcourt Brace Jovanoich, Basal readers], and until that time we had three reading groups... . It was mandated that we were going to change and everyone was going to be in the same book.” In this case the teachers initially had to have three homogenous ability groups, until the district switched reading programs, seemingly pulling out ability matched reading all together. However, sometimes these forced decisions may be limited to our general education students. A special education teacher from the same district is quoted as saying she is allowed to use, “whatever works best for the children” (Moody and Vaughn, 1997).

Requirements from the administration can limit the use of ability grouping as much as require it. Poole (2008) quotes Chorzempa & Graham (2006) saying that “some teachers report they are forbidden to use ability grouping” (p. 228).

Using the strict 3-tiered ability groups as is common with basal reading programs, also creates its own set of proponents and critics. On the supporting side, proponents feel the three

groups offer a “reasonable size” of groups, that reduce “the range of reading abilities which teachers must plan for and deal with” (Lasswell, 1967, p. 1). On the opposing side beliefs include that only three levels of groups, “will not adequately meet the needs of all class members” (Lasswell, 1967, p. 1). It may also become difficult for students to move from one group to the next, when the jump is too big for them (Schooley, 1994, p. 6-7). Lasswell quotes Macdonald et al. (1966) who feel that homogeneous ability grouping “is an artificial competitive situation which deliberately invites the immature appraisal and judgment of the performance of selves and others in the peer group context” (p. 2).

When grouping strategies are chosen by the administration, and not the teachers in the classroom who are working directly with the students, they cannot be customized to the needs of the students at hand. This can create a further complication for building situations where students can excel in their own strengths, particularly when looking at issues of equity. One grouping method may suit a particular group of students, and place another at a disadvantage. Due to the implementation of grouping strategies, whether required by administration, or chosen by the teacher, it is important to understand how they affect the students being forced to work and excel within them. Through the rest of this chapter I will be introducing the fields of thought across varying grouping decisions, based off the available research. In the following chapter I will provide a more in depth review of specific studies on grouping.

Grouping Methods Reviewed

Whole class instruction

When teachers use whole class instruction, students are taught one lesson as an entire class. Generally speaking, “there is an emphasis on uniformity of instruction rather than on

diversity” (Lou et al., 1996, p. 424). The most common implementation of this method involves the teacher delivering instructions for the assignment to the class, followed by individual seat work. Emphasis during instruction is primarily based on “teacher explanations and encouragement” to promote student learning.

There are proposed benefits to this instruction method. One involves the use of time. Because there is one lesson for the entire class, less time is spent preparing and developing differentiated instruction and lesson plans for various ability levels. This can be useful to introduce and practice a single-set of objectives. Downsides include a set pacing level, leaving students that work faster with bored time on their hands, or students who work slower feeling rushed, left behind, or simply not being able to finish.

Whole-class instruction also favors equality in education, rather than equity. As Lou et al. (1996) describes whole-class instruction, “all students may be exposed to the same learning opportunities, emphasizing the open, democratic principles of the educational system and the realities of a life in a world that operates according to the survival of the fittest” (p. 425). If this idea of survival of the fittest, is appealing to you, then you have nothing to fear, but with an eye towards equity and not equality, it leaves many gaps. With whole class instruction, there is lack of differentiation in instruction, pacing, and individual student needs concerning instructions, questions, and class management. This means that, when used in literacy instruction, a student who is above or below the mean reading level of the class, may not be having their education needs met for their optimal level for learning growth.

As alluded to by Lou et al. (1996), whole class instruction also strengthens and favors an individual survivalist system, where those who have strengths and abilities suited for the task

make it, and others don't. This is in direct contrast to the social-cultural research done on the workings of education and society by Rogoff (2003). Rogoff exposes the actual interconnectedness of our society, arguing that it takes groups of people working together to get a job done. She uses the work of Ed Hutchins (1991) as an example of distributed cognition in society. In these studies professional work, often argued as being successful due to an individual's strengths or abilities, such as maneuvering a large ship or selling girl scout cookies, are shown to instead be dependent upon a group of people working together, distributing the cognition across them all to get the job done. In these cases, along with millions of other examples where people work together, it is group skills that get the job done, skills that are commonly introduced and improved upon during small-group lessons in school.

Small-Group Instruction

Small group instruction involves breaking students into separate pods containing at least three students. In the most broadest sense small group instruction could be implemented with any arrangement of students including across-grade, across-class, or within-class. Here I will only be looking at the use of within-class, small-group instruction, as across-class, or across-grade grouping commonly leads into tracking systems. Throughout this review small group instruction will be used only to refer to within-class groups of three or more students, where there are three or more groups in a class. Small group instruction is used interchangeably with intra-class grouping or within-class grouping.

The use of small-group instruction is applied so that different lessons can be used per group. This allows for more extensive scaffolding and differentiation for individual needs. Small

groups are commonly formed based off ability level, where students are either placed into same-ability groups, or cross-ability groups.

Three groups in a single classroom is a common arrangement in literacy instruction due to the popularity of basal reading programs with three tiered levels of reading materials. The implementation of class instruction around small group use varies. It may start with teachers providing a class-wide set of instructions, or alternatively teachers may give individual instructions to each group. Either case is followed by separate seatwork per group where the results may be individual or cooperative. Whether the end product is based off individual accountability or cooperative interdependence group work places the emphasis on peer to peer interaction and support, with helping generally encouraged (Lou et al., 1996, p. 425). This is in contrast to the direct assistance of the teacher as in whole-class arrangements, where students are expected to wait for teacher assistance when questions or difficulties arise.

One of the primary benefits and reasons for implementing small-group instruction is that instruction can be tailored to the ability level of individual students or small groups of students. With the use of homogeneous groups pacing can move at the speed of the group, and concepts or skill sets can be introduced to those who need, or are ready for them. With the small-groups' emphasis on peer-to-peer support the teacher can also be left free to assist students in need of greater portions of time, or attend to other duties. Small groups also introduce important social and group work skills that students would not otherwise learn, including debating and discussion of topics, cooperation, and peer tutoring and assistance. Small groups also encourage the application of Vygotsky's ZPD in a peer to peer interaction, particularly with the use of heterogeneous groups. Small groups also focus on the application of social skills important for

further development as an adult in the work force, where group work is needed to complete a job (Rogoff, 2003, p. 270-271). This emphasis on cooperative gains (Lou, 1996, p. 425), rather than competitive gains, is in direct opposition to equality, where the fittest survive the best.

Homogeneous vs. Heterogeneous Grouping

The use of small groups can be divided into two categories; homogeneous and heterogeneous. Homogeneous groups are designed with students of the same ability placed into the same group. Also known as same-ability groups, or sometimes, more simply just ability groups, they are contrasted with heterogeneous grouping strategies. Heterogeneous or mixed-ability grouping is the placement of students with varying content or concept ability levels in the same group. Heterogeneous groups are not designed to be random, but are carefully chosen to include students of high, middle, and low ability levels working together.

When placing students into groups teachers generally utilize a combination of methods including; a wide assortment of tests such as basal tests, administration chosen standardized tests, informal tests, along with personal judgment, and the previous years' placement (French, Robbins, Oliver, 1991, p. 7). Students may also be placed in low-ability groups for reasons such as "home literacy experiences, immaturity and inattentiveness" (Schooley, 1994, p. 6). These strategies used utilized to determine the student's ability level in comparison to other students in the class. Students are then sorted into a tier of three or more ability levels, and then into either homogenous or heterogeneous groups. Both arrangements have possible strengths and weaknesses.

There are many arguments suggesting that either heterogeneous or homogeneous grouping models are superior due to a combination of academic and social implications. One of

the biggest criticisms against homogenous ability groups deals directly with issues of equity. Many believe that homogeneous groups, particularly where the ability level of the students is known in the classroom, serve high ability students over low, learning disabled, and possibly even middle ability students, both academically and socially. Thus homogenous grouping may hold low-placed students back, while providing greater opportunity and instruction for the high-placed students, expanding the already developing reading gap between high and low reading students. The research and opinions supporting this, claim that homogeneous grouping creates an environment that is “elitist and destructive” to the classroom community (McCoach, O’Connell, Levitt, 2006, p. 339).

Part of this is the expected level of achievement by students. In this argument teachers may come to expect less and place reduced levels of standards on low-leveled students. These reduced standards may result in a lower quality of instruction time, less concepts or skills being covered, and lower quality of reading and instruction materials, thus denying the low ability students the much needed opportunities to learn and advance. Because students achieve to provided levels of expectations, particularly when those expectations are defined based off the materials and lessons provided by classroom teachers, this could be detrimental to the quality of learning of some students. McCoach, O’Connell, and Levitt (2006) quote Lou et al. (1996) saying that “opponents of ability grouping fear students in lower ability groups may be denied appropriate opportunities to learn and advance academically” (p. 339).

Implemented, this divided expectation level would further the gap between low and high ability students, pushing the high ability students even further up the academic achievement scale, and the low ability students “further and further behind” (McCoach, O’Connell, Levitt,

2006, p. 339) . When taking into account mandated grouping strategies and reading materials this decision may not even be up to the teacher, though the effects on the students would be the same. It may also have a lasting consequence, where a student who is placed in a low-ability group, even if done so wrongly, stays there for their academic career, an implementation of tracking that starts with small ability groups.

However, for each opinion, there is another to contrast it. In support of homogeneous grouping, another position argues that students should be exposed to reading materials and instruction that is at their level. This proposal believes that children who can read at a higher level than others, should be placed in groups where they will receive more of a challenge, while children who read at a lower comprehension level should read in groups where they don't feel left out or buried by material that is too difficult. Under this standpoint, using heterogeneous groups would leave low-ability students incapable of fully accessing materials that may be too difficult, or high-ability students with materials and lessons that aren't challenging enough, stunting the literacy development of one or both ability levels.

To contrast with that, some feel that higher ability students should be placed in heterogeneous groups where they can serve as peer role models for lower-ability students, and thus provide much needed scaffolding, as is demonstrated in Vygotsky's ZPD. In this scenario the lower ability children receive guidance from the higher ability, and the higher ability gain from teaching to the lower ability. The proponents of this strategy argue that by mixing ability levels, lower ability students learn more from the higher ability students, while the higher ability students tutor the lower level, creating learning for all. Schooley, (1994) discussed a study by

Cunningham, Hall, and Defee (1991) where this strategy was used, and “students in the low group learned much more and did not hinder the high group” when grouped together (p. 6).

When looking at social implications of ability based grouping, contradictions also occur. Some feel that a student placed in a low-ability group will self-identify as a low-reader, and thus never live up to their full potential. These students’ low self-perceptions as a reader may be self-made based off labeled identification, or taken from low standards by the teacher. The belief here is that homogeneous ability grouping is “detrimental to students’ self-concept and potential achievement in the low group” (Schooley, 1994, p. 5), and that students who get placed in this group often suffer from low self-esteem. This aspect of self-esteem and happiness is important to learning due to the complexity of the brain and intertwining of emotions and memory or thought. Because of this, it cannot be overlooked. These low standards may even be implied from the strategies and lessons used in the classroom or group work. Students placed in a low reading group may also feel a social sting from others, including students labeled as high-ability, teachers, and other adults. As a result of all these implications these students could suffer a drop in motivation.

Arguments are also made that friendships are fostered through the use of small groups in class. Some argue that student friendships in a same-ability group are made stronger, based off the commonality of a similar reading ability, creating longer connections. Others differ in opinion feeling that friendships are limited by homogeneous ability groups, and thus prevent students from making friends outside their perceived ability level.

Summary

With all of these varying options and conflicting opinions and research for each option, it is easy to see why teachers and educational professionals alike can get confused, and why further research and reviews are needed. As I review the following studies I hope to wade through the research out there, looking for grouping methods that provide the optimal learning environments for all of the students involved.

CHAPTER 3: A CRITICAL REVIEW OF THE LITERATURE

Introduction

This literature review examines the effectiveness of grouping methods to create equity in student learning. In chapter one the importance of being exposed to a literature rich environment was discussed, as an avenue to success in a society where the demand on literacy skills in constantly rising. Chapter one also discussed the methods and theories of learning, decidedly pointing out that neither emotions nor the student's ability level can be overlooked, or optimal learning will be blocked. Chapter 2 discussed the use and development of ability groups in the history of education as the most common and popular form of grouping, extending as back as far as the nineteenth century. It went on to show that more recently teachers and administration have begun turning toward heterogeneous grouping methods. The critiques and opinions of various grouping formats were also discussed, pointing to the amount of planning time required, and ability to individualize instruction for the needs of the students across whole-class, small-group, along with the implications of students learning with the use of heterogeneous versus homogeneous grouping strategies, when looking at issues of equity. Chapter three reviews the research on grouping. It is divided into five sections: Group sizes, Small group dynamics, Students needs, Teacher's Practices, and Student Perceptions.

Group Sizes

This section examines the effects of varying group sizes, including whole class instruction, the use of small groups, and teacher to student ratios with 1, 3, or 10 students per teacher. The implications of group sizes on learning disabled and general education students is also included. The size of the group is an important aspect of grouping, as the dynamics of

instruction and social roles in a group are changed greatly based off the size of the group. A group with only two members functions quite differently, then one with 3 or 10 members. When looking at teacher led groups, a difference in the number of students per group may influence the outcomes of the students involved. It is also important to examine the use of groups at all, comparing them to a whole-class arrangements where student learning and instruction happens differently. By asking should small-groups happen at all, and if so, at what size are students producing the best academic results, we can get closer to finding the optimal grouping placement for students.

Harris and Harrison (1988) conducted this study as a means to explore the use of a whole-class arrangement in place of homogeneous ability groups. This method is studied as an alternative to the traditional basal based, three-tiered reading program for first grade reading instruction. With the intent of gaining insight into the struggles teachers dealt with in a whole-class approach, the researchers asked 4 questions:

1. To what degree will teachers use whole class instruction as an instructional strategy if it is incorporated in a materials-based first grade reading program?
2. What are some of the concerns, problems and issues associated with using whole class teaching in first grade reading instruction?
3. What are the teachers' perceptions concerning use of whole class instruction as a classroom instructional strategy (e.g. ease of use, effectiveness, etc.)?
4. What are some of the effects of whole class instruction on the first grade students?

(p. 8)

Two first-grade classrooms were included in the study, based off their willingness to participate, both in a middle-class section of Orem, Utah. Both classrooms had been using a basal program as their primary reading instruction tool, and both had their classes divided into 3 ability-based groups. The basal program used for this study was Companion Reading Program, which incorporated peer tutoring as a crucial component. The use of this program served as a pilot implementation of the reading program. Using a naturalistic inquiry method, classrooms observations, interviews, and documents were collected as data. Interviewees included; teachers, principals, and the reading program author. Field notes were summarized and analyzed following the completion of the study.

Because Harris and Harrison (1988) recognize that teachers commonly modify basal reading programs, deeming them as inaccurate, they sought to discover to what degree teachers would follow the Companion Reading Program without modification. To this they found that teachers modified the program after about 5 months of use, making changes to the step-by-step instructions in the teacher's guide. One teacher reduced the number of concepts introduced to the children, saying the concepts were too difficult, and would be introduced to students in second grade. Both teachers continued to follow the scope and sequence of the program.

Harris and Harrison (1988) also sought to discover how frequently the teachers would be using whole-class instruction, finding that it was used every day for about 20 minutes during their 2 hour reading instruction time. Peer tutoring, individual seat work, silent reading, and mastery checking were also used. During whole-class instruction it was found that teachers had to deal with several issues and concerns. They had difficulty maintaining all students' attention, particularly when accounting for the varying levels of reading ability. Students would respond to

questions out of turn. Teachers had difficulty being aware of students who were less inclined to respond during group exercises. It was difficult to get the class to respond in unison when a group question was asked. Teachers also felt concern when over correcting shy or low-self-concept children in front of the class. Teachers reported that they found it difficult to accommodate individual children's differences, due to the large group size.

When looking at teacher's perceptions of whole-class instruction over three-tiered, homogeneous, ability-based groups, teachers were found to prefer whole-class. This was due primarily to the fact that they spent less time preparing individual lessons. They also found it to be an effective method to reach the desired outcome, against their initial, previous fear that students would be penalized based off reading ability.

When looking at how the whole-class instruction effected the students, teachers and principals alike felt the students learned more, though this may not be due to the grouping procedure, but rather the reading program. Students at all ability levels were reported as having a higher level of confidence, and teachers "really liked not having to place students into ability groups," and had concern that they "really had a negative effect on the low ability students" (Harris & Harrison, 1988, p. 15). Students also learned more concepts, were better readers, and made more progress, even though the previous year's students were more advanced. These perceptions were supported by the teachers and principal.

This qualitative study reported the experiences of two teachers as they employed a new basal reading program which included the use of whole-class instruction as a segment of the teaching, eliminating the requirement of homogeneous, ability based groups. This study adequately identified several issues these two teachers dealt with while enacting whole-class

instruction, as part of the reading program, all of which were directly related to classroom management. Because no quantitative data was compiled during the study, and no comparison was made, it is impossible to say whether these results were more or less frequent than the use of small-group lessons. The researcher's theoretical grounding, including a belief in the negative effects of ability grouping, and incompleteness of basal reading programs, were identified at the beginning of the study.

In this study, Harris & Harrison (1988) sought to gain insight into the struggles teachers deal with while using a whole-class approach (p. 8). Their results showed that the use of whole-class instruction, through a new reading program, was connected to exceptional student growth in reading. Due to the implication of this new reading program, which included several innovative teaching and learning methods, the differences in the students' learning outcomes cannot be concluded to have been a sole result of whole-class instruction. The new instructional elements including peer tutoring, home involvement, and consistent mastery checking across the class, may all have played a pivotal role in the higher learning outcomes of the students, as acknowledged by the teachers. The results cannot be transferred to other students, or classes, as this was a study involved around a specific reading program. Because student and teacher variables are not taken into account, and due to the pilot implementation of the Companion Reading Program, the results cannot be transferred to other classes. However, even with this aside, it cannot be disputed that the teachers saw positive social aspects to eliminating the use of the traditional three-tiered homogeneous ability based groups. When using the previous teaching methods including tiered homogeneous groups, and basal readers, teachers reported that they worried about the effects on low-ability students, to the extent that they preferred whole-class to

eliminate this effect. It is impossible to say if the same effects would have been seen in a heterogeneous grouping strategy.

In this meta-analysis Lou, Abrami, Spence, Poulsen, Chambers, and d'Apollonia (1996) explore the effects of small-group use in the classroom through three questions:

1. Does placing students into small groups facilitate learning and other outcomes, including attitudes and self-concept?
2. Which factors explain variability in study findings?
3. Which type of grouping is best and under what conditions? (p. 424)

Their search for studies resulted in 3,000 published articles. Those were narrowed through a selection process resulting in 66 studies. Selection criteria for inclusion included studies that took place in the classroom only, in an elementary, secondary, or postsecondary level, and involved within-class ability grouping, either homogenous or heterogeneous, that was taught by the student's regular classroom teacher. Groups had to have a minimum of 2 students, a maximum of ten, and had to take place for more than 1 day. Any training included as part of the studies had to be administered to all group members, and outcomes from both experimental and control groups had to be reported. Studies where achievement, attitudes, and self-concept were included in the meta-analysis, but each variable was analyzed separately. Any research where the focus was on students with a learning disability or enrichment programs for gifted students, was excluded, as were studies that focused on peer-tutoring groups. Reciprocal tutoring, where students were paired up and took turns tutoring each other were included.

Following the collection of the final 66 studies coding was performed independently by 2 coders, with an agreement rate of 88.24%. All disagreements were solved through discussion and

further review of the study. Two major sets of analysis were performed, effect sizes that compared and contrasted the use of grouping to a lack of grouping, and the effect of group ability on student achievement.

When looking at the results of grouping versus no grouping, analysis 1, the average result revealed that “students learning in small groups within classrooms achieved significantly more than students not learning in small groups” (Lou et al., 1996). However, the results were widespread, and further exploration was needed. When looking at outcome features the measure source was significant, with tests made by either the researcher or teacher leading to a significantly larger effect size than standardized tests. The type of measurement was also significant, with effect sizes being larger with measures that were specifically geared toward what was taught. There was no significant difference when comparing homogenous grouping to no grouping versus heterogeneous grouping to no grouping.

Nine methodological features were looked at, under the first analysis, with overall design quality not contributing to larger effect sizes. Studies where experimental control, or a statistical control were used, showed a significant relation to higher effect sizes in small-group classes. Overall instructional equivalence, was also significantly related to higher effect sizes. Instruction type was also analyzed, and achievement gains were seen during within class grouping for studies that provided “different instruction” to the experimental group, as was teacher training, more or different materials for student use, and the application of reward strategies (Lou et al., 1996, p. 440).

Substantive features that led to a significantly higher effect size for grouping versus non-grouping included group composition being based off mixed sources, rather than standardized

tests or teacher opinion alone. In addition, when the groups' composition was based off specific or general ability in addition to other factors, the effect size was higher. Groups that were sized at 3 to 4 members, also contributed to a significantly higher effect size than groups of 5 to 7 students. Group stability made no significant difference. Teacher training made a significant difference with a +0.57 for teachers who had received extensive training or experience. Types of small-group instruction was divided down into cooperative and non-cooperative, with cooperative learning and outcomes gaining significantly higher results, +0.28, than non-cooperative groups. Small groups also fared significantly higher when compared to control groups where traditional, frontal teaching methods, or mastery teaching was taking place. However, not when compared to experimental teaching and learning methods. The goal structure, competitive or individualistic, made no significant difference. The ability of students, high, low or medium, was also taken into account, with all three groups gaining from being placed in small groups. The results, however, were not uniform, with the low-ability students making significantly higher gains than the medium ability. Content also made a significant difference, with math and science gaining the largest effect size, when compared to reading, language arts, or other courses. The length of the study made no significant difference, while the frequency of treatment during the week did. When looking at the effects of small group work compared to no groups student attitudes were also affected. In grouped classes students had a significantly more positive attitude toward the content, with no difference toward the instructional approach. Student self-confidence was higher in small group classes at +0.16, while no significant difference was found in non-grouping classes.

In analysis 2, homogenous grouping versus heterogeneous grouping, the mean effect showed a “slight superiority” to homogenous grouping techniques. However, this finding was not consistent and required further analysis. Deeper analysis looked into the type of small-group instruction, cooperative or not, which made no significant difference. Ability level of the students, and content area being taught, both made significant differences in the effect size. When looking at ability level of the students, low-ability students did significantly better in heterogeneous groups, while medium-level students did significantly better in homogeneous groups. High-level students, however, showed no significant preference for either grouping method. Heterogeneous groups also made a significant difference in the effect size when applied in reading groups, while no significant difference was shown for math or science.

This quantitative meta-analysis effectively brought together 66 studies to look at the effects of within-class grouping on student achievement, self-concept, and student attitude. Their selection criteria remained consistent, with a low level of disagreement across coders. Their final selection of studies maintained criteria which reflected their study questions, keeping with small-group classrooms in the school. Their results showed that, “students placed in small groups achieved more, held more positive attitudes, and reported higher general self-concept than students in nongrouped classes” (Lou et al., 1996, p. 446). However, with this being said, the results also pointed to a lack of a grouping strategy that “*uniformly*” [italics authors] suited all students (Lou et al., 1996, p. 450).

In this meta-analysis small-group formats showed several positive benefits for students including; a higher academic growth, positive attitudes for students, and higher self-esteem, when compared to students in non-grouped classes. Because this meta-analysis reviewed studies

across multiple grouping methods, we have no indication as to how grouping methods may affect equity in the classroom, except to say that no single method suited everyone uniformly.

However, we can conclude that the use of small-group formats are more effective than whole-class.

Easton, Muirhead, Fredrick, and Vanderwicken (1979) conducted this quantitative study to examine student involvement in the classroom as correlated to whole class or small group (2 or more in the class) organization and teacher instruction. The study was conducted in 39 Chicago elementary schools, and had 3 objectives; (a) to determine “whether student attention is higher when the teacher directs instruction to the entire room than when the teacher directs instruction to a group of students,” (b) “to discover if there is a difference in student attention” when instructional groups are organized as whole class, or multiple groups, and (c) “to investigate the relationship between student involvement and various school level characteristics” (Easton et al., 1979).

Schools were selected in a systematic fashion, and are a representation of Chicago school schools in relation to size, poverty, and other variables, though the selection criteria or other variables is not discussed. Two classes were used in each school, one with student ages of 6 to 8, and one with student ages of 9 to 11, were used in 35 of the schools, while the remaining 4 schools only had one participating classroom. The classrooms were chosen randomly, and a total of 74 classrooms were observed.

Observations were conducted on random days, and the entire reading session was observed. During observations teacher and student behavior was coded and classified. Students were observed for 4 seconds each, and a decision was made if they were on task or not, based off

the four-second observation. Observers agreed on these decisions 75% of the time. This decision seems to be made at the discrepancy of the observer, as no data is provided to define what constitutes on task, or not. The number of students on task, and in the classroom were recorded every 5 minutes.

Classrooms were separated by teacher orientation into 3 categories; classrooms where the teacher shifted between the class and group, classrooms where the teacher was always orientated to the class or to a group, and classrooms where the teacher was sometimes orientated to the class and sometimes to a group. Student attention rates were compared within the same room, based off instruction orientation, and across all rooms. Poverty was found to play a role on student achievement in reading based off the Iowa Test of Basic Skills (ITBS), and was controlled for when calculating percentages of student's attention.

In regard to student attention in correlation with teacher instruction to the class orientation, the researchers found that a lower level of student attention was taking place by students in groups, rather than in whole-class instruction. However, when student attention was compared across whole-class instruction versus group instruction, there was very little difference within the same room.

The definition of a group is not discussed in this study, in relation to either size (pairs, teams, two large groups, several small groups, individuals), nor as an indication of whether the students were divided into group based activities, or if the teacher was simply addressing a group of students with instructions, while the remaining students work individually. The definition of whole class instruction is also not discussed, so it is not known whether during these sessions the teacher is giving directions to the whole class, and then they are working individually, or if they

are all orientated on the same task as a whole-class, for example listening to a book being read by the teacher. Because of the lack of definitions I find it hard to draw many inferences about group work effectiveness or student attention from these findings. However, I do respect the attention drawn to the need to be “aware of all students and not regard any activity as having secondary importance” (Easton et al., 1979).

Vaughn, Linan-Thompson, Kouzekanani, Bryant, Dickson, and Blozis (2003) sought to weigh out the effects of group size used in a intensive reading program offered in schools for struggling readers. Due to the high expenses of these programs an ideal balance, where group size is as large as possible without reducing students’ growth as a reader, is preferred. In this study Vaughn et al. (2003) compared different group sizes, where students worked one-on-one with a teacher in three different formats; 1:1, 1:3, and 1:10.

Ninety students were included in the study from ten Title I elementary schools in 2 neighboring urban school districts from the southwest. Seventy percent of the students in the schools qualified for free or reduced lunch. Participating students had to be identified as a struggling reader, and “at-risk for referral” to special education classes or “failure to read” (Vaughn et al., 2003, p. 303). Along with that students must have failed the state imposed, standardized second grade reading benchmark, Texas Primary Reading Inventory (TPRI), 1998. The parent and child must also have agreed to participate. Of the 90 students, fifty were monolingual English speakers (MES), and forty were English language learners. Due to students leaving the school, 77 students comprised the final sample set.

Reading instruction was treated the same across all groups, and included five elements for building reading skills: fluency building, development of phonological awareness, vocabulary

building, comprehension strategies, and word analysis. Comprehension included three strategies. The first of these was previewing and building connections to the text prior to reading. The second identifying the main idea, and third summarizing through the generation and answering of questions. Instructional awareness during group time also included phonological awareness, particularly phoneme blending and segmenting. Word study was comprised of the use of decoding and word patterns. Reading fluency was implemented through repeated reading, modeling and previewing.

Five female teachers with 22 hours of training taught all the groups. Two of the teachers were bilingual (Spanish\English), and 3 were monolingual. All had been teaching for a minimum of 1 year. The teachers were trained in all methods of instruction, and required to demonstrate the methods before beginning the sessions. During the intervention period they received feedback and support from observations and weekly meetings.

Students were grouped (low, medium, high) based off their ability in phoneme segmentation through the *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS) test, and assigned to group sizes of 1, 3, or 10. In the 1:1 groups there were 12 high ranking students, 8 medium, and 10 low. In the 1:3 student groups there were 12 high, 14 medium, and 4 low. In the 1:10 groups there were 8 high, 14 medium, and 8 low. Assignments were matched to students in accordance with school confounds. Groups met for 30 minutes, 5 days a week, for 13 weeks, making a total of 58 sessions.

A series of tests were given to the students to determine their reading proficiency on a variety of levels. Testing times included pretest, prior to the start of the group work, posttest, following all 58 sessions, and postintervention, following 4 weeks of schooling after the end of

the sessions. Tests were also administered during the treatment period as a monitor of student achievement. All tests were conducted by trained research assistants, blind to both the conditions and hypothesis of the study. The TPRI was used as both a pretest, and post test, and is used to identify students who need “additional interventions for reading” (Vaughn et al., 2003, p. 304). Two subtests, Word Attack, and Passage Comprehension, were used from the *Woodcock Reading Mastery Test-Revised*. These individually administered tests were performed as a pre- and posttest measurement, along with being a post-intervention measure at 4 to 6 weeks following the treatment. They are intended to be a test of the students’ basic reading skills. The *Test of Oral Reading Fluency*, also individually administered, was given at each of the three assessment points. The DIBELS phoneme segmentation test was administered as a pretest, a progress monitor, an outcome measure, and as a follow-up. The *Woodcock-Munoz Language Survey* was administered to ELL students both in English and in Spanish as an indicator of their oral language proficiency. Over the course of the 13 week intervention period each teacher was observed nine times, and an internal validity checklist was filled out. The checklist was used to record the materials, time and activities used with each strategy, along with their student monitoring strategies.

Results showed that all three groups showed statistically significant growth over the course of the study, with greater effects in the 1:1 & 1:3 group size compared to the 1:10 group size. These results were consistent across outcome measures including; passage comprehension, phoneme segmentation, reading fluency, and reading screening measures. The 1:1 group size was not found to be superior, based off the predetermined statistical significance of .05, when compared to the 1:3 group size in any of the previously mentioned outcome measures. Word

Attack was the only measure that yielded no statistically significant results at all. Minimal differences were found between ELL and MES student growth when compared across the same outcome. In phoneme segmentation and reading fluency outcomes a statistically significant difference was found when comparing the 1:1 group to the 1:10 group, but not from the 1:1 group to the 1:3 group, nor the 1:3 group to the 1:10 group, indicating a small growth difference from one group to the next. In reading proficiency more than twice as many students in the 1:1 and 1:3 groupings acquired proficiency at 40 words per minute, a critical benchmark, when compared to the 1:10 group during follow-up testing. When looking at the posttest results of the TPRI phenomenal results were also seen, with 39% of the 1:1 group students, 46% of the 1:3 group students, and only 7% of the 1:10 grouped students passing.

This quantitative study used a posttest/pretest design without a control group. The wide variety of tests implemented continued to show the same results across the varying grouping strategies, strengthening their reliability. Because no control group was employed it is hard to judge how much growth would have happened over the 13 weeks, had the students not been receiving the additional treatment. The lack of equal sample sizes for the three group types (1:1, 1:3, 1:10) could create uncontrolled variables. However, it can be gained from this study that the 1:1 and 1:3 group sizes achieved greater gains than the 1:10 group size. When using small-groups in the classroom, this could be an important part of the decision making process. A larger group, even with the assistance of a teacher, will not achieve the same high level of growth as the smaller groups.

This study was conducted as a supplement to the aforementioned study by Vaughn, Linan-Thompson, Kouzekanani, Bryant, Dickson, and Blozis (2003), looking at the effects of

group size in supplemental reading programs for struggling readers. Because Vaughn et al. (2003) saw a significant difference between the 1:10 group, and the 1:1 group, but not the 1:1 and 1:3 groups, here, Helf, Cooke, and Flowers (2009) sought to further investigate the difference between the 1:1 and 1:3 teacher to student ratio groups.

Helf, Cooke, Flowers (2009) sought to reduce uncontrolled variables that were not accounted for in the original study. By writing scripted lessons, and using the same tutors for both grouping conditions, the effects of teaching methods and experience on the student's outcome are minimized. The 9 tutors used in this study were selected by the principles of the participating 3 elementary schools. The schools taught K-5th in a large urban district from southeast United States.

Across the three schools, 54 randomly selected 1st grade students who were (a) placed in the instructional level labeled *strategic-additional intervention*, (b) had no identifiable disability, and (c) were in classrooms using Open Court (2000) primary reading program were selected. Student demographics included: 55% male, 45% female; 2% American Indian or multiracial, 4% Asian, 63% African American, 25% Hispanic, and 6% White. Five students transferred schools during the course of the study. Their place was filled with another student of a similar level in the supplemental program, to ensure integrity of the group size, but data on these 5 students was not collected.

The Early Reading Tutor (ERT) program is a supplemental reading intervention designed for use with students who are performing below expected standards. It is implemented in kindergarten and first grade with students who have been identified at the “supplemental level of a three-tiered reading model (i.e. core, supplemental, intensive levels of instruction) that the

Reading First initiative recommended” (Helf, Cooke, Flowers, 2009, p. 114). ERT follows the formats and instructional methods of *Direct Instruction Reading* by Carnine, Silbert, Kame’enui & Traver, (2004) (Helf, Cooke, Flowers, 2009, p. 114). This program was used because it has scripted text for all 110 lesson plans, and is believed to minimize instructor differences. Daily sessions included the implementation of these lessons, followed by recommendations by the tutor for the student to continue working on. Those recommendations of specific skills were completed as additional sessions immediately following the regular instructional session. In 1:3 groups students responded chorally to group questions, and a wrong answer (even if made by only one student) was followed through by a correction procedure for the entire group, including remediation on specific skills. All supplemental sessions were timed and recorded by tutors.

Students took the Dynamic Indicators of Basic Literacy Skills (DIBELS) test as a pretest and posttest. The DIBELS test included three subtests; Phoneme segmentation fluency (PSF), Nonsense word fluency (NWF), and Oral reading fluency (ORF). Student progress was monitored weekly on NWF & ORF, but not PSF because it is a key skill for Kindergarten. Each student had a progress-monitoring booklet. All tutoring groups used identical management strategies for student behavior, utilizing a star system, and acknowledging when students swayed from prescribed behavior (including group answering of questions) as outlined by ERT. Timing was recorded for each session, and allowed for individual duration at a student’s need. Each tutor was observed 20% of the time, with specific check-lists filled out designed to “measure the integrity of delivering ERT to a small group,” with additional measure to check for student opportunities to answer individually (Helf, Cooke, Flowers, 2009, p. 115).

Results showed all students scoring higher on the posttest, when compared to the pretest, regardless of grouping condition, as expected. Posttest results also showed that 59% of all the students were at a passing level, and on track for meeting end-of-year benchmark goals. Students of the 1:3 grouping condition completed 39 lessons, and students in the 1:1 groups completed 40. When comparing grouping conditions for efficiency of time it was found that the 1:3 condition was higher. The mean number of remedial lessons was not found to be statistically significant across the two conditions. When looking at student behavior no statistically significant differences were found across the two conditions. There was no significant difference in reading levels achieved in the individual tutoring (1:1) compared to the small group (1:3) tutoring. There was also no significant difference in the time spent on instruction, per group, the number of lessons covered, nor the amount of inappropriate behavior observed. This leads researchers to believe that small-group instruction is more efficient than one-on-one tutoring, because students are learning at equal growth rates, and less time is spent per-student.

As a follow up teachers were asked which method of grouping they preferred. Seven of the nine chose the 1:1 format, feeling they could monitor student's growth, and assess their needs and strengths more accurately. When asked which method produced the most effective instructional time six of the nine preferred the 1:3, noting it took less time to instruct three students than 1. One teacher had no preference, while the other two noted student pacing as an instructional difficulty. When asked about student performance five teachers preferred the 1:1, saying the students seemed to progress faster, and it was easier to track student responses. One teacher felt the addition of the other students in the 1:3 group benefited the group. The remaining teachers had no preference. Teachers were also asked about difficulties with the ERT program,

with complications only being expressed in the 1:3 group due to choral answering. When asked about student behavior, only two teachers had a preference, noting the small group format created distractions.

This quantitative study was designed as a true experimental, pretest\posttest group design, with random group assignment. Because this study was conducted with very specific, direct instruction, scripted teaching methods, using only phonics based literacy instruction the results would not transfer to more whole-language teaching methods, or whole language literacy learning. The use of cooperative group lessons may also create different results. Because there is no control group, student history and maturation, including reading growth made during regular class time and at home, could not be accounted for. There is also an accounted for discrepancy between the teacher's accounts of being able to better track and monitor student growth, needs and strengths with the 1:1 group than the 1:3 group. With only 39 and 40 sessions taught respectively, this added understanding of the student's reading ability and need for continued growth could affect what direction further teaching would take. Over a longer period of time this could produce greater reading ability in these students, then those in a small group. However, when looking at a small group, using direct instruction teaching methods, these results could be applied, concluding that students of similar ability level will learn just as much in a small group of 3, as in individual teaching.

Several different studies examining the effects of various grouping sizes were reviewed here, from whole-class to small group, including the implementation of varying group sizes, and considerations for low ability or learning disabled students. In the first study whole class instruction was favored over the use of small group homogeneous ability groups. This was due to

the negative effects of the leveled ability groups in the previously used basal reading program. Teachers found the homogenous ability groups limiting to the lower ability students in particular. When those groups were done away with, not only the low ability students, but the entire class showed a greater reading growth than previous classrooms. While this result cannot be attributed to the use of whole class instruction alone, it is a noteworthy finding.

In the second study Lou, et al. (1996) conducted a meta-analysis across various small-group methods. Here 66 studies were reviewed for commonalities in academic growth, student self-confidence, and positive attitudes toward reading. Results showed that when compared to classes where small-groups were not used, students in small-group instruction showed a significant advantage over non-grouped students in all three areas. While no one method of grouping could be found that served all students equally, there is considerable evidence that the use of grouping provided a significant growth for all students, over those not implementing small groups.

In the third study, Easton, Muirhead, Fredrick, & Vanderwicken (1979) sought to determine if there was a correlation between student attention and grouping methods, or the instructional delivery used. While the results showed that less attention was paid during small-group implementation, it is hard to apply these results directly to small-groups, without being provided definitions or descriptions of the small group or classroom activities. The nature and setup of the groups and activities can make a significant difference in student attention. However, it is important to acknowledge their finding that attention levels need to be considered as an important aspect of successful grouping.

In the last two studies the student to teacher ratio was looked at in small group lessons. In the first, Vaughn et al. (2003) found that a 1 to 10 ratio suffered when compared to a 1:1 or 1:3 ratio. Helf, Cooke, & Flowers's (2009) study supported the findings of the first, concluding that the differences between a 1:1 and 1:3 ratio group was inconclusive. This shows that in homogenous ability groups, where lessons are tailored to the group, particularly if using direct instruction methods, a smaller group is favored, but no smaller than 3 students to a teacher.

These results show that, when outside variables are controlled for, the use of small groups have a greater learning affect for the class. While a single grouping method could not be isolated as most effective for all, the strength of the small groups to create higher levels of learning could not be denied. The results also show that groups should be small, but don't need to be smaller than a 3:1 student-teacher ratio, when using ability-based groups, with students who are closely matched.

Small Group Dynamics

In this section the dynamics that play out in small groups will be analyzed. Study topics include examining the amount of reading done by various abilities of students in a group, the affect of ability grouping on kindergarten reading skills, how ability groups impact student attention, along with two studies looking at the moment by moment factors the converge during the use of small group guided reading lessons. The dynamics of a group, including the use of heterogeneous versus homogeneous groups, along with instructional strategies, make a significant difference in a student's learning. A student who is left unchallenged, or feels swamped by material they can't understand, will not be able to achieve as much. Group

dynamics also play a role in student distraction, affecting the amount of material covered and absorbed or understood by the students.

Allington (1982) conducted this study to look at the proportion of reading done in a homogeneous group setting by students who were placed in low ability groups, when compared to that of student placed in high ability groups. The participants included schools from several states, including New York, at three varying grade levels; 1st, 3rd, and 5th. Classrooms were based off volunteer teacher participation, and had to meet two criteria. First, they had to assigned to a grade level as specified above. Second, they had to use classroom based ability grouping as part of their reading instruction. Across all grades, schools, and classes, there were a total of 600 reading groups.

To give the most accurate account of reading instruction, the number of words read per session was counted. This was considered by the researcher to be the most sensitive and accurate method of recording the amount read by the students, particularly when compared to previous studies that implemented the number of pages. The purpose of this study was not divulged to the participating teachers in order to ensure a greater accuracy of results, and curtail possible inflation. Teachers were sent a letter with instructions, and index cards, previously prepared so they could be filled out in less than a minute following each reading session. After they had collected information for five consecutive days, the cards were mailed back to the researchers.

Data collected included the date, group ability ranking, the grade level, the material being read (including publisher, title, and copyright date), the pages read, the mode of reading (oral, or silent), and a sentence “indicating the general affective responses to the material” (Allington, 1982, p. 9). All data was compiled into a teacher log for further analyzation by the researchers.

Reading materials were also collected by the researchers, and a word count made, including titles, headings, captions, and text. Words were either counted individually, or based off a ten-line estimation method. From the teacher log 7 sets of numerical data was compiled for analysis; (1) total words read during the group session across all five days, (2) the total pages read across all five days, (3) the number of words read orally across all five days, (4) the words read silently across all five days, (5) unique text read, or text that was read for the first time during reading group, and not reread, (6) unique pages read, again excluding pages reread, (7) days reading, or the number of days spent reading during the five day period. All data was collected in the last half of the school year.

Results showed several statistically significant findings. First of those, though not unexpected, was that students in higher grades read more than those in lower grades. Teachers were also found to schedule reading nearly every day, regardless of grade. When looking at the difference across regions (New York vs. Other), no significant difference was found. When looking at the total number of words read, it was found that better readers covered more material than students in lower reading groups. This was true of the total pages and words read, along with the amount of silent reading. No significant difference was found for oral reading. Low ability groups also reread considerably more (about 25%) of the material than those of higher ability, but only in first grade. This number was dropped to 15% in third grade, and about 5% in fifth grade. The resulting effect is that students of a low reading ability are not only reading less material, but reading even further less new material. Students of a higher ability are being introduced to more new words. In first grade 1010 unique words were read in the high group compared to 288 words in low group. Third grade saw 4048 unique words in the high group,

while 2145 were read in the low group, and 6513 unique words were read in the 5th grade high group, and 4115 in the low group. Similar findings resulted across unique pages read. Pacing was also found to vary, with students in the high group reading and completing a story during each session, “poor reader groups often simply read pages while good readers invariably read complete stories with some good reader groups regularly completing more than one story per day” (Allington, 1982, p16). When comparing across grade levels a trend was shown that low ability readers nearly matched those of the high ability readers in the lower grade, giving some credit for the “hypothesis that teachers are simply instructing poor readers in much the same way as younger better readers are taught,” according to Allington (1982, p17). With this large gap in the amount being read by students the achievement gap is sure to continue to widen (Allington, 1982, p. 18).

This qualitative survey-based study surveyed a total of 60 schools across multiple states, and provided findings showing that the amount of reading done by low ability students varies quite significantly from the high ability students. The high number of participants in the study, and wide range of schools that it pulled from make these findings reliable. The analysis would have been strengthened had it taken place over a longer period, including multiple weeks. The amount of reading done per student is very significant when looking at grouping methods for reading. A student who reads more, deals with more problems and concepts, in a real life way, creating higher literacy skills and abilities. A student who reads slowly, but is still provided the same amount of reading time as a student who reads faster, is being limited in their instructional coverage. Students who read slower require longer reading times to cover the same material,

along with different pacing levels. When looking at issues of equity and not equality, this creates an issue of contingency.

This study, conducted by McCoach, O'Connell, Levitt (2006), sought to assess the effects of within-class ability grouping on reading growth across the kindergarten year, and identified other school and instructional factors associated with the growth in kindergarten students' reading achievement. The original study sample included 620 schools, with an average of 17.7 students in preschool. No data is provided for what was used, beyond the actual student count. 10,191 kindergarten students were included who (a) were included in the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K) conducted by the National Center for Education Statistics, (b) were first-time kindergarten students, (c) were not identified as learning disabled or limited English proficient (d) remained in the same school across the base year, (e) were from a school that had more than 5 participating students.

The ECLS-K investigated early literacy and reading skills, based off an item response theory (IRT)-scaled cognitive assessments on; (a) printed word recognition, (b) sound identification, (c) word reading, (d) vocabulary & reading comprehension. The scores were collected at the beginning and ending of the school year, subtracting the beginning scores from the ending ones, to get the end gain. The scores were then aggregated per school to provide an estimated average growth for the year. Variables were coded for schools, teachers, and students. The school-leveled variables were coded and looked at under multiple deviations to create extremely high-gain or low-gain schools. Teachers were surveyed concerning their frequency in use of ability grouping, school climate, and school-supported professional development. Logistic regression was used to determine the final outcome of high versus low-gain schools, controlling

for the percentage of Asian-Pacific Islander students, variables that did not make a significant difference contributing to high-gain or low-gain schools were discarded. The remaining variables were; (a) frequency of use of ability grouping (b) principals rating of teacher's success at meeting instructional goals.

In the final multi-level model for individual gain scores; gender, fall reading scores, and socioeconomic status were used as child predictors. At the school level; public versus private, half-day versus full-day, the percentage of Asian-Pacific Islander students, the success rating (as determined by the principal), and the aggregated frequency that Kindergarten teachers used ability grouping were the final predictors. The results showed that "kindergarten teachers' use of ability groups in reading was related positively to their students' early literacy skills and reading growth from the fall to the spring of their kindergarten year" (p. 344).

There were many variables that contributed to the growth of individual student's reading level, some that were accounted for, but many more that were outside this study. Limitations concerning selection biases, and self-reporting questionnaires for both teachers and principals were used, though they may not be entirely accurate. The frequency and length of ability grouping, was not otherwise verified. There is also a lack of data reporting how ability groups were used in the class, nor is there a report of quality of instruction in group work. The relationship found between the use of ability groups and an increase in school gains, may be purely causal, as no in-depth data was collected. This relationship cannot be extended to individual students' needs, nor can it be applied to low-achieving or high-achieving readers, as individual scores are neither supplied, nor evaluated. However, with all that aside, the end results still showed a positive correlation between the use of homogeneous ability groups, and

reading levels at the kindergarten level. Unfortunately, because we do not have data about the use and methods of the groups, nor the instruction method, we cannot make correlations concerning equity across the students, nor if heterogeneous groups would be more, less, or equally successful.

This quantitative study by Flemlee and Eder (1983) was designed to examine the effects of student attention in ability based groups, particularly within-class, homogeneous groups. Flemlee and Eder (1983) used a shift in the student's attention as the unit of analysis, tracking time with each shift, from being attentive, on topic, to being inattentive, or off-topic, with a particular focus on what makes a student inattentive.

The subjects of this study came from a single classroom in California. There were 23 students included from the classroom, primarily from a middle-class background. Four, relatively equal sized, ability groups were assigned by the teacher at the beginning of the school year. The students' ability was based off kindergarten teacher recommendations, and the current teacher's observations. No formal testing was done. There were 13 students divided between the two high groups; 8 male, and 5 female, and 10 student in the two low groups; five each of male and female. It is not specified in the study if the high or low groups were leveled, so that there were four levels of ability identified, or if there were only two levels of ability specified, and then divided into two groups each. The groups worked for oral reading lessons each day for about 15 to 20 minutes, during which each student was instructed to read in turn, with everyone else following along silently.

For purposes of this study group sessions were videotaped, and later coded. Two sessions of each group was recorded in the fall, and two in the spring, making a total of 16 recorded

sessions. Flemlee & Eder, (1983) reported that the students had previous experience with videotapes, and that the videotaped lessons were typical of group behavior. Attentive behavior was defined as looking at either what was being read, or taught during the lesson. All other behavior was considered inattentive behavior. Intercoder agreement was 89%, based off 4 group sessions. Variables were divided into 3 categories; individual characteristics, group characteristics, and time-dependence. Inside group characteristics group ability, reading length, and reading errors were coded as variables. Group ability was coded as either high or low, and was expected to make a significant difference in the results based off two processes; reading turns (including turn length, number of errors), and peer influence. An inattentive student was expected to distract others. Reading length and errors were recorded as separate variables in group characteristics, but still expected to reflect the group ability level. It is questioned whether reading errors would focus students on the reading task at hand, or get them off task, causing them to lose focus. Variable for individual characteristics included student gender, maturity level, reading aptitude, socioeconomic status, and past individual inattention. Time-dependence was a measurement of the length of time the student stayed in an attentive or inattentive state.

Results for individual characteristics were largely insignificant. Gender, maturity level, and past individual inattention were found to be not statistically significant. Reading aptitude though, was found to make a significant difference, with students who had a higher reading aptitude also having a lower inattentive rate. This relationship was negated, however, and found to be insignificant when group variables were included in the model. That is the effects of the group outweighed the effects of the individual.

When looking at group characteristics the group level had a highly significant impact. Students in the high ability groups became inattentive at one-half the rate of the low-ability groups, even when controlling for the individual variable of reading aptitude. Reading length and reading errors also created a significant difference in attentiveness. Longer reading turns created more inattentiveness, however a higher level of reading errors created less inattentiveness, and actually increased attentiveness. This was seen by Flemlee & Eder, (1983) to be caused by bringing the students focus back to the task at hand, reminding students where they were at in the course of the assignment.

Flemlee & Eder, (1983) also included an analysis between fall and spring observations. A significant difference was found in the fall results, where the levels of inattentiveness could not be accounted for by any significant variable, with the exception of attention duration. Stated simply the more a student was attentive in the fall, the less they were found to be inattentive (Flemlee & Eder, p. 84, 1983). In the spring there was a strong relationship between group characteristics, and inattentiveness, group ability being one of the most significant. Students in the high group became inattentive at about one-third of the rate of the students in the low group.

This quantitative study sought to discover the causes of student inattentiveness during the specific social context of homogenous ability groups. It affectively separated variables relating to individual students and group effects, providing different analysis for each. Because they applied the tests to only this type of student instruction arrangement, it cannot be transferred to other types of social contexts or instruction. All of the observations also took place during oral reading lessons only, also preventing transfer to other types of small-group instruction, including silent

reading or other content areas. This study used ability-based homogeneous groups only, preventing transfer to mixed-ability or heterogeneous groups.

The results showed a significant difference between results in the spring and fall, which may be due to the mood of the students at different times of the year. There could also be a causal relationship with group standards and rules, that would have been settled on during spring, and not fall. There is a lack of data provided for the days of the week, and time of the day during which the observations took place on. Both could lead to different classroom and student moods, causing greater inattention on Fridays, or at the end of the day, or before significant events (i.e., recess, lunch, field trips). Because the sample base was so small, only two days for each group, and with only one class as subjects, the possibility of the class mood affecting the results is higher. However, in light of all this, the study showed an example of the groups' characteristics dominating that of the individuals. Here the group mood and influence created more or less attention in the students, keeping them on or off task. This is an important characteristic of group dynamics that could lead to success or failure within a group.

The topic researched in this microanalysis was to “examine the social and cognitive consequences of silent reading, as compared to oral reading, in small-group lessons” (Wilkinson & Anderson, 1995). This study was done as part of a program of studies seeking to find a comprehensive understanding of the factors converging at a given moment as they occur during a small-group reading lesson in elementary school classrooms. This moment by moment account of what is happening during the lesson, is why Wilkinson and Anderson define the study as a microanalysis. Their goal throughout the series of studies is to discover what factors contribute to students comprehending, learning, and remembering information covered in a small-group

lesson. With a sociocognitive perspective on learning the researchers sought to analyze factors that contribute to group dynamics during small-group reading lessons. Specifically Wilkinson & Anderson (1995) looked at three factors; pacing, student engagement, and emphasis on story meaning. Their guiding questions were;

1. Does the type of lesson affect the group processes?
2. Does the type of lesson affect comprehension?
3. What role do group processes play in contributing to comprehension? (p. 716)

Variables including student, group, and material characteristics, were also recorded as part of the study, but not manipulated.

This study took place in two schools, one in a working-class neighborhood, and one in a middle-class neighborhood of an industrial city. Four 3rd grade classes participated with a total of 100 students. All classes were had grouping set up by the teacher, with three homogeneous ability groups in each class. Four basal stories were selected from the Economy *Keys to Reading* program, which was not being used in either school. Each story was 12 pages long, and intended to be covered in one lesson. Researchers developed 2 teacher's manuals for each story, one for silent reading, and one for oral. Teachers were instructed to ask the students a question, of their discretion, at the end of each page. During silent reading students were to wait quietly until the remaining students finished reading, then a discussion of the book was to take place. During oral reading, students were to remain quiet, following along, while one student read. When everyone had read an equal portion, the group would engage in discussion about the book. The study was conducted in fall in one class, and spring in the other. Stories were paired by interest (boy, girl) and difficulty, with one set for each condition (silent, oral). Stories were assigned to groups, so

that in one day, only one group read a particular story. Following each lesson each student was taken out of the room, and asked to retell the story to a research assistant, “so that a second grader who has not read the story will be able to understand it – even the hard parts” (Wilkinson & Anderson, p. 718, 1995). The median time students had to wait before retelling the story was 14 minutes. Students were also asked to engage in one other activity, based off the day so that each student engaged in each task once. The tasks, by day, included; (1) pseudoword list one, (2) Gray Oral Reading Tests-Revised, Passage, (3)pseudoword list 2, (4) Gray Oral Reading tests-Revised, passage two.

Stories were used in a lesson taught for two days in week 3, and two days in week 6. Prior to these lessons a socialization period was established, so that groups had “appropriate group norms governing teacher and student behavior” established (Wilkinson & Anderson, p. 717, 1995). Five separate tests were used to establish the student’s reading ability; the comprehension subtest from the SRA Achievement Series, the comprehension subtest from the Illinois Goal Assessment Program(IGAP), teacher’s ratings of comprehension on a 6-point Likert scale, Gray oral reading Tests, and finally two lists of pseudowords adapted from Stanhope and Parkin (1987) and Stanovich, Cunningham, and Freeman (1984) (Wilkinson & Anderson, p. 717, 1995).

Measures used for analysis included; recall, time use, student attention, and teacher-student discussion. Recall was based off a point scale for concepts and propositions repeated by the students. Time use was based off both teacher and student use of time, and was intended to monitor pacing. It included 5 measure; lesson time, reading time, waiting time, page-by-page discussion time, and pre-and postreading discussion time. Discussion time measures included

time spent during both the assignment and teacher management. Student attention was based off a moment-by-moment account of where the student was physically looking, and defined as “looking at the appropriate page(s) of the story or looking at the teacher if she was speaking to the group, without engaging in contact with other children that was not directly related to the story being read.” (Wilkinson & Anderson, p. 719, 1995). If, at any time, students were not following this prescribed order of attention, they were considered to be in a state of inattention. Scoring was based off a videotape of the lesson, and based off a percentage of time paying attention, or inattention. Teacher-student discussion was recorded due to the possible influences on comprehension when concepts or propositions were restated by the teacher during discussion, presumably relating to an increase in recall.

Results for time use suggested that pacing was affected by both the type of the lesson (silent or oral), and ability. Silent reading lessons were recorded to require a longer lesson time, though individual students spent less time reading, they spent more time waiting for the remaining students to finish reading. They also spent slightly more time discussing, though this time was considered insignificant at $p=.012$. On average silent reading lessons lasted 26 seconds in reading, 13 seconds waiting for others to finish, and 57 seconds in discussion for each page. Oral reading lessons lasted 33 seconds reading, no time waiting, and 50 seconds discussing the story per page. High ability groups spent less time reading, less time waiting, and less time in discussion.

When looking at student attention, “children were significantly more attentive during silent reading lessons than they were during oral reading lessons” (Wilkinson & Anderson, p. 724, 1995). Attention was also influenced by the average group fluency in oral reading lessons,

however it was documented that students were prone to distraction during silent reading lessons, from students who had already finished reading. There was also a decline in attention over the course of the study.

When looking at teacher-student discussion results showed that “substantially more story-related information – and important information- was discussed during silent reading lessons than oral” (Wilkinson & Anderson, p. 727, 1995). However, this focus seems to have taken place primarily in the students, not in the teacher’s reinstatement of concepts or propositions.

Comprehension showed no significant difference between the type of the lesson, however attention did. When returning to the original study questions it was found that silent reading lessons encouraged children to be more attentive, engage in more detailed conversations about the text, and recall more of the story. They also resulted in a slower pace of the lesson. When looking at the second question of comprehension, results showed no difference between silent and oral reading lessons. The third question, what role did group processes play in contributing to student comprehension, showed a positive correlation between silent reading lessons and both discussion and student attention. The researchers found that this result was negated by the slower pace of the lessons, caused by waiting time.

The results showed that students discussed more of the story during a silent reading lesson. This could be due, not to a greater understanding of the story, but to a desire by the students to share what had been read. The result of a higher level of engagement during silent reading is based purely off where the students were looking. During oral reading this result was confounded by listening to others reading, where mistakes may have been made that distracted the students. Students may have also read at a faster or slower pace than the student reading

orally, causing them to lose attention. There may have also been other factors at play, dropping student engagement, during the lessons, that was not obvious by the direction they were looking.

This study was designed as a within-subjects, quantitative study with a control or comparison group. Because the lessons were scripted, and there was a counterbalance for lesson order and the stories included the results cannot be due to variables outside of this study such as classroom climate, teaching style or experience, the characteristics of the stories, or student variables. Group norms were also accounted for, and both practices of silent and oral reading were previously existing parts of the students' routine. Students and teachers both had 2 to 3 weeks to get use to the lessons, before the treatment lessons took place.

Wilkinson and Anderson also found that, when time was taken into consideration, particularly waiting time, the benefits of reading silently, (greater engagement, discussion and recall) was nullified, though I call question to this. A student who more deeply discusses the story stands greater potential to understanding and recalling the story later. The result found in this study may have been nullified due to the low-level of complexity found in the stories being used. A simple basal reader, completed in a matter of seconds, does not leave much to discuss. It is difficult to predict how different the results would have been between the two lesson types, had they been reading a more in-depth story. The effect of waiting time, and group composition also should be reexamined. We know that the groups were preexisting, and based off ability, but no data was given (with the exception of one class) whether the groups were homogeneous or heterogeneous, and what considerations were taken into account to compose the groups. Reading speed or fluency being a consideration may create a change in the effects of this study. With that being said, Wilkinson's and Anderson's (1995) suggestion can be agreed with, that "teachers

will need to modify their instruction to accommodate changes in teacher and student behavior that accompany these lessons... and capitalize on group processes” during the lessons to make best use of silent reading (p. 736-737).

This quantitative study was done as part of a comprehensive series of studies on small-group, guided reading lessons, to better understand what “factors ... converge at a given moment” that will better inform the students for comprehension, and later recall (Anderson, Wilkinson, Mason, p. 419, 1991). The idea individual factors converging at a given moment, is the motivation for calling this study a microanalysis. Anderson, et. al, (1991) are interested in the “page-by-page and even proposition-by-proposition transaction” that occur during the lesson (p. 419). This study focused down to two issue of investigation. The first was whether an emphasis on meaning making while reading leads to better student performance, when compared to an emphasis on word analysis through oral reading. The second issue was to discover whether being the “*turntaker*” at a given moment, creates deeper learning or recall, when compared to students who aren’t currently reading or contributing to the discussion, the “*nonturntakers*”.

This participants of this study were drawn from 6 third-grade classes across two schools. A total of 149 students took part. The students were given a standard reading test during the study, and scored slightly higher than the national average with a stanine score of 6.6, compared to the national score of 5. The study took place during one week in the spring, utilizing 4 different previously selected stories. All 4 stories were utilized uniformly across all 6 rooms, and selected to be slightly easy for the students, with a Fry grade level score of 2.8. The stories were approximately 12 pages each, and read on Monday, Tuesday, Thursday and Friday, with a “battery of tests” taking place on Wednesday. Before each story was read the teachers delivered

the same instructions, telling students that they will be taking turns reading the story aloud, and that they may not talk during the other student's reading turn. Following the completion of each page a question was to be asked by the teacher, and tasks were to be completed following the reading of the book. Group arrangement was the same as was normally utilized during the school year, as set up by the teacher, all teachers had three groups, and all were based off ability. Originally one teacher had 4, and merged one of the middle ability groups into two of the other groups.

During reading lessons, the beginning and ending time for all lessons was tracked, and students took turns reading in a fixed order, randomly selected at a previous time. Two types of lessons took place; meaning-emphasis and surface-emphasis. Two lessons of each type took place on two different days. During a meaning-emphasis lesson student errors were not tracked in the book, but instead teachers were told to ignore them, unless they changed the meaning of the story, then teachers were to orally correct the mistake, and have the child repeat the sentence with the correct word. During reading questions made by the teacher included prediction questions about plot, and characters' actions. During a surface-emphasis lesson teachers were instructed to correct every mistake made during oral reading, even stumbles or self-corrections, making the student repeat the word promptly. While reading teachers asked questions relating to pronunciation of words on the page, including locating words that use *s* to indicate plurality, or words that use a long vowel and silent *e*. Following reading two tasks were completed by the students to measure comprehension and memory of the story, including writing a letter to a friend retelling the story. Short answer questions were also administered to the students, with one question for each page of the book.

Two tests were administered to measure the students' comprehension and fluency. Comprehension was based on each subjects' raw scores from the reading subtest of the 1971 SRA Achievement Series, Primary II, Form E, which were converted into national stanine scores. Fluency scores were computed by timing the children, in tenths of a second, while they read four lists of 15 words out loud. Recall scores were based off the four books read during the lessons. The stories were divided into propositions, based off an independent clause, and rated for importance in the story. Students' recall of these propositions in final assignments following reading, were scored. Oral reading were also scored, based off the mistakes made by the students during group work. Following all lessons the students were also asked to rank the four stories in order of preference, from 1 to 4, most to least interesting respectively.

Two units of analysis were completed to determine the final results, using the child and the group as a unit. Results indicated that the students recalled more of the story when there was a meaning-emphasis, rather than a surface-emphasis, both in a group, and with individuals. Turntakers also made significantly fewer errors while orally reading, during meaning-emphasis lessons, when compared to surface-emphasis lessons. This remained true when looking at interest in the book too, stories read with a meaning-emphasis scored significantly higher than those read with a surface emphasis.

When looking at the group, this average became smaller as the group comprehension increased. In turntaking, children were seen to recall more when they had a turn, then when they didn't. The pages of the book also made a difference in recall, with students remembering the beginning of the book more, then the last of the book. Group fluency had an influence the recall, with a greater difference on the early pages, then the later pages. A relationship was also found

between an individual's comprehension and fluency, and the group's average fluency. When looking at the length of the lessons, two factors were found to make a significant difference. Group fluency led to a shorter lesson, as did lessons that focused on meaning-making. The meaning-emphasis lessons led to an increase in students' scores in every measure conducted in this study including, "more complete recall of propositions, superior recall of important story elements, and more correct answers to short-answer questions They also rated the stories as more interesting, made fewer oral reading errors, and finished the lessons more quickly" (Anderson, et al., 1991).

This study was designed as a counter-balanced, within-subjects design, reducing the chance of results being swayed by reading levels of the students, or classroom atmosphere. Because the lessons were scripted, including all the questions to be delivered by the teachers, and following activities, discrepancies from teacher experience, training, or teaching methods are reduced. There is however, possible variables in interpretation for what should be marked as errors while students read aloud. The detail provided on the set-up of the groups is sparse. While we know they were all based off ability, we do not know if they were homogeneous or heterogeneous, what size were they, what the selection criteria was that the teacher considered for placing students together. Tests were administered on Wednesday, but what tests, and to what purpose? Did the administration of those tests affect the temperament of the students for the subsequent reading lesson? We also have no data providing which reading lessons took place on which days, or if they happened consecutively or not. However, the results still stand.

This study, like the previous one, provided another example of group dynamics affecting the student's learning. Here the group's fluency outweighed the student's fluency for recalling

and comprehending the story. This is a very important finding suggesting that student's with low fluency stand to take more away from an oral reading lesson when they are placed with higher ability students in their group.

This section has reviewed several examples of small group dynamics playing a significant role in student learning, even weighing out as a higher influence than the student's personal ability. These accounts have included the influence of ability levels, student attention and interest, and reading orally or silently. Allington (1982) looked at the amount of reading done by varying students during group work. Results showed that lower ability students who read slower have less opportunities to read. This may be due to time constrictions in the classroom. When student reading, whether individual or group based, is monitored not by the amount read, but by the amount of time spent reading, students who read slower will read less, thereby encountering less material, practice fewer reading skills, and test fewer concepts. In the second study McCoach, O'Connell, & Levitt (2006) looked specifically at homogenous ability groups in kindergarten reading lessons, and found a positive correlation between reading levels and ability-based groups. However, these results could not be correlated with equity across the classroom, nor could the results show that homogeneous groups were superior to heterogeneous groups. Flemlee and Elder (1983) reviewed student attention within group work. In this study two important factors concerning group dynamics were pointed out. First of those was that students who a high aptitude and ability for reading stayed on task more frequently. The second was that individual student attentiveness was swayed by the group's attentiveness and ability more than the individual's. So, if a group's ability is brought up, the group as a whole will do better. These findings suggest a heterogeneous group would create academic benefits for the

learners, provided ability was not spread far enough to create inattentiveness in high or low ability students. The results of the fourth and fifth studies were very similar. In the fourth, Wilkinson and Anderson (1995) found that students discussed a silently read book more thoroughly than an orally read book. This may potentially be due to the desire of students to share what they had read with their peers. In the fifth study Anderson, Wilkinson, & Mason (1991) found that the group's fluency that created more comprehension and better recall of a story than the individual's. Both of these studies have results suggesting a heterogeneous group may be more advantageous for learning.

Across the section several results were found suggesting that the group's dynamics affected individual student learning more heavily than an individual student's ability. Attentiveness, amount of material encountered, and reading orally or silently were all factors of consideration. While, results favored the use of heterogeneous groups, homogeneous groups were also found to be favorable, when they increased student attentiveness.

The Effect of Grouping on Student Needs

In this section I will be examining the affects of group use on various student needs. I hope to get closer to issues of equity by looking at the needs and results of students who are considered low, average, and high ability, along with those regarded learning disabled (LD), and gifted in various group formats. Many of the studies included look at the social aspects of grouping formats.

The first of the studies looks at the teacher's perceptions, and how they feel homogenous or heterogeneous grouping practices affect the students, both socially and academically. The second study sought to determine how students fare in a classroom with differentiated instruction

as the norm. The third study looks at inclusion classrooms, and how they affect LD and Language Learning Disabled (LLD) students. The fourth study seeks to determine if the negative stigmatizations that are associated with homogenous ability groups are isolated to those groups, or if they still exist in heterogeneous groups. The fifth study looks at the social environments in classrooms where homogeneous groups are used, under the belief that the teacher alone does not create the social atmosphere, but rather all members create it together. In the sixth study the social effects of homogenous grouping are also looked at, with the added element of examining how the teacher treats the leveled groups, and how the level of the group affects the students. In the last study homogenous and heterogeneous pairs are compared using paired writing.

This study was the first of two reviewed in this paper, under the same title. It was done as part of a series of studies addressing the need to better understand the academic, social, and instructional outcomes of grouping practices in reading instruction, as they are used in general education classrooms. The review following this is also from this series. This study seeks “to learn more about the teachers’ grouping practices as well as their perceptions of the effect of homogenous and heterogeneous groupings on the academic and social progress of students across a range of achievement groups likely to be represented in their classrooms: students with LD, low-, average-, and high-achieving students; and students identified as gifted” (Schumm, Moody, Vaughn, 2000, p. 478).

Twenty-nine 3rd grade teachers from an urban school district in southeastern US, were selected for this survey based study. A purposive sampling procedure was used so that the students represented in each teachers’ class matched the district’s average identified student population. The teachers used for this study were further limited by nomination of the

representing principal as being interested in the study, and containing at least one learning disabled (LD) student in the classroom. Third grade was selected as a pivotal year due the students continued early literacy instruction, and being one of the earliest years where teachers are “likely to have large numbers of students with reading problems who have been identified with LD” (Schumm et al., 2000, p. 479).

Data was collected over the course of one school year, from September to June, and based off a combination of interviews and observations. Interviews were conducted twice during the year, once at the beginning, and once at the end, lasting for 30 to 60 minutes. Teachers were questioned about their “perceptions about the social and academic effects of various grouping formats,” along with “descriptions of the teacher’s reading program and grouping practices within the program” (Schumm et al., 2000, p. 479). Observations took place 3 times over the course of the year, in the beginning, middle, and end of the year, and included an entire reading or language arts lesson, lasting for approximately 90 minutes. Field observers took notes using the Classroom Climate Scale (CCS), providing information on student and teacher behavior, “with a particular focus on the role of students with LD” (Schumm et al., 2000, p. 479). The scale also included specific information regarding grouping strategies, which were categorized as: (a) whole class; (b) groups with three or more; (c) pairs of 2 students; (d) students working independently. Following observations teachers were given a Teacher Self-Report, a checklist that contained five questions with possible answers listed below. Teachers were instructed to check all answers that applied concerning grouping strategies, leadership, and material use.

Observations concluded that the teachers were “in general...accepting of students (i.e., were fair/impartial and rarely used sarcasm or ridicule),” and were described as “average or

above average” by Schumm et al. (2000, p. 479). In grouping, whole class was found to be dominate, based off data from classroom observations, followed by independent activities, group activities, and finally pairs. Schumm et al. found that teachers reported differently in their interviews, when compared to these results. During interviews teachers revealed that about half used mixed-ability grouping, in either a whole class or small group format, and the other half using same-ability groups. In 21 out of 29 of the classrooms, teachers used whole class instruction with undifferentiated materials and instruction for all students, “whether they could decode it or not” (Schumm et al., 2000, p. 481). Three teachers had “permanent, same-ability groups (primarily for low-achieving readers)” (Schumm et al., 2000, p. 481) which included differentiated reading materials. However, only one of these teachers used differentiated instruction more than 95% of the time, and she disbanded the same-ability group by midyear due to “difficult[ies]” with “classroom management problems and lack of appropriate materials” (Schumm et al., 2000, p. 481).

Four additional teachers used mixed-ability groups “on an ongoing basis” (Schumm et al., 2000, p. 481), however all four teachers had an additional trained adult in the classroom, either a special education teacher, a Title I teacher, or a teacher’s aide. These four teachers did not provide differentiated instruction, but instead used the grouping strategy to “lower the student: teacher ratio” (Schumm et al., 2000, p. 481). The final teacher of the 29 is described as an “experimenter” (Schumm et al., 2000, p. 481), using a variety of grouping techniques that changed over the year. In this district Schumm et al. found that the traditional 3-tier ability grouping of low, average, and high, had disappeared along with differentiated instruction, to be

replaced by “whole class, undifferentiated instruction” (Schumm et al., 2000, p. 481), regardless of the students’ abilities.

Schumm et al. (2000) also reported that during interviews teachers explained their use of small groups for more intensive instruction, based off skill need, but did not find this during observations. Teachers also reported to have used mixed-ability groups for practice and reinforcement of skills, which was observed a total of 11 times in non-teacher-led groups.

The rationale for this shift is explained in two parts; school decisions, either from administration or the school council, and “limited access to materials” (Schumm et al., 2000, p. 481). Teachers were not “uniformly enthusiastic” about the use of these teaching methods. While some teachers sited the ease of preparation and management, others felt the opposite saying that whole class “introduced behavior problems” (Schumm et al., 2000, p. 482). Teachers reported a wide range of opinions, some saying it was school policy, while others intended to change it the following year. Some reported that it had a negative impact on students to be classified in the lower groups, while others felt they had more confidence using materials they could be successful at (Schumm et al., 2000, p. 482). Some teachers also reported that it helped students’ self-esteem to receive instruction from another student who may use different language, making the directions more accessible, while others said it did nothing for the higher students, who did the tutoring (Schumm et al., 2000, p. 483).

Schumm et al. (2000) identified word analysis as a particular need in differentiated instruction. They found that only two of the teachers provided this instruction for students who needed it, specifically citing students who were very low readers, or nonreaders in every classroom. They also paid attention to the reading instruction provided to students with an LD,

and found two trends. In cases where the student was pulled out of the classroom for reading instruction they worked on non-reading related work including homework, art activities, or math. Students with an LD who stayed in the class for reading instruction 25% of the time or more, participated in the same reading activities as the rest of the class, “whether the student can decode the material or not” (Schumm et al., 2000, p. 484).

Schumm et al. (2000) called question to the reading ability of students either with an LD, or based off perceived decoding abilities, yet performed no measurements of the students reading ability. They also concluded that teachers were incorrect in evaluating their own classroom procedures when looking at grouping strategies, yet spent only 3 separate days observing their language arts instruction. This limited observational sample could lead to gaps in their observations to what teachers were actually employing. Teachers were also reported as having confused questions concerning grouping strategies in the original interview with classroom arrangement choices (Schumm et al., 2000, p. 481). Due to the possible school policies concerning grouping strategies and limited materials these results cannot be transferred out of the district sampled. However, this study calls good attention to practices being employed by teachers with a wide range of ability levels in their class, including students with an LD.

This is the second study reviewed as part of the series conducted by Schumm, Moody, Vaughn (2000). As mentioned above, this series sought to better understand the academic, social, and instructional outcomes of grouping practices in reading instruction, as they are used in general education classrooms. Following the first study Schumm et al. (2000) were left with the question, “how do students fare in classrooms where differentiated instruction is the norm?” (p.

484). From that they derived the following qualitative study intending to assess “students perceptions of grouping practices for reading instruction” (p. 478).

A total of 146 students from the first study were used, but only those from the 21 teachers who used whole-class, mixed-ability grouping as their primary reading instruction strategy. Teachers were asked to nominate students based off 4 achievement categories selected by the researchers. Two students from each class were nominated, except in cases where two students were not possible due to class composition. The categories of high achievement (HA), average achievement (AA), and low achievement (LA), were based off students’ grades, scores of recent standardized tests, and classroom performance. Students placed in the learning disabilities (LD) group were based off district identification.

Students were given three separate tests both in the fall and spring, including the Kaufman Test of Educational Achievement (KTEA), the Peirs-Harris Children’s Self Concept Scale, and the Elementary Reading Attitude Survey (ERAS). The KTEA reading decoding and reading comprehension subtests were administered individually by researchers to each student. The Peirs-Harris Scale is a “self-repost questionnaire designed to assess how children feel about themselves,” and includes 80 ‘yes’ or ‘no’ statements from 6 cluster scales: Behavior, Intellectual and School Status, Physical Appearance and Attributes, Anxiety, Popularity, and Happiness and Satisfaction (Schumm et al., 2000, p. 484). The ERAS was administered in small-groups of 4 to 8 students, with all items read aloud. It used a scale of 1 to 4, asking students to “rate items related to their feelings about reading-related activities in the school and home” (Schumm et al., 2000, p. 484).

When comparing the KTEA test results from fall to spring a significant difference was found across all 4 achievement levels. High achieving students were found to have received the “most substantial progress in both decoding and comprehension,” with a gain of 0.7 in decoding, and 1.0 in comprehension (Schumm et al., 2000, p. 485). Average-achieving students made growth in decoding, at a rate of 0.5, but “made much less progress in comprehension,” with a growth of 0.2. While the reading progress of low-achieving students and students with an LD was “minimal,” receiving a gain of 0.4 and only 0.1 by the LA and LD students respectively in decoding, and 0.3 and 0.2 respectively in comprehension (Schumm et al., 2000, p. 485).

The Piers-Harris self concept test, provided “no statistically significant change over time” for any of the achievement groups (Schumm et al., 2000, p. 485). In the ERAS no significant difference was found across achievement groups, however with the students pooled from all four groups a significant decline in “attitudes toward reading” in school and home was found. Schumm et al. (2000) found that while “the average and high-achieving students made moderate academic progress in decoding and comprehension, the students with reading and learning disabilities demonstrated minimal gains. In general, students’ self-concept about themselves plateaued, while their attitudes toward reading declined” (Schumm et al., 2000, p. 486). Based on these results Schumm et al. (2000) recommend moving education for students with learning disabilities “beyond ‘one size fits all’” (authors quotations) and to provide “more intensive and explicit instruction, aimed at meeting students’ specific reading needs” (Schumm et al., 2000, p. 487). To implement this they recommend preservice and inservice teacher training to include a better understanding of grouping methods and procedures that account for growth and benefits both academically and socially, along with necessary resources and administrative support.

These results were based the practices of one school district, and could not be transferred to other school districts. Because there is no comparison group there is no account for the maturity or history of the students. There is also no comparison to other grouping strategies. However, it does point to solid evidence that equity in growth is not being received across ability levels. Schumm et al.'s (2000) results provided proof that a whole-class, mixed-ability grouping instruction method for reading and language arts best suited those students who already succeed in the education system, and provided no support for the low achieving or learning disabled students. As a means for equity these findings are very consequential.

This qualitative study notes that while “inclusion may be A student’s right, it may not be *right* for all students” [emphasis author’s], and places particular focus on the effect of undifferentiated instruction (all students, regardless of ability, participating in the same reading practices) for students with a language learning disability (LLD) (Silliman et al., p. 266, 2000). Researchers asked whether undifferentiated reading instruction, fails to be adequate for “included” students with severe reading problems. They feel an important question to ask is “whether these children are precluded from being academically successful by the very nature of the reading activities in which they are asked to engage,” and point out that “few studies in either general of special education have examined how the scaffolding of instruction influences children’s success or failure as readers” (Silliman et al., p. 266, 2000). Because children with an LLD may have added complications with inference capacities, the researcher’s wonder if the student’s abilities and differences in sociocognitive, linguistic, and discourse development influence the kinds of scaffolding they benefit from. This study compares students with an LLD to typically developing students, and seeks to discern a) the discourse patterns of two small-

group reading activities to define scaffolding sequences and degree which patterns reflected are supportive or direct instruction, and b) whether differentiated instruction characterized interactions with two children with an (LLD) contrasted with two, typically developing, younger children.

This study took place in a primary-level inclusion classroom in west-central Florida, and included 2 students with an LLD, and 2 typically developing, younger students. The two typically developing students are Bobby and Tim. They are both Caucasian males in first grade. Jimmy and Jerry, the students with an LLD, are both in 3rd grade, Caucasian, and male. Jimmy had education goals that included (a) word recognition, (b) vocabulary development for describing real-life situations in small & large group activities, and (c) increasing appropriate social interaction with peers. Jerry's education goals were (a) improving decoding skills, (b) increasing use of inventive spelling, (c) increasing ability to deal with frustration and teacher redirection, (d) improving verbal and written language production, including pronouns. Tim and Jerry were in group 1, with a general educator for a teacher, and Jimmy and Bobby were in reading group 2, with a special education teacher. Both groups were observed for a total of 13 sessions during the 1998-1999 school year, over an 8-week period. All observations took place in the natural setting of the classroom.

As an inclusion based classroom, with continuous progress, "emphasis is placed on the individual student's developmental level" (Silliman et al., 2000, p. 268), so student's of different ages may be at the same level, and students are supported at their current developmental level, regardless of grade. The classes are also multi-age, meaning that one class may have students 3 or 4 grade levels apart. The classrooms also utilize a theme-based curriculum through team

planning. Themes are selected that integrate various subjects across content “in a manner that will promote critical thinking through problem solving” (Silliman et al., 2000, p. 268).

The study looked at two types of instruction, and used as a lenses to categorize teacher assistance and student scaffolding in small groups; a) Knowledge Transmission and Direct Instruction, alongside b) Instructional Conversation and Strategy Instruction, based off the work of Vygotsky. Phonological Awareness Test (PAT), a standardized test, was conducted prior to the study. Jimmy and Jerry, were the only students to perform below standard. Student portfolios containing progress in spelling and writing over time were made available. Observations were made weekly, including videotaped sessions, and running records. Each session was mined for scaffolding sessions, and coded into one of four categories; (a) modeling, (b) offering explanations, (c) soliciting verbal participation, and (d) verifying/clarifying understanding. Scaffolding sessions were also grouped as either; (a) supportive assistance, or (b) directive, skills-orientated.

Because learning activities in this inclusion based classroom were based off a critical thinking philosophy, an expectation was made that a “problem-solving approach” would be used in the classroom, particularly in reading instruction. However, when looking at the discourse patterns of the teachers, “a gap emerged between beliefs and actual discourse practices” (Silliman et al., 2000, p. 275). Out of the 231 scaffolding sessions recorded, less than 1% of these were supportive, instead teacher assistance modeled direct instruction, and not “analytical thinking about phoneme-grapheme relationships” (Silliman et al., 2000, p. 271). During reading sessions discourse styles for both teachers were primarily consistent with the “gentle inquisitions found in a knowledge transmission model where IRE sequences predominate,” as defined by

Eeds & Wells, 1989 (Silliman et al., 2000, p. 275). An IRE sequence is a three-step process defined as Inquiry, Response, Evaluation, where the teacher asks the student a question, they respond, and the teacher confirms or denies their response. This type of discourse is part of the knowledge transmission, or direct instruction model (Silliman et al., 2000, p. 266). This idea of “container ... competence,” as is reflected in the direct instruction and transmission knowledge model, was also shared in assessment methods across both teachers.

The study reports the LLD students, Jimmy and Jerry, as receiving “interactional time that ... was at least equal to or greater than the time for Bobby and Timmy” (Silliman et al., 2000, p. 276). However the SLD teacher provided an additional 7 scaffolding sessions with Bobby, the TD student, than Jimmy, the LLD student, a difference of 35%. Across both groups, “instruction was essentially undifferentiated for Jimmy and Jerry (Silliman et al., 2000, p. 276). That is, they didn’t receive any special treatment when compared to the typically developing students. However, according to Salisbury et al. (1994), as quoted by Silliman et al., “inclusive schooling, multilevel instruction means that the amount of support and curriculum adaptation necessary to meet individual needs should vary” (p. 275). There is also residing questions about the level of instruction occurring in the groups, and whether it was “outside” of the LLD student’s “zone of proximal development, and, therefore, too difficult” (Silliman et al., quoting Salisbury et al., 2000, p. 276). Yet, in this study, the LLD students did not receive any differentiated instruction, leaving this paradox; “the expectation appears unwarranted that Jimmy and Jerry can succeed with the identical direct instruction provided to their younger peers in the group” (Silliman et al., 2000, p. 276). Silliman et al. (2000) suggest a number of measures to be “continuously evaluated” for successful academic outcomes for LLD students:

1. Teachers need to constantly monitor the strategies they are teaching, and children must learn to take over responsibility of these for themselves.
2. Individual differences among children must be continuously monitored, for problem-solving responsiveness and long-term academic effects.
3. Educational teams must engage in “risk-taking,” and recognize when “supportive scaffolding is appropriate for particular instructional goals”. (p. 277)

Silliman et al. (2000) pose a larger question, not of inclusion or special education, but of “scaffolding practices employed that best meet children’s individual needs, engage them in real learning, and assist them to realize their potential as contributing members of a community of learners” (Silliman et al., 2000, p. 277).

This study brought up some important questions that need to be asked in order to really understand the dynamics at play in grouping strategies across ability levels. Namely, does the type of “scaffolding practices employed ... best meet children’s individual needs, engage them in real learning, and assist them to realize their potential as contributing members of a community of learners” (Silliman et al., 2000, p. 277)? While this study does not seek to answer this question, it does beg of us to review our own practices and the needs of our students so that we can best serve them all.

In this study Poole (2008) seeks to determine if homogeneous grouping leads to a wider achievement gap between students, so that lower-ability students will be stigmatized as inferior, creating a negative effect on self-esteem, and a loss of motivation to learn, or if they still exist in a heterogeneous groups. This and similar assessments of level grouping methods have led many teachers and school administrations to forbid ability-grouping within the classroom. The goal is to “focus more attention on the instructional practices currently serving in place of traditional ability grouping in an effort to understand whether these practices may be inadvertently fostering the conditions they are designed to prevent” (Poole, 2008, p. 229).

This study was conducted with two groups of fifth graders in Highland Elementary, a public school located in a large urban area of Southern California. The school is predominantly Latino/Hispanic (87%), with 80% of all students qualifying for free or reduced lunch. The two groups used were built as mixed-ability, with a careful selection of one high-ability (HA), one low-ability (LA), and one average-ability (AA), with two additional students of any level. The groups observed in the is study were formed for instructional purposes in the classroom, and not just for the purposes of the study.

Of the ten students participating there were three monolingual English speakers; David, Shirley, and Keisha. The other students all spoke Spanish as a first language, and were labeled as “limited English proficient” either at a previous date, or currently. Three of the included students were considered low-ability students or struggling readers, however only two are highlighted during this study; Keisha, and Shirley. The third is not included because Keisha and Shirley represent two “distinct types of struggling readers,” when considering their reactions and participation in the reading group, while the third student is very similar to Shirley. Both Keisha

and Shirley were monolingual English speakers, “who exhibited more difficulty with reading than did six of the bilingual students” (Poole, 2008, p. 232). Both girls had been assessed using an Individualized Educational Program (IEP), however, according to their current teacher, Ms. Acedo, nothing was found that qualified the students for the additional resources provided by an IEP. However, in the year following this study, when the girls went into 6th grade, an intensive reading interaction was taken with both of them, regardless of the assessed need.

Data was collected based off two observations of each reading group instruction from two separate days. Each group met for 20 to 25 minutes, and finished their entire reading selection during that time. Each session was video and audio recorded, with two separate video cameras, and two supplementary audio recorders. These recordings were later transcribed to “capture as much of the participants’ verbal contributions as possible, along with relevant movement, gestures, and/or facial expression” (Poole, 2008, p. 232). Repeated analysis of the recordings focused on the participants interactions and resulted in patterns and regularities in the following phenomena; turn-taking, overlapping talk, pausing, and correction or assistance given during peer students’ reading turns.

Reading group sessions followed an instructional sequence where students took turns read a short section of text aloud, followed by teacher-initiated questions about the content of the text, or associated illustrations. The student’s level of engagement was high in both groups, “in part from the vivid illustrations of dinosaurs that appeared on each page of the text” (Poole, 2008, p. 233). Throughout all interactions the teacher, Ms. Acedo, remained “relaxed, patient, and humorous... she encouraged participation, asked numerous follow-up questions that prompted students to expand their verbal contributions, and made a concerted effort to engage

the students in meaningful interaction surrounding the text” (Poole, 2008, p. 233). Interactions were made up of the typical initiation-response-evaluation sequences, and were marked by conversation characteristics (Poole, 2008, p. 233). Analysis focuses on the read-aloud segments, and not the following interactional segments, because they offer the “clearest picture of the differentiation process” (Poole, 2008, p. 233).

Poole (2008) found significant interactional details in the two reading groups that showed “the same phenomena that have been previously found to characterize low-ability groups in general also affected individual students in these mixed-ability contexts” (p. 244). In fact, results showed that in the heterogeneous group, not only did the least efficient readers read the fewest words, but it was also a “powerful setting for identifying and maintaining a student hierarchy” based off the students’ reading ability (Poole, 2008, p. 245). This was seen repeatedly in student-to-student and student-to-teacher interactions.

Poole (2008) provides a clear explanation of data gathering methods, decision-making procedures, and analysis leading to the derived patterns. Triangulation was not used, nor was member-checking. This study cannot be transferred to other settings, due to a limited data set, but identifies issues that need large scale research. These results provide an example of the struggles low-ability readers deal with in read-aloud instruction in a mixed-ability group, and proved that the social stigmatism associated with homogeneous groups also exists in heterogeneous groups. This is an important finding. Social stigmatism needs to be avoided to create a classroom based off equity where every student can succeed. If that stigmatism is creating a situation where teachers or administration are choosing an alternative instruction method, yet the stigmatism is carried over with it, the students needs are not being served.

This qualitative study sought to analyze social environments in the classroom, where homogenous ability groups are in use. Grant and Rothenberg (1981) separated themselves from other studies on the same topic through the implementation of two beliefs in their study. First, that social environments are not the product of the teacher alone, but of “ongoing, contingent social relationships among all members” (Grant & Rothenberg, 1981, p. 5). Second, that these experiences cannot be understood by “summing and comparing micro-bits of activities” (Grant & Rothenberg, 1981, p. 5).

Research was conducted through ethnographic observations in 8 classrooms. Five classrooms were labeled as schools who served blue-collared communities, and three as white collared communities. The classrooms includes a college attendance rate by seniors of 30 percent by the blue collared community, along with aggregate reading scores that are below the Michigan reading standards. The white collared community has reading scores in the 95th percentile. All of the blue-collar classrooms were first grade. In the while-collared classes, one is first grade, one second, and one is first-second combination.

Observations were completed over a five or six month period for each classroom, with total observation time lasting for 15 to 30 hours. Individual observations sessions lasted for 20 to 80 minutes. Observation of reading group time accounted for 3 to 8 hours of total observation time in each classroom, with at least 3 sessions of each ability level group being observed. Observers took detailed ethnographic notes covering each session, recording data on task activities and patterns, classifying them as strong, moderate, or weak autonomy tasks. A strong classification required a larger proportion of the task to be completed individually, rather than as a group. Interruptions during group work were coded under four classifications; (IC)

interruptions from children not in the group, seeking the help of the teacher, (IT) interruptions initiated by the teacher to help a student outside the group, (ITD) interruptions initiated by the teacher to discipline children outside the group, and (ID) interruption where the teacher disciplined students inside the group. Praise, conversation, and feedback was also recorded, as were personal relationship between the students and teacher. Status of personal relationships were based off the frequency of conversations coded as either chatting, a personal exchange, or agenda-setting, a student's attempt to create or alter the agenda for the day's activity.

Notes were coded into four major types of activities; "quantity and quality of time devoted to each group; discipline patterns in each group; patterns of praise, criticism, and stated expectations for performance of group; and personal relationships among children and teachers" (Grant & Rothenberg, 1981, p. 7). Coding was designed to differentiate between variations in meanings, where the activity was similar. Group activities were coded off two criteria; (1) "the degree to which tasks offered opportunities for individual performance as opposed to performance as part of a group, and (2) the degree to which tasks promoted student autonomy as opposed to dependence on the teacher"(Grant & Rothenberg, 1981, p. 9). Assignments completed by students along with other group members, that were controlled by the teacher, were also included in the group tasks. Data was also drawn from formal and informal interviews with teachers, the researcher's observations of other interactions, both in and out of class, and the perspectives of the researcher's themselves, as "peripheral members of ongoing social systems" (Grant & Rothenberg, 1981, p. 9).

Results showed that only 7 students, four percent of the total, moved between ability groups during the course of the study. Of those, only one moved up, while the rest moved down.

When comparing across ability groups, the results found that students in high-ranking ability groups had significant advantages under seven themes, both academic and social.

Academic advantages include learning more academic skills. Children in high-level groups not only covered more material, but covered more complex material with fewer interruptions. High ability student also benefited from greater opportunities to demonstrate their academic competence. Because low ranking groups completed more group work, and were commonly required to complete it in unison, instead of finishing individual assignments, they had less opportunity to show individual growth. High ability students engaged in more autonomous, self-directed learning. This was found to be especially true in the white-collar classrooms, where high-ability groups worked independently more frequently. They were also afforded a higher level of chat and agenda-setting conversations, another form of autonomy. Finally, high ability grouped students were provided more opportunities and experiences to enhance expectations for future academic success. This included the use of praise, criticism, and expectations, with low-ability students receiving negative criticism and low expectation levels.

Under the social themes, higher grouped children enjoyed more trust from their teachers, had a higher frequency of equal status relationships and interchanges with teachers, and experienced more openings to develop close personal relationships with their teachers. The social themes of advantage primarily extend from conversations such as chat or agenda-setting dialog, and provide observable patterns that demonstrate teachers as being more accessible to students in higher ranking ability groups. This study showed patterns of social engagement across 8 classrooms in Michigan. Examples of dialogue between the teacher and student are

included, along with 5 tables of numerical data, though no statistical analysis or comparison is made across groups or community classification.

Schools are labeled as serving white or blue collared communities, and at times the researchers draw distinctions in the result between the two classifications. However, these classifications seem to be based off the researchers opinion, without providing any community or socioeconomic data for either the schools, district, or students. No data was provided as to whether the classrooms were in the same schools, or different, nor is there data relating to districts. Though I do not go into great detail in this literature review, the researchers separate the results of the white and blue collared classrooms, drawing separate findings from each side. Because only the white-collared had second grade students, this may have created a significant difference in the results when comparing the two groups to each other, swaying the outcomes unjustly. In conjunction with the rest of the study they do not use statistical analysis to support the differences found between the white or blue collared schools.

There is a wide range in the recorded observation times spent per site, including the length of the study, overall time in the classrooms, individual sessions, and amount of time observing group activity alone. This study would have greatly benefited from longer and more consistent observations. With some observations lasting only 20 minutes, and only 3 sessions of a group being observed, very little group work or dialogue could be accounted for.

This study showed significant results pointing to different treatment of students based off perceived achievement level, however these results to be consistent with the school or classrooms, not a separation of treatment between students in a single classroom. While, these treatment differences may be a social factor, rather than a response to a favored group, it is still a

significant consideration to be accounted for. Students who are treated like they can achieve less, will do so. The same is true for students who feel believed in.

In this quantitative study Gamoran, (1984) looks at homogeneous grouping under a model designed to “detect the social effects of ability grouping” (p. 4). With controls in place for individual reading aptitude he investigates two questions. First, does a student’s group rank affect their learning when controlling for individual ability and instructional differences. Second, does the teacher’s method of application of ability grouping effect student’s learning. To answer these questions Gamoran, looks at social grouping in two kinds of environments labeled egalitarian and elitist. Under the first label the teacher enacts a conscious effort to minimize the awareness of reading ability across groups among the students. To contrast with that teachers under the elitist label, enact instruction and behavior practices that make the ability level of students a conscious awareness across the classroom.

The data for this study was collected during the 1981-1982 school year from six schools that varied in socioeconomic status across three Chicago-area school districts. A total of 12 first-grade classrooms participated, two from each school, bringing 280 students into the study. Research was based off classroom observations that were conducted twelve times during the year, lasting for the duration of the school day. Observations were done at approximately three-week intervals. Social and instructional organization from each class was collected, along with student data from district and school records. Teacher interviews were also conducted. The students participated in aptitude tests in September, followed by additional testing in December, March, and May. Aptitude testing included the Barr-Ribby Word Learning Task, where testers taught words to groups of 6 students at a time, and recorded their retention 2 days later. The

Wide Range Achievement test was also conducted, where students required to read a list of words out loud. In December, March and May testing consisted of a sample of words taught to the reading group during class time, measuring the words learned.

Grouping strategies in classrooms were all designed and implemented under the teacher's usual manner. However, all were small, homogeneous groups. To isolate instructional from social effects of grouping a ranking system was developed for the purposes of this study. This rank was designed to be measured separately from the instruction each group received. This rank was recorded 12 times during the school year for each student, and was based off group placement made by the teacher. It took into account movement between groups. Data sets were analyzed in three waves; September to December, September to March, and September to May. Individual student background variables included; age, gender, socioeconomic status, and reading aptitude prior to the start of first grade. Because reading instruction in the first grade is centered on the control and introduction of words, along with beginning phonics skills, Gamoran (1984) included absences as a variable (p. 11).

The data was analyzed across all classrooms and students as a whole data set, prior to looking at elitist or egalitarian classrooms. The results examined the effects of homogenous ability grouping. Of the background variables, initial aptitude made the most consistent significant effect, and age had a negative effect, due to students repeating the first grade. With these variables accounted for data looking at group rank concluded that ranking had a small, though significant, effect on the learning of students in the fall. With an unstandardized regression coefficient of just under three words, some classes saw a difference of an additional 12 words learned by students in the high group when compared to the low group. This effect

minimized over the course of the year, with the smallest group rank coefficient in the spring, and an insignificant result in both winter and spring. Words taught makes the “most powerful influence” on the number of words learned by the students, and the unstandardized coefficients remained stable over the year, with students learning approximately three-quarters of the words they were taught, when other variables were accounted for. Gamoran went into further results of this phase of the study in a separate article, however the central finding of the results concluded that “what students learn depends on what they are taught, rather than on the rank of their reading group” (Gamoran, 1984, p. 14). This result was used as a baseline for the following research and results on egalitarian versus elitist classrooms.

Three classrooms were selected under each category, egalitarian or elitist. The distinction was based off existing teacher protocol implemented in the classroom from three factors; teacher ideology, discreteness of reading groups aptitude, and discreteness of reading group instruction. The basis of teacher ideology was concluded from sources such as interviews, remarks to students and parents, and names given to the reading groups. Discreteness of both aptitude and instruction was based off gaps between groups; were there large ability gaps between groups, or did groups have ability levels that overlapped; were the words taught to each group the same, or did it vary greatly. For a class to be categorized under either label it had to meet at least two of the above listed protocol. For example, classes under the elitist label named groups based off a ranking system, assigned by the teacher, so that entire class knew and felt the distinction in ability. They also announced separate instructions and expectations of different groups to the class, so that all the students knew there were lower expectations of lower ranked groups, and higher for higher ranked groups. An egalitarian class example allowed students to name their

own groups, and downplayed the ability level of groups when talking to students and parents, at times telling them they were of the same ability.

Results concluded a definitive differentiation between the two types of classrooms. In the elitist classrooms, the mean aptitude of every reading group was separated by at least one standard deviation in the fall, while groups in the egalitarian classes were only separated this extremely half of the time. The amount of instruction received by the groups also varied across class distinctions, with elitist classrooms showing a difference of as much as 60 percent less of the words being taught to one group than the next higher group. In egalitarian classrooms the same aspect averaged 70, 79, and 92 percent between three homogeneous groups.

Egalitarian classrooms also had a minimum of one pair of reading groups who received identical instruction. In egalitarian classrooms group rank was shown to not affect the words learned, even in the fall, with unstandardized coefficients of .97, .91, and .72 in the fall, winter, and spring, respectively, compared to .75, .74, and .74, in the full sample. Phonics did not contribute to learning in these classes until spring when its coefficient was 1.03, similar to the full model. Aptitude had an effect of “just over half” when compared to the full model, similar in March, and greater in May.

Group ranking in elitist classrooms had a “strikingly different” effect on the students’ learning when compared to both the egalitarian and full sample. The result for the full sample in the fall was just under 3 words, and in the egalitarian classrooms, there was no effect of group ranking found, as stated above. However, in the elitist classrooms there was an unstandardized regression coefficient of 7.72 words learned based off group rank, more than twice that of the full sample. This result also continued to grow the year, in contrast to the full sample, where the

results showed a waning, with a group rank of 20.69 words in the winter, and 36.82 words in the spring. In classes with three ability groups this resulted in students in the high group learning 74 more words than the students in the middle group, and 148 more words than students in the low group. These results hold student aptitudes and instruction constant across all groups, meaning that these show that “when teachers intensify the classroom stratification system, grouping has an effect on student learning that goes beyond differences in instructional content” (Gamoran, 1984, p. 22).

With these results in mind, Gamoran (1984) asked which groups are being affected. Are the students in the high group being benefited by their high status, and learning more, or are the students in the low groups achieving less due to feeling less capable of learning? With the group ranks set at a fixed score Gamoran (1984) looked at student achievement in reading. Results showed that students in the high groups stayed consistent in the elitist classrooms, when compared to students in other samples. However, students in the low and middle groups achieved less, learning only 222.1, and 295.7 words respectively. This is significantly lower than the average student in the full sample who learned 371.2 words in the low group, a difference of 149.1 words, and 374.2 words in the middle group, a difference of 78.5 words.

The final conclusion of this study indicates that homogenous ability grouping is not “inherently detrimental,” but rather it depends on how the grouping is structured and utilized in the classroom (Gamoran, 1984, p. 24). Generally speaking, student learning was found to be affected by the content covered by the teacher, however, “under certain circumstances learning is greatly constrained by the social consequences of the classroom stratification system,” with students in lower reading groups suffering the heaviest (Gamoran, 1984, p. 24).

This study sought to discover the effects of homogenous ability grouping in the classroom, and found significant results. It made use of a large data base, and controlled for many different effects, providing a sufficient look at the social effects of ability grouping on students. The study took place over the course of an entire school year, and included a look at the words learned by each student. While words learned alone does not completely equate the reading ability of a student, the results provided a consistent measure for learning over the year. The implication of egalitarian and elitist labels on classrooms was based primarily off “speculative” measures (Gamoran, 1984, p. 15). However, the application of each label was strengthened by the descriptive measure of the classroom environment provided, particularly with the elitist classrooms. The use of triangulation and member checking would have strengthened the results of this study, and the separation of the classroom type. Background data for the students and community in the elitist and egalitarian classrooms was not specified, and could have contributed to the results of students learning, particularly in the case of socioeconomic status.

This study was an investigation into the effects of Paired Writing (a specific, structured writing process) on the quality of creative writing and attitudes with 8-year-old children.

Sutherland & Topping (1999) focused on 4 questions for study:

1. “The evaluation of a clearly structured and therefore readily replicable intervention method (Paired Writing); which incorporated brief for participants, the collection of process data on implementation integrity and of outcome data through the analysis of writing products;

2. Effecting this with young children, where the applicability of collaborative writing is in greater doubt;
3. Comparing the effectiveness of Paired Writing with that of regular classroom teaching; and
4. Comparing same-ability reciprocal role and cross-ability fixed role pairing”. (p. 166)

This study was conducted as a true experimental design; pretest\posttest control group. Two classes were utilized; class A) cross—ability paired writers; class B) same-ability paired writers. Each class contained 2 groups of 16 students, 32 students total for the class, and 64 students that participated in the study. This break down of the two classes into 2 groups left 4 groups: Cross-Ability Experimental Group, Cross-Ability Control Group, Same-Ability Experimental Group, and Same-Ability Control Group. There is no data provided outlining diversity, location, or whether the school was urban, suburban, or rural.

Writing lessons were guided by topics, not titles, that were chosen during an open class discussion. From those topics were also chosen to be used as topics for the pre and post test assessments. Participant groups participated in two training sessions of the Paired Writing method, and all students were given a flow-chart of the steps. Sessions occurred twice a week, at 40 minutes each across all groups. A control group was in place in each class, that did not use the Paired Writing method, and instead worked individually. Participants were placed in either group by alternate alphabetic placement, and then groups were randomly allocated as experimental or control. Writing ability of students was ranked by the teacher, according to “perceived pre-project writing ability” (p. 166), and not based of pre-test. The cross-ability group was divided in

half; the most able group became the Helpers, the remainders were the Writers, and they were then matched into pairs with the most able helper to the most able Writer, and so on with parallel matching. In the same-ability group the most able participant was paired with the next most able participant, and so on. A completed writing project was expected every week. In case of student absences in the pair groups the remaining student worked alone, but this was rare, and evenly dispersed. Teachers were asked to not be involved with the groups, however they found this difficult. Sessions took place for 8 weeks, and were preceded and followed by pre and post tests of individual writing. Direct observations and informal discussions with the students and teachers took place throughout. A post-project feedback questionnaire was completed by participants, followed by group discussions. The 5-14 *National Curriculum Guidelines* (Scottish Office Education Department, 1991) was used to assess writing quality. The pre and post test was an individual writing piece, and were completed immediately before and after the project. Analysis of collaborative writing was based off the pair's choice of their best piece.

Concluding the study researchers found that students had difficulty with the vocabulary and concepts in step 2, drafting, during the training. Students also tended to follow the prompts in step 1, ideas, in the order they were written on the flow chart, instead of relevance of the task on hand. Groups did not progress through flowchart as fast as predicted, and the classroom was bound by time, so some groups did not complete finished pieces, until they became more familiar with the process. Time was also a factor in the final step, peer evaluation. This step was sometimes omitted, but frequency of this occurring is not given. Step 3, read the first draft aloud, was commonly omitted by students, unless prompted. Editing commonly focused on spelling or

punctuation errors, as opposed to focusing on meaning and order, as outlined in the Paired Writing process.

Test results showed that the writing quality was not significantly different across groups prior to the test. In the pre-test compared to post-test the only significant difference was found in the Cross-Ability Experimental Group, with an improvement of .036, while the Same-Ability Experimental Group still had score improvement, the Same-Ability Control Group had an insignificance scores decrease. Both classes showed improvement, but it was more evident for cross-ability pairs, then same-ability pairs.

When looking at the effects of the pre-test compared to the collaborative writing the Cross-Ability Experimental Group showed a positive difference, but more so for writers (lower-ability) then helpers (most-able). The Same-Ability Experimental Group showed a .014 significant improvement. When looking at the collaborative writing scores compared with post-test scores the Cross-Ability Experimental Group showed a positive difference, more so for writers (lower-ability) then helpers (most-able), but not statistically significant. While the Same-Ability Experimental Group showed statistically significant lower scores with a drop of .044. The cross ability writer's showed great improvement as a group, however in individual scores this improvement was primarily amongst the less able students in the class. The same-ability group showed a decrease in scores at the end of the study, and may be the result of anomaly or other uncontrolled variable.

This study leaves gaping holes when looking at the teaching methods inside the two classrooms. This variable is not discussed at all in the article, nor is there any attempt to try and control for variables in teaching styles or other classroom differences. While they asked teachers

to not interfere with group work, so as to minimize teaching style impact, the teacher's "found this passivity difficult" (p. 167). Their involvement may have provided more scaffolding to some groups than others. A pivotal element of the Paired Writing method was also omitted from the students' process at times. This final step, peer evaluation, could very well prove to be a key element of the learning process in this method, and may have made a significant difference in the learning of the experimental groups, particularly that of the helpers (the more-able peers) in the cross-ability experimental group, who showed less growth than the Writer's of the same group.

The same-ability groups showed either no statistically significant difference in improvement, or their scores actually decreased. I suspect there may be a outside variable not being accounted for, or this is an anomaly. This may be due to problems with the Paired Writing program, or the implementation of it, including leaving out the final step. There may have also been other variables occurring during the time when the students were completing their final work, that distracted them. Alternatively, it is also possible that the same-ability groups did not receive the challenge provided in the mixed-ability groups, as being either the tutor or the learner. Without further research, this question cannot be answered.

This study effectively showed the growth of cross-ability pairing in creative writing, with more growth happening as a group when working in cross-ability pairs, and more growth within that group for the less-able individuals. Alternatively, there was less growth for the more-able writers, in that group work. When taking into account equity it is hard to say why the high ability students demonstrated less growth. Had they already mastered the material at hand and need more complex concepts, or were they at a point in their learning where practice was needed, and

their growth amount was appropriate for that stage? Further research would be needed for this answer.

When deciding on grouping methods for equity, we must consider all student's growth, and not only those of the low ability, or high ability. By looking at how students of varying ability levels are fairing within different grouping arrangements, we can better understand what grouping methods will serve all students equally. In the first study Schumm, Moody, & Vaughn (2000) looked at teachers' use of scaffolding and differentiation in instruction for varying ability levels in their class. They also sought to understand how frequently homogenous or heterogeneous groups were used, and to what extent they helped students succeed, based off teacher's perceptions. While there were several teachers who behaved differently than the majority of the teachers surveyed, most were found to use whole class instruction without differentiating instruction for varying students' needs, including those identified with an LD. In the second study similar questions were explored by Schumm, Moody, & Vaughn (2000), focusing on student perceptions and academic success. Results here showed that only those students who were already successful in the system were served. That is, high performing students found continued success in a whole-class, mixed-ability instruction model, but other students, particularly low performing students and students with reading and learning disabilities. In the third study Silliman et al. (2000) sought to identify the scaffolding used in small-groups with LD and LLD students. Their findings supported those of Schumm et al. (2000) in the first two studies, showing that no differentiation or scaffolding was supplied to the students in need, even with the use of small groups. In the fourth study Poole (2008) sought to discover if the negative social stigmatism that is associated with homogenous ability groups is isolated to those

groups, or if they also play out in heterogeneous groups. His findings supported the later, showing the heterogeneous groups not only exposed the ability level of students, but led to creating a powerful social hierarchy that carried with the students outside the group. This is a very significant finding when deciding on grouping methods. If heterogeneous groups provide the same social hierarchy that homogenous groups provide, avoiding social stigmatism by using one method over the other is not possible. To create a situation of equity, and to allow every student to achieve to their best potential, these social roles need to be avoided. The indication of a social stigmatism needs to be considered when choosing reading programs, and when implementing them in the classroom.

In the next study Grant and Rothenberg (1981) looked at the treatment of students based off expected achievement level, and found that students who were expected to achieve more received more encouragement. The same was true of students who were expected to achieve less. While these results may be due to social differences across schools, it is still worth considering when making grouping decisions. In the sixth study Gamoran, (1984) looked at the affect of reading group rank in homogenous classrooms on students. He found startling results, indicating that the more a teacher emphasized the student's group ranking, the more it affected lower ranked students, dropping their academic reading scores. This same result was not found for high ranking students. That is, only the lower ranked students suffered under the homogenous ability group ranking system, while the high ability students remained relatively stable. While considering equity in the classroom, this is a very important result, providing examples of how to utilize and not utilize homogenous grouping systems while still creating equity.

In the final study Sutherland & Topping (1999) compared student achievement across homogenous and heterogeneous pairs. Their findings concluded that students who were considered less-able received greater growth while paired with a more-able student. Unfortunately, we cannot say if the higher ability students were suffering, or if the lower ability students were simply making large leaps. Either way, both groups grew academically when working in cross-ability pairs, but no growth was received at all in same-ability pairs.

Results across this section show the extent to which treatment of students affect their ability to succeed, along with evidence demonstrating ways higher ability students are severed, while low-ability, LD, and LLD students can be forgotten. Results also concluded that the treatment of student ability levels makes a lasting effect on the student's perceptions, and their learning ability. Classrooms where low and high ability levels are accentuated create a greater learning gap between the two. The opposite is also true, a de-emphasis on student ability levels can have a positive impact, shrinking the learning gap, and bringing the ability level of low-level students up, closer to high performing students.

Teacher's Practices: Instruction and Organization

This section reviews various studies examining what teachers are actually doing in the classroom. In the included studies questions seek to cover varied aspects including organizational strategies, practices, and teacher perceptions of grouping methods. In some of the studies the implications of these methods on student learning is also looked at. Three of the included studies have a focus on the diversity made available for ability groups in the classroom, including quantity and size of the groups.

This quantitative study sought to look at the relationship between teacher's assessment methods for ability grouping, and the "interrelations, correlations, and predictive power" of five different testing methods (French, 1991, p. 7). It was based off a single classroom of 26 third grade students, with one teacher. The classroom is described as homogeneous, low-to middle socioeconomic status, rural, and self-contained. No students were classified as learning disabled or in a Chapter 1 reading setting. The reading instruction across the class was based off a basal reading program, though the name of the program is not provided by the researcher.

The five tests included in this study were as follows. The Cloze test required students to provide missing words from a 4.1 grade level basal. Scores were recorded as a percentile based on exact answers, with synonyms or semantically acceptable responses "not accepted" (French, 1991, p. 8). The second test was Oral accuracy. The students read a 4.1 level basal book. The number of errors were counted, excluding self-corrections. An accuracy percentile was assigned to students based off the difference between words read, and misread words. Retelling was scored based off "Clark's method (1982) for free recall" (French, 1991, p. 9). A basal test from the program implemented was the fourth test, and a standardized test assigning a total reading percentile score was the final test. Along with these tests, the teacher's ability-group ranks, 1 to 3, along with student ranks of 1 to 26 (high to low) made by the teacher, were included in the study. Analysis was conducted two ways. First, across the five tests, and second, between the teacher assigned rank held against the test results.

Results showed that a high level of variation took place across the student rankings and the test results. The three "more traditional" assessment tests, were considered to be the Cloze, the basal placement test, and the standardized reading test (French, 1991, p. 14). These three tests

most closely matched the teacher's ranking of the students. The oral accuracy and retelling tests "appear not to ultimately contribute to the predictive power of the model" (French, 1991, p. 15).

This study evaluated the differences between the ranking of students by the teacher, and the results of five commonly used reading ability tests, and effectively showed where this one teacher's opinion varied from that of the tests. The researcher did not provide the time of the year the tests were administered, not how close to each other they were taken. This could have produced an effect on the student's results if the tests were all taken following each other, producing either an expectation or exhaustion in test taking. The time of the year the tests were could also be related to the correlation of the tests to the teacher's ranking, if the teacher had already administered any of the tests, particularly that of the basal program test. Because the teacher's reasoning for placing students in the groups as they were was not provided, we do not know if previous tests results were part of this consideration, or if the slight crossing of tested reading-ability and group level was intention to either minimize the effects of homogenous ability grouping, or classroom management considerations. This study would have also been strengthened by including a wider data sample. Because only one class was used, which implemented a basal reading program, the results may have been different in a classroom which utilized a whole-language method of instruction.

This study sought to examine teachers' perceptions, and uses of varying instructional grouping practices implemented in both general education, and special education settings for reading instruction, as well as the teachers' views concerning effectiveness with these grouping strategies. Forty-nine third grade teachers, 29 from general education (GE), and 20 from special education (SPED), were included in this study. Of the special education teachers, 17 were

teaching in a resource room, and 3 were teaching in a general education classroom. All teachers had at least one student with a learning disability, who was included in mainstream education, in their classroom to participate. The school district where they worked was a large metropolitan area in the South East United States with a high ethnic diversity, 54% to 98% minority students. School management happened in the school, and teachers were allowed to chose reading materials between 3 different district adopted basal reader programs, or use trade books.

Two series of interviews were conducted. The individual interviews were the focus of the study, which everyone participated in. The focus group interviews were voluntary, and only 29 teachers participated; 16 general education, and 13 special education. There is no mention which SPED teachers participated; resource room or general ed. class. Interview questions were based off previously conducted literature-based research on grouping from sources in education journals and research databases. The session was intended to be more of a conversation than a question and answer session, so questions were designed to be open-ended. All interviews were initially recorded, but later typed, and kept anonymous. This initial interview lasted for 30 to 60 minutes and focused on a) the teacher's background & perception of ability grouping during reading instruction b) factors that contributed to their decisions in practice c) their perceptions of social and academic effects of various grouping practices. The focus group interviews were done in groups of 3 to 7 teachers at a time, with 6 separate groups. These sessions were designed to follow the same formula as the individual sessions, with participants talking freely and comfortably.

Following the completion of the interviews qualitative data was pulled from the conversations consistent with the comparative method. Big ideas were identified based off;

words & ideas that dominated the conversations, intensity of responses, and nonverbal communication. Big ideas were considered hypotheses, not definite findings. Salient pieces of information, a few words to several sentences, were identified as units, and categorized with names. Related units were grouped in a common category. Following the identification of the units and big ideas, themes were generated. General and special education groups were analyzed separately, to compare and contrast the findings of the two groups. Neither triangulation, nor member checking was used during this process.

Three categories of teacher's perceptions were revealed. The first of these groups, "control over grouping," was predominately a frustration felt by general education teachers. They felt their decisions on how to group students for reading instruction was "influenced by decisions made by both district-and school-level administrators and that pressure in then applied for them to conform" (Moody & Vaughn, 1997). Special education teachers did not feel this same pressure, one teacher quoted the principal as saying "I wash my hands of you... as long as the children are being taught and I don't have them in the office, and I don't have any problems with the parents, then you can do what you want". Another participant expressed that the administration's stance toward special education classes is to leave it up to the teacher, doing "whatever is best for the children" (Moody & Vaughn, 1997).

The second theme that emerged was "whole-class instruction versus varied grouping formats". While special education teachers predominately used ability grouping in their classrooms, taking into account multiple issues including the child, the material, the goals, and in one case the IEP, general education teachers did not, and used predominately whole class instruction for teaching literacy. Different reasons were expressed for why they preferred this

method, but the majority reported that it was for management reasons. While this was the chosen method implemented by the majority of the general education teachers, several teachers expressed concern over the equity in learning for all students with this arrangement, saying the lower students fall behind or achieve less. The third theme that emerged was “achievement levels of students in groups; same or mixed?”. This theme pointed to the use of mixed-ability groups by general education teachers, when grouping was used, and same-ability groups by special education teachers. The general education teachers preferred this method due to the extra help it gave the lower-academic students.

While this study brought forward some valuable information about teachers opinions, no exact numbers were given, and opinions conflicted, so there is no way to know which ideas were more prevalent in what the researchers found. There was also a lack of triangulation or member-checking. It seems all the information collected and validated by the same two researchers, and teachers were never asked if they agreed with what was collected about them. All data was collected from two series of interviews, and no observation data was collected from the teacher’s classrooms or their students to verify what the teachers were reporting. This data was collected from one school district, and could not be transferred to other locations, particularly due to the general education teachers reporting mandated literacy curriculums dictating reading groups.

This qualitative ethnographic study examines the frequency and use of various heterogeneous grouping methods across 32 second and third grade classrooms. Guided by two questions; “(a) how prevalent are various instructional and organizational strategies to accommodate diversity in highly heterogeneous classes?, and (b) what are their consequences, if any, on student learning,” (Emmer, 1984, p. 2) researchers sought to establish to what extent

student achievement in mathematics and reading are related to heterogeneous grouping strategies.

The classrooms used in this study were a subset of from two previous studies. Following exclusions 32 of the original 42 classes were used, from 12 schools in a large, southwestern school district. Eligible classes had to be from second and third grade classes, be heterogonous based off two district-wide standardized tests, and have available student achievement data. Once the included classes were determined they were divided into three categories, based off their range of heterogeneity; (a) most heterogeneous, including 5 classes across all 5 levels of student reading ability; (b) heterogeneous high, including 14 classes with students across 4 levels of student ability, but not the lowest; and (c) heterogeneous low, including 13 classes with students across the 4 lowest levels of ability, but none in the highest.

Data included in this study was gathered from classroom observations and teacher interviews, as they were conducted and recorded from the original two studies. In the original studies observations included 15 to 17 hours in each classroom in 2 to 2 ½ increments over the course of the school year. Half of the observations were conducted at the beginning of the year. In one study teacher interviews were done twice a year, in October and May. In the other only one was performed. Interviews covered information regarding teachers' use and basis for grouping practices, instructional strategies and management, and management of student behavior.

Coding for heterogeneity use in the classroom was also devised by the previous studies, and was based off the interviews and observations. A series of themes was pulled from the data and used to make a questionnaire. Themes included; the teacher's instructional grouping

practices, whole-class instruction procedures, supplemental instruction programs, pull-out programs, teacher teaming, and other practices. Data for this study was gathered from the responses made on those questionnaires, and the extensive notes taken during observations and interviews.

Three variables were determined from the data set. First of those was organizational practices. This included the degree of use of small groups, the resource room, pullout programs, team teaching, individualized instruction, and peer or adult assistance utilized in the classroom for math and reading. Use of all instruction methods were coded. The second variable was modification of whole-class instruction. Under this five types of modifiers were coded on a dichotomous scale; extensive individual assistance, supplemental instruction for lower students, changes to seating arrangements so lower achieving students sat closer to the teacher, differentiated whole-class assignments, and extensive direction or explanations. The final variable was recorded as other characteristics. This included three variables which were expected to increase achievement, based off previous research; level of student success, emphasis on grades and accountability, and feedback on performance. The extent of free time was also included as a variable expected to negatively affect student performance.

When looking at organizational strategies results showed that only one teacher did not group students for reading instruction. That lone teacher had an extensive individualized instruction program. Of those using grouping, implementation of three student groups was the most common, however a few teachers used 2 or 4 student groups for reading. Pullout programs were used by an average of 8 students in 14 classes. Team teaching was used extensively in reading by 7 teachers, and moderately by 5. In math 7 teachers used it. Individualized instruction

was implemented in reading extensively by 11 teachers, and moderately by 3 teachers. A combination of strategies was used frequently. Of the 32 classrooms, only 6 used small-group instruction as their sole tactic for heterogeneity. In nine classes, special reading instruction was used outside the classroom. A combination of small group instruction with individualized instruction was used in 3 classes, with another seven classes adding group instruction to those methods. In 12 classrooms peer assistance was used on a regular basis, where students were instructed to ask other students questions, or for assistance. This was most common in reading instruction. Cooperative learning was not observed in any of the classrooms. However, this wide variety of organizational patterns and strategies is noteworthy.

For reading, significant gains were found when comparing whole-class student achievement and the use of individualized instruction, along with team teaching, the quantity of students utilizing a resource program, use of other adults in the class, the extent of explanations and assignment directions. A negative correlation was found when looking at free-time. In math, few significant variables were found. This study used effective sources for data, including triangulation. Data collection, and exclusions are clearly stated. A wide range of classrooms and teachers were observed, providing for a range of subjects. Because all of the classes were inside of the same school district, there is a possibility that district decisions, and resource or program availability could be influencing results. Member checking was not employed in this study.

This study was conducted as a follow up to a study conducted by Wilkinson and Townsend (2000) where four best practice teachers were interviewed and assessed based off their grouping strategies for reading instruction. The results found by Wilkinson and Townsend (2000) were tested across all teachers, in contrast to four carefully selected teachers, and to

include writing as well as reading. Rubie, Wilkinson, Parr, Townsend (2000) sought to determine if the results could be more broadly generalized by looking for answers to three questions;

1. How were groups formed?
2. What information was used to guide the formation of groups?
3. What are teacher's expectations about ideal group characteristics? (p. 5)

Subjects for this study were drawn from New Zealand public schools. Thirty different schools were included with a wide range of socioeconomic status and ethnicity. Among those, ninety classes were utilized, and of those 49 had children of four mixed ethnic groups, and 4 had only one ethnic group. Of the included schools 21 were contributing primary, 7 were full primary, and 2 were intermediate. Forty of the classes were composites, and one was vertically grouped for years 4, 5, 6. All primary grade levels, 1-8, were represented.

Data was based off observations and reports provided by student teachers placed in the classroom by the University of Auckland Schools, and a survey was completed by teachers. Student teachers were asked to write a report based off their observations, beginning on day one of the 1999 school year, and interviews with the teacher. These reports were to include information gathered from the teacher about the processes and procedures used to make "initial judgments about the literacy development of individual children and how they used this knowledge to organize their classes to cater for children's differing needs and abilities" (Rubie, C., Wilkinson, I., Parr, J., Townsend, M., 2000, p. 5). Student teachers were to first, describe the context of their placement, including the class demographics. Second, they were asked to describe how the class was organized for instruction in reading and writing. Student teachers were asked to answer the following questions (Rubie et al., 2000):

1. When classroom structures (e.g., reading groups) were established?
2. What were the structures (e.g., groups, whole class)?
3. What information did teachers use to make initial judgments of children's abilities and to make decisions about placing children in groups or other structures?
4. What form did this information take (e.g., reading levels, colors on the colour wheel)?
5. What importance did the teacher attach to the sources of information?
6. How flexible were the structures? (p. 5)

Student teachers were also asked to include and discuss any other factors that teachers may have considered when organizing their classes for reading and writing instruction.

The results of this study were "broadly consistent" with the findings of the Wilkinson and Townsend (2000) case study, with very few exceptions. As in the Wilkinson and Townsend (2000) study teachers formed reading groups based largely off ability. Writing groups, however, were primarily heterogeneous. Teachers defined their optimal groups as containing 8 or less students, with 4 or 5 groups in the class. However, in practice they often had to work with more.

When creating ability based reading groups teachers relied on three sources of information; the student's previous year's formal assessment conducted by a former teacher, their own informal classroom observations, and standardized tests. Teachers used the three assessment methods to varying degrees, but some consistencies and trends maintained across all teachers. Standardized tests were "only used to a limited degree, mainly at the intermediate level" (Rubie et al., p. 17, 2000). The prior formal assessments were the primary factor when creating reading groups, with personal observations used to a lesser degree, while the opposite

was true for writing groups. Writing instruction programs also started sooner in the year, sometimes in the first couple weeks, and generally by the fourth week for all teachers. Reading, however, took longer for teachers to start, and some delayed start of the reading program until the second term.

Other factors, beyond ability were also considered when forming groups, including organizational and management needs such as; how well students worked well together, who would provide peer support, timing, and where students were seated. Teachers also gave consideration to ethnicity, language use and acquisition, social factors, gender, and whether the class was single level or composite. Even with these considerations being accounted for in grouping, ability was still considered with the greatest weight.

The study provides a thorough explanation of grouping practices used in New Zealand, as they are perceived by teachers. This study does not provide evidence as to how effective these methods are, or what the students' perceptions of them are. This study could not be generalized to areas outside of Auckland, New Zealand, where the study was conducted.

This study is “premised on the belief that much could be learned about excellent beginning reading instruction by observing and interviewing excellent beginning reading teachers, extending research on effective education with respect to a critical educational task – teaching children to read” (Pressley, et al., 2001, p. 36). As such they sought to answer the question; what is the nature of excellent beginning reading instruction in contemporary America? Using a quasi-experimental, time series design researchers compared highly effective teachers to teachers who produce typical results of the district when looking at student achievement through the use of observations, interviews, and ongoing analysis.

Thirty first-grade classrooms were recruited from varying districts in New York, New Jersey, Madison, Wisconsin, the Dallas-Fort Worth area, and California. At each school the administration was asked to nominate both high achieving and normal achieving teachers, in pairs. Effectiveness of teachers were based off local standards and expectations. Administration was suggested to base their recommendations off standard tests for that area, knowledge of student performance in writing, student enthusiasm in the classroom, teaching styles that were the best representative practices as perceived by their district, involvement in professional development, and teacher creativity.

The nominations were not taken for granted to be correct. The researchers observed and evaluated all the classrooms to determine the highest and lowest achieving teachers at each school, and conducted their final report and study on them. However, all teachers were observed for the full length of time to determine which teachers should be used in the study. Fifteen pairs of teachers were initially studied and observed, each pair consisting of a highly effective teacher, and a typical teacher, and following reduction, 10 pairs were analyzed in detail for the study.

The teachers all had typical education including undergraduate certificate programs, graduate work, and continuing education. They were all experienced elementary teachers with varying lengths of experience from 2 to 30 years, and a medium of 16 years teaching experience. They all also had experience teaching in the first grade (range=2-21 years; M=11 years). There was variance in the approaches to teaching they utilized, including materials, and literacy instruction in the classroom, however, all were “well within the range of what should be expected in an American first-grade classroom in the mid-1990’s” (Pressley et al., 2001, p. 39).

Schools were selected from five different states in various regions, and spanned multiple school districts in each region. They included: upstate New York (4 school districts, 8 classrooms); urban New Jersey (3 school districts, 6 classrooms); Madison, Wisconsin (2 schools, 4 classrooms); the Dallas-Forth Worth area (4 school districts, 8 classrooms); and northern California (2 school districts, 4 classrooms).

No student data is supplied by the researchers covering diversity at the schools. Socioeconomic status is reported as being varied, but no further data is given. Some of the school districts were specified as being urban, though there is no data reporting any further detail on this (i.e. whether the schools were suburban or inner-city) there is also no indication of rural schools being included. The participating school districts seem to be locale to the varying Universities who participated in the study, though no discussion is provided in the article. Methodological triangulation was used to generate conclusions about the teaching methods of each teacher, utilizing observations, interviews, and a minimum of two observers visiting each classroom, and agreeing with the results. Some of the results were also confirmed by the teacher, and generally agreed with.

Observations were conducted five or more times at each site, with half-day or longer periods in each classroom. The exact frequency and length of observations per site is not given. Observations were conducted with a “privileged observer approach,” trying to be unobtrusive, and minimize their interactions with both students and teachers. The observers felt they were “seeing business as usual” (Pressley et al., 2001, p. 40). The number and length of observations completed varied per site, with 2 to 5 observers visiting schools, and 15 to 30 hours observed in the classrooms. Limitations due to financial and personnel resources, and travel time created

variances in the amount of observations, resulting in fewer observations at Wisconsin and California. The observers in the remaining sites lived in close proximity to the districts. All of the individuals conducting the observations had a background in reading education, ranging from professors to graduate students, including 3 graduate students working on dissertations, and another 3 working on master's degrees. There was a wide variance in beliefs regarding beginning reading instruction.

During observations the instructional methods of reading and writing were documented, and observers tried to “construct grounded theories” following the procedures of Strauss and Corbin (1998) (Pressley et al., 2001, p. 6). Data collection and analysis occurred simultaneously, and conclusions became more focused over time. Observations paid attention to teaching processes, types of materials used, and student's reading and writing performances and outcomes. Researchers also looked for information confirming or dispelling previous conclusions at each observation, taking note of new events, and checking them with previous observations for change or refinement of conclusions. Final conclusions were made based off multiple observations and confirmed by multiple observers.

An interview was designed by the researcher following the observations, and questions were written with the conclusions in mind designed to clarify their findings, and be informative about aspects of the classroom that the observer did not fully understand (e.g., “Can you describe your philosophy of literacy development at this level? Can you explain how your reading groups work?”). When necessary changes were made to conclusions, although those adjustments were generally very slight. In some locations the conclusions were presented to the teachers for review and comment, though specific locations and numbers are not given. Teachers agreed with the

conclusions offered, with few disagreements over minor issues. Those issues are not discussed in the article.

Teachers were defined as being either most or least effective based off their students' performance during class time, and their end of year performance. A standardized reading test was also administered as pilot data to six students in each class chosen by the teacher with 2 high-achieving, 2 mid level, and 2 low-achieving. Definitions for most-effective-for-locale and least-effective-for-locale teachers included the following. For a most-effective teacher 90% or more of the students had to be recorded as being engaged in productive reading and/or writing 90% of the time or more. At the end of the year these students also had to be reading books that were designed for end-of-year first grade students (i.e. picture books with several sentences on each page, or chapter books), and writing compositions that were over a page in length, with attention to correct usage of capitalization, correct spelling of high-frequency words, good use of inventive spelling, and sentence punctuation. For least-effective teachers less than 90% of the students were engaged in reading or writing less than 90% of the time, where observers notes included estimates of less than 2/3rds of the class being productive at the time of calculation. At the end of the year students were still reading books with very few words per page, and compositions were short (i.e. few sentences). Capitalization, sentence punctuation, and spelling were all "less impressive" (Pressley et al., 2001, p. 42).

Following data collection, teaching components in the most-effective-for-locale classrooms was identified. To do so researchers developed a summary of each teacher's methods, with detail strongly encouraged, commenting on 7 core topics: (a) the daily schedule of the class, (b) the nature of reading instruction and types of reading, (c) the nature of writing

instruction and types of writing, (d) how skills developed, (e) the extent and nature of opportunistic teaching, (f) the extent and nature of across-content connections, and (g) methods and effectiveness of classroom management. These summaries were all analyzed by one researcher who developed an exhaustive list 221 teaching behaviors and characteristics of the most-effective teachers. The list was double checked by the individual researchers who conducted the observations, and finally used to produce a list of 103 behaviors that occurred in all 5 of the most-effective teacher's classrooms. They were organized under the following arching themes: (a) excellent classroom management; (b) positive, reinforcing, cooperative environment; (c) skills explicitly taught (word-level, comprehension, and writing skills); (d) literature emphasis; (e) much reading and writing; (f) match of accelerating demands to student competence, with a great deal of monitoring and scaffolding; (g) self-regulation encouraged; (h) strong connections across the curriculum. To separate these characteristics from matching characteristics of least-effective teachers, the researchers at each locale were asked to indicate which behaviors were also demonstrated in the least-effective classrooms. A resulting list of 11 characteristics were compiled.

The results of the study were found to be very similar to Wharton-McDonald et. al.'s study conducted in upstate New York, and can be generalized to be the characteristics of effective literacy teaching in first grade across contemporary America in the 1990's. Findings showed that there was no significant correlation between socioeconomic level and most effective vs. least effective classrooms. New York and Wisconsin had equal classes. New Jersey and Texas showed a slight advantage in socioeconomic status in the most effective classes, while in California the least-effective class had the socioeconomic advantage.

The results of this study generated two lists of behaviors and characteristics that are utilized by Most-Effective teachers. These two lists are included as appendixes in the article; appendix A: Teaching Behaviors and Characteristics Typifying all Five Most Effective Teachers, and appendix B: Teaching Behaviors and Characteristics Distinguishing All Five Most-Effective-for-Locale Teachers From their Least-Effective-for-Locale Counterparts.

Overarching conclusions of effective teaching included some broad themes. The first was excellent classroom management. This included additional aspects not found in typical classrooms; (a) one or more assistants to the teacher, where coordinated instruction provided by paraprofessionals and special teachers maintained the integrity of the curriculum set by the teacher, and (b) teachers monitored reading, assisted with book selections when needed, insisted that students actually read the book they chose, & included lots of problem solving strategies. The second theme was extremely positive environment, where all students believed they were capable of being a 'reader'. The third theme was balancing skills instruction with whole language. Under that theme teachers did not hold true to a specific theory of teaching, but were adaptive to the needs of the students. They utilized opportunistic teaching and reteaching, mini-lessons when needed, teaching of word letter-sound analysis, immersion in literature and writing, and lots of planning, drafting, revising, and publishing of student writing. Teachers also taught students lots of problem-solving strategies, and to utilize multiple cues without priority. The fourth theme was the use of great deal of scaffolding, and keeping students on the edge of their ZPD, along with increasing expectations in writing, and being ready to provide assistance when needed, but never providing the answer. The fifth theme was expected self regulation and

independence, so that students monitored themselves and performed more reading and writing during class time. The final theme was strong cross-content connections and integration.

This study concluded that excellent First-Grade teaching is very complex, and requires a lot of elements. It showed that a purity in theoretical teaching strategies is not part of effective teaching, and that literacy is best acquired through a “focus on engagement, actual reading, and actual writing” (Pressley et al., 2001, p. 50). While observations were frequent enough to draw conclusions about the daily habits of the teachers, more observations along with consistent observations at the varying sites, would have added strength to the findings. The results are generalizable to the nation, and apply as general classroom teaching techniques for first-grade literacy programs. No acknowledgement is given to diversity of the varying schools, nor is attention given to individual student’s achievement, low-performing students’, or high-performing students’ growth. Possible variance by socioeconomic status is attended to, but not greatly. A more in-depth look, or reporting on economic and cultural variances between the districts would have benefited the study. The ongoing documentation performed by the observers provided progressive subjectivity. Triangulation was used through varying observations performed by multiple researchers, and interviews. Member Checks were used at some of the sites, providing a confirmation of the conclusions drawn from the classroom and teaching techniques. The history, or events that occurred during the study, is not addressed.

Inside appendix A: Teaching Behaviors and Characteristics Typifying all Five Most Effective Teachers, several indications toward the use of grouping show up. One of these is the seating of students in groups of three to six at each table, “to encourage interaction between

students” (Pressley et al., 2001, p. 53). There is also indication of group activities in reading and writing through buddy reading, and choral reading.

Literacy instruction is complex, and involves many elements. The strategies the teacher employs in the classroom make a significant difference in the learning of the students. Extensive use of scaffolding is important to student learning. Independent, self-regulated work by students increases student’s efficiency of class time. Actual reading and writing, along with engagement, are the most important aspects of a literacy program. If group work could exemplify some of the above mentioned methods, it would increase student learning.

In this study Paterson, Henry, O’Quin, Ceprano, and Blue (2003) looked at the effectiveness of the Integrated Learning Systems, a set of locally networked hardware and software that constitutes a complete curriculum delivered through computer-based instruction. Specifically the research focused on the use of the Waterford Early Reading Program. It was implemented through a title III technology and literacy federal grant in a large urban school in New York. Following its implementation Paterson, et al. (2003) sought to answer four questions concerning its effectiveness.

1. Did the Waterford Early Reading Program have a positive effect on literacy learning in the targeted classrooms? Was the learning significantly better than that of children in comparable classrooms without the Waterford program?
2. How was literacy in general enacted in all of the classrooms observed, and were there any patterns related to measureable differences in literacy learning among classrooms?

3. Were there behavioral and attitudinal changes toward literacy learning for children (a) in classrooms using the Waterford program and (b) when compared to children in classrooms not using the program?
4. Did skills from the computer-based learning transfer to classroom literacy learning?

The study took place across 30 urban schools in western New York state, with a total of 3,604 teachers. Of those 79% were Caucasian, 14% African American, 5% Hispanic >1% Native American or Asian American. There was a total of 45,902 students; 71.6% minority, 18% disabled, and 6.1% ELL. The socioeconomic rate of the schools included 43% of the schools with a poverty rate of 70%, and 67% of the students on free or reduced lunch. The 1999-2000 academic year had a class size of 3,478 kindergarten students, and 3,715 first grade students.

The Waterford program was implemented in 25 Kindergarten classes, and two 1st grade classes, due to high percentages of at-risk children. Eight kindergarten and one 1st grade class using the Waterford program, along with a matching number of non-Waterford program users were selected for inclusion in the study. Classes were chosen based off volunteer, and matched to comparable socioeconomic status, and style of teacher's instruction and classroom management. Research design was based off mixed methodology (both quantitative & qualitative), with a qualitative survey administered to answer one question, while quantitative data was gathered for the other three. A true-experimental design was implemented with a pre-test/post-test control group.

Observations were conducted by 14 graduate assistants, who were also certified teachers, and was based off Durkin's (1978-1979) method of timing certain instructional sequences. Time was used as a variable to study teacher directiveness, child-centered learning, time on each

literacy activity and classroom management. Observations were typically 2 & ½ hours, in the morning (as requested by the teacher), and were done 2 or 3 times in each classroom.

Observation field notes were completed with structured questions on a form to provide focus on previously established interactions they believed would provide the best sources of information for qualitative questions. Those notes were compiled to determine the best survey questions for teachers. Initial observations were done with two observers to check for clarity of definitions, and agreement was made at 100%. Observation notes were later analyzed by researchers, who categorized child and teacher actions. Student and teacher actions were discussed by the researchers, and a uniform code developed. They were used for multiple iterations of analysis, and the original forms referred to when needed.

Several additional methods were used to gather data. Brigance screens were administered to incoming kindergarten students by the district, and was utilized as a covariate, rather than write or administer a separate test. Clay's observation survey was given to a random selection of 6 to 8 students in each classroom. The survey took approximately 40 minutes and was administered to students individually by individual researchers in all 16 classrooms during May and June. The Clay survey included a running record, with-on-the-spot assessment, testing in letter identification, concepts about print, word tests, hearing and recording sounds in words in dictation, and the writing vocabulary test.

In fall of the year preceding the observations, participating teachers were asked to complete a survey to ascertain "how they perceived their literacy program" (p. 188). Teachers were asked to rate on a scale of 1-3 how important 12 different literacy components were to their program. Topics ranged from word study to reading aloud, but a complete list is not provided.

Open-ended questions were also to ascertain teachers areas of pride and goals for the future in their programs. They were also asked for specific information regarding the best features and biggest concerns of the Waterford program. These surveys were used to corroborate teacher's instructional style, as a basis for discussion during teacher interviews, and as a method of triangulation with observational data for teacher profiles through qualitative analysis.

End-of-year interviews with teachers were conducted in spring either in person, or by phone when not possible to meet. Interviews were semi-structured. Teachers were asked to elaborate on and clarify their beliefs on all questions asked by researchers. Questions involved how critical were varying components of their literacy program, what are they most proud of, goals for improvement, how do they accommodate wide range of abilities, and if they used the Waterford program, how affective was it. Teachers using the Waterford program were also asked to elaborate on community and parent response to the program, their own experiences, and what they would need to use the program more effectively. Results were used to construct teacher profiles, to provide data on attitudes toward the Waterford program, and to triangulate observational and survey data, so that conclusion could be drawn about the way teachers practices reflected their beliefs about technology and literacy.

Teacher and classroom profiles were developed, so that variables caused by classroom management and literacy facilitation could be accounted for. Teacher Profiles were developed through triangulation of observations, interviews and surveys. Discrepancies were found between what teachers identified as highly important and actual use observed in the classroom. Classroom Profiles were developed based off observations, and a matrix of 19 common elements was created. From those 3 primary themes were created as teacher variables;

- 1) Teacher control; teachers were labeled as being
 - a. Directive: where the teacher dominated activities
 - b. Facilitative: focused on student solicited ideas
 - c. Mixed: not predominately either, but utilized both
- 2) Literacy facilitation; Teachers were also categorized as being either:
 - a. High-literacy facilitators
 - b. Moderate-literacy facilitators
 - c. Low-literacy facilitators
- 3) Instructional time versus management time; based off duration and number of occurrences of classroom management compared to incidences and duration of instruction
 - a. High-instruction to low-management
 - b. Balanced instruction and management
 - c. High-management and low-instruction

The nature and quality of literacy events in each classroom were categorized based off; (a) duration, (b) quality of student to teacher interaction, (c) creativity, (d) student engagement, (e) group work, (f) materials.

Findings showed that the Waterford Program *did not* produce any significant effects on reading or literacy learning in kindergarten or first grade classrooms. Waterford students were not superior in reading and writing achievement to non-Waterford students. Waterford does not make up for early literacy at home, but rather children's scores on the Brigance Screen test remained a "strong predictor of year-end performance" (Paterson et al., 2003,

p. 199). There was no evidence that Waterford bridged development gaps for children who entered kindergarten or 1st grade with less literacy skills. Waterford claims that their computer system can individualize instruction for students in a more timely manner than teachers, but this “alleged superiority” was not supported.

When little significant difference was found with Waterford program classrooms, and the control group, the researchers turned to find out what could be concluded about the other forms of literacy instruction or conditions of learning that did have a positive effect? Findings showed that what did create a significant difference was “classrooms where teachers facilitated children’s active engagement in instruction” (Paterson et al., 2003, p. 199), including a greater proportion of time on instruction rather than management. Children in classrooms that demonstrated a number of best practices in teaching literacy as laid out by Cunningham and Allington, showed gains in reading and writing. “Only the classroom variables directly relating to teachers, as opposed to the Waterford program, produced noticeable differences in achievement for kindergarten and first-grade students at the end of one year of instruction” (p. 201).

This study provided continued support of what good teachers look like, and what strategies produce the best literacy in a classroom. Triangulation was achieved within each individual test method, and across methods. Teacher surveys, observation notes, and teacher interviews were triangulated. A control group was used. The researchers accounted for variables in teaching methods, strategies, and classroom management to the benefit of the research. The researchers did not spend enough time doing field observations in the classroom. Two or three observations could have created an anomaly. Doing further observations may have also dispelled

some of the discrepancies found with triangulation of the teacher's stated importance of features in their literacy program, and observed features.

Teacher instruction, modeling, and use of class time make a great difference in students' ability to acquire literacy skills. The use of group work could be included in that, provided that it led to increased time spent on instruction, and not class-room management issues. Use of class time being fully instructional, and not focusing on management issues is very important in student literacy learning. This could be an important aspect of grouping strategies and methods. The observations were low, and no specific data about activities taking place in the classroom was reported. The study did not look at grouping strategies or methods in literacy instruction.

This section looked at what was currently being done in classrooms including; grouping practices being utilized by teachers, their perceptions of those practices, the implementation of grouping strategies by administration, methods of determining student ability for group placement, and variations between considerations made for general education versus special education, along with other aspects. In the first study French (1991) sought to determine the similarities and differences between teacher's placement of students into homogeneous groups, versus the use of testing to determine their ability level, for group placement. While the results of the five different tests showed a high level of variation across the students, three of the more traditional tests closely matched the teacher's ability ranking of the students.

Moody & Vaughn (1997) sought to examine teacher's perceptions and uses of grouping practices, both in general education and special education, along with their perceptions of these practices. Their results showed threads of perceptions across the teachers. The first of these was a frustration from outside control over grouping made by administration in the school and

district, however this was not felt by special education teachers. In the second theme it was found that general education teachers predominately used whole class instruction, while special education teachers used ability grouping. The final theme showed the use of mixed-ability groups by general education teachers, and same-ability groups by special education teachers.

In the next study Emmer (1984) looked at what accommodations are made in the classroom to create diversity for highly heterogeneous classes, and if there are any consequences to those accommodations. Results showed that the use of three levels of ability-based groups were the most common, however, a large number of other strategies were also implemented. These included individualized programs, 2 or 4 student groups instead of three, team teaching, pullout programs, small-group instruction, special reading instruction, peer assistance, and a wide variance of combinations of these strategies. Strengths in student scores were found with the use of individualized instruction, team teaching, the use of resource programs, additional adults in the classroom, and thorough directions and explanations. Lower student achievement was found in classrooms that had greater free-time.

Rubie, Wilkinson, Parr, Townsend (2000), also looked at the formation and use of groups by teachers. They looked at how groups were formed, what information was used to guide that formation, and what were the teacher's expectations for ideal group characteristics. Conducted in New Zealand, the results showed that reading groups were nearly consistently based off ability, while writing groups were heterogeneous. Preferred groups sizes consisted of 8 or less students, with 4 to 5 groups in the classroom. To determine the groups students belonged in teachers relied on the formal assessment done by the previous teacher, standardized testing, and informal observations. These methods of judging ability were used to varying degrees by different

teachers, and were coupled with social and cultural factors to determine which group each student belonged in.

Pressley et al. (2001) sought to determine what excellent beginning reading instruction looked like today by comparing highly effective teachers to standard teachers. From this study two lists of behaviors and characteristics were generated, typifying the most effective teacher's practices. Of those classroom management was predominant, including coordination with classroom assistants, special education teachers, and monitored reading. The second strong theme was an extremely positive environment, where all students believed in themselves as a reader. In the third theme skills instruction was balanced with whole language, creating instruction that was balanced for the needs of the students, instead of being bound by a specific method. The use of scaffolding, and keeping students at their Zone of Proximal Development, was the fourth theme, and the fifth theme was student self-monitoring, so that they stayed on task for longer periods of time. The final theme was cross-content connections, and integration.

In the final study of this section, Paterson, Henry, O'Quin, Ceprano, and Blue (2003) focused on the use of the Waterford Early Reading Program, a locally networked computer-based instruction program. Through examining four questions about its effectiveness on student achievement, along with teacher perceptions of the program, findings showed that the program failed to produce significant gains in literacy for students in kindergarten or first grade classrooms, when compared to standard teacher-based instruction. It also failed to show evidence of a claimed superiority that the program could make up for development gaps derived from children who enter with less literacy skills.

Results across this section varied greatly, showing the high diversity level of grouping strategies that are being employed by teachers. It also showed that these choices are not always made by the teachers, but sometimes come from administration.

Student Perceptions

In this section the research centers around what student perceptions are of grouping strategies; what kinds of groups are they in, how does it affect them, what do they prefer. These perceptions are important to understanding how students' emotions over their group placement affect their learning. Student perceptions, rather than teacher perceptions, may also show evidence of social factors caused by ability level that could go unseen by teachers and administration. There are four studies reviewed in this section. The first looks at the effectiveness of homogeneous grouping towards students' attitudes of reading. The second study looked at the implications of friendships in homogeneous groups. The next study investigated the connections between students' enjoyment of reading, and their self-perception as a reader, with that of the ability group they are placed in. The final study reviewed in this section reviews student perceptions of six different grouping methods, with the hypothesis that students would be more concerned of peer perception in the classroom in cases of LD inclusion, and the use of homogeneous groups.

This study investigated the effects of homogeneous ability grouping in a third grade classroom in central Virginia, and its effects on attitudes of the students toward reading. It is a qualitative study based off a combination of observations and a questionnaire. The classroom is referred to as "nontraditional" due to the teacher's use of ability grouping and treatment of the students based off their ability, as in comparison to past research done on classrooms where

heterogeneous ability grouping was used (Schooley, 1994, p. 15). The students are described as “mostly middle class,” and “ethnically diverse” (Schooley, 1994, p. 15). There are 23 students total; 9 girls, and 14 boys. The classroom reading instruction is based off a basal series, and uses three levels of reading groups; low, middle, high. The teacher places a high value on reading, both in attitude and architectural layout, including resources, posters, and furniture placement. The teacher has approximately 30 years of experience, and maintains professional development, “keeping abreast of the latest trends” (Schooley, 1994, p. 16). She is described as “caring in the classroom and provides a lot of positive reinforcement for her students” (Schooley, 1994, p. 17).

Individual Reading Inventories are conducted at the beginning of the year, and used in conjunction with the teacher’s judgment to place students into homogeneous ability groups. Groups included 4 boys in the low ability; 6 girls, and 2 boys in the middle ability group; and 4 girls, and 6 boys in the high ability. During group work students read a book at their level while the rest of the class works on another assignment at their desk. Primarily these other activities include reading or writing in their journal, but sometimes other activities take place, such as completing an art project. The teacher works with each reading group during this time, providing them undivided attention, including supportive gestures and encouraging words. Each group met and worked together every day, and included time where the teacher read passages to the group, followed by questions about the text including word meaning, predictions, or character connections. Students were allowed to discuss answers freely. Students also read passages during each session, followed by engaging in discussion. Following each session students were asked to write a response to what they had read in their journals. The teacher commented on their writings before the next day. Students did not read chorally.

The survey was conducted at the beginning of the year, and consisted of 18 questions. They were all read aloud to the students, who answered by circling an appropriate emotional response represented by images of Garfield. There were four possible answers, and each was scored with varying points for later analysis; happy (4 points), somewhat happy(3 points), somewhat sad(2 points), and sad(1 point). Questions alternated between feelings about reading while at home to while at school. Before starting the survey students were asked two sample questions, and explained that everyone had different answers, and that was okay. They were also told they were not being tested, and that the test was to see how they felt about reading, not for a grade. Analysis included results for feelings about home, feelings about school, and points for each question. Observations were conducted for the first seven weeks of school, and again in March for 1 week. Observations included watching the teacher-to-student interactions, lessons being taught, teacher expressions, and corrections. Observations did not result in any quantitative data analysis.

The results are reported to be different from what the researcher expected (Schooley, 1994, p. 21). Schooley's expectations were that students of low ability would rate their appreciation of reading low, and student of high ability, high. The results showed only a small correlation. The low reading group did "generally fall in the bottom half," but not consistently. A prime example is the lowest ranking student in the class, whose reading appreciation ranks in the middle. (This student is pulled out of the class for extra reading time, in addition to group work in the class.) The students who were in the high ability group were spread all over the scale, though the top three ranks were held by students in the high ability group. Gender was also found to be a factor in reading appreciation in the survey. The attitudes of girls was found to be a

predicting factor in reading ability, but not with boys. So girls who appreciated reading did better, and girls who didn't appreciate reading did worse. But this was not consistent with boys. When comparing home and school reading, 6 students preferred reading at school over home, one was an even tie, and the remaining students liked reading at home better. Schooley predicts that the cause of this is the students get their choice of books. This idea is reflected by their appreciation of liking to go to the library and choose books.

Observations showed that the teacher spent more time with the low-level groups and students than with the high-level. She also spent extra time with four boys who ranked in the lowest reading group. During low level reading groups she is reported as being affectionate and supportive of the students including putting her arm around their shoulder, and making encouraging remarks (Schooley, 1994, p. 24). She also makes considerable eye contact and smiles frequently. She would also follow up reading times with comprehension, and reflective questions, including questions that connected the student to the main characters. "Word calling" was not allowed during group work, and students were rarely interrupted. Phonics was not included in group work activities. When not in reading groups students were treated equally, regardless of ability level. Students at a low reading level were not asked to read text to the class that the teacher considered to be difficult, but they did read writing workshop passages, and text at their level to the whole class. All students were allowed to ask questions equally, and were called on equally.

Schooley found that the attitudes of the readers did not correlate with the students' reading abilities due to the classroom structure, and positive emphasis on reading (1994, p. 27). Schooley attributes the general tendency of lower ability students to being toward the bottom of

the reading appreciation scale due to the added difficulty these students have when reading. That is, they are not very good at reading, and would naturally not enjoy doing it as much (1994, p. 27). However, they are not consistently the lowest ranking on the appreciation scale due to the teacher treating all the students as though they were high-level students. Schooley also says the dispersion of the high ability students is “most likely due to peer pressure” (Schooley, 1994, p. 27).

This study gave us an adequate look into one classroom, and one teacher’s method of enacting homogenous ability groups, using a three-tiered system. The results cannot be transferred to other classrooms due to the unique circumstances being employed by the teacher. No data is provided why this class was utilized during the study. The researcher’s theory is clearly stated, and counter-proved in the results. Progressive subjectivity is employed through the study, and discussed in the findings. Member checks and triangulation would have strengthened this study. Observations were conducted over an extended period in the beginning of the year, and followed up in the spring, providing for an adequate period to discover any salient issues.

Hallinan and Sorensen (1985) sought to research their preexisting hypothesis, that students placed together in a homogeneous ability group would build friendships. Their belief was that the use of ability groups would bring together children who had commonalities, and that the longer these groups stayed in tack, the more likely and closer these friendships would become.

Data for this study was taken from a data set of 1,477 students in 48 classes, across 10 elementary schools in northern California. The data set was narrowed by use of homogeneous ability groups as part of the classroom routine, and the final data set contained 104 reading

groups in 32 classes. Most classes were based off three ability groups, and teachers were asked to provide names of all students in each group, and the ability level of each group. Gender and race of all students was also recorded as a possible variable. Students were surveyed 6 times over the year, and asked who in their class was their friend. For each survey the student was given a list of names of everyone in the class, followed by 4 columns; “best friend,” “friend,” “know,” “don’t know,” and “my name”. Students were asked to circle the appropriate response for each name in the class, and encouraged to name as many friends as possible in the class, or to not name any, if their friends were not in the class. The data was analyzed for three factors; density, clique, and logit analysis. Density was aimed at determining whether students would have a greater proportion of friends over time. Clique was used to determine if there was an overlap between cliques and ability group members over time. Logit analysis was used to determine if there was a significant relationship between a two-student friendship and ability groups.

The findings found the original hypothesis to be true, “that membership in the same ability group increases the likelihood that students will become best friends” (Hallinan & Sorensen, 1985). Across each method of analysis results pointed to students becoming friends while being in the same ability-based group. This was true in density, while controlling for the growing ‘friendliness’ of all the students over the year, regardless of group size, with a longer group stability, and larger group size leading to a stronger effect on friendship. It was also true in clique analysis, where a clique was defined as a group of 4 or more students, who all named the same student as a friend. And it was true in the logit analysis, with controls for gender and race variables. Over the course of the year Hallinan & Sorensen (1985) found that students became better friends with the students in the same-ability group.

This study provided a quantitative look at the growth of friendships across students within the same classroom over the course of one school year. This study incorporated a large sample base, with mixed races, gender, teaching methods, and classrooms, providing for a wide diversity in subjects. There was also a lack of looking at friendships outside of the existing classroom, and students may have felt peer pressure to fill out the surveys with particular answers while taking them. There was no other data sources collected confirming the student friendships, only that of the student filled in survey.

Because it did not compare the findings of ability based homogeneous groups to heterogeneous groups, it is impossible to draw the conclusion that the same effects would not have been found in mixed-ability groups. The study also lacked to follow these friendships over the following year. Had the students truly been bonded to the degree the researchers claimed, over their same-ability status in reading groups, this bond would have had lasting effects on the students over more time, and may have disintegrated when the students ability levels moved away from one another.

This qualitative study sought to investigate the connections between the reading group level students were placed at in a homogeneous reading group across two variables. First, the enjoyment of reading by the students, and second, the student's self-perception as a reader. To do this Lasswell (1967) asked 5 questions:

1. Do children in the primary grades accurately perceive the level of their reading group placement?
2. Does accuracy of perception of reading group placement increase as children advance through the primary grades?

3. Is accuracy of perception of reading group placement related to level of group placement?
4. Will children in the high reading groups report greater enjoyment of reading than children in the low reading groups?
5. Will children whose perceived level of group placement is consistent with their actual placement report greater enjoyment of reading than children who over-estimate or under-estimate their placement? (p. 3)

To answer these questions individual interviews with 71 students from the Corvallis school district were conducted. A randomly selected classroom was used from each grade, first through third. The interview included 5 questions, and was presented under the title “Finding out how you feel about some things” (Lasswell, 1967, p. 4). The first question (“What is your favorite color?” and “How do you like reading?”) was an introduction to the interview, and not used as evidence in this study. The next two questions (were designed as a forced-choice situation, where they had to answer “very much,” “a little bit,” or “in between”. Question four was designed to ask the child to rank themselves as a reader compared to the rest of their class. To do this they were presented with a sheet of paper titled “My Class,” followed by an equal number of blank lines as there were students in the class. The students were then asked to point to where they would fall, if all the students in their class were listed there, from best reader to worst. The fifth question (“What would you like to be when you grow up?”) was intended to alleviate any nervousness the students might have felt over questions three and four. The interview was conducted twice with each participating student; once in September, 1966, and

again in January 1967. The data from these interviews, along with classroom data from teachers was compiled for analysis. The results were concluded based off each of the research questions.

When looking at children's perceptions of their reading ability group placement, about 35% of all the students accurately placed themselves at the level of reading as their group placement (low, middle, or high), in September. This number increased to 43% in January. Between those two test dates 7 students moved between ability groups. Of those, only one changed his perception of his reading ability in accordance with group movement. The second question (Does accuracy of perception of reading group placement increase as children advance through the primary grades?) showed a "tendency toward increased accuracy" from first to second grade, with 29% accurate in first grade, and 37% accurate in both second and third. In January these results grew slightly to 40% of first graders, 48% of second graders, and 43% of third graders.

Findings showed that when looking at a comparison between the student's group level placement, and accuracy of perceived reading level group placement, there was a positive correlation with high-ability students. That is, low ability students did not agree that they were low, nor did students grouped in the middle ability level. However, students grouped in high ability level primarily did. When looking at children's enjoyment of reading by grouped ability level across all three grades, the result was found to be insignificant, both in September and January. When looking at the same question only across third grade a significant difference was found. The students who had been placed in the high reading group were found to enjoy reading more than those placed in the low or middle reading group, however this result disappeared in the January interviews.

To look at the last question, (“Will children in the high reading groups report greater enjoyment of reading than children in the low reading groups?”) student’s answers were divided into two groups; agreement between the perceived ability level and actual ability level, and disagreement. For this question only those students who were available for both interviews were considered. Results were found to be “independent of agreement or disagreement” with placed and student perceived ability levels (Lasswell, 1967, p. 11).

This study looked at the correlation between placed reading group level, and student’s perceptions, both of their own level, and how that affected their enjoyment of reading. The use of two surveys provides a good comparison of student’s perceptions about reading from the beginning to the middle of the school year. However, no data was provided in the course of this study describing the grouping methods of the students. With the exception of knowing that 1st and 2nd grade were within class, and 3rd grade was across class, no other data is found to verify that the students were low, middle, or high ability readers. Because the third grade class was using grouping strategy that was different than the lower two grades, it may have some affected the third grade student’s enjoyment or perception of reading differently than the first or second grade students. The inclusion of only one class per grade, and all from the same school, may have also affected the results, due to administration imposed reading programs, or grouping guidelines. A larger subject base would have strengthened this study. The students may have also been intimidated by the researchers, who they most likely were completely unfamiliar with, and provided answers either just to get through the test quickly, or to inflate their ego. Because no testing was done on the students ability level, and only three levels were provided as part of the studies available rankings for students (low, middle, high), that students had a better perception

of their reading ability then was available in three groups. This could have put a student who perceived themselves as being between two groups, into one or another that would not have been as accurate according to an outside grouping strategy.

However, across these three classrooms there is still a trend to be found. Namely, that student's perceptions of their ability level in reading may vary quite a lot from that of teacher's or a placement test. Because we do not know what placement method was at use here, it is hard to implement any results. This could be a significant finding in the course of a student's path to literacy. A student who believes they are of a higher ability than a test or teacher, and receives a markedly low comment or score, could take it as an emotional blow. This defeat to their belief could affect them as a learner.

In the past, or in schools where some students are still removed from the general education classroom, students with an LD would receive their reading instruction "in a resource room reserved for students with special needs" (Elbaum, Moody, Schumm, 1999, p. 61). As more students with an LD are being included in general education classrooms, teachers find themselves with a wider range of ability levels in the classroom. This diversity has led both teachers and researchers to reconsider grouping methods. As a source of possible of guidance for how to best organize group work Elbaum, Moody, Schumm (1999) wanted to investigate students' perceptions of grouping strategies within school.

Students' concerns, as studied by Blumenfeld (1992) and Doyle (1977), and cited by Elbaum, Moody, Schumm, (1999), "have been shown to influence their motivation for learning as well as their classroom behavior" (p. 62). Thus, what students experience in grouping methods, both academically and socially, is important, and should be kept in mind when

planning activities in the classroom. Elbaum, Moody, Schumm (1999) hypothesized that students “would be more concerned with how they are perceived by peers and how their classmates behave when the class is divided into groups” for instructional reading activities (p. 62). To conduct this qualitative study they conducted an initial questionnaire as a preliminary study. The initial questionnaire was administered to 500 students from 3rd, 4th, and 5th grade. Twenty-three of those students were identified as students with a learning disability (LD). The subsequent follow-up study was conducted with 55 third-grade students, 27 of which were identified as students with an LD.

The initial questionnaire asked students about their perceptions of six different grouping formats for reading instruction; whole class, mixed-ability groups, same-ability groups, mixed-ability pairs, same-ability pairs, and independent work. Mixed ability groups and pairs drew the highest approval rating, while same-ability groups and working independently scored the lowest. The two open-ended questions, when responded to, indicated that students preferred group work in reading for promoting cooperation, and enhancing reading achievement, particularly so if the students were grouped to represent a range of ability levels. No specific data, or students responses, is provided concerning these results.

The follow-up study, built off the original questionnaire, was based off one-to-one interviews to get a better view of students perceptions of grouping formats for reading instruction. In this study the subjects were solely 3rd graders, nearly half of which, 27 out of 55, were students identified with an LD. Interviews composed of 5 open-ended questions with relevant probes to “elicit students’ ideas about different grouping formats” (Elbaum, Moody, Schumm, 1999, p. 62). Questions investigated what students liked the most, least, and why

when considering grouping methods. The various grouping methods being discussed, whole class, small group, pairs, and individual work, were described and defined to the students with the assistance of visual aids. Interviews lasted approximately 1 hour, and encouraged students to provide examples of various grouping methods used in class.

Through multiple reviews and analysis of the resulting tapes and transcripts, “emerging patterns were grouped according to the concerns students voiced about small-group work in reading” (Elbaum, Moody, Schumm, 1999, p. 63). Qualitative data was provided about these themes, but no data or results were provided about the other grouping methods. Students voiced the same concerns, as a typically developing student, or LD student, and drew into several themes. “Working conditions” (Elbaum, Moody, Schumm, 1999, p. 63) among students, including noise and disruptions, was a major concern for students. Elbaum, Moody, Schumm, (1999) state that these distractions may be “particularly detrimental to students with LD, who often require a calmer, quieter, environment in order to be successful at challenging academic tasks” (p. 63). Students also complained that getting assistance from a teacher is more difficult. Students felt that with the additional groups to go between, the teacher forced other groups to wait for questions to be answered, to the extent that some preferred whole class instruction. Students preferred whole class to avoid anxiety of not knowing the answer, or being made fun of. Some higher-achieving students complained about mixed-ability groups feeling that their progress was slowed to help the others. These results match general themes found in the initial questionnaire, where higher ability students also expressed opinions that they liked helping other students.

Elbaum, Moody, Schumm, (1999) draw four conditions that are necessary to control for to provide positive student experiences; working conditions, getting help from the teacher, ridicule, and pace of learning (p. 64). They make suggestions for specific strategies including, managing both the social and academic interaction of the group, providing opportunities for students to seek help from the teacher, controlling student ridicule by teaching students to apply constructive criticism, and providing multiple grouping strategies so that all students are challenged at their own pacing level . Many of their suggested strategies are group facilitation strategies including assigning specific group roles, rules, and procedures for asking questions. Elbaum, Moody, Schumm, (1999) also suggest that flexible grouping, or using a wide arrangement and varying their grouping formats, both in general education and special education classrooms, will better provide for students' needs, particularly "when students with disabilities are included in the general education classroom" (Elbaum, Moody, Schumm, 1999, p. 65).

Not all students could provide examples of the various grouping formats, and may not have experienced them, resulting in possible hypothetical answers (Elbaum, Moody, Schumm, 1999, p. 63). There was also a striking lack of data provided about students' perceptions of other grouping methods. Researchers used peer debriefing to reduce any possible biases.

Four different studies looking at student perceptions of homogeneous grouping were reviewed above. The first of those found that the students' attitudes toward reading did not correlate with their reading ability due to classroom structure, or use of grouping techniques. Instead Schooley (1994) found that when teachers treated all the students like high-level students, it raised the appreciation level of reading. The second study looked at the implications of friendships in homogeneous groups. Halliman and Sorensen (1985) found their hypothesis to

be true, concluding that students who were in the same reading homogeneous reading group became friends. In the next study Lasswell (1967) investigated the connections between students' enjoyment of reading, and their self-perception as a reader, with that of the ability group they are placed in. Findings showed that the students' perceptions of their ability level was quite different from that of the group they were placed in, with the exception of high-ability students. Student enjoyment of reading only showed a difference in the beginning of the school year with low and middle level students, but disappeared by January.

Elbaum, Moody, Schumm (1999) reviewed student perceptions of six different grouping methods. This study is conducted with an awareness toward inclusion of learning disabled students in general education, the resulting growth of diversity across ability levels in the classroom, and with the hypothesis that students will be more concerned of peer perception with the use of homogeneous groups. Resulting trends showed no difference in answers between answers from LD students or general education students, and four conditions were identified to control for positive student experiences; working conditions, teacher assistance, ridicule, and pace of learning. To control for these conditions varying grouping formats, group facilitation, controlling student ridicule, and opportunities to get teacher assistance are all suggested.

Results across this section point to a glaring difference in student perceptions of their reading ability, and teacher perceptions of their reading ability. Students who teachers may identify as low-ability, or learning disabled, may not see themselves in the same light. Results also emphasized that the ability level of the students did not create more or less enjoyment in reading by the students, nor better friends.

Summary

Chapter three reviewed research about grouping strategies for teaching literacy. Findings of these studies were analyzed, and summarized, based off the conclusions provided in each study. The research was examined for the effects on equity across all students of the classroom, in learning literacy.

When looking at group sizes, the results concluded that small groups were more effective than whole-class instruction, however this result could be dismantled by outside factors such as negative implementations from ability-based stigmas. Groups sizes should also be small, but don't need to be smaller than 3:1 student-teacher ratio, when students were closely matched in ability. The effect of small group dynamics favored the use of heterogeneous groups, but also showed some positive correlations for homogeneous groups. Student attentiveness, amount of material covered, and group fluency were all considered important factors in creating student learning. When looking at the Effect of Grouping on Student Needs, the results showed the in cases of homogeneous grouping, the teacher's emphasis or de-emphasis can make a significant difference in the success of students. This effect can create a growth in the learning gap between ability levels. Results also concluded that in cases of whole-class instruction low performing students, and students with reading or learning difficulties and not being provided the additional scaffolding they need. The review on teacher's practices showed that while a high level of varying decisions are made, these decisions do not always come from them. Sometimes there is a top-down decision made by administration for a particular method to be used. In Student Perceptions, results showed the differences in perception held between students and teachers. Students who may be classified as low-ability, or learning disabled, by teachers, may not

reciprocate that classification. Findings also concluded that student ability level did not directly correlate to enjoyment of reading. That is students at all reading levels could enjoy or dislike reading equally. Chapter four will conclude and summarize the findings of Chapter three, with considerations toward group sizes, small group dynamics, student needs, teacher's practices, and student perceptions. It will go on to propose classroom considerations, and suggestions for further research.

CHAPTER 4: CONCLUSION

Introduction

Chapter 1 examined the importance of a high level of literacy in our society, and its place in our society. As technology and society continue to grow, the demand for individuals to acquire and maintain higher literacy ability also continues to grow. Literacy is imperative for student's growth not only academically and professionally, but also for personal and social growth as well. Chapter one also explains how imperative it is for students to acquire the building blocks of a literature rich environment at a young age, and how the learning gap between more able and less able students continues to grow. In chapter one the many decisions to made when grouping students for reading instruction are discussed. Group dynamics, student interest, and ability level can all change the effectiveness of a literacy group. Because of the biological level of learning happening in the brain, the interference of emotions, and the need for students to stay optimally challenged at their Zone of Proximal Development (ZPD), these consideration in grouping can be convoluted.

In Chapter two the use and development of grouping in education is explained, along with the controversies around different grouping methods. Whole-class instruction may leave students far below or above their ZPD, and favor an individualist society steeped in equality and not equity. Small-group instruction allows for more diversity and individualized lessons, but may point out differences in ability to peers, leaving students to feel stupid or incapable, varying with the use of heterogeneous or homogeneous grouping strategies.

Chapter three reviewed the research of grouping methods, and was divided into 5 sections; Group Sizes, Small Group Dynamics, The Effect of Grouping on Student Needs,

Teacher's Practices, and Student Perceptions. Each section was summarized and analyzed based off the findings provided. All research was reviewed to examine the effects of grouping in the classroom to create equity across all student's learning.

Chapter four will conclude this literature review. It will revisit the guiding question; what grouping strategies and methods are most effective at creating equity in learning literacy across ability levels. It will use the summary of findings from the five sections of chapter 3 to answer the question, propose implications for classroom use, and suggest further research.

Summary of the Findings

What grouping strategies and methods are most effective at creating equity in learning literacy across ability levels? This question is immensely influential when you look at the vast number of variables that teachers and schools have to deal with. Across the five sections from chapter 3; group sizes, small group dynamics, student needs, teacher's practices, and student perceptions, significant results were reviewed and evaluated. The results from each of those sections is summarized below under the same headings.

Lou et al. (1996) found that classrooms who used small-groups held higher levels of academic growth, student self-confidence, and positive reading attitudes, when compared to classrooms who did not use small groups of 2 to 10 students. These results did not hold consistent to one particular method of grouping, but rather concluded that small-groups were more effective. Easton, Muirhead, Fredrick and Vanderwicken (1979) found that student attention, within the same classroom, was not affected by small-group or whole-class arrangement. Vaughn et al. (2003) and Helf, Cooke, Flowers (2009) determined that an optimal student to teacher ratio is less than 10:1, but need not be smaller than 3:1. These results conclude

that small groups can and do provide for optimal learning opportunities, including academic growth, self-confidence, and positive attitudes. While Harris and Harrison (1988) found that teachers preferred the use of a whole-class instruction method, over the use of the previously used basal program implementing homogeneous groups, this finding cannot be transferred to all whole-class instruction techniques. It does prove as a shining example that whole class instruction can be better than basal-based reading programs implementing homogeneous grouping techniques. Considering the results on group sizes, the use of small-groups in teaching literacy can easily be seen as favorable to optimize learning. Those groups need not be minimized to individual or pair instruction, to receive optimal learning. Results also show that whole-class can be favorable to some methods of basal, leveled reading programs.

When looking at small group dynamics Allington (1982) found that students who are at a lower reading level, have less opportunities to read. While this result may be due to time constraints in the classroom, it continues to inhibit the amount of material, concepts and skills these students encounter and test. Three studies in this section resulted in findings that suggest heterogeneous groups may be beneficial to lower ability students. Flemlee and Elder (1983) found that students with a high reading aptitude stay on task more frequently, and that students in small-groups have a higher level of attentiveness when the over-all group's level of reading includes higher fluency ability. Wilkinson and Anderson's (1995) results showed that discussion of a silently read book was more extensive than that of an orally read book, in small groups. It also concluded that higher fluency when reading orally, maintained higher student attentiveness. Anderson, Wilkinson, and Mason (1991) found similar results, showing that the group fluency, not individual fluency, created higher comprehension and better recall of the text being read. All

of the aforementioned studies concluded that the group's ability level was the larger dictator of the learning happening, then the individual's. While the results of McCoach, O'Connell, and Levitt's (2006) study found that homogeneous groups related to higher reading levels, the uncontrolled for variables, and limitations of the study, as discussed in Chapter 3, prevent these results from further comparison or transferability. They cannot be compared to the use of heterogeneous groups, not can the results provide further evidence concerning equity of students. The evidence from this section strongly suggests that a group with an overall higher ability will create a higher level of learning for the members of that group. The use of heterogeneous groups would create this situation. By placing higher level students with lower level students, higher comprehension, better recall, and student attentiveness could take place.

Schumm et al. (2000) found that most teachers use whole-class instruction, without differentiation or scaffolding for student needs, particularly those identified with an LD. They furthered that research to find out how students fared in these situations, and found that high performing students, already successful in the system were the only group being served. Silliman et al. (2000) found the same results when looking at small groups including LD or LLD students. Poole's (2008) results showed that the negative social stigmatism associated with homogeneous groups, was also found in heterogeneous groups. In heterogeneous groups the ability of the students were exposed, and led to a social hierarchy that went beyond the groups. The results of Grant and Rothenberg's (1981) study showed that students who are expected to achieve more, receive more encouragement to do so, while those who aren't expected to achieve as much, are not encouraged as much. Gamoran (1984) found that when homogeneous groups are used in the classroom, the more heavily the teacher emphasized the ranking, the more drastically it affected

the students in lower ranking groups. The same was not true for higher ranking students, who felt no difference, and remained stable. Lower and middle ranking students, however, suffered academically from the social stigma of the elitist treatment of the groups. In classrooms where an egalitarian treatment method was employed this drastic difference was not felt. These results show us that homogeneous grouping is not inherently detrimental, but is dependent on the structure and treatment of the groups in the classroom. Sutherland and Topping (1999) concluded that heterogeneous pairs, of one more-able student, and one less-able student, provided greater growth to the less able student. While the more-able student still grew academically, the less-able student made considerably more growth. In same-ability groups no growth was seen for either student. This finding lends confidence to the use of heterogeneous groups where high ability students grow from teaching, and low-ability students grow from peer interaction. Across this section the findings bring forth some striking results as to how student needs are being met. In cases where whole-class instruction is being used, no differentiation or scaffolding is taking place to include the needs of LD or LLD students. Classrooms where homogeneous groups are being used it is the employment of the teacher's strategies to either treat the groups in an egalitarian manner, or elitist manner, that creates additional stress, and reduces learning in low and middle leveled students. Findings also show that the ability level of the students will not be automatically hidden by using heterogeneous groups strategies. Heterogeneous groups were also shown to increase student learning, particularly that of lower ability students, when compared to homogeneous groups.

French (1991) found that teacher's placement of students into ability levels closely matched 3 of the 5 tests utilized, though the student level varied across all the tests. Moody &

Vaughn's (1997) results showed the perceptions of grouping felt by general and special education teachers. One of these results was frustration from general education teachers in dealing with outside control over the grouping method they used. Special education teachers however, did not feel this frustration, as administration had a hands-off stance in their classroom. They also found that general education teachers used predominately whole-class instruction, but also employed heterogeneous groups, while special education used homogeneous grouping. The results of Emmer (1984) found that heterogeneous classrooms where a high spread of ability was found employed 3 levels of homogeneous groups. A large number of other methods were also employed, and of those individualized instruction, team teaching, the use of resource rooms, additional adults in the classroom, and explicit directions and explanations all produced higher academic scores in students. Lower student scores were found in classrooms with higher frequencies of free-time. Rubie, Wilkinson, Parr and Townsend (2000) found that teachers in New Zealand nearly consistently used ability-based groups when teaching reading, and heterogeneous groups for writing. These teachers also preferred groups of 8 or less, but commonly had to deal with higher numbers, and used 4 to 5 groups in the classroom. Results also found what practices were used by excellent reading teachers, that standard teachers did not employ. One of these practices was a highly positive environment, where all students believed themselves to be a reader. Highly effective classroom management, including coordination with classroom assistants, special education teachers, and monitored reading was also found to create higher academic scores for students, according to Pressley et al. (2001). Paterson, Henry, O'Quin, Ceprano, and Blue (2003) found that a locally networked computer based instruction program, the Waterford Early Reading Program, failed to provide better education or make up

for inexperience with reading prior to starting school. Concluding results for teacher's practices showed that a wide range of methods have been employed, sometimes by the choice of the teacher, and sometimes mandated by administration choices. It also found that testing and computer based programs did no better at assessing or teaching students than the teachers.

The literature on student's attitudes toward reading found that attitude did not correlate with reading ability (Schooley, 1994). That is, students at all levels of reading ability enjoyed and disliked reading, and a lower reading ability did not predict less enjoyment in reading. Results also showed that when teachers treated all students like high-ability readers, the level of appreciation of reading raised. Lasswell (1967) found that student's enjoyment of reading, and their perception of themselves as a reader was disconnected to the ability group they were placed in. Instead results showed that students perceived of themselves to be in different levels of ability, than those they were placed in. The only exception to this was high-ability students, who accurately placed themselves as high-ability. Any differences in reading appreciation between reading ability levels disappeared by mid-year altogether. When looking at student perceptions of grouping strategies Elbaum, Moody, and Schumm (1999) found that no differences in preferences existed between LD and general education students. Both groups identified four conditions that should be controlled by the teacher to make the experience more positive for students; working conditions, teacher assistance, ridicule from peers, and pace of learning. The findings on student perceptions significantly point out that reading ability does not dictate student appreciation or enjoyment of reading. With the exception of high-ability readers, students also are not likely to not see their ability in the same light as their teachers.

Implications for Teaching

This literature review has researched what grouping strategies and methods are most effective at creating equity in learning literacy across ability levels. The implications for teaching from these findings are substantial. Small groups have been found to be favorable to whole class instruction, providing better learning. In some cases these findings were so significant that the grouping method didn't affect the results. That is a small group arrangement, whether it be homogeneous or heterogeneous didn't matter as much as the use of small groups.

This finding bears further credence when we look at the function of literacy as a method of communication. Being able to work with other students during literacy instruction, particularly with the use of discussion, is very important. As a reader, students need to be able to discuss what they have read in order to continue their meaning making from the text. The NCTE acknowledges these interactions saying that they “expand and strengthen [the readers] comprehension and interpretation” (1999, ¶ 3). Put simply, by having students talk about what they have read with other students, they will all become better at reading, taking more away from the text. Not only that, but they will also build better critical thinking skills. As the NCTE puts it, “in these interactions, we learn to read critically, to question what we read, and to respond in a certain way” (1999, ¶ 3). That is, without the interaction of other readers, as is provided in group work, students would fail to fully evaluate the meaning of a given text.

The method of assigning groups, either heterogeneously or homogeneously, has been argued in both directions for years. The findings in this literature may add some light to this discussion. When looking at large sample bases, the method of grouping made no significant difference to student success, only that small groups were preferred over whole-class (Lou et al.,

1996). In other studies, where smaller sample sizes were examined differences were found. Harris and Harrison (1988) found that whole-class instruction was favored over the previously used leveled, basal program. McCoach, O'Connell, and Levitt (2006) found that in the kindergarten classes they studied homogeneous reading groups were correlated with higher reading levels. Poole (2008) found that a negative social stigmatism and hierarchy was started in heterogeneous groups, based off student reading ability. Sutherland and Topping (1999) observed two classrooms, and found heterogeneous groups to be favorable for academic growth. And the list of contradictions goes on. However I believe if we look to the work of Grant and Rothenberg (1981), and Gamoran (1984) we may find out why.

Both Grant & Rothenberg (1981) and Gamoran (1984) studied the effect of grouping on student needs, particularly how teachers treated students. Grant & Rothenberg (1981) examined the social environments of classrooms where homogeneous groups were used. These classrooms were looked at in terms of white and blue collar communities. The results here showed that students who were expected to achieve more were treated as such, given more opportunities, encouragement, and support, while students not expected to achieve did not receive the same level of support. These even included incidences of poor treatment toward expected low-achievers, including additional negative criticism, and low expectations. Students expected to achieve were also afforded more autonomy in self-directed learning and leniency, including higher levels of visiting and agenda setting conversations allowed. Grant and Rothenberg (1981) see this as a result of the society, and social relationships, and not the result of one or more individuals, however the results are still striking. Whether it be from society, or the individuals

that make it up, a lower level of opportunity, encouragement, and expectation will keep even the best of readers from achieving their full potential. Gamoran (1981) found similar results.

In the study conducted by Gamoran (1984) classrooms where homogeneous grouping was regularly used were divided into two categories; egalitarian and elitist. In an egalitarian classroom the teacher makes a conscious effort to minimize the awareness of reading ability among students. Teachers in an elitist classroom do the opposite, bringing emphasis to the ability level of students and groups. Methods of hiding or minimizing the student or groups ability level include remarks and attitude toward students and parents, groups names, classroom arrangement and environment, and discreteness of aptitude and instruction. The amount of learning happening in those classrooms was found to be strikingly different. Classes where an elitist treatment of the students and groups was used, significantly brought down the reading levels of all students, except the high performing. The exact opposite was true in classes where ability levels of groups were hidden or down played. These classes saw the reading gap between low and high performing students shrink, bringing the low ability students up with the high ability students, and creating more academic growth for all. This study shows that homogeneous groups are not inherently detrimental to student learning, but are greatly influenced by teacher attitude.

Both of these studies may explain why such large differences are found across heterogeneous and homogeneous groups. Should a teacher and environment treat every student as though they have every bit of potential as the next, each student will achieve to the best of their ability. Pressley et al. (2001) found the same to be true when identifying traits of excellent reading teachers. Among other traits, these teachers maintained a strongly positive classroom, where every student had high self-confidence in reading. This finding closely matches the

learning theories of Zull (2002), in that emotions are tied up with learning. A student who feels lousy about themselves as a reader, will consequently achieve less. This also rings true when looking at issues of identity presented by Gee (2003), “Learning involves taking on and playing with identities”. That is a learner has to agree with the identity of the learning to acquire it. A learner who does not identify themselves as a good reader, will never become one. By creating an environment where all student believe in themselves as readers, and believe they are of high-ability in reading, students will achieve more.

Teachers may not always have a choice in the grouping method used in the class, but teachers do have a choice in how those groups are treated. By downplaying the ability levels of the groups and students, and creating a strong sense of self-belief, students need not be subjected to the possible detrimental effects of exposed reading levels. Other considerations for the classroom have also been found in this research.

LD and LLD students are being more commonly included in general education classrooms, and often times appropriate scaffolding for their needs are not being included in instruction. Schumm et al. (2000) found that in the majority of classes, including those with LD and LLD students, teachers used primarily whole-class instruction, but did not provide scaffolding or differentiation. Silliman et al. (2000) found the same result where LD and LLD students were concerned. Teachers did not employ differentiation or additional scaffolding for students, whether in homogeneous or heterogeneous small-groups. This method of teaching, providing instruction to one group, but not providing the necessary stepping stones to those who are less familiar with it goes against Vygotsky’s Zone of Proximal Development. These findings

suggest that teachers need to make sure they are providing the necessary scaffolding for students of all levels and needs.

Suggestions for Further Research

The research covered in this literature review has sought to identify the grouping strategies and methods that are most effective at creating equity in learning literacy across ability levels, and many important findings have been brought forth, and discussed. These findings have important implications for classroom practice. First that small-group work provides more effective learning, then whole-class instruction, however not when homogeneous groups are used, and ability levels of students and groups are treated in an egalitarian manner. Results also showed that teachers need to ensure that scaffolding and differentiation are taking place for all levels of students.

These results were concluded from the findings of five overarching threads of research; what is the most effective group size, how do small group dynamics effect student learning, how does grouping affect student needs, what are teacher's practices concerning instruction and organization, what are student perceptions about grouping. To continued this research, and better understand what grouping methods would most appropriately serve all students cooperative learning should be explored. This method of grouping asks students to be concerned not only with their own learning, but with all members of the group. As an additional grouping method, separate from traditional heterogeneous and homogeneous groups where each student is concerned with only their own learning, and turns in independent work, this method may provide for a higher level of equity in the classroom. It may also work towards creating an egalitarian

system where student's ability levels aren't what define them. Further research in this area would need to be done before drawing any conclusions.

Further studies in classrooms defined as egalitarian or elitist should also be done. While the study by Gamoran (1981) provided strong results, additional studies in the same area could provide beneficial evidence, and a deeper understanding of how to apply these methods to the classroom.

Conclusion

Chapter one presented the focus of this literature review; what grouping strategies and methods are most effective at creating equity in learning literacy across ability levels. It explained the rationale of this significant question, and the importance of reading on our society. Reading, and being highly literate can be a determining factor in how successful we are not only professionally and academically, but can also determine how we read situations, communicate, and are able to critically reason. Chapter one also explained how the ever growing demands of our society, and increasing technology continually require a higher and higher level of literacy. As this demand increase across our society, the gap between successful and struggling readers continues to grow. Chapter one also covers the possibilities of grouping in the classroom. When not used as a method for tracking, grouping can be homogeneous, heterogeneous, and vary in sizes. Group dynamics, and student perceptions should also be considered. These aspects are important when looking at the work of Zull (2002), and considering the connection of emotions and learning. Keeping students appropriately challenged must also be considered, as can be seen in Vygotsky's Zone of Proximal Development.

Chapter two discusses the origins of groups in school as beginning in the late 1800's, and continuing under different pretences to today. Until recently homogeneous grouping has been preferred, but more recently a negative view of this grouping method has driven teachers and administration to prefer heterogeneous groups, or even whole-class instruction. These preferences have come from proposed negative associations with homogeneous grouping. Namely, that it pushes high-level students higher, and lower students lower. This elitist treatment of students is said to have a lasting negative affect particularly on low-level students, both socially and academically. These grouping methods are commonly driven by leveled basal reading programs, sometimes required by administration in the school, and forcing teachers to employ a three-tier system of grouping.

In chapter three a critical review of the literature takes place. It is divided into 5 sections; Group Sizes, Small Group Dynamics, The Effect of Grouping on Student Needs, Teacher's Practices, and Student Perceptions. Under each section research is reviewed and findings summarized. When looking at group sizes results showed that small-groups fared better results for student learning than whole-class instruction. Group sizes, however, need not be minimized to a lower than 3 to 1 student-teacher ratio. When looking at small group dynamics the evidence suggested that students will retain more in a group that is of a higher level, than a group of a lower level. Creating heterogeneous groups could create this ideal situation, and build higher comprehension, better recall, and student attentiveness. When looking at student needs findings showed that differentiation and scaffolding for LD or LLD students is not taking place, whether in cases of whole-class or small-group instruction. It was also found that classrooms where homogeneous groups are being used it is the employment of the teacher's strategies to either

treat the groups in an egalitarian manner, or elitist manner, that creates additional stress, and reduces learning in low and middle leveled students. When heterogeneous groups are used in a classroom findings showed that the ability level of the students will not be automatically hidden. Heterogeneous groups were also shown to increase student learning, particularly that of lower ability students, when compared to homogeneous groups. When researching teacher's practices a wide range of methods was found to be employed, sometimes by the choice of the teacher, and sometimes mandated by administration choices. It was also found that testing and computer based programs did no better at assessing or teaching students then the teachers. Student perceptions of significantly pointed out that students of lower and middle level abilities did not perceive themselves in the same ability categories as they were placed. Results also proved that students of all levels can enjoy or dislike reading, and that ability is not a predictor of enjoyment.

Chapter four is the conclusion of this literature review. A summary of the findings for each section covered in chapter 3 was reviewed and evaluated. The implications of these findings in the classroom is also covered. Small groups should be used to teach literacy, preferably with the advantages of discussion. If grouping requires a homogeneous set up, the levels of the groups, and abilities of individual students should be down played, and hidden. All students should be treated as though they have the capability to reach the highest level of reading possible. All students should also be made to feel fully capable of reading, and be able to identify themselves as a reader. In the suggestions for further research a continued look at the implications of egalitarian and elitist groups should be explored. Continued work on cooperative learning in groups would also be beneficial.

The findings of this research have brought forth some important implications when looking at effective methods of creating group work for equity in the classroom. By employing the findings of the included research in the classroom a higher level of equity may be provided for all students of every level.

REFERENCES

- Allington, R. (1982, November). *Content coverage and contextual reading in reading groups*. Paper presented at the annual meeting of the National Council of Teachers of English, Washington, DC. (ERIC Document Reproduction Service No. ED228604).
- Anderson, R., Wilkinson, I., & Mason, J. (1991). A microanalysis of the small-group, guided reading lesson: Effects of an emphasis on global story meaning. *Reading Research Quarterly, 26*(4), 417-441.
- Easton, J., Muirhead, R., Fredrick, W., & Vanderwicken, S. (1979, April). *Relationship among student time on task, Orientation of teachers, and instructional grouping in elementary reading classes*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service No. ED169503).
- Elbaum, B., Moody, S., & Schumm, J. (1999). Mixed-ability grouping for reading: What students think. *Learning Disabilities Research & Practice, 14*(1), 61-66.
- Emmer, E., (1984). *Management and instruction strategies for heterogeneous elementary school classrooms*. (R&D Report No. 6009). Washington, DC, National Inst. Of Education. Austin: Texas University, Research and Development Center for Teacher Education (ERIC Document Reproduction Service No. ED251431).
- Felmlee, D., & Eder, D. (1983). Contextual effects in the classroom: The impact of ability groups on student attention. *Sociology of Education, 56*(2), 77-87.

- French, M. (1991). Placing students in instructional reading groups: A comparative investigation of five assessment techniques. *Ohio Reading Teacher*, 25(2), 7-16.
- Gamoran, A. (1984, April) *Egalitarian versus elitist use of ability grouping*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED245821).
- Gee, J. (2003) *What video games have to teach us about learning and literacy*. New York, NY: Palgrave MacMillan.
- Grant, L., & Rothenberg, J. (1981, April). *Charting educational futures: Interaction patterns in first and second grade reading groups*. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, CA. (ERIC Document Reproduction Service No. ED200902).
- Hallinan, M., & Sorensen, A., (1985). Ability grouping and student friendships. *American Educational Research Journal*, 22(4), 485-499.
- Harris, B., & Harrison, G. (1988). *A naturalistic study on whole class instruction in first grade reading*. Paper presented at the California Educational Research Association, San Diego, CA. (ERIC Document Reproduction Service No. ED311422).
- Helf, S., Cooke, N., & Flowers, C., (2009). Effects of two grouping conditions on students who are at risk for reading failure. *Preventing School Failure*, 53(2), 113-128.
- Huetra, G., (2009). *Educational foundations: Diverse histories, diverse perspectives*. Boston: Houghton Mifflin.

- Lasswell, A. (1967). *Reading group placement – Its influence on enjoyment of reading and perception of self as a reader*. American Educational Research Association, Washington, DC (ERIC Document Reproduction Service No. ED011816).
- Lou, Y., Abrami, P., Spence, J., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, 66(4), 423-458.
- McCoach, D., O'Connell, A., & Levitt, H. (2006). Ability grouping across kindergarten using an early childhood longitudinal study. *The Journal of Educational Research*, 99(6), 339-346.
- Moody, S., & Vaughn, S. (1997). Instructional grouping for reading. *Remedial and Special Education*, 18(6), 347-357.
- National Council of Teachers of English (1999). *NCTE Position Statement on Reading*. Retrieved from <http://www.ncte.org/positions/statements/positiononreading>.
- National Council of Teachers of English (2004). *A Call to Action: What we know about adolescent literacy and ways to support teachers in meeting students' needs*. Retrieved from <http://www.ncte.org/positions/statements/adolescentliteracy>.
- Paterson, W., Henry, J., O'Quin, K., Ceprano, M., & Blue, E., (2003). Investigating the effectiveness of an integrated learning system on early emergent readers. *Reading Research Quarterly* 38(2), 172-207.
- Poole, D. (2008). Interactional differentiation in the mixed-ability group: A situated view of two struggling readers. *Reading Research Quarterly*, 43(3), 228-250.
- Pressley, M., Wharton-McDonald, R., Allington, R., Block, C., Morrow, L., Tracey, D., Baker, K., Brooks, G., Cronin, J., Nelson, E., & Woo, D., (2001). A study of effective first-grade

- literacy instruction. *Scientific Studies of Reading*, 5(1), 35-58. Notre Dame, IN: Dept. of Psychology, University of Notre Dame.
- Rogoff, B. (2003). *The cultural nature of human development*. New York: Oxford University Press, USA.
- Rubie, C., Wilkinson, I., Parr, J., & Townsend, M. (2000). *Sizing up and organizing for instruction in the first weeks of school*. Paper presented at the International Reading Association World Congress on Reading (ERIC Document Accession No. ED448429).
- Schooley, F. (1994). *Within class ability grouping and its effect on third grade attitudes toward reading*. (Tech. Report) University of Virginia, The Curry School of Education.
- Schumm, J., Moody, S., & Vaughn, S. (2000). Grouping for reading instruction: Does one size fit all? *Journal of Learning Disabilities*, 33(5), 477-488.
- Silliman, E., Bahr, R., Beasman, J., & Wilkinson, L., (2000). Scaffolds for learning to read in an inclusion classroom. *Language, Speech, and Hearing Services in Schools*, 31, 265-279.
- Smith, F. (1985). *Reading without nonsense*, (2nd ed.). New York: Teachers College Press.
- Sutherland, J., & Topping, K., (1999). Collaborative creative writing in eight-year-olds: Comparing cross-ability fixed role and same-ability reciprocal role pairing. *Journal of Research in Reading*, 22(2), 154-179.
- Vaughn, S., Linan-Thompson, S., Kouzekanani, K., Bryant, D., Dickson, S., & Blozis, S., (2003). Reading instruction grouping for students with reading difficulties. *Remedial and Special Education*, 24(5), 301-315.
- Wells, G. (2000). Dialogic inquiry in education: Building on the Legacy of Vygotsky. In C. Lee & P. Smagorinsky (Eds.), *Vygotskian perspectives on literary research: Constructing*

meaning through collaborative inquiry (pp. 51-85). New York: Cambridge University Press.

Wilkinson, I., & Anderson, R. (1995). Sociocognitive processes in guided silent reading: A microanalysis of small-group lessons. *Reading Research Quarterly*, 30(4), 710-740.

Zull, J. (2002). *The Art of Changing the Brain*. Sterling, VA: Stylus Publishing.