Multi-Grade Versus Single-Grade Classrooms: Student Performance on the Iowa Assessments

Ja-Ronika D. Veldheer
B.S., University of Missouri, 2007
M.S., University of Missouri, 2008

Submitted to the Graduate Department and Faculty of the School of Education of Baker University in partial fulfillment of the requirements for the degree of Doctor of Education in Educational Leadership

Susan Rogers
Susan K. Rogers, Ph.D.
Major Advisor

Sharon Zoellner
Sharon Zoellner, Ph.D.

Sister Elizabeth Youngs
Sister Elizabeth Youngs, Ph.D.

Date Defended: April 26, 2019

Abstract

Creating the best instructional practice for students is the goal of all educators. The first purpose of this study was to determine whether there is a difference in the Iowa Assessments in reading, written expression, mathematics, social studies, and science scores between third through eighth-grade students enrolled in multi-grade level classrooms and third through eighth-grade students enrolled in single-grade classrooms. The second purpose was to determine whether the difference in the Iowa Assessments reading, written expression, mathematics, social studies, and science scores between third through eighth-grade students enrolled in a multi-grade level classroom and third through eighth-grade students enrolled in a single-grade classroom is affected by student grade level. The independent variables used in this study included the classroom configurations in the schools and the grade levels. The dependent variables in this study were archived student achievement scores in reading, written expression, mathematics, social studies, and science on the Iowa Assessments from school years 2016-2017 and 2017-2018. The average Iowa Assessments reading, mathematics, social studies, and science scores for students enrolled in a multi-grade classroom were not different from the scores of students enrolled in a single-grade classroom. However, the average Iowa Assessments written expression score for students enrolled in a multi-grade classroom is lower than the average Iowa Assessments written expression score for students enrolled in a single-grade classroom. When Iowa Assessments scores were compared at each grade level, no significant differences were found between students in multi-grade and single-grade classrooms. This study has implications for district administrators, building administrators, teachers, and parents. As the results indicated that students are equally

successful in either classroom configuration, this study has offered insight into the feasibility of using multi-grade classrooms. Recommendations for future research include conducting the study over a multi-year period to monitor the full impact of the same group of students who advance from grade 3 to grade 8 to track their achievement. Another recommendation for future research would include public schools as a comparison group in the study.

Dedication

This work is dedicated to my family. First, to my husband, Tommy, who without his love, support, and continuous encouragement over many years, this doctoral work would have never happened. Next, to my daughter Brea who showed me daily that her love and support of me during this process was unwavering. Your sacrifice has not gone unnoticed, and for that I am grateful.

Also, this work is dedicated to all of my past and present students that I taught over many years. I have always wanted to be an educator and further my love of learning. I thank you for allowing me to continue to do what I love.

Acknowledgments

I knew as a little girl that I wanted to reach the highest milestone in my educational journey, and I cannot believe that I have finally reached that goal. I am forever grateful to all who listened to my doubts and encouraged me to reach my lifelong goal. It is with deep appreciation that I thank all the individuals who have helped me through this process. Without these committed and caring people, this work would not have been possible.

I want to thank God for allowing me to wake up every day and experience this life journey. You knew I would make it through even when I did not. Thank you for your unconditional love.

To my husband, Thomas, who has never faltered in his support of my completion of this degree. His statement "get it done babe" echoed in my ear every time I doubted myself and wanted to quit. I truly value his constant support and encouraging words. I feel so honored to be your wife, and I know that you are so proud of everything I do. I am so thankful that you put your goals on hold and allowed me to reach this goal and never doubted me. It took a lot longer than we hoped, and you have witnessed my lowest and highest points in this journey, but I could not imagine having anyone else to walk with me through this crazy life. I love you!

To my daughter Brea, mommy finally did it! I started this journey when you were three and watched you grow up into a beautiful girl while completing my doctorate. I know that you did not understand why mommy had school at night, was not home in the evenings, and missed school activities, but as you got older, the more you understood that I was doing

this for you. I hope that you see my dedication and hard work as a model for your life. Let this achievement be proof that you can accomplish any goal you set in your life.

To my family, you have been a source of strength and support even before this process began. You were my cheerleaders in my decision to begin this monumental accomplishment. Dad, mama, and Du-Wan, thank you for not talking me out of this decision when I shared with you that I wanted to earn my doctorate. You have always encouraged me every step of the way. It was my goal for my grandpa to see me graduate and he will be there in the front row!

To my girl tribe, Stephanie, Desiree, Erin, Joanna, and Lexi you girls are my rock! You called me crazy at first, but without you, the venting texts, long phone calls, and girls' trips, I would not have made it. When I wanted to quit, each one of you told me that I had gone this far and cannot stop now. You helped me escape my writing when I needed a break even when I should have been writing, but overall, I am so thankful for my girl tribe! Desiree, I want to thank you especially because you were the one that gave me the idea for this research study and your connections allowed me to reach my goal. I value your professional experience as an outstanding educator, and I am so thankful for our friendship.

I want to thank Sister Elizabeth Youngs and the Diocese that allowed this meaningful research to take place. Sister Youngs, your support and encouragement have meant the world to me. I would also like to thank you for joining the dissertation committee and providing insight, information, and encouragement along the way. Thank you for allowing me to study your Diocese, and I hope that the results of my research are able to aid your current and future students.

I would like to thank Dr. Peg Waterman for her incredible calm brilliance with data analysis and your encouraging posture throughout this process. Dr. Zoellner, your feedback and guidance to assist in the refinement of this study were immeasurable. To my major advisor, Dr. Susan Rogers, I owe you a million thanks for your support and encouragement. I needed you, and you came at just the right time in my life. Your gentle nudges pushed me out of my comfort zone at times, but you knew I needed that. I am so thankful for the friendship that has grown during this journey, and there are not enough caramels in the world to show how much I appreciate you! Thank you for being you and showing me that I can complete my goal of being a doctoral graduate!

To my Baker classmates, thank you for the laughs, encouragement, and support along this journey. Marlaine, thank you for talking me into starting this journey, and we have finally reached the end together. Thank you for being my sounding board and really understanding how I feel because you were right there with me.

Finally, I would like to thank the hundreds of students I have had the privilege of serving over my 12 years in education. Each one of you has brought me perspective and new-found wisdom. You have made this work meaningful and relevant every step of the way, and for that, I am forever indebted.

Table of Contents

List of Tables

Chapter 1

Introduction

All parents want the best learning environment for their children. Teaching strategies are used to engage students in their learning and as a way to promote student achievement. In our current educational environment, schools place students in the grade level appropriate to their chronological age (Carter, 2005). Factors such as enrollment size, school resources, state funding, and personnel can determine how a school designs the grade configuration or structures the educational setting (Carter, 2005). "Thus, the multi-grade classroom still holds a significant place in schools, especially in small isolated rural districts" (Miller, 1990, p. 1). There are differing viewpoints on the effectiveness of multi-grade classrooms. Elkind (1993) purported that multi-age grouping is more natural and educationally beneficial than the rigid single-age grouping that dominates our schools. In a single-age classroom, students do not have the flexibility to progress naturally through social and cognitive development that students in a multi-grade configuration have (Carter, 2005; Song, Spradlin, & Plucker, 2009). "Proponents of multi-age education report that students benefit academically, emotionally, and socially by being in a mixed-age classroom" (Proehl, Douglas, Elias, Johnson, & Westsmith, 2013, p. 422). However, according to Miller (1990) "many teachers, administrators and parents continue to wonder whether the multigrade organization has negative effects on student performance" (p. 2).

Background

The schools involved in the study are from a private, parochial Catholic diocese in a Midwestern state. The schools are in rural, suburban, and urban areas within the state.

The schools have the same curriculum and use the same standardized testing platform. The multi-grade schools have combined two consecutive grade levels and assigned one teacher to the combination classroom. Superintendent of Schools Diocese X (personal communication, June 13, 2018) stated:

> Schools that have multi-grade classrooms do so because of economic/resource reasons. These are very small schools and single grade classrooms are not an option. Principals have no choice in determining if their school will be a single or multi-grade school. If enrollment increases or decreases in a school, there will be conversations with the pastor and advisory groups. The principal would make the case for classroom configuration changes. Parents who enroll their children in these schools rarely comment on the options that they have or don't have. The level of centralization is low with our system.

Table 1 includes the 13 Diocesan schools that employed a multi-grade classroom configuration. Additionally, the number of students included in each grade configuration is presented. Enrollment information for the 2016-2017 and 2017-2018 school years is shown. From the data below, on average student enrollment is higher in the younger grades and gradually decreases as the students advance through the grade levels.

Table 1

Enrollment in Multi-Grade Schools (2016-2017 and 2017-2018 School Years)

	Grade					
	3rd-4th		5th-6th		7th-8th	
School	Y1	Y2	Y1	Y2	Y1	Y2
A	15	9	0	6	0	0
B	10	10	5	5	9	9
C	11	11	19	19	9	9
D	14	15	17	20	11	11
E	12	15	8	6	13	12
F	10	9	12	12	11	12
G	15	15	9	12	13	7
H	17	17	9	9	8	8
I	14	10	8	8	4	4
J	9	7	10	7	9	8
K	17	19	10	13	9	6
L	20	15	18	17	16	17
M	12	12	13	10	6	6

Note. Y1 = 2016-2017, Y2 = 2017-2018. Adapted from *Multi-Grade Schools*, by Diocese X, 2018a.

Table 2 includes the 24 Diocesan schools that used a single-grade classroom configuration. Additionally, the number of students included in each grade configuration is presented. Enrollment information for both 2016-2017 and 2017-2018 is shown. As the data in Table 1 and Table 2 indicate, there are more schools in the Diocese that use a single-grade configuration than multi-grade configuration.

Table 2

Enrollment in Single-Grade Schools (2016-2017 and 2017-2018 School Years)

	Grade											
	3rd		4th		5th		6th		7th		8th	
School	Y1	Y2	Y1	Y2	Y1	Y2	Y1	Y2	Y1	Y2	Y1	Y2
N	24	23	16	24	25	16	20	26	25	21	20	25
O	71	75	75	73	75	73	72	73	72	71	70	71
P	11	17	11	10	9	10	12	10	13	13	10	12
Q	24	18	13	23	14	14	23	14	22	24	10	22
R	13	19	12	14	14	13	11	13	14	11	8	14
S	26	38	33	27	36	30	37	34	31	34	30	31
T	46	45	53	50	39	50	55	37	40	54	42	36
U	46	58	64	47	40	64	52	41	52	49	40	53
V	15	22	12	13	17	12	12	17	15	11	19	15
W	17	15	9	17	13	9	19	11	11	19	13	11
X	12	6	9	12	4	8	13	4	4	11	3	4
Y	10	4	6	9	9	6	5	5	5	5	3	4
Z	23	20	23	20	17	20	12	14	11	10	14	9
AA	25	11	14	21	18	15	14	14	15	13	15	13
BB	10	17	15	9	11	16	14	11	15	13	9	14
CC	17	20	13	17	17	11	11	17	16	12	21	17
DD	24	23	19	21	22	19	16	18	16	14	19	17
EE	19	23	20	19	19	21	25	20	23	25	20	23
FF	15	14	11	12	9	11	11	8	9	9	16	8
GG	24	23	22	27	31	20	26	24	25	28	30	25
HH	17	17	22	22	13	13	20	20	26	26	12	12
II	13	15	11	14	14	11	15	11	8	16	12	9
JJ	26	26	14	14	21	21	15	15	8	24	12	16
KK	17	24	24	17	32	24	24	34	29	25	25	29

Note. Y1 = 2016-2017, Y2 = 2017-2018. Adapted from *Single-Grade Schools*, by Diocese X, 2018b.

Superintendent of Schools Diocese X (personal communication, June 18, 2018) stated:

> The Iowa Assessment is administered every fall in October, and all schools in the Diocese use the same testing window. The assessment is administered through paper and pencil, and schools must create schedules for students in grades 3-8 to complete the assessment. The Diocese chose the Iowa Assessment when the standardized assessment instrument they had been using became outdated and was no longer supported by the publisher. The school office (which consisted of a superintendent and an associate superintendent) researched available products and interviewed two providers. One was a broker who had several options, and the other was HMH Riverside, the publisher of the Iowa Assessments. The superintendents felt that the Iowa Assessments offered the highest quality product and the best support for schools.

Statement of the Problem

"One of the core managerial decisions made at the level of the school district is how to organize students into schools" (Jacob & Rockoff, 2011, p. 12). Educators in rural areas often think of ways to teach all students with limited resources and staff. One of the solutions has been to combine students in a manner they see fit with their resources. This educational grouping is different from a traditional single-grade setting. The question all educators ask is whether their approach will influence student performance. Advocates for multi-grade classroom instruction believe this method allows the teacher a chance to build relationships with students over a longer time and allows for more cooperative instruction. "Multi-grade grouping enables a teacher to use

the knowledge she or he has gained about a child during the first year to plan learning experiences for the next year" (Elkind, 1993, p.11). Determining the best classroom configuration for schools to maintain student achievement is a problem many educational leaders face. The importance of determining this will help school leaders decide what best meets the needs of students.

Purpose of the Study

The purpose of this study was to determine if a multi-grade educational setting was beneficial for student achievement. More specifically the first purpose was to determine whether there is a difference between the Iowa Assessments reading, written expression, mathematics, social studies, and science scores of grades 3-8 students enrolled in multi-grade level classrooms and grade 3-8 students enrolled in single-grade classrooms. The second purpose was to determine whether the difference between the Iowa Assessments reading, written expression, mathematics, social studies, and science scores of students enrolled in multi-grade level classrooms and third through eighth-grade students enrolled in single-grade classrooms was affected by student grade level.

Significance of the Study

The significance of this study is the contribution to the field of education for possible improvements to the classroom configuration of schools. This study included schools that are from the same diocese but make individual decisions within their single-grade and multi-grade schools. "There is very little comparison among the school communities. Teachers are somewhat mobile and, in some cases, move from one school to another because of a choice in the classroom configuration" (Superintendent of Diocese X Schools, personal communication, June 13, 2018). Studies have been

conducted like the current study; however, the schools in this study were private Catholic schools. Few studies have been conducted to determine which teaching configuration is best for student achievement (Mulryan-Kyne, 2004). Ricard, Miller, and Heffer (1995) recommended that future research be conducted to include students who are non-impaired and non-referred in mixed-grade classrooms. The current study includes students who are non-impaired and non-referred in mixed-grade classrooms. The current study might add to the body of literature related to combination classrooms. The results of the current study may also help the diocesan schools determine what class configuration would be beneficial for student achievement. MacDonald and Wurster (1974) recommended selecting similar populations as the participants and conduct some further studies of achievement in all cognitive areas, including reading.

Gorrell (1998) suggested conducting the study over a longer period of time and with larger sample groups, possibly using data from several schools that have both single and multi-grade classes. The current study responds to Gorrell's suggestion by including schools from one diocese that has multiple schools that employ either a traditional or multi-grade configuration. Additionally, two years of data were analyzed.

Delimitations

Delimitations have been defined as "self-imposed boundaries set by the researcher on the purpose and scope of the study" (Lunenburg & Irby, 2008, p. 134). For this study, the researcher established the following delimitations:

1. The schools selected for this study were Catholic schools.
2. The grade levels selected for this study were third through eighth grades.

3. The achievement data for this study included reading, written expression, mathematics, social studies, and science scores collected from the Iowa Assessments for school years 2016-2017 and 2017-2018.

Assumptions

Lunenburg and Irby (2008) defined assumptions as the "postulates, premises, and propositions that are accepted as operational for purposes of the research. Assumptions include the nature, analysis, and interpretation of the data" (p. 135). The following assumptions were made in this study:

1. The participants completed the Iowa Assessments to the best of their ability.
2. The participants in the study understood the vocabulary and concepts associated with the Iowa Assessments.
3. The participants were taught the content that was assessed by the Iowa Assessments.
4. The data collected measured the knowledge and skills of the participants in the subjects of reading, written expression, mathematics, and social studies.

Research Questions

Creswell (2009) noted that research questions (RQs) should "shape and specifically focus the purpose of the study" (p. 132). The following research questions were used for this study:

RQ1. To what extent is there a difference between the Iowa Assessments reading scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

RQ2. To what extent is the difference between the Iowa Assessments reading scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level?

RQ3. To what extent is there a difference between the Iowa Assessments written expression scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

RQ4. To what extent is the difference between the Iowa Assessments written expression scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level?

RQ5. To what extent is there a difference between the Iowa Assessments mathematics scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

RQ6. To what extent is the difference between the Iowa Assessments mathematics scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level?

RQ7. To what extent is there a difference between the Iowa Assessments social studies scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

RQ8. To what extent is the difference between the Iowa Assessments social studies scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level?

RQ9. To what extent is there a difference between the Iowa Assessments science scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

RQ10. To what extent is the difference between the Iowa Assessments science scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level

Definition of Terms

A definition of specific terms is provided so the reader will understand the word or phrase in the context of the study. Creswell (2009) stated that a term should be defined "if there is any likelihood that readers will not know its meaning" (p. 39). The following terms are defined for this study:

Iowa Assessments. According to the University of Iowa (2018), "the Iowa Assessments measure student achievement and growth from kindergarten through grade twelve as designed by the Iowa core" (para. 1). The subjects tested are reading, written expression, mathematics, social studies, and science.

Multi-grade classroom. Veenman (1995) defined a multi-grade classroom as, "One where the same instructor teaches students from two or more grades at that same time primarily as an administrative device to consolidate school staffing, thus addressing declining enrollments or uneven class size" (p. 319).

Multi-age classroom. Stone (1996) indicated that a multi-age classroom is a mixed-age group of children who stay with the same teacher for several years. The children are randomly selected and balanced by age, ability, and gender. "This grouping

is deliberately made for the benefit of children, not for reasons of economics, curriculum, or convenience" (Stone, 1996, p. vii).

Nongraded education. Gaustad (1994) defined nongraded education as, "the practice of teaching children of different ages and ability levels together in the same classroom, without dividing them or the curriculum into steps labeled by 'grade' designations" (p. 2).

Single-grade classroom. Mason and Stimson (1996) defined a single-grade classroom as with "one teacher [who] specializes in the curriculum of one grade level" (p. 441).

Single-age classroom. Bailey, Werth, Allen, and Sutherland (2016) defined a single-age classroom as being when "students are grouped by grades based on the chronological age of the student and moved through the system in an assembly line fashion" (p. 240).

Organization of the Study

Five chapters are contained in this study: introduction, review of literature, methods, results, and interpretations and recommendations. Chapter 1 included the background information, statement of the problem, the purpose of the study, and the significance of the study. Also, included in Chapter 1 were the delimitations, assumptions, research questions, and the definition of terms. Chapter 2 includes the history of classroom configurations, pros and cons of multi-grade instruction, teacher and parent perceptions of multi-grade classrooms, and student achievement differences between multi-grade and single-grade classrooms. Presented in Chapter 3 are the specifics of the methodology used in the study. The results of the analyses are reported

in Chapter 4. Chapter 5 includes a study summary, findings related to the literature, and conclusions.

Chapter 2

Review of the Literature

Lunenburg and Irby (2008) defined the review of literature as, "The basic rationale for your research from which will emerge the statement of the problem, research questions or hypotheses, and design of your study" (p. 137). The review of literature creates a framework for this study and allows for further research. This chapter includes a review of the history of classroom configurations, the pros and cons of multi-grade instruction, perceptions of multi-grade classrooms, and student achievement differences between multi-grade and single-grade classrooms.

History of Classroom Configurations

The multi-grade classroom has been a vital part of education since formalized education began. During the 1500s, receiving an education was largely restricted to privileged classes being taught by tutors or teachers in an individualized setting. Public education in the United States was introduced in the early 1800s in order to teach large populations of children skills training most efficiently. Economics helped the multi-age configuration, as small agricultural communities required the aid of children during planting and harvest. Furthermore, teachers and facilities were scarce and expensive; formal education was not generally perceived as necessary for day-to-day existence (Kolstad & McFadden, 1998). Initially, the one-room schoolhouse was the type of configuration in which an education was received. Minimal student enrollment necessitated that arrangement (Kolstad & McFadden, 1998). The graded concept of education was introduced by Horace Mann in the 20th century; graded instruction was considered the normal way of school by the latter half of the 1800s (Yarborough &

Johnson, 2000). The increase in student population caused the shift from multi-grade classrooms to single-grade classrooms in the education field. Single-grade classrooms have increased as a result of increased teacher population; however, the concept of multi-grade classrooms is a model that is still used.

Grouping by grade level originated because it was assumed that it would enhance learning and make teaching easier (Ansah, 1989). "In 1918, there were 196,037 one-room schools, representing 70.8 percent of all public schools in the United States" (Miller, 1991, p. 1). The 1960s saw the rise of considerable interest in the non-graded movement advocated by Goodlad and Anderson (1963). Teaching children of different ages and ability levels in one classroom is called nongraded education. "In the 1970s, that trend continued, even grew, until the early '80s, when there was a decline of momentum in the development of non-graded schools" (Yarborough & Johnson, 2000, p. 42). Less than 1,000 multi-grade classrooms remained by 1980.

"The multiage education philosophies have been supported by much of the historical research and adopted by many schools all over the world" (Song et al. 2009, p. 1). The idea of grouping students based on Vygotsky's (1978) zone of proximal development was discussed by Chapman (1995). Vygotsky (1978) stated that each student has an "attained" developmental level. The competency zone is when a student feels comfortable and can do what is required in a learning situation. Chapman (1995) reiterated:

> Advocates of multiage classes suggest that teachers can apply Vygotsky's theory by capitalizing on a wider age and ability range than we would normally find in

> single-grade classrooms, so that older or abler students can act as mentors of younger or less able children. (p. 416)

Bailey et al. (2016) shared that the foundation of a multi-age classroom can be built upon the framework of theorists Piaget (1959), Bandura (1977), and Vygotsky (1978). "These theorists believed that the environment to which students were subjected impacted their academic and social development" (Vygotsky, 1978, p. 240). The multi-grade configuration is in direct accordance with John Dewey's goal of "child-centered learning" (Trusty & Beckenstein, 1996).

Approaching the 21st century, the need for change became apparent when educators tried meeting the needs of our diverse population (Britt, 1997). Nongraded primary education was one response to these needs (Gaustad, 1994). "This structure can be reminiscent of the one-room schoolhouse, which provided for a multiage approach to education" (Britt, 1997, p. 3).

Some schools use a multi-grade configuration because it is appropriate for school enrollment. These schools also use a multi-grade configuration because of the availability of teachers and for administrative purposes. A multi-grade approach is still prevalent in schools, especially in small isolated rural districts and parochial schools (Miller, 1990).

Benefits, Advantages, and Strengths of Multi-Grade Classrooms

Stephney (1970) acknowledged that multi-grouping was an innovation that could be tried by today's educators to meet the needs of today's children. He also believed this arrangement enhanced friendships, personal and social adjustments, and self-confidence. Those social connections with peers enhance learning and peer relations. Gaustad

(1994), Kolstad and McFadden (1998), Pavan (1992), and Stephney (1970) supported multi-grade instruction. These researchers purport that multi-grade instruction is beneficial for students.

Pavan (1992) conducted a meta-analysis of 64 previous studies of nongraded or similar classroom configurations. Pavan (1992) reported, "On mental health and school attitudes, 52 percent of the studies indicated nongraded schools as better for students" (p. 23). The results of the meta-analysis indicated that nongraded schools were favored by administrators and teachers over graded schools. Pavan (1992) found that "In most cases, students in schools organized in a nongraded style do as well or better than students in traditional self-contained classes in terms of both academic achievement and mental health measures" (p. 7). Additionally, the multi-grade grouping could allow for a broader pool of candidates from which to develop friendships for students who, lacking social confidence, do not find friendships within their own age group. Pavan (1992) stated, "We now know that the most natural learning environment for children calls for heterogeneous multi-age groupings, within which all sorts of homogeneous and heterogeneous subgroupings can be created as needed" (p. 36).

Many studies have been conducted on classroom configurations with mixed results. Gaustad (1994) stated that "despite inconclusive studies, however, research evidence generally supports the effectiveness of nongraded programs" (p. 2). Gaustad (1994) supported nongraded programs by making the connection that these programs are setting students up for real-world experiences. "America's population becomes more heterogeneous with each passing year. Employers seek employees with problem-solving

and teamwork skills, technical knowledge, and, above all, the ability to adapt and learn as changing conditions make old information obsolete" (Gaustad, 1994, p. 2).

Teachers also benefit from multiyear classes. Gaustad (1994) shared that "Multiyear classes also enable teachers to accumulate knowledge of learners for several years, instead of starting anew with all their students each fall, and to build stronger bonds with parents" (p. 17). Furthermore, Gaustad (1994) stated that "Research has found nongraded programs to be at least as successful academically as graded programs, even when assessed by means of the standardized tests to which graded programs are geared, and superior socially and emotionally" (p. 2). Gaustad (1994) supported Pavan (1992) that "the nongraded groups performed better than (58 percent) or the same as (33 percent) the graded groups on measures of academic achievement" (p. 23).

Grant, Johnson, and Richardson (1996) discussed how to implement a multi-age program in a district or school system. In their notebook, they shared their best advice on ways to make a successful transition. They noted that grouping of children as individuals working alone, or in small groups, requires students to take responsibility for their own learning and their own academic success. In addition to academic advances, students in multi-grade classrooms serve as peer-teachers or peer tutors, which gives them a sense of intrinsic pride and success when empowered to act as a teacher for another student or group of students. Grant et al. (1996) said multi-age classroom configurations help students' personal development. Children who develop slower can use the continuous progress classroom to help them adapt and develop at their own pace with extra time.

Kolstad and McFadden (1998) took a critical look at research findings related to social development in the multi-age classroom concept and delineated the pros and cons

of such a teaching strategy. Educators reported advantages in the areas of academia, self-esteem, and socialization skills. Within the multi-age classroom structure, students often work alone or in large-group or small-group settings. Kolstad and McFadden (1998) stated that research had shown overwhelmingly and uniformly that students make considerable advancement in social and affective skills in the multi-age classroom. They went on to state that teachers who enter willingly into this endeavor are more likely to be successful in the program. Kolstad and McFadden (1998) discussed that individual teachers should be able to decide if they want to partake in a multi-age classroom setting. Kolstad and McFadden (1998) indicated that teachers who find themselves teaching in a multi-age classroom need to observe and collaborate with experienced teachers. There is a need for additional planning and preparation time for instruction; curriculum must be designed and aligned to fit the different educational needs and levels which will exist in a multi-age classroom. Administrators of such programs must be prepared to supply teachers with basic supplies and materials which might mean a reallocation of funds.

Students in multi-grade classrooms demonstrate increased positive attitudes toward school, greater leadership skills, greater self-esteem, and increased pro-social and fewer aggressive behaviors, compared to peers in traditional graded classrooms (McClellan & Kinsey, 1997; Veenman, 1995). Nye, Cain, Zaharias, Tollett, and Fulton (1995) discussed the positive social and cognitive outcomes of non-graded programs. Trusty and Beckenstein (1996) with Veenman (1995), McClellan and Kinsey (1997), and Nye et al. (1995) concluded that there is no significant difference in the cognitive development of elementary students between the two classroom configurations.

Veenman (1995) conducted a meta-analysis and discussed the reason for the review, which was "to highlight the reasons for the increasing number of multi-grade classes, the concerns and perceptions about multi-grade classes, and the claimed advantages of multi-age grouping" (p. 320). He analyzed 56 studies dating from 1949 to 1995. Veenman (1995) shared that advocates of multi-age grouping claim there are cognitive and noncognitive benefits of this grouping. Stronger relationships, individualized instruction, secure teacher-student relationships, increased cooperation, and prosocial behaviors are some of the positive outcomes of multi-age grouping. Opportunities to learn from each other, modeling, and leadership are other positive multi-age outcomes.

Nye et al. (1995) conducted a longitudinal school success study over six years. The purpose of this research was to identify successful school practices that may exist in both non-graded and graded programs. Programs, where students are not categorized in one specific grade, are called non-graded programs. The study was conducted to determine if non-graded programs have cognitive and social benefits for elementary-school students and professional-practice benefits for teachers and school administrators. Fifteen hundred elementary students in grades K-4 from seven Tennessee schools were in the non-graded group, and five comparison schools where all students were in single-grade classes totaled 2,250 students were included in the study. The Tennessee Comprehensive Assessment Program and the Tennessee Holistic Writing Assessment were used to test the outcome measures of academic achievement. The researchers concluded that students from non-graded classes significantly outscored those from traditional classes.

McClellan and Kinsey (1997) conducted a study to test children's prosocial, aggressive, and friendship behaviors comparing children in a mixed-age classroom versus a same-age classroom. The study was conducted at two suburban middle-class elementary schools in greater Chicago and two Milwaukee schools in the inner city. The study participants included 312 mixed-age classroom students and 325 same-age classroom students. Twenty-nine classroom teachers rated the students on their prosocial and friend behavior using a rating scale. Based on the results of a teacher survey, McClellan and Kinsey (1997) concluded, "Mixed-age grouping is highly significant in predicting increased levels of children's prosocial behavior and friendship behavior" (p. 12). McClellan and Kinsey (1997) also found that "mixed-age grouped classrooms are also highly significant in predicting reduced levels of negative and aggressive behaviors among classroom children" (p. 12). Based on teacher survey results, McClellan and Kinsey (1997) further concluded that "Mixed-age grouping is highly significant in predicting increased levels of children's prosocial behavior and friendship behavior" (p. 12). McClellan and Kinsey (1997) also found that "mixed-age grouped classrooms are also highly significant in predicting reduced levels of negative and aggressive behaviors among classroom children" (p. 12).

Aina's (2001) findings supported Trusty and Beckenstein's (1996) findings related to teachers feeling overworked and lacking planning time. Aina (2001) conducted a qualitative study whose purpose was to determine the child's, parent's, and teacher's perspectives of the multi-age classroom. The research sample consisted of teachers, parents, and children involved in the program. The researcher observed many benefits from multi-age classrooms. Aina (2001) found that "In multi-grade classrooms, children

progress at their own pace, individuality is valued, labels are not used to identify students at different levels of academic achievement, and competitiveness is deemphasized" (p. 219). According to Aina, students can progress to different skills when they are ready instead of waiting until the next school year when that material would be taught. Aina suggested that teachers need to understand the pedagogy (teaching) of multi-age and the developmental process first in order to implement this teaching practice successfully. "As schools progress toward more developmentally appropriate programs, multiage and nongraded environments will flourish" (Aina, 2001, p. 224).

Challenges, Disadvantages, and Weaknesses of Multi-Grade Classrooms

With any classroom configuration, there are times where there could be some challenges, disadvantages, and weaknesses. Some researchers in the field shared their findings about multi-grade classrooms. Ansah (1989) shared some negative results for multi-grade classrooms. There was less time for class discussions because of the setup of the day which resulted in students doing more individual work. Teachers felt the curriculum of multiple grades was too difficult to combine. Teachers did not have time to work closely with their classes, resulting in both grade levels receiving inadequate instruction.

Veenman (1995) agreed with Ansah (1989) and analyzed some problems and concerns of multi-grade classrooms relating to teachers and parent perceptions. Kolstad and McFadden (1998) agreed with Ansah (1989) and reported weaknesses primarily involving parents, educators, and administrators. Parents of students placed in a multi-age environment are concerned about the quality of academics their children will receive compared to students in single-grade classrooms.

Veenman (1995) conducted a meta-analysis and presented a justification for the review. "The reason for the review is to highlight the reasons for the increasing number of multi-grade classes, the concerns and perceptions about multi-grade classes, and the claimed advantages of multi-age grouping" (Veenman, 1995, p. 320). He analyzed 56 studies dating from 1949 to 1995 and shared that studies conducted in different countries and at different times share similar disadvantages and concerns with multi-grade classes. Veenman (1995) stated:

> Lack of time for teaching the required content, greater workload, lack of time for individual attention and remediation, lack of adequate classroom management skills, lack of teacher training, inadequate materials, and parental concerns about academic achievement for their children are disadvantages and concerns of multi-grade classes. (p. 324)

Teachers are under pressure to judge learners based on what is expected for that grade. "Parents, teachers, students, and other staff members have expressed their concerns with multi-age classrooms. Student involvement, teacher training, and standards are other issues concerning parents of multi-grade students" (Veenman, 1995, p. 325). Child development in a multi-grade class is another issue of concern for parents. Parents feel that students who are in their second year in a multi-grade classroom will not be challenged and will not continue to grow academically (Trusty & Beckenstein, 1996). The main concern of parents was that their children were not being challenged in the multi-age classroom, while parents of older students felt that the curriculum would be watered down to be more on a level with the younger students (Gutloff & National Education Association, 1995).

Trusty and Beckenstein (1996) conducted a study "to address the concerns of educators and parents who believe the multi-graded classroom is detrimental to student development" (p. 21). They believed that "this additional year of instruction is intended to better assist the students with mastery of crucial material, increasing curriculum expectations and standards of learning" (p. 3). They conjectured that increased time teachers have with students might allow more time to assess student progress. Kindergarten/first, first, and second-grade teachers at two schools in Virginia were the participants in this comparative study. The researchers used two types of assessments, standardized assessments assigned by the county that tested cognitive literacy levels and a teacher questionnaire that addressed the students' social skills within their school grade configuration. Findings showed that there was "no substantial difference in the cognitive development between the students in the multi-graded kindergarten/first-grade class and those students in the single-graded kindergarten or first-grade class" (Trusty & Beckenstein, 1996, p. 18). The findings also showed that "there are no substantial differences in the social/affective skills of students who were taught in a multi-graded classroom and those students who were taught in a single classroom" (p. 18).

"The key message here is planning and taking into account all the many factors and facets that are constantly evolving" (Aina, 2001, p. 223). Educators feel they are ill-prepared or equipped to teach multi-age classrooms. In addition to the lack of time for actual instruction, "Teachers also report a lack of planning and preparation time. Administrators in schools with multi-age classes report difficulties with parents, scheduling, and budget constraints. Multi-age classrooms do not fit neatly into the traditional organization plan for schools" (Aina, 2001, p. 224).

Perceptions of Multi-Grade Classrooms

Administrator, teacher, and parent perceptions of multi-grade classrooms have been collected and analyzed. Through surveys, these individuals have shared their perceptions about this classroom configuration. This section focuses on the perceptions of these groups related to multi-grade classroom configurations.

Britt (1997) conducted a qualitative study by gathering and transcribing field notes from on-site observations and in-person structured and unstructured interviews. He wanted to examine nongraded elementary classrooms from the perspective of the principals, teachers, and parents involved. The two schools involved were in a small urban town located in north central Mississippi.

The principal of school one liked the idea of the multi-age program "and was pleased with the results she has seen so far in her students, school, and parents" (Britt, 1997, p. 7). The principal has personal connections to multi-age education. Principal one also believed that cooperative working relationships with parents, teachers, and students made the multi-age configuration work.

The principal of school two was interested in this configuration setting after hearing other schools were also doing this. Principal two felt that authentic assessment was the key to making the multi-age configuration work. The teachers from both schools were excited to try something different and change their teaching. They received training through attendance at numerous workshops and conferences. Teachers from school two were excited about the change but were not as self-assured of their abilities probably due to their lack of experience in actual practice with this new process. The parents of students in school two were very supportive of this approach. Some concerns voiced by

parents were mixing of the sexes, math curriculum being short-changed, siblings in the same room, giving up a grade system for assessment, and using a different assessment system. The newer parents of students in school two had more problems with multi-age configuration than the current parents in that building. Britt (1997) concluded that the multi-age classes that "will experience the greatest success are those in which the teachers do not feel threatened and are given the freedom to operate as they feel is appropriate" (p. 16).

Penney (2005) conducted a mixed method research study by using observations and quantitative research methods for obtaining quantifiable results from a survey administered to teachers and administrators. The purpose of this study was to examine administrator, teacher, and parent attitudes about different classroom organizations. The research sample for the qualitative part consisted of two multi-age Canadian elementary classrooms at two schools within a metropolitan city and surrounding suburb Penney referred to as North School and South School. The North School was still in the growing stages of transitioning from a single-grade configuration to a multi-grade configuration. In contrast, the South School was very experienced with multi-grade grouping and is well established in this philosophy. Schools in rural Canada are shifting from single-grade instruction to multi-grade due to the decreased student enrollment and economic changes. Penney (2005) indicated that "Disapproval of the multi-grade arrangement has been evident in the consistently negative perceptions and attitudes often displayed toward multi-grade classroom by both parents and educators" (p. 3). Penney (2005) determined that "The administrators are very supportive of multi-age grouping philosophy" (p. 243). Affective and social development were major advantages of multi-grade grouping.

Regarding teacher perceptions, teachers in the North School who have more experience with multi-grade grouping had a positive attitude toward teaching this way. Most teachers in the South School who had many years of experience in multi-grade grouping favored the philosophy. Some of the advantages of this grouping found by Penney (2005) were more creativity and flexibility with the curriculum and stronger and positive relationships.

Disadvantages of this grouping shared by teachers were the time constraints, lack of resources and training to teach in this manner, and math being a subject that is difficult to teach in this grouping due to it being skill-based. Parent perceptions and attitudes toward multi-age grouping were varied. Parents with less experience with multi-grade classroom configurations hold more negative attitudes and perceptions than parents who have more experience in the classroom configuration. In the North School, many parents preferred single-grade grouping for their children. "In the South School, parents report very strong positive perceptions and attitudes toward multi-age grouping. Parents from the South School indicate a preference to have their child attend a multi-age grouping classroom" (Penney, 2005, p. 246). Recommendations suggested by the researcher are having schools take a deeper look at their current practice and provide more professional development to train teachers with this grouping practice.

Eichacker (2008) conducted a study using a survey that queried teachers and parents about their perceptions of multi-age classrooms. The research sites were in Mitchell, Rapid City, Sioux Falls, and Salem, South Dakota School Districts. The study included parents and guardians of students enrolled in grades 3-5 and grades 3-5 students and teachers. The participants from the student sample in the study were 267 students in

multi-age classrooms and 527 students in traditional, single-grade classrooms. Eichacker (2008) wanted to investigate the perceptions of parents and teachers of the participants with regard to teacher-student relationships, parent communication, and student performance within those classroom environments. The parents of students in the multi-age classroom were generally satisfied with their child's performance, the teacher-student relationship, and parent communication. Data analysis from Eichacker (2008) revealed that teachers in the multi-grade classrooms had higher teacher-student relationships and parent communication than traditional classroom teachers. Parent responses suggest that parents of students in multi-age and traditional classrooms feel positive about their child's achievement, relationships, and parent-teacher communication.

Ramrathan and Ngubane (2013) conducted a qualitative research study that explored teacher's experiences with multi-grade teaching in rural areas. The interpretation of the situation included the need for multi-grade classrooms for this school were due to geography, low population, lack of facilities and resources, and location. Due to these hardships, teachers in this area are still able to be resilient and have success with their students. Data were collected through interviews and observations of teachers instructing students. The school is in a deeply rural area with no other surrounding schools nearby. There were two primary teachers, including the principal with 55 students attending. From interviews, Ramrathan and Ngubane (2013) reported "Young learners could be easily motivated by older ones and work up to the best of their abilities to be competent as well" (p. S101). When teachers are faced with challenges including lack of training, lack of resources, no support from stakeholders, and no professional development they are still able to do their best despite hardships. When teachers attend

professional development, they have to adjust that information to fit in a multi-grade configuration. Teachers use strategies to help students learn effectively. Ramrathan and Ngubane (2013) concluded:

> Instructional leadership taken by these multi-grade teachers to adapt to their teaching environment has enabled them to overcome the challenges of teaching in a poorly resourced, neglected environment of multi-grade classes, suggesting that it (instructional leadership) is the key driver for teachers teaching in multi-grade classrooms. (p. S104)

Kivunja and Sims (2015) conducted a study using narrative inquiry. Data collection relied on the written or spoken words of the Zambian stakeholders. The researchers understood the Zambian culture and indicated that participants were most likely comfortable with the concept of story-telling. This study was based on the interpretive paradigm and a social constructivist epistemology. The study was conducted in two neighboring rural schools in Zambia. The purpose of this study was to examine the "different underpinning approaches to multi-grade teaching (a philosophical commitment versus a practical action) and the research that investigates these approaches and then discuss the perceptions of a wide range of stakeholders involved in multi-grade education" (Kivunja & Sims, 2015, p. 10). The researchers wanted to study the perceptions of stakeholders in rural Zambia with multi-grade classrooms about multi-grade teaching as an education strategy. The sample consisted of 13 teacher educators from a key teacher education institution as well as six teachers working in a multi-grade school in rural Zambia. Also participating in the research were 20 student teachers studying primary education at the University of Zafunda and three school principals of

multi-grade schools. Parents were invited to participate by sharing their experience, but only four agreed to participate. The parents who were interviewed were under the impression that the single-grade classes were normal practice in their school and multi-grade was only a substitute if single-grade was not an option. Parents discussed that they would have enrolled their child into a single-grade classroom in contrast to a multi-grade classroom. Kivunja and Sims found that the principals and teachers interviewed agreed that the reason for setting up multi-grade classrooms was based on practicality, although they had issues with student class sizes and the lack of teachers at the school. Kivunja and Sims (2015) acknowledged that multi-grade schooling was undesirable, but a necessary solution to low student and teacher populations in rural and regional areas.

According to Kivunja and Sims (2015), there were positive perceptions about multi-grade instruction. Despite the universal perception among interviewees that multi-grade was the only option for schooling for their children; there was a clear understanding that multi-grade schooling was filling an important gap in the education of their children. The analysis of the interviews showed that teachers supported multi-grade instruction. In conclusion, the researchers found teachers' perceptions of multi-grade classes varied significantly depending on their training. The approach, which focused on the supply of single-grade teachers and the provision of curriculum and resources only designed for single-grade education, exacerbated the challenges faced by teachers in multi-grade settings. The perceptions that teaching multi-grade is "more work" than teaching single-grade also need to be acknowledged.

Bailey et al. (2016) wanted to explore, analyze, and describe the impact of transitioning from a single-age to multi-age classroom on students, parents, and teachers.

"Emphasis was placed on providing details of parents' and teachers' perceptions of transitioning into a school-wide multiage classroom design, as this was identified as a gap in the current literature regarding multiage classrooms" (Bailey et al., 2016, p. 243). In the study, there were two elementary schools located in a small district in the Northwestern United States during the 2009-2010 school year. The schools in this study lacked opportunities for teachers to have collaborative partners so they could share strategies and ideas (Bailey et al., 2016). Classes were usually divided by age; however, some overflow students from an unusually large grade-level population were placed in a multi-grade classroom (Bailey et al., 2016). Deep concern was expressed about the time wasted at the beginning of each year when a teacher had spent "much of the first month instructing the students in general classroom procedures as well as getting to know the individual learning capabilities of each student" (Bailey et al., 2016, p. 240). Findings showed that parents supported the transition to a multi-age design. According to Bailey et al. (2016) "teachers were significantly more neutral than the parents in several areas, such as family-school relationships, class size stability, teacher assignment stability, and overall ability of students to do well in the multi-age classroom" (p. 256). In this study, parents supported the multi-age design more than teachers.

Student Achievement between Multi-Grade and Single-Grade Classrooms

Multiple researchers have conducted studies comparing student achievement in multi-grade and single-grade classrooms. The following studies addressed the results of student achievement in the multi-grade and single-grade configurations. The results related to student achievement are mixed.

MacDonald and Wurster (1974), Eames (1989), Ricard et al. (1995), Veenman (1995), and Gorrell (1998) reported no differences in student achievement between multi-grade and single-grade classrooms. MacDonald and Wurster (1974) conducted a quantitative causal-comparative study to determine if the separation of first-grade children from second and third-grade children resulted in improved vocabulary skills and reading comprehension skills for these children when they started second grade the following year. The curriculum of both configurations was kept constant. The two groups of 20 students were first graders who were taught the same reading curriculum. They began first grade in either 1971 or 1972 and remained in attendance through the end of the school year in 1974. The students came from middle and upper-middle socioeconomic communities with few minority representatives in the population at a Tempe, Arizona elementary school. The Gates MacGinitie Reading Test Primary Form B was administered to the beginning second graders at the school. The results of the data analysis indicated there was no difference in assessed reading skills of both groups. "Therefore, it must be concluded that, for the specific population included in this study, the organizational plans, multiple grade primary team and segregated first-grade team, did not affect first-grade reading progress" (MacDonald & Wurster, 1974, p. 30).

Eames (1989) conducted a quasi-experimental design study to determine if fourth-grade students in a multi-age class scored higher on word achievement tests than their peers in a traditional single-grade classroom. The site of research was a public school district in the Dutchess County area of New York. The participants were 22 fourth- grade students in a multi-age class of combined fourth and fifth graders and 22 fourth graders in a traditional single-grade classroom. The purpose of this study was to determine

"whether significant pupil cognitive achievement differences exist between two systems of vertical school organization, namely, the traditional and multi-age plans, as measured by the STB/McGraw-Hill's Level F, Form U, 1981, in total reading" (Eames, 1989, p. 21). The findings indicated there was no significant difference between averages of total reading scores for students enrolled in both the traditional and multi-age teaching designs.

Ricard et al. (1995) examined the developmental trends relationship between students' adjustment to school and academic achievement for elementary school children in primary mixed-age classrooms. Teacher perceptions of school adjustment of students were also studied. The study was quantitative and included 191 kindergarten through second-grade students and 10 teachers. The Weschler Individual Achievement Test was administered to the students in mathematics, reading, and spelling. Teachers completed the Walker- McConnell Scale of Social Competence and School Adjustment. Ricard et al. (1995) indicated that standardized scores for reading, mathematics, and spelling as well as nonverbal reasoning ability were appropriate relative to age-based norms and did not significantly differ for boys or girls. Student adjustment ratings by teachers were lower for kindergarten students when compared to second graders. It was anticipated that girls would be rated by teachers as more adjusted to the demands of mixed-age classrooms; however, no gender differences were found in mean scores for school adjustment ratings.

Veenman (1995) reviewed 56 studies in his best-evidence synthesis comparing single-grade to multi-grade and multi-age classes. Multi-grade classrooms are classes that have two to three grade levels combined. In contrast, multi-age classrooms have

students of varied ages combined in a classroom. He stressed that "these classes are simply no worse, and simply no better than single-grade or single-age classes" (p. 367). The overall finding of multi-grade and single-grade classrooms relating to achievement are very consistent. Veenman (1995) concluded from these studies that students in multi-grade classrooms learn equally to their single-grade classroom peers.

Gorrell (1998) compared the effects of multi-age classroom strategies to those of traditional classroom strategies on the academic achievement of fourth-grade students in reading and math. The purpose of the quantitative study was to "examine the effect of the traditional and the multi-age classroom instructional setting on the reading and math standardized test scores of two groups of fourth graders" (Gorrell, 1998, p. 34). The researchers used the Stanford Achievement Test, 9th edition from the 1996-1997 and 1997-1998 school years. The study was conducted in a large grade school in Clarksburg, West Virginia. The scores from 40 students randomly chosen from seven traditional fourth grade classes and two third-fourth grade multi-grade classroom were analyzed in the study. Gorrell (1998) revealed there were no significant differences in reading and math scores for the students enrolled in the traditional fourth-grade classrooms and the multi-age classrooms.

Mariano and Kirby (2009) conducted a study to examine how multi-grade students would perform if they had been placed in single-grade classrooms. "This study used a quasi-experimental approach to examine the effect of being assigned to multi-grade classrooms on student achievement" (Mariano & Kirby, 2009, p. 14). For each subject, they examined the effects on five different multi-grade outcomes: third graders in a grade 2-3 configuration; third and fourth graders in a grade 3-4 configuration; fourth

and fifth graders in a grade 4-5 configurations. Monograde students in the same grade as the multi-grade students served as a comparison group. The participants were self-contained students in grades 3 through 5 in the Los Angeles Unified School District from 2002 to 2007. Approximately 3.8% of students in the school district were identified as being enrolled in a multi-age classroom. The California Standards Test in English language arts and mathematics were the student achievement measurements. Mariano and Kirby (2009) found "consistently small and negative effects on student achievement, regardless of grade or subject, even controlling for teacher characteristics" (p. 14). The negative findings could be due to the lack of teacher training and preparation. Based on the findings, the researchers recommended that instructional practices need to be more definitive in order to fully assess what happens in multi-grade classrooms in order to determine the pros and cons of this educational grouping.

Baukol (2010) conducted a quantitative study to compare the academic performance in reading and mathematics of students in a multi-grade classroom and single-grade classroom who attended the same school district from kindergarten through fifth grade. Participants in the study included 250 kindergarten through fifth-grade students who attended school in Minnesota and had Minnesota Comprehensive Assessment scores. Only students who participated in kindergarten through second grade multi-age classroom for all three years were included in the multi-grade group. Multi-grade classrooms were only offered for grades kindergarten through second grade with parents having a choice if they wanted their child enrolled in the multi-grade or single-grade configuration in the district. Assessment data were obtained from the district for third-grade and fifth-grade students only.

Baukol (2010) reported that there was no significant difference in reading and mathematics academic performance of students in multi-age classrooms compared to traditional classrooms in grades three and five. There was also no significant difference in reading and mathematics academic performance between students in multi-age classrooms and traditional classrooms based on gender. Baukol (2010) concluded that regardless of gender, multi-age and traditional instructional groups produce equivalent academic performance in reading and mathematics among third and fifth-grade students.

Slavin (1986), Kinsey (2000), and Eichacker (2008), all reported statistically significant differences in student reading and mathematics achievement. Slavin (1986) conducted a best-evidence synthesis of student achievement and grouping. After reviewing 10 studies regarding nongraded grouping in reading and mathematics, Slavin (1986) concluded that student achievement gains were higher in nongraded grouping than in other classroom groupings. "Results of nongraded plans are remarkably strong. Both randomized studies found positive effects on student achievement" (Slavin, 1986, p. 46). A study conducted by Kinsey (2000) supported Slavin's (1986) work by suggesting a relationship between multi-age classrooms and academic achievement outcomes existed.

Kinsey (2000) compared the effect of multi-age and single-grade classrooms by comparing the academic achievement of second-grade children who participated in a multi-age classroom for two years with second-grade children who participated in single-grade placements for two years. The participants included 261 second-grade students from a suburban middle-class school district who were enrolled in either a multi-age or single-grade classroom over two years. Students in the multi-age classroom were enrolled by parent choice. Results from the data analysis showed that outcomes on the

California Achievement Test in reading and mathematics were greater for students who were enrolled in a multi-age classroom than the students enrolled in a single-grade classroom. Students in multi-age classrooms also received higher gains on report cards in reading, writing, spelling, and mathematics. "This study lends support to the multi-age model as a viable educational alternative which contributes to both academic and affective outcomes for students" (Kinsey, 2000, p. 139).

Eichacker (2008) wanted to examine the reading performance of elementary students in the multi-age classrooms as compared to the single-grade classrooms. Data collection was from the Dakota STEP scores for reading for the 2004-2007 school years. The research site was in Mitchell, Rapid City, Sioux Falls, and Salem, South Dakota School Districts. The study included students in grades 3-5. The participants in the study were 267 students enrolled in multi-age classrooms and 527 students enrolled in single-grade classrooms. Eichacker (2008) reported, "there was a significant difference in reading performance with the multi-age classroom compared to the traditional classroom" (p. 82). Students in multi-age classrooms performed better than their single-age peers on the Dakota STEP reading assessment.

The goal of classroom instruction is student learning, and the unwavering objective of student learning is achievement. Resources, funding, enrollment, and parent and teacher perceptions and are all factors to analyze when deciding the best classroom configuration for schools. As educators attempt to find ways to make learning meaningful and engaging, they may want to consider different classroom configurations to engage learners and advance student achievement.

Summary

Chapter 2 provided a review of the literature about the history of classroom configurations. Included in the review were past studies about the benefits and challenges of multi-grade classrooms, perceptions of multi-grade classrooms, and student achievement differences between multi-grade and single-grade classrooms. Presented in Chapter 3 are the research design, selection of participants, measurement, and data collection procedures. The data analysis and hypothesis testing for the study are described as well as the limitations of the study.

Chapter 3

Methods

The purpose of this study was to determine if a multi-grade educational setting was beneficial for student achievement. The first purpose was to determine whether there is a difference between the Iowa Assessments reading, written expression, mathematics, social studies, and science scores of third through eighth-grade students enrolled in multi-grade level classrooms and third through eighth-grade students enrolled in single-grade classrooms. The second purpose was to determine whether the difference between the Iowa Assessments reading, written expression, mathematics, social studies, and science scores of third through eighth-grade students enrolled in a multi-grade level classroom and third through eighth-grade students enrolled in a single-grade classroom was affected by student grade level. Included in this chapter are the details of the design of the study and descriptions of how each research question was addressed. This chapter includes an explanation of the research design, the selection of participants, measurement, data collection procedures, hypothesis testing, and the limitations of the study.

Research Design

A quantitative causal-comparative research design guided this study. A causal-comparative design was most appropriate for this study because two comparison groups were used. According to Lunenburg and Irby (2008), in causal-comparative research, the independent variable is not manipulated because it already occurred and cannot be controlled. The independent variables used in this study included the classroom configurations in the schools and the grade levels. The dependent variables in this study were archived student achievement scores in reading, written expression, math, social

studies, and science on the Iowa Assessments from school years 2016-2017 and 2017-2018.

Selection of Participants

The participants of this study were students enrolled in grades 3-8 in a multi-grade classroom configuration and a single-grade classroom configuration in Diocese X in a Midwestern state. A purposive sampling procedure was used to select all students from grades three through eight enrolled at the 37 elementary schools in Diocese X. Lunenburg and Irby (2008) defined purposive sampling as "selecting a sample based on the researcher's experience or knowledge of the group to be sampled" (p. 175). A student's data were included in this study if the following criteria were met:

1. The student attended Diocese X during the 2016-2017 and 2017-2018 school years.
2. Each student was enrolled in grades 3 through 8.
3. The student received a valid Iowa Assessment score in reading, written expression, mathematics, social studies, and science.

Measurement

Scores from the reading, written expression, mathematics, social studies, and science sections of the Iowa Assessments were used to measure student achievement. The Iowa Assessments use a vertical score scale known as the national standard score. Welch and Dunbar (2014) discussed the scoring of the Iowa Assessments as "a metric that ranges numerically from 80 to 400 and spans a developmental continuum from kindergarten to grade 12 in major content domains such as reading, mathematics, science, and written expression" (p. 2). The Iowa Assessments use subscores, which are called

standard scores when determining achievement levels. In grades 3-8, the standard scores range from 150-350. Welch and Dunbar (2014) stated:

> The standard score is a number that describes a student's location on an achievement continuum or scale. As students grow and learn, it is expected that their standard score will continue to increase. In Iowa, Proficient/Not Proficient status is defined in terms of Standard Scores. (p. 4)

The reading subtest on the Iowa Assessments is administered in two parts. Each part is 30 minutes in duration and includes the areas of literary text, informational text, vocabulary, explicit meaning, implicit meaning, key ideas, and author's craft. The written expression subtest is 40 minutes in duration and includes the areas of usage, grammar, sentence structure, planning, organization, and appropriate expression. The mathematics subtest is administered in two parts, 30 minutes each. The subtest includes number sense, operations, algebraic patterns, data analysis, probability and statistics, geometry, and measurement. The social studies subtest is 35 minutes in duration and includes history, geography, economics, civics, and government. The science subtest is 35 minutes in duration and includes life science, physical science, earth science, and space science by grade level for grades 3-8. The number of items in each subtest is included in Table 3.

Table 3

Number of Items on Iowa Assessments

	Grades					
Subject	3	4	5	6	7	8
Written Expression	35	38	40	43	45	48
Mathematics	50	55	60	65	70	75
Social Studies	30	34	37	39	41	43
Reading	41	42	43	44	45	46
Science	30	34	37	39	41	43

Note. Adapted from *Forms E and F Scope and Sequence for Complete and Core Batteries,* by Houghton Mifflin Harcourt, 2016. Retrieved from https://itp.education.uiowa.edu /ia/documents/ Iowa_Form_E_F_Scope_and_Sequence.pdf

Lunenburg and Irby (2008) stated that validity is "the degree to which an instrument measures what it purports to measure" (p. 181). The Iowa Assessment has been tested for content and concurrent validity. Cronbach (1971) made the point that validation is the task of the interpreter:

> In the end, the responsibility for valid use of a test rests on the person who interprets it. The published research merely provides the interpreter with some facts and concepts. He has to combine these with other knowledge about the person he tests. (p. 445)

The steps in the development of the Iowa Assessments are as follows: internal review stage one, external review, internal stage two, item tryout, data review, operational forms construction, and forms review. Test validity was measured using concurrent validity, which is "the degree to which scores on one test correlate to scores

on another test when both are administered at about the same time" (Lunenburg & Irby, 2008, p. 181). University of Iowa (2015) shared the concurrent validity of Iowa Assessments:

> Concurrent validity coefficients are presented in the form of correlations between scores on the *Iowa Assessments* Form E and (1) scores on *Cognitive Abilities Test* (*CogAT*) Form 7 and (2) scores on the *Iowa Tests of Basic Skills* (*ITBS*) and *Iowa Tests of Educational Development* (*ITED*) Form A. (p. 45)

The highest correlation is represented by the *CogAT* Composite score or the score from the Verbal Battery (University of Iowa, 2015). The lowest correlation, indicating the least overlap between achievement and the cognitive skills measure, involves the skills tests in the Iowa Assessments and the *CogAT* Form 7 Nonverbal Battery (University of Iowa, 2015). Average correlations with the *Iowa Assessments* Levels 5/6-17/18 Complete Composite and *CogAT* Form 7 are 0.77 for the Verbal Battery, 0.71 for the Quantitative Battery, 0.64 for the Nonverbal Battery, and 0.80 for the *CogAT* Form 7 Composite (University of Iowa, 2015). "The relationship is substantial in all cases; however, the correlations are not so high as to suggest that the achievement and ability measures lack discriminant validity" (University of Iowa, 2015, p. 45).

Test reliability is "the degree to which an instrument consistently measures whatever it is measuring" (Lunenburg & Irby, 2008, p. 182). The University of Iowa researchers (2015) employed the internal-consistency estimates the Kuder-Richardson Formula 20 (K-R 20) to determine the reliability of the instrument. Reliability coefficients derived by this technique were based on data from national comparison samples and are reported for both fall and spring administrations (University of Iowa,

2015). The reliability coefficients based on K-R 20 Spring *Iowa Assessments* Form E in reading for grades 3-8 ranged from .90 through .91 (University of Iowa, 2015). The reliability coefficients based on K-R 20 Spring Iowa Assessments Form E in mathematics for grades 3-8 ranged from .86 through .93 (University of Iowa, 2015).

The grade levels used in the study are grades 3-8 in every school in Diocese X. The multi-grade classrooms consisted of two grade levels (3/4, 5/6, and 7/8) in one classroom with a single teacher. The classrooms in the single-grade schools consisted of one teacher with students in the same grade. For example, third-grade students enrolled in a multi-grade classroom were compared to third-grade students enrolled in single-grade classrooms.

Data Collection Procedures

Before data collection, the Diocesan Superintendent of Schools gave written consent for this study to be conducted in April 2018 with the condition of having the study approved by Baker University's Institutional Review Board (IRB). A consent form was signed by the Diocesan Superintendent of Schools who collected the archival test score data and student sample data (see Appendix A). On July 17, 2018, a request for permission to conduct the study was submitted to Baker University IRB committee which was approved on July 19, 2018 (see Appendix B). Six Excel worksheets were sent to the researcher. The data was coded to ensure the anonymity of the students. The worksheets were merged into one file and imported into IBM SPSS Statistics Faculty Pack 25 for PC for data analysis.

Data Analysis and Hypothesis Testing

Data from the Iowa Assessments were analyzed to address each research question in this study. Two-factor analyses of variance (ANOVAs) were used to test the hypotheses. The following format of the data analysis and hypothesis testing includes the research question, hypothesis, and the data analysis.

RQ1. To what extent is there a difference between the Iowa Assessments reading scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

H1. There is a difference between the Iowa Assessments reading scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom.

A two-factor ANOVA was conducted to test H1 and H2. The two categorical variables used to group the dependent variable, Iowa Assessments reading scores, were classroom configuration (multi-grade and single-grade) and grade level (grades 3-8). The two-factor ANOVA can be used to test three hypotheses including a main effect for classroom configuration, a main effect for grade level, and a two-way interaction effect (Classroom Configuration x Grade Level). The main effect for classroom configuration was used to test H1. The level of significance was set at .05.

RQ2. To what extent is the difference between the Iowa Assessments reading scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level?

H2. The difference between the Iowa Assessments reading scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom is affected by grade level.

The interaction effect from the first ANOVA (Classroom Configuration x Grade Level), with Iowa Assessments reading scores as the dependent variable, was used to test H2. The level of significance was set at .05.

RQ3. To what extent is there a difference between the Iowa Assessments written expression scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

H3. There is a difference between the Iowa Assessments written expression scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom.

A second two-factor ANOVA was conducted to test H3 and H4. The two categorical variables used to group the dependent variable, Iowa Assessments written expression scores, were classroom configuration (multi-grade and single-grade) and grade level (grades 3-8). The two-factor ANOVA can be used to test three hypotheses including a main effect for classroom configuration, a main effect for grade level, and a two-way interaction effect (Classroom Configuration x Grade Level). The main effect for classroom configuration was used to test H3. The level of significance was set at .05

RQ4. To what extent is the difference between the Iowa Assessments written expression scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level?

H4. The difference between the Iowa Assessments written expression scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom is affected by grade level.

The interaction effect from the second ANOVA (Classroom Configuration x Grade Level), with Iowa Assessments written expression scores as the dependent variable, was used to test H4. The level of significance was set at .05.

RQ5. To what extent is there a difference between the Iowa Assessments mathematics scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

H5. There is a difference between the Iowa Assessments mathematics scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom.

A third two-factor ANOVA was conducted to test H5 and H6. The two categorical variables used to group the dependent variable, Iowa Assessments mathematics scores, were classroom configuration (multi-grade and single-grade) and grade level (grades 3-8). The two-factor ANOVA can be used to test three hypotheses including a main effect for classroom configuration, a main effect for grade level, and a two-way interaction effect (Classroom Configuration x Grade Level). The main effect for classroom configuration was used to test H5. The level of significance was set at .05

RQ6. To what extent is the difference between the Iowa Assessments mathematics scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level?

H6. The difference between the Iowa Assessments mathematics scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom is affected by grade level.

The interaction effect from the third ANOVA (Classroom Configuration x Grade Level), with Iowa Assessments mathematics scores as the dependent variable, was used to test H6. The level of significance was set at .05.

RQ7. To what extent is there a difference between the Iowa Assessments social studies scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

H7. There is a difference between the Iowa Assessments social studies scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom.

A fourth two-factor ANOVA was conducted to test H7 and H8. The two categorical variables used to group the dependent variable, Iowa Assessments social studies scores, were classroom configuration (multi-grade and single-grade) and grade level (grades 3-8). The two-factor ANOVA can be used to test three hypotheses including a main effect for classroom configuration, a main effect for grade level, and a two-way interaction effect (Classroom Configuration x Grade Level). The main effect for classroom configuration was used to test H7. The level of significance was set at .05

RQ8. To what extent is the difference between the Iowa Assessments social studies scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level?

H8. The difference between the Iowa Assessments social studies scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom is affected by grade level.

The interaction effect from the fourth ANOVA (Classroom Configuration x Grade Level), with Iowa Assessments social studies scores as the dependent variable, was used to test H8. The level of significance was set at .05.

RQ9. To what extent is there a difference between the Iowa Assessments science scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

H9. There is a difference between the Iowa Assessments science scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom.

A fifth two-factor ANOVA was conducted to test H9 and H10. The two categorical variables used to group the dependent variable, Iowa Assessments science scores, were classroom configuration (multi-grade and single-grade) and grade level (grades 3-8). The two-factor ANOVA can be used to test three hypotheses including a main effect for classroom configuration, a main effect for grade level, and a two-way interaction effect (Classroom Configuration x Grade Level). The main effect for classroom configuration was used to test H9. The level of significance was set at .05

RQ10. To what extent is the difference between the Iowa Assessments science scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level?

H10. The difference between the Iowa Assessments science scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom is affected by grade level.

The interaction effect from the fifth ANOVA (Classroom Configuration x Grade Level), with Iowa Assessments science scores as the dependent variable, was used to test H10. The level of significance was set at .05.

Limitations

Lunenburg and Irby (2008) stated, "limitations are factors that may have an effect on the interpretation of the findings or on the generalizability of the results" (p. 133). While the researcher does not have control over the limitations, explicitly stating them can assist in preventing misapprehensions (Lunenburg & Irby, 2008). Limitations of this study included:

1. Teachers may have different years of teaching experience and different levels of experience in multi-grade classrooms.
2. Teachers may have participated in different amounts and types of professional development in education.
3. External or individual differences affecting student achievement like motivation, absences, or preparation might exist.
4. Student experiences in multi-grade or single-grade classrooms prior to the school years involved in this study are unknown.

Summary

Chapter 3 included information regarding the methodology utilized in this study. The topics covered included the research design, the selection of participants,

measurement, data analysis and hypothesis testing, data collection procedures, and the limitations of the study. Chapter 4 contains the descriptive statistics and results of the analysis of the data collected that addressed the research questions of this study.

Chapter 4

Results

The first purpose of this study was to determine whether there is a difference between the Iowa Assessments reading, written expression, mathematics, social studies, and science scores of students enrolled in a multi-grade level classroom and students enrolled in a single-grade classroom. The second purpose was to determine whether the difference between the Iowa Assessments reading, written expression, mathematics, social studies, and science scores of students enrolled in a multi-grade level classroom and students enrolled in a single-grade classroom was affected by student grade level. This chapter includes the descriptive statistics and the results of the data analysis.

Descriptive Statistics

The frequency table (see Table 4) was included to aid in a clear representation of the descriptive statistics. The frequency table is used to describe the number of students in each classroom configuration. This table includes information regarding the number of students at each grade level during both school years, organized by multi-grade and single-grade level configurations.

Table 4

Frequencies for Grade Configurations for 2016 and 2017 by Grade Level

	f Multi-grade		f Single-grade	
Grade	2016	2017	2016	2017
3	103	94	534	548
4	82	91	518	529
5	85	71	505	492
6	63	76	533	477
7	64	57	516	517
8	54	61	475	505

Hypothesis Testing

The results of the hypothesis testing are included to address each of the research questions. The researcher conducted two-factor ANOVAs to determine the differences in Iowa Assessment scores of students enrolled in multi-grade classrooms and students enrolled in single-grade classrooms. The interaction effect was used to determine whether the differences between students in multi-grade classrooms and single-grade classrooms were affected by grade level. Each research question is followed by the corresponding hypothesis, data analysis, and results.

RQ1. To what extent is there a difference between the Iowa Assessments reading scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

H1. There is a difference between the Iowa Assessments reading scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom.

A two-factor ANOVA was conducted to test H1 and H2. The two categorical variables used to group the dependent variable, Iowa Assessments reading scores, were classroom configuration (multi-grade and single-grade) and grade level (grades 3-8). The two-factor ANOVA can be used to test three hypotheses including a main effect for classroom configuration, a main effect for grade level, and a two-way interaction effect (Classroom Configuration x Grade Level). The main effect for classroom configuration was used to test H1. The level of significance was set at .05.

The results of the analysis indicated there was not a statistically significant difference between the means, $F = 2.010$, $df = 1, 7031$, $p = .156$. See Table 5 for the means and standard deviations for this analysis. The average Iowa Assessments reading score for students enrolled in a multi-grade classroom ($M = 216.46$) is not different from the average Iowa Assessments reading score students enrolled in a single-grade classroom ($M = 222.57$). H1 was not supported.

Table 5

Descriptive Statistics for the Results of the Test for H1

Classroom	*M*	*SD*	*N*
Multi-grade	216.46	38.67	899
Single-grade	222.57	35.12	6144

RQ2. To what extent is the difference between the Iowa Assessments reading scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level?

H2. The difference between the Iowa Assessments reading scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom is affected by grade level.

The interaction effect from the first ANOVA (Classroom Configuration x Grade Level), with Iowa Assessments reading scores as the dependent variable, was used to test H2. The level of significance was set at .05.

The results of the analysis indicated a statistically significant difference between at least two of the means, $F = 3.950$, $df = 5, 7031$, $p = .001$. See Table 6 for the means and standard deviations for this analysis. A follow-up post hoc was conducted to determine which pairs of means were different.

Table 6

Descriptive Statistics for the Results of the Test for H2

Classroom	Grade	*M*	*SD*	*N*
Multi-grade	3	177.53	19.48	195
	4	197.40	22.71	173
	5	214.62	23.10	156
	6	228.42	26.36	139
	7	250.40	29.70	121
	8	263.47	32.98	115
Single-grade	3	184.30	18.90	1080
	4	203.09	21.15	1046
	5	214.71	21.35	995
	6	228.35	24.75	1010
	7	248.48	26.92	1033
	8	260.25	27.59	980

The Tukey's Honestly Significant Difference (HSD) post hoc was conducted at $\alpha = .05$. A large number of the differences between the means were greater than the critical difference, HSD = 13.171. H2 was supported. However, the post hoc results relevant to this study, which involved comparisons of the reading scores between students enrolled in the two types of classroom at each grade level, indicated no significant differences (see Table 7).

Table 7

Differences in Mean Reading Scores by Grade Level and Classroom Configuration

Grade	Multi-grade	Single-grade	Difference
3	177.53	184.30	-6.76
4	197.40	203.09	-5.68
5	214.62	214.71	-0.10
6	228.42	228.35	0.06
7	250.40	248.48	1.91
8	263.47	260.25	3.22

RQ3. To what extent is there a difference between the Iowa Assessments written expression scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

H3. There is a difference between the Iowa Assessments written expression scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom.

A second two-factor ANOVA was conducted to test H3 and H4. The two categorical variables used to group the dependent variable, Iowa Assessments written expression scores, were classroom configuration (multi-grade and single-grade) and grade level (grades 3-8). The two-factor ANOVA can be used to test three hypotheses including a main effect for classroom configuration, a main effect for grade level, and a two-way interaction effect (Classroom Configuration x Grade Level). The main effect for classroom configuration was used to test H3. The level of significance was set at .05

The results of the analysis indicated there was a statistically significant difference between the means, F =7.657, df = 1, 7027, p = .006. See Table 8 for the means and standard deviations for this analysis. The average Iowa Assessments written expression score for grades 3-8 students enrolled in a multi-grade classroom (M = 220.66) is lower than the average Iowa Assessments written expression score for grades 3-8 students enrolled in a single-grade classroom (M = 228.97). H3 was supported. The effect size, as measured by partial eta squared = .001, indicated that .1% of the variability in written expression scores is explained by classroom configuration, indicating a small effect, according to Cohen's (1988) conventions.

Table 8

Descriptive Statistics for the Results of the Test for H3

Classroom	*M*	*SD*	*N*
Multi-grade	220.66	44.062	898
Single-grade	228.97	42.427	6141

RQ4. To what extent is the difference between the Iowa Assessments written expression scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level?

H4. The difference between the Iowa Assessments written expression scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom is affected by grade level.

The interaction effect from the second ANOVA (Classroom Configuration x Grade Level), with Iowa Assessments written expression scores as the dependent variable, was used to test H4. The level of significance was set at .05.

The results of the analysis indicated there was not a statistically significant difference between at least two of the means, $F = 1.116$, $df = 5, 7027$, $p = .350$. See Table 9 for the means and standard deviations for this analysis. A follow-up post hoc was not warranted. H4 was not supported.

Table 9

Descriptive Statistics for the Results of the Test for H4

Classroom	Grade	*M*	*SD*	*N*
Multi-grade	3	176.38	18.212	194
	4	198.70	23.358	173
	5	216.63	28.286	156
	6	238.63	32.632	139
	7	261.34	37.411	121
	8	269.34	37.543	115
Single-grade	3	181.73	17.478	1077
	4	204.22	23.696	1047
	5	220.99	28.170	994
	6	238.86	30.922	1010
	7	260.56	33.441	1033
	8	271.91	34.797	980

RQ5. To what extent is there a difference between the Iowa Assessments mathematics scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

H5. There is a difference between the Iowa Assessments mathematics scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom.

A third two-factor ANOVA was conducted to test H5 and H6. The two categorical variables used to group the dependent variable, Iowa Assessments mathematics scores, were classroom configuration (multi-grade and single-grade) and grade level (grades 3-8). The two-factor ANOVA can be used to test three hypotheses including a main effect for classroom configuration, a main effect for grade level, and a two-way interaction effect (Classroom Configuration x Grade Level). The main effect for classroom configuration was used to test H5. The level of significance was set at .05.

The results of the analysis indicated there was not a statistically significant difference between the means, $F = 2.264$, $df = 1, 7022$, $p = .132$. See Table 10 for the means and standard deviations for this analysis. The average Iowa Assessments mathematics score for students enrolled in a multi-grade classroom ($M = 208.90$) is not different from the average Iowa Assessments mathematics score students enrolled in a single-grade classroom ($M = 212.51$). H5 was not supported.

Table 10

Descriptive Statistics for the Results of the Test for H5

Classroom	*M*	*SD*	*N*
Multi-grade	208.90	35.39	895
Single-grade	212.51	34.10	6139

RQ6. To what extent is the difference between the Iowa Assessments mathematics scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level?

H6. The difference between the Iowa Assessments mathematics scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom is affected by grade level.

The interaction effect from the third ANOVA (Classroom Configuration x Grade Level), with Iowa Assessments mathematics scores as the dependent variable, was used to test H6. The level of significance was set at .05.

The results of the analysis indicated there was not a statistically significant difference between at least two of the means, $F = 1.647$, $df = 5, 7022$, $p = .144$. See Table 11 for the means and standard deviations for this analysis. A follow-up post hoc was not warranted. H6 was not supported.

Table 11

Descriptive Statistics for the Results of the Test for H6

Classroom	Grade	*M*	*SD*	*N*
Multi-grade	3	174.71	14.287	195
	4	188.19	16.577	172
	5	205.07	19.179	155
	6	217.34	21.062	138
	7	243.03	26.140	120
	8	257.30	31.259	115
Single-grade	3	175.62	12.416	1077
	4	190.36	15.376	1046
	5	203.20	18.109	996
	6	215.61	20.137	1010
	7	238.78	24.661	1031
	8	255.35	28.185	979

RQ7. To what extent is there a difference between the Iowa Assessments social studies scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

H7. There is a difference between the Iowa Assessments social studies scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom.

A fourth two-factor ANOVA was conducted to test H7 and H8. The two categorical variables used to group the dependent variable, Iowa Assessments social

studies scores, were classroom configuration (multi-grade and single-grade) and grade level (grades 3-8). The two-factor ANOVA can be used to test three hypotheses including a main effect for classroom configuration, a main effect for grade level, and a two-way interaction effect (Classroom Configuration x Grade Level). The main effect for classroom configuration was used to test H7. The level of significance was set at .05.

The results of the analysis indicated there was not a statistically significant difference between the means, $F = 0.061$, $df = 1, 7035$, $p = .805$. See Table 12 for the means and standard deviations for this analysis. The average Iowa Assessments social studies score for students enrolled in a multi-grade classroom ($M = 219.70$) is not different from the average Iowa Assessments social studies score students enrolled in a single-grade classroom ($M = 225.05$). H7 was not supported.

Table 12

Descriptive Statistics for the Results of the Test for H7

Classroom	*M*	*SD*	*N*
Multi-grade	219.70	40.385	901
Single-grade	225.05	39.289	6146

RQ8. To what extent is the difference between the Iowa Assessments social studies scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level?

H8. The difference between the Iowa Assessments social studies scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom is affected by grade level.

The interaction effect from the fourth ANOVA (Classroom Configuration x Grade Level), with Iowa Assessments social studies scores as the dependent variable, was used to test H8. The level of significance was set at .05.

The results of the analysis indicated there was not a statistically significant difference between at least two of the means, $F = 1.228$, $df = 5, 7035$, $p = .293$. See Table 13 for the means and standard deviations for this analysis. A follow-up post hoc was not warranted. H8 was not supported.

Table 13

Descriptive Statistics for the Results of the Test for H8

Classroom	Grade	*M*	*SD*	*N*
Multi-grade	3	178.36	19.344	197
	4	200.54	22.373	173
	5	218.66	23.214	156
	6	233.11	28.679	139
	7	254.28	34.000	121
	8	268.14	34.118	115
Single-grade	3	182.59	18.152	1080
	4	202.47	22.257	1046
	5	218.60	24.045	997
	6	233.10	29.215	1010
	7	252.84	33.384	1033
	8	264.93	33.241	980

RQ9. To what extent is there a difference between the Iowa Assessments science scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom?

H9. There is a difference between the Iowa Assessments science scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom.

A fifth two-factor ANOVA was conducted to test H9 and H10. The two categorical variables used to group the dependent variable, Iowa Assessments science scores, were classroom configuration (multi-grade and single-grade) and grade level (grades 3-8). The two-factor ANOVA can be used to test three hypotheses including a main effect for classroom configuration, a main effect for grade level, and a two-way interaction effect (Classroom Configuration x Grade Level). The main effect for classroom configuration was used to test H9. The level of significance was set at .05.

The results of the analysis indicated there was not a statistically significant difference between the means, $F = 2.478$, $df = 1, 7035$, $p = .115$. See Table 14 for the means and standard deviations for this analysis. The average Iowa Assessments science score for students enrolled in a multi-grade classroom ($M = 222.71$) is not different from the average Iowa Assessments science score students enrolled in a single-grade classroom ($M = 226.10$). H9 was not supported.

Table 14

Descriptive Statistics for the Results of the Test for H9

Classroom	*M*	*SD*	*N*
Multi-grade	222.71	41.219	901
Single-grade	226.10	38.790	6146

RQ10. To what extent is the difference between the Iowa Assessments science scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom affected by grade level?

H10. The difference between the Iowa Assessments science scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom is affected by grade level.

The interaction effect from the fifth ANOVA (Classroom Configuration x Grade Level), with Iowa Assessments science scores as the dependent variable, was used to test H10. The level of significance was set at .05.

The results of the analysis indicated a statistically significant difference between at least two of the means, $F = 3.234$, $df = 5, 7035$, $p = .006$. See Table 15 for the means and standard deviations for this analysis. A follow-up post hoc was conducted to determine which pairs of means were different.

Table 15

Descriptive Statistics for the Results of the Test for H10

Classroom	Grade	*M*	*SD*	*N*
Multi-grade	3	182.40	20.291	197
	4	204.64	26.038	173
	5	221.43	27.099	156
	6	234.70	32.323	139
	7	255.64	32.608	121
	8	271.57	35.473	115
Single-grade	3	185.69	19.669	1081
	4	207.89	25.261	1046
	5	218.38	27.287	996
	6	234.71	29.832	1010
	7	249.24	32.683	1033
	8	264.73	34.404	980

The Tukey's HSD post hoc was conducted at $\alpha = .05$. A large number of the differences between the means were greater than the critical difference, HSD = 13.164. H10 was supported. However, the post hoc results relevant to this study, which involved comparisons of the science scores between students enrolled in the two types of classroom at each grade level, indicated no significant differences (see Table 16).

Table 16

Differences in Mean Science Scores by Grade Level and Classroom Configuration

Grade	Multi-grade	Single-grade	Difference
3	182.40	185.69	-3.30
4	204.64	207.89	-3.24
5	221.43	218.38	3.05
6	234.70	234.71	-0.01
7	255.64	249.24	6.40
8	271.57	264.73	6.85

Summary

This chapter included descriptive statistics related to the frequencies for grade configurations for 2016 and 2017 by grade level. The results of the hypothesis testing analysis were also shared. Chapter 5 includes the study summary, the findings related to the literature, and the conclusions.

Chapter 5

Interpretation and Recommendations

The purpose of this study was to determine whether there is a difference between the Iowa Assessments reading, written expression, mathematics, social studies, and science scores of students grades 3-8 enrolled in multi-grade and students grades 3-8 enrolled in single-grade classrooms and whether those differences were affected by grade level. The students included in the study were enrolled in schools from a Midwestern Catholic Diocese. Included in this chapter are the study summary, the findings related to the literature, and the conclusion.

Study Summary

This section provides a summary of the current study. The summary includes an overview of the problem concerning the classroom configuration that would meet the best educational needs for students and maintain high levels of student achievement with limited resources and staff. Also included in this section are the purpose statement and research questions, and an overview of the methodology. Finally, the major findings of the study are explained.

Overview of the problem. Meeting the needs of students is a goal that all school communities continually work to reach. Educators often think of ways to teach all students with limited resources and staff in rural areas and small private schools. These school communities need to decide if their classroom configuration benefits student achievement as this is a problem those schools face. Those in favor of multi-grade classrooms believe this configuration builds relationships among students and increases cooperative groups. The results of previous studies were mixed when single-grade

classrooms were compared to multi-grade classrooms. In this study, the student achievement of students enrolled in single-grade classrooms was compared to that of students enrolled in multi-grade classrooms.

Purpose statement and research questions. The purpose of this study was to determine if a multi-grade educational setting was beneficial for student achievement. More specifically the first purpose was to determine whether there is a difference between the Iowa Assessments reading, written expression, mathematics, social studies, and science scores of students grades 3-8 enrolled in multi-grade classrooms (3/4, 5/6, and 7/8) and students grades 3-8 enrolled in single-grade classrooms. The second purpose was to determine whether the difference between the Iowa Assessments reading, written expression, mathematics, social studies, and science scores of students enrolled in multi-grade level classrooms and students grades 3-8 enrolled in single-grade classrooms was affected by student grade level. To address the purposes of this study 10 research questions were posed, and 10 hypotheses were tested.

Review of the methodology. A quantitative causal-comparative research design guided this study. The independent variables used for the study included the classroom configurations in the schools and the grade levels. The grade levels used in the study are grades 3-8 in every school in Diocese X. The dependent variables in this study were archived student achievement scores in reading, written expression, math, social studies, and science on the Iowa Assessments from school years 2016-2017 and 2017-2018 for students in grades 3-8. Data from the Iowa Assessments were analyzed to address each research question in this study. Two-factor ANOVAs were conducted to test the hypotheses.

Major findings. The average Iowa Assessments reading, mathematics, social studies and science scores for students enrolled in a multi-grade classroom were not different from the scores of students enrolled in a single-grade classroom. However, the average Iowa Assessments written expression score for students enrolled in a multi-grade classroom was lower than the average Iowa Assessments written expression score for students enrolled in a single-grade classroom. The differences between the Iowa Assessments written expression, mathematics, and social studies scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom were not affected by grade level. The difference between the Iowa Assessments reading and science scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom were affected by grade level. However, the comparisons of the reading and science scores between students enrolled in the two types of classroom at each grade level indicated no significant differences.

Findings Related to the Literature

The findings from this study related to the literature on student achievement between multi-grade and single-grade classrooms are included in this section. The current research focused on the differences in Iowa Assessment scores. In 2018-2019 when this study was conducted, no research was found related to the effect of grade level on the differences in student achievement between students enrolled in multi-grade and single-grade classrooms. Although differences were found between the Iowa Assessment written expression scores of grades 3-8 students enrolled in a multi-grade classroom in a single-grade classroom, no research was found to compare these results. Similarly, no

differences were found between the Iowa Assessment social studies and science scores of grades 3-8 students in multi-grade and single-grade classrooms. Only previous research related to the differences in student achievement for reading and mathematics between students enrolled in multi-grade and single-grade classrooms were located.

The findings of the current study indicated there was not a difference between the Iowa Assessments reading scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom. MacDonald and Wurster (1974) conducted a quantitative causal-comparative study to determine if reading comprehension skills improved if students were in a multiple grade primary team. The authors reported that "for the population included in the study, the organizational plans, multiple grade primary team and segregated first-grade team, did not affect first-grade reading progress" (MacDonald & Wurster, 1974, p. 30). Eames's (1989) study was designed to determine if fourth-grade students in a multi-age class scored higher on word achievement tests than their peers in a traditional single-grade classroom. The findings indicated there was no significant difference between averages on achievement tests between students in different classroom configurations. Gorrell (1998) conducted a study to examine the effect of traditional and multi-age classroom instructional setting on reading standardized test scores of fourth graders. Gorrell's findings revealed there were no significant differences in reading scores for students enrolled in traditional classrooms and multi-age classrooms. Baukol (2010) conducted a study to compare the academic performance in reading of students in multi-grade and single-grade classrooms. Baukol (2010) found no significant difference in reading academic performance of students in multi-age classrooms compared to traditional classrooms. The results from the current

study support the findings of MacDonald and Wurster (1974), Eames (1989), Gorrell (1998), and Baukol (2010).

Slavin (1986) conducted what he called was a best-evidence synthesis of student achievement and grouping. After reviewing 10 studies regarding reading, Slavin (1986) concluded that "student achievement gains are higher in nongraded groupings than in other classroom groupings" (p. 46). Kinsey (2000) compared the effect of multi-age and single-grade classrooms by comparing the academic achievement of second-grade children who participated in a multi-age classroom for two years with second-grade children who participated in single-grade placements for two years in reading. Kinsey (2000) concluded outcomes on the California Achievement Test in reading were greater for students who were enrolled in a multi-grade classroom than the students enrolled in a single-grade classroom. Eichacker (2008) studied the reading performance of elementary students in the multi-age classrooms compared to traditional classrooms. Eichacker (2008) reported, "there was a significant difference in reading performance with the multi-age classroom compared to the traditional classroom" (p. 82). Reading performance was higher within the multi-age classroom compared to the single-grade classroom. The findings from the current study are in contrast to the findings of Slavin (1986), Kinsey (2000), and Eichacker (2008).

The findings from the current study indicated there was not a difference between the Iowa Assessments mathematics scores of grades 3-8 students enrolled in a multi-grade classroom and grades 3-8 students enrolled in a single-grade classroom. Gorrell's (1998) study examined the effect of traditional and multi-age classroom instructional setting on mathematics standardized test scores of fourth graders. Gorrell (1998) and

Baukol (2010) both found no significant difference in mathematics academic performance of students in multi-age classrooms compared to traditional classrooms. The findings of the current study support Gorrell (1998) and Baukol (2010).

Slavin (1986) reported that there was a significant difference in mathematics achievement between multi-grade classrooms and single-grade classrooms. Positive effects on student achievement were found. Kinsey (2000) compared the effect of multi-age and single-grade classrooms of mathematics achievement of second-grade students. Kinsey (2000) concluded outcomes on the California Achievement Test in math were greater for students who were enrolled in a multi-grade classroom than the students enrolled in a single-grade classroom. The findings of the current study are in contrast to the findings of Slavin (1986) and Kinsey (2000).

Conclusions

The subsections that follow provide conclusions drawn from the current research on the impact of student achievement of students enrolled in multi-grade and single-grade classrooms. Implications for actions and recommendations for further research are included. This section closes with concluding remarks.

Implications for action. Based on the findings of this study, school systems that employ multi-grade and single-grade classroom configurations should analyze student achievement scores to determine if this is the best practice for the school system. This study has implications for district administrators, building administrators, teachers, and parents. First, for district administrators, this study offers insight into the discussion of classroom configurations and how it could affect student achievement. It is also an option for schools with low enrollment at individual grades and in rural areas where there

is not enough student population to have single-graded configurations. District administrators should be aware of the training and support that building administrators and classroom teachers require in order to teach students in a multi-grade classroom and improve resources that those teachers would need in order to teach successfully and promote high engagement. However, district administrators must be cognizant of the data from the current study that reported minimal differences between multi-grade and single-grade student achievement. Given this data, it would be recommended that the Diocese should continue to use the classroom configurations. District administrators should continually analyze testing data to ensure that students in various classroom configurations are continuing to achieve. For written expression that showed a statistically significant difference between multi-grade and single-grade classrooms, the superintendent and principals should evaluate curriculum and look into what could be affecting that skill.

Based on the results of this study, written expression was the only subject that had a statistically significant difference in student achievement. The Superintendent of the Diocese should meet with teachers to review the curriculum to determine if there could be changes to the curriculum or pedagogy to obtain a different result between classroom configurations and student achievement in written expression. The Diocese should also continue to allow multi-grade classroom configurations for students. The results of this study indicated that students would achieve similar testing results in either classroom configuration. Lastly, the Diocese should keep teaching styles and curriculum constant and monitor the assessment data regularly.

Recommendations for future research. The purpose of this study was to determine if there were differences in student achievement between students enrolled in multi-grade classrooms and students enrolled in single-grade classrooms. Thus, the first recommendation that stems from this study would be to conduct the study over a six-year period to monitor the full impact of a group of students who advance from grade 3 to grade 8. In this way, achievement could be tracked for individual students over the course of the study.

A second recommendation would be that a study be conducted that includes public schools that also have multi-grade classrooms. Future research could compare differences in student achievement based on classroom type in both public and private schools. This recommendation is derived from the current study only representing Catholic schools, and the future study can include the public school data and whether the different configurations yielded different results.

A third recommendation would be to add the Missouri Assessments as the measure of student achievement to see if there are similarities or differences with the results. Another recommendation for future research would be to change the type of study that was conducted. A qualitative study could be conducted in which parents, teachers, or students are interviewed to find out their perceptions of multi-grade and single-grade configurations. It would also be a recommendation that the study be conducted across public and private schools. This sample size would help give a better comparison between the different types of schools regarding student achievement and classroom configurations.

Concluding remarks. Creating the best instructional practice for students is the goal of educators. Teaching approaches are ever changing to meet the needs of students. Educators in rural areas often think of ways to teach all students with limited resources and staff. One of the solutions has been to combine students in a manner they see fit with their resources. The question all educators ask is whether their approach will influence student performance.

The results of the present study provided information on the difference between student achievement on the Iowa assessments between students enrolled in multi-grade and single-grade classrooms. The current study explored this question and examined data to determine what was best for this Diocese. It is in the examination of these varied classroom configurations that educators could ultimately be equipped to support all students in their educational journey.

References

Aina, O. E. (2001). Maximizing learning in early childhood multiage classrooms: Child, teacher, and parent perceptions. *Early Childhood Education Journal, 28*(4), 219-224. Retrieved from ERIC database. (EJ631437)

Ansah, V. (1989). *Multi-grouping and academic achievement* (Information Analyses 070). Retrieved from ERIC database. (ED315163)

Bailey, G. J., Werth, E. P., Allen, D. M., & Sutherland, L. L. (2016). The Prairie Valley project: Reactions to a schoolwide, multiage elementary classroom design. *School Community Journal, 26,* 239-263. Retrieved from ERIC database. (EJ1104399)

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Baukol, D. (2010). *The impact of multi-age instruction on academic performance in mathematics and reading.* (Doctoral Dissertation). Retrieved from ProQuest Dissertations and Theses. (UMI No. 3434225)

Britt, P. (1997, November). *Inside the one room schoolhouse: A look at nongraded classrooms from the inside out.* Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Memphis, TN. Retrieved from ERIC database. (EJ415217)

Carter, P. (2005). The modern multi-age classroom. *Educational Leadership, 63*(1), 54-58. Retrieved from ERIC database. (EJ725901)

Chapman, M. L. (1995). Designing literacy learning experiences in a multiage classroom. *Language Arts, 72*(6), 416-428. Retrieved from ERIC database. (EJ513266)

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* New York, NY: Routledge Academic.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Los Angeles, CA: SAGE Publications.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 443–507). Washington, DC: American Council on Education.

Eames, F. H. (1989). *A study of the effectiveness of instruction in multi-age grading vs. traditional single-grade level organization on the reading achievement of fourth graders* (Master's thesis). Retrieved from ERIC database. (ED309388)

Eichacker, D. J. (2008). *Reading achievement and perceptions regarding the multi-age classroom environment* (Doctoral Dissertation). Retrieved from ProQuest Dissertations and Theses. (UMI No. 3333955)

Diocese X. (2018a). *Multi-grade schools.* Available by request from the Superintendent of Schools. In possession of the author.

Diocese X. (2018b). *Single-grade schools.* Available by request from the Superintendent of Schools. In possession of the author.

Elkind, D. (1993). Multiage grouping. In D. Sumner (Ed.), *Multiage classrooms: The ungrading of America's schools*. Peterborough, NH: Society for Developmental Education.

Gaustad, J. (1994). Nongraded education: Overcoming obstacles to implementing the multiage classroom. Special issue. *OSSC Bulletin, 38*(3-4), 1-97. Retrieved from ERIC database. (ED379744)

Goodlad, J. I., & Anderson, R. H. (1963). *The nongraded elementary school.* New York, NY: Harcourt, Brace, & World.

Gorrell, J. L. (1998). *A study comparing the effect of multiage education practice versus traditional education practices on academic achievement* (Master's thesis). Retrieved from ERIC database. (ED424008)

Grant, J., Johnson, B., & Richardson, I. (1996). *Our best advice: The multiage problem solving handbook.* Peterborough, NH: Crystal Springs Books.

Gutloff, K., & The National Education Association. (1995). *Multi-age classrooms. NEA teacher-to-teacher books.* Retrieved from ERIC database. (ED402035)

Kinsey, S. J. (2000) *The relationship between prosocial behaviors and academic achievement in the primary multiage classroom* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (UMI No. 9969115)

Kivunja, C., & Sims, M. (2015). Perceptions of multi-grade teaching: A narrative inquiry into the voices of stakeholders in multi-grade contexts in rural Zambia. *Higher Education Studies, 5*(2), 10-20. Retrieved from ERIC database. (EJ1075120)

Kolstad, R. K., & McFadden, A. (1998). Multiage classrooms: An age-old educational strategy revisited. *Journal of Instructional Psychology, 25*(1), 14-18. Retrieved from ERIC database. (EJ565426)

Lunenburg, F. C., & Irby, B. J. (2008). *Writing a successful thesis or dissertation: Tips and strategies for students in the social and behavioral sciences*. Thousand Oaks, CA: Corwin Press.

MacDonald, P. A, & Wurster, S. R. (1974). *Multiple grade primary versus segregated first grade: Effects on reading achievement.* Retrieved from ERIC database. (ED094336)

Mariano, L. T., & Kirby, S. N. (2009). *Achievement of students in multigrade classrooms.* Retrieved from https://www.rand.org/content/ dam/rand/pubs/working_papers/2009/ RAND_WR685.pdf

Mason, D. A., & Stimson, J. (1996). Combination and nongraded classes: Definitions and frequency in twelve states. *The Elementary School Journal, 96*(4), 439-452. Retrieved from https://www.jstor.org/stable/1001866?origin=JSTOR-pdf&seq=1#page_scan_tab_contents

McClellan, D.E., & Kinsey, S. (1997, April). *Children's social behavior in relationship to participation in mixed-age or same-age classrooms.* Paper presented at the Biennial Meeting of the Society for Research in Child Development, Washington, DC. Retrieved from ERIC database. (ED418771)

Miller, B. A. (1990, Fall). A review of the quantitative research on multi-grade instruction. *Research in Rural Education, 7*(1), 1-8. Retrieved from ERIC database. (EJ418886)

Miller, B. A. (1991). *A review of the qualitative research on multigrade instruction.* Paper presented at the Rural Education Symposium, Nashville, TN. Retrieved from ERIC database. (ED342563)

Mulryan-Kyne, C. (2004). Teaching and learning in multigrade classrooms: What teachers say. *The Irish Journal of Education/Iris Eireannach an Oideachais, 35*, 5-19. Retrieved from http://www.jstor.org/stable/30077492

Nye, B. A., Cain, V. A., Zaharias, J. B., Tollett, D. A., & Fulton, B. D. (1995, April). *Are multiage/nongraded programs providing students with a quality education?* Paper presented at the meeting of Create the Quality Schools Conference, Oklahoma City, OK. Retrieved from ERIC database. (ED384998)

Pavan, B. N. (1992, April) *School effectiveness and nongraded schools.* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA. Retrieved from ERIC database. (ED346608)

Penney, A. J. (2005). *Multi-age grouping perspectives* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses. (UMI No. NR07860)

Piaget, J. (1959). *The language and thought of the child* (3rd ed.). London, UK: Routledge & Kegan Paul.

Proehl, R. A., Douglas, S., Elias, D., Johnson, A. H., & Westsmith, W. (2013). A collaborative approach: Assessing the impact of multi-grade classrooms. *Catholic Education: A Journal of Inquiry and Practice, 16*(2) 417-440. Retrieved from ERIC database. (EJ1005662)

Ramrathan, L., & Ngubane, T. I. (2013). Instructional leadership in multi-grade classrooms: What can mono-grade teachers learn from their resilience? *Education as Change, 17*(S1), S93-S105. doi:10.1080/16823206.2014.865995

Ricard, R. J., Miller, G. A., & Heffer, R. W. (1995). Developmental Trends in the Relation between Adjustment and Academic Achievement for Elementary School Children in Mixed-Age Classrooms. *School Psychology Review, 24*(2), 258-270. Retrieved from ERIC database. (EJ587374)

Slavin, R. E. (1986). *Ability grouping and student achievement in elementary schools: A best-evidence synthesis* (Report No. 1) Retrieved from ERIC database. (ED348174)

Song, R., Spradlin, T. E., & Plucker, J. A. (2009). The advantages and disadvantages of multiage classrooms in the era of NCLB accountability. *Center for Evaluation & Education Policy: Educational Policy Brief, 7*(1), 1-8. Retrieved from ERIC database. (ED504569)

Stephney, V. (1970). Why multiage grouping in the elementary school? *The National Elementary Principal*, *45*(3), 21-23. Retrieved from ERIC database. (EJ014783)

Stone, S. J. (1996). *Creating the multiage classroom.* Parsippany, NJ: Good Year Books.

Trusty, E. M., Jr., & Beckenstein, S. (1996). *A comparative study of single-graded versus multi-graded classrooms*. Retrieved from ERIC database. (ED417014)

Veenman, S. (1995). Cognitive and noncognitive effects of multigrade and multi-age classes: A best-evidence synthesis. *Review of Educational Research*, *65*(4), 319-81. Retrieved from ERIC database. (EJ522378)

University of Iowa. (2015). *Forms e and f research and development guide.* Retrieved from https://itp.education.uiowa.edu/ia/documents/Research-Guide-Form-E-F.pdf

University of Iowa. (2018). *The clear choice for Iowa's educators: Accurately measure the Iowa core, growth, and readiness with a proven assessment.* Retrieved from http://itp.education.uiowa.edu/ia/documents/Iowa-Assessment-Informational-Flyer.pdf

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Welch, C. J., & Dunbar, S. B. (2014). *Measuring growth with the Iowa Assessments. ITP research series.* Retrieved from https://itp.education.uiowa.edu/ia/documents/Measuring-Growth-with-the-Iowa-Assessments.pdf

Yarborough, B. H., & Johnson, R. A. (2000). Nongraded schools: Why their promise has not been realized and should be reconsidered. *Contemporary Schools, 71*(3), 42-48. Retrieved from ERIC database. (EJ193476)

Appendices

Appendix A: Permission to Acquire Data

Consent Form to Participate in the Study

You are invited to participate in research conducted by Ja-Ronika Veldheer related determining whether there is a difference between the Iowa Test of Basic Skills reading, language, math, social studies, and science scores of students enrolled in a multi-grade level classroom and students enrolled in a single grade classroom. The data that will be collected in contained in the following example:

Student Number	School year	Student Grade level	ITBS Reading Total Standard Score	ITBS Language Total Standard Score	ITBS Math Total Standard Score	ITBS Social Studies Standard Score	ITBS Science Standard Score	Single grade or Multi-grade classroom
Example:								
100	2017-18	6	274	245	264	252	280	Multi

Your signature on this form grants me, as the investigator, permission to receive this data from you. The investigator will not use this for any other reason than those stated in the consent form without your written permission.

If you have read this form and have decided to participate in this project, please understand your **participation is voluntary** and you have the **right to withdraw your consent or discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled. The alternative is not to participate.** The individual privacy of your students and schools will be maintained in all published and written data resulting from the study.

If you agree to participate in this research, please sign your name on the line below and indicate the date of your permission.

[redacted] 4/17/2018

Date

Appendix B: Baker IRB Approval Letter

Baker University Institutional Review Board

July 19th, 2018

Dear Ja-Ronika Veldheer and Susan Rogers,

The Baker University IRB has reviewed your project application and approved this project under Exempt Status Review. As described, the project complies with all the requirements and policies established by the University for protection of human subjects in research. Unless renewed, approval lapses one year after approval date.

Please be aware of the following:

1. Any significant change in the research protocol as described should be reviewed by this Committee prior to altering the project.
2. Notify the IRB about any new investigators not named in original application.
3. When signed consent documents are required, the primary investigator must retain the signed consent documents of the research activity.
4. If this is a funded project, keep a copy of this approval letter with your proposal/grant file.
5. If the results of the research are used to prepare papers for publication or oral presentation at professional conferences, manuscripts or abstracts are requested for IRB as part of the project record.

Please inform this Committee or myself when this project is terminated or completed. As noted above, you must also provide IRB with an annual status report and receive approval for maintaining your status. If you have any questions, please contact me at npoell@bakeru.edu or 785.594.4582.

Sincerely,

Nathan Poell, MA
Chair, Baker University IRB

Baker University IRB Committee
Scott Crenshaw
Erin Morris, PhD
Jamin Perry, PhD
Susan Rogers, PhD