

2009

# A comparison study of Montana's intermediate and K-8 schools with regard to student performance on a criterion referenced test, incidence of at-risk behaviors, and perceptions of educators from both grade configurations

Robert Jeffrey Watson  
*The University of Montana*

Follow this and additional works at: <http://scholarworks.umt.edu/etd>

---

## Recommended Citation

Watson, Robert Jeffrey, "A comparison study of Montana's intermediate and K-8 schools with regard to student performance on a criterion referenced test, incidence of at-risk behaviors, and perceptions of educators from both grade configurations" (2009). *Theses, Dissertations, Professional Papers*. Paper 1274.

This Dissertation is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Theses, Dissertations, Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact [scholarworks@mail.lib.umt.edu](mailto:scholarworks@mail.lib.umt.edu).

A COMPARISON STUDY OF MONTANA'S INTERMEDIATE AND K-8 SCHOOLS  
WITH REGARD TO STUDENT PERFORMANCE ON A CRITERION REFERENCED  
TEST, INCIDENCE OF AT-RISK BEHAVIORS, AND PERCEPTIONS OF  
EDUCATORS FROM BOTH GRADE CONFIGURATIONS

By:

ROBERT JEFFREY WATSON

B.S., Montana State University, Bozeman, MT, 1993

M.Ed., University of Alaska, Anchorage, AK, 1998

Dissertation

presented in partial fulfillment of the requirements  
for the degree of

Doctor of Education  
in Educational Leadership

The University of Montana  
Missoula, MT

May 2009

Approved by:

Perry Brown, Associate Provost for Graduate Education  
Graduate School

Dr. John Matt, Committee Chair  
Department of Educational Leadership

Dr. Merle Farrier, Committee Member  
Department of Educational Leadership

Dr. William McCaw, Committee Member  
Department of Educational Leadership

Dr. John Lundt, Committee Member  
Department of Educational Leadership

Dr. David Aronofsky, Committee Member  
Legal Counsel Office

## Abstract

Watson, Robert, Ed.D., May 2009

Educational Leadership

A comparison study of Montana's intermediate and K-8 schools with regard to student performance on a criterion referenced test, incidence of at-risk behaviors, and perceptions of educators from both grade configurations.

Committee Chair: Dr. John Matt

Within the last ten years, middle schools have been criticized for focusing too much on social development while ignoring academics. Several large school districts around the country are transforming their middle schools into K-8 schools. The purpose of this mixed methods study was to compare K-8 schools with intermediate schools in the State of Montana. Quantitative research examined the relationship between school grade configuration and student performance on a criterion referenced test. Descriptive analysis examined the relationship between school grade configuration and incidence of at-risk behavior among students. Additional data was collected through 1-1 comparison interviews with educators from both grade configurations.

A factorial ANOVA was used to compare student performance on the CRT. Test data from 2005 included 11,690 students attending 58 K-8, 156 elemiddle, and 57 intermediate schools. Interacting variables of school size and socioeconomic status were also analyzed. Summary results found that there were a limited number of important and consistent differences among the means of the various sub-groups. The majority of the differences were related to size and socioeconomic status.

To compare incidence of at-risk behavior among students attending K-8 and intermediate schools, results from the 2005 Youth Risk Behavior Survey were examined. The sample included 1,971 students attending 11 intermediate schools and 13 K-8 schools. Interacting variables of school size and socioeconomic status were also considered. Summary results found that there were a limited number of YRBS differences between K-8 and intermediate schools. Important differences were found among non-school related behaviors.

To compare the perception of educators working in K-8 and intermediate schools, 1-1 comparison interviews were conducted with 22 professionals. Interview results showed similar responses for barriers to success and characteristics of successful students. Overall, educators agreed on advantages and disadvantages for both configurations.

Findings suggest that students attending K-8 and intermediate schools are more similar than different. From these results it would be erroneous to recommend one grade configuration over another.

Copyright 2009, Robert J. Watson

## Acknowledgements

Robert J. Watson  
Department of Educational Leadership  
University of Montana

I have grown both professionally and personally from this dissertation process. I would not have completed this project without the help and guidance of the University of Montana faculty as well as the support of my colleagues and family. I would like to formally acknowledge those who have been instrumental in my studies.

First, and foremost, I have appreciated the guidance and support from my committee chair, Dr. John Matt. From the beginning, Dr. Matt believed in the premise of my study as well as my abilities to finish. He encouraged me when I was frustrated and redirected me when I got off track. He made every effort to help me comprehend the recommendations of my committee and incorporate their suggestions into the final project. I also wish to thank the members of my committee: Drs. Farrier, McCaw, Lundt, and Aronofsky. Their ideas and advice helped to create an improved final product.

Second, I wish to thank the education professionals working on the front line who helped me to complete this project: the data experts from the Office of Public Instruction who developed a method to release the test scores, the administrators from around the State who shared their YRBS results, and the teachers and administrators from Missoula County who shared their perceptions in the one-on-one interviews. I would also like to thank all the outstanding professionals from the Missoula County Public Schools District who supported me through kind words, flexibility, and subtle reminders. Through their actions, they have shown that the MCPS district is truly dedicated to the professional and educational growth of their employees.

Finally, I will be forever indebted to my family. I would not have completed this process without their encouragement and love. I am thankful for Joann and Gordon who instilled the values of life-long learning and continuous self-improvement. I am grateful for Sheree who was there to pick me up when I fell down, as only a big sister can do. To my daughters, Emma and Isabel, I owe the gift of love and time for all the moments that they have spent without me over the past few years. Most importantly, to my loving wife Heidi, I owe an eternal debt of admiration. She has been there for every success and failure. She provided inspiration when I needed it and desire when I did not have it. As always, her love and self-sacrifice has been the glue which kept our family together and on-track to complete this process.

I dedicate this work to Sarah Jo. Though I only knew her for short time, she inspired me to learn, love, and live.

## Table of Contents

CHAPTER ONE - Introduction .....	1
<i>Background of the Study</i> .....	1
<i>Statement of the Problem</i> .....	3
<i>Purpose of the Study</i> .....	5
<i>Research Question</i> .....	5
<i>Significance of the Study</i> .....	6
<i>Definition of Terms</i> .....	7
CHAPTER TWO – Review of Literature .....	11
<i>History of Grade Span Configurations</i> .....	11
<i>The Early Years: Pre-1910</i> .....	12
<i>The Junior High: 1910 – 1960</i> .....	14
<i>The Middle School: 1960 – 1995</i> .....	17
<i>Revival of the K-8 Configuration: 1995 – 2007</i> .....	20
<i>Middle School Criticism</i> .....	24
<i>No Clear Purpose</i> .....	25
<i>Partial Implementation of Middle School Components</i> .....	28
<i>School-to-School Transition</i> .....	31
<i>Lack of Academic Rigor</i> .....	39
<i>Variables for Consideration</i> .....	46
<i>Primary Independent Variable</i> .....	47
<i>Interacting Independent Variables</i> .....	48
<i>Potential Dependent Variables</i> .....	57
<i>Relevant Comparison Studies</i> .....	62
<i>Large Scale and Meta-Analysis Studies</i> .....	63
<i>Transitional Grades: Fifth and Sixth</i> .....	68
<i>Eighth Grade Research</i> .....	74
<i>Longitudinal Studies and Influences on High School</i> .....	80
<i>Social and Emotional Research</i> .....	88
<i>Summary</i> .....	92
CHAPTER THREE - Methodology .....	95
<i>Mixed Method Research Design</i> .....	95
<i>Quantitative Methodology</i> .....	95
<i>Quantitative Research Questions</i> .....	96
<i>Quantitative Population</i> .....	97
<i>Quantitative Sample</i> .....	99
<i>Quantitative Variables and Level of Data</i> .....	100
<i>Null Hypothesis for Student Performance on CRT</i> .....	102
<i>Hypothesis for Incidence of Student At-Risk behavior</i> .....	102
<i>Quantitative Statistical Procedure</i> .....	105
<i>Threats to Validity</i> .....	106
<i>Qualitative Methodology</i> .....	109
<i>Qualitative Research Questions</i> .....	109
<i>Qualitative Population</i> .....	109
<i>Qualitative Participants</i> .....	110

<i>Qualitative Data Analysis</i> .....	114
<i>Role of the Researcher</i> .....	115
<i>Integration of Qualitative and Quantitative Data</i> .....	116
<i>Delimitations</i> .....	117
<i>Limitations</i> .....	120
CHAPTER FOUR - Results.....	123
<i>Analysis of Criterion referenced Test (CRT) Data</i> .....	124
<i>Data Categories and Sub-groups</i> .....	124
<i>Sample Size</i> .....	125
<i>Test for Homogeneity of Variance</i> .....	128
<i>CRT Reading Data as Defined by Configuration and SES Level</i> .....	128
<i>CRT Math Data as Defined by Configuration and SES Level</i> .....	131
<i>CRT Reading Data as Defined by Configuration and School Size</i> .....	133
<i>CRT Math Data as Defined by Configuration and School Size</i> .....	136
<i>Null Hypothesis Analysis for CRT Scores</i> .....	139
<i>Analysis of the Montana Youth Risk Behavior Survey Results</i> .....	156
<i>Sample</i> .....	157
<i>Data Handling Procedure</i> .....	159
<i>Comparison Tables and Percent Differences</i> .....	160
<i>Hypothesis Analysis for Incidence of At-Risk Behavior</i> .....	169
<i>Results of One-on-One Comparison Interviews</i> .....	175
<i>Participants</i> .....	176
<i>Interview Analysis</i> .....	179
<i>Organizational Structure</i> .....	180
<i>Barriers to Student Success</i> .....	184
<i>Characteristics of Successful Students</i> .....	187
<i>Advantages of Both Configurations</i> .....	193
<i>Disadvantages of Both Configurations</i> .....	200
CHAPTER FIVE - Conclusion.....	212
<i>Introduction</i> .....	212
<i>Summary of Null Hypotheses for CRT Scores</i> .....	213
<i>Summary of Hypotheses Related to the YRBS</i> .....	217
<i>Summary of Comparison Interviews</i> .....	219
<i>Integration of Quantitative and Qualitative Conclusions</i> .....	222
<i>Recommendations</i> .....	225
<i>Implications of the Results</i> .....	230
<i>Suggestions for Further Research</i> .....	234
References .....	236
Appendix A .....	249
Appendix B .....	252
Appendix C .....	256
Appendix D .....	258
Appendix E .....	260

## List of Tables

Table 4.1	<i>Overview of Data and Analysis</i> .....	124
Table 4.2	<i>Descriptive Statistics for Categories of Sample Schools</i> .....	126
Table 4.3	<i>Descriptive Statistics for Subgroups Defined by SES Level</i> .....	127
Table 4.4	<i>Descriptive Statistics for Subgroups Defined by Size</i> .....	127
Table 4.5	<i>Descriptive Statistics for CRT Reading Scores (SES sub-groups)</i> .....	129
Table 4.6	<i>Tukey/Kramer’s Post Hoc Test for CRT Reading Scores Defined by SES</i> ....	130
Table 4.7	<i>Descriptive Statistics for Math CRT Scores (SES sub-groups)</i> .....	131
Table 4.8	<i>Tukey/Kramer’s Post Hoc Test for CRT Math Scores Defined by SES</i> .....	133
Table 4.9	<i>Descriptive Statistics for CRT Reading Scores (school size sub-groups)</i> .....	134
Table 4.10	<i>Tukey/Kramer’s Post Hoc Test for CRT Reading Scores (size sub-groups)</i> .....	135
Table 4.11	<i>Tukey/Kramer’s Post-Hoc Test for CRT Reading Scores (size sub-groups)</i> .....	136
Table 4.12	<i>Descriptive Statistics for CRT Math Scores (size sub-groups)</i> .....	137
Table 4.13	<i>Tukey/Kramer’s Post Hoc Test for CRT Math Scores (size sub-groups)</i> ....	138
Table 4.14	<i>Tukey/Kramer’s Post-Hoc Test for CRT Math Scores (size sub-groups)</i> ....	139
Table 4.15	<i>Demographic Data for Population and Sample of YRBS Schools</i> .....	158
Table 4.16	<i>Demographic Data for Categories of YRBS Sample Schools</i> .....	159
Table 4.17	<i>Comparison Table: State of Montana and Research Sample</i> .....	161
Table 4.18	<i>Comparison Table: All Intermediate and K-8 Schools in Sample</i> .....	162
Table 4.19	<i>Comparison Table: All High SES and Low SES Schools in Sample</i> .....	163
Table 4.20	<i>Comparison Table: All Large and All Small Schools in Sample</i> .....	164
Table 4.21	<i>Comparison Table: High SES Intermediate and High SES K-8 Schools</i> ....	165
Table 4.22	<i>Comparison Table: Low SES Intermediate and Low SES K-8 Schools</i> .....	166
Table 4.23	<i>Comparison Table: Large Intermediate and Large K-8 Schools</i> .....	166
Table 4.24	<i>Comparison Table: Small Intermediate and Small K-8 Schools</i> .....	167
Table 4.25	<i>Demographic Information: Schools Participating in Interviews</i> .....	177
Table 4.26	<i>Demographic Data: Subjects Participating in Comparison Interviews</i> .....	179
Table 4.27	<i>Organization of Schools Participating in Comparison Interviews</i> .....	181
Table 4.28	<i>Summary of Advantages and Disadvantages</i> .....	211
Table 5.1	<i>Summary of Null Hypothesis Comparisons and Conclusions</i> .....	215
Table 5.2	<i>A Summary of Important Differences for the YRBS Results</i> .....	218

## CHAPTER ONE

### Introduction

#### *Background of the Study*

For the past 10 years, in schools across the country, there has been a subtle movement away from the middle school configuration towards a K-8 configuration (George, 2005; Pardini, 2002, Weiss & Kipnes, 2006). The Commissioner of Education in Colorado has endorsed the K-8 school as the next big idea in school reform (Yecke, 2006). In 1998, Cleveland Public Schools, serving 77,000 students, began to phase out middle schools and return to the K-8 model. A \$1 billion dollar bond was passed to reconfigure schools. Educators in Cleveland believed that middle schools were not adequately serving the needs of students age 10 to 14. Under the middle school model, sixth grade students in Cleveland not only demonstrated low test scores but also showed high absence and suspension rates (Pardini, 2002). After converting some middle schools into K-8 schools, Cleveland sixth graders demonstrated higher test scores. Students attending the K-8 schools scored 7.43 points higher on the reading proficiency test and 9.41 points higher on the math proficiency test given to all students. These differences represented a moderate effect size of .29 and .38 respectively (Poncelet, 2004).

Other school districts around the country have cited similar reasons for abandoning their middle schools in favor of a K-8 model. At a cost of \$530 million, in 2001 Oklahoma City started a seven year plan to renovate every school and change from middle schools to a K-8 configuration. Educators in Oklahoma City believed the conversion would help reduce the high school dropout rate. Similar situations have occurred in other cities. Citing parent dissatisfaction, in 2000, educators in Cincinnati



completed a five-year transition to K-8 schools. Everett, Massachusetts has also converted all five elementary school there to K-8 schools. In converting to the K-8 grade configuration, education policy makers in Massachusetts believed that the K-8 schools provided a better environment for middle-level education (Pardini, 2002).

When questioned about the reasons for the conversions, many school administrators cited perceived problems with the middle school model. For example, some have described middle school environments as large and impersonal. This environmental issue has led to poor test scores and discipline problems (Gewertz, 2004). On the other side of the argument, some believed that the criticism against the middle school model is misplaced. Some believed that middle schools are taking the blame for problems which were never created or affected by the middle school philosophy (Swaim, 2004).

In theory, middle schools differ from other schools, both in grade configuration, and instructional practices. In general, a middle school would contain several of the following organizational structures: interdisciplinary teaming, an advisory program, flexible blocks of time, exploratory courses, and transition programs. Unlike the departmentalized junior high, middle school teachers are divided into interdisciplinary teams with each member on the team serving the same group of students (*This We Believe*, 2003).

When schools implement these practices with fidelity and longevity, research has shown that test scores and environmental factors have improved (Lee & Smith, 1993; Mertens & Anfara, 2006). Many schools have fallen short on their commitment to implement fully the recommended middle school practices (Neill, 1999). Therefore, the

public criticism of the middle school model may be misdirected. It may be unfair to condemn an entire philosophy which was never implemented properly (Swain, 2004). Before abandoning middle schools completely, perhaps it would be appropriate to examine potential relationships between student performance and grade configuration.

#### *Statement of the Problem*

Throughout the history of American education there have been varying opinions about the best way to educate young adolescent students (Weiss & Kipnes, 2006). These philosophical shifts have resulted in changes to school configuration. Early in America there were small, rural schools based on a one-room, K-8 philosophy (Calhoun, 1983; Pulliam, 1987). As the nation became more industrialized and population became more urbanized, the young adolescent was educated in large junior high schools with an increased focus on preparing them for high school. In the 1960's the American culture shifted its focus. People became less interested in capitalism, and more interested in humanity and social issues. Educators answered the call with a shift in philosophy, as well. Starting in the late 60's and continuing into the late 80's the middle school configuration was implemented to help students develop both socially and emotionally (Jackson, Davis, Abeel, & Bordonaro, 2000).

Within the last ten years, middle schools have been criticized for focusing too much attention on social development while failing to develop the academic individual (Yecke, 2006). Now several school districts around the country are returning to the K-8 structure. Some districts have claimed that the K-8 learning environment provides a better atmosphere than the large, impersonal middle schools (Beane & Lipka, 2006).

Ironically, some of the same reasons were cited in the mid-1960s when junior highs were being transformed into middle schools (Calhoun, 1983; Cuban, 1992).

Other districts have claimed that the K-8 model improves student achievement as measured by standardized test scores (Abella, 2005; Offenber, 2001). A standardized test is typically used to measure student acquisition of knowledge and skills in areas such as reading and math. A standardized test typically does not measure the social and emotional development of an individual. If the ultimate goal of the middle school concept is to nurture the developmental needs of the young adolescent, it may be inappropriate to use a standardized test to measure the success of middle school. However, school boards and administrators are making decisions to abandon middle schools based on community dissatisfaction related to low test scores (McEwin, Dickinson, & Jacobson, 2004).

Some middle school experts believe that the decision to revert back to the K-8 model is based on faulty reasoning. The director of the National Middle School Association (NMSA), Susan Swaim, believes that the middle school concept works well when all the components are fully implemented. She also believes that some of the districts which have decided to return to K-8 schools did so without fully analyzing the status of middle school implementation (Swaim, 2004).

Regardless of the reason, many city districts have decided to return to the K-8 model (Gewertz, 2004). These decisions can be costly in terms of teacher training, building renovations, and curriculum adaptation (Pardini, 2002). With the implementation of the No Child Left Behind Act (NCLB), the federal government has increased academic accountability for schools (U.S. Dept. of Education, 2002).

Restructuring the configuration may be associated to meeting NCLB standards. The research related to the connection between configuration and student achievement is inconclusive (Paglin and Fager, 1997). However many districts are converting their middle schools to K-8 schools without a clear understanding of what influence this change will have on student performance (Beane and Lipka, 2006). Recent research has shown that the move to eliminate middle schools may be premature as there is no clear indication that the change to K-8 would translate to automatic increases in achievement (Weiss and Kipnes). Before investing tax dollars, it would be important to investigate how factors such as school configuration, school size, and socioeconomic status relate to achievement and social development for students in the middle grades.

#### *Purpose of the Study*

The purpose of this two-phase, sequential mixed methods study was to obtain statistical, quantitative results from a sample and then follow up with selected individuals to explore those results in more depth. In the first phase, quantitative research questions or hypotheses addressed the potential relationship among school grade configuration, student performance and behavior variables with middle-grade students in the State of Montana. In the second phase, qualitative interviews were conducted with educators to probe consistent student performance results by comparing characteristics of schools with various grade configurations.

#### *Research Question*

What was the relationship between the grade configuration of a school and student success for eighth grade students attending an intermediate level school, as compared with those attending a K-8 school?

### *Significance of the Study*

School boards and school leaders make decisions related to school expansion, closure, consolidation, and configuration. In this era of increasing accountability and changing enrollments it would be important for decision makers to know how school characteristics, like configuration, relate to student success. There exists a need for such research. David Hough (2004), editor of the *National Middle School Research Journal*, has stated that no empirical, large-scale study to date has examined the relationship between grade configuration and student achievement. Similarly, middle school researcher Kenneth McEwin (2004) has stated: “There is as yet almost no definitive research to answer the question of whether young adolescents in K-8 schools achieve better than they do in middle schools” (p.1).

The middle school research which does exist is focused on analyzing the practices that promote developmentally responsive schools. Because the middle school philosophy is grounded in the social/emotional context, research studies tend to examine issues related to the affective domain. There is very little middle school research related to student achievement and the research that does exist is inconclusive (Heller et al., 2003).

In the absence of research, school leaders across the country will be making decisions based mainly on local pressure from parents and school boards. These configuration changes have shown to be costly (Pardini, 2002). Before abandoning the middle school and returning to the K-8 model, it was important to examine how school configuration relates to student success.

### *Definition of Terms*

*Criterion referenced test (CRT).* A criterion referenced test measures an individual student's performance against a predetermined set of standards which are based on the curriculum (Blair, 2007). In Montana, the CRT is also known as the Montana Comprehensive Assessment System (MontCAS).

*Elemiddle school.* A school that includes both primary and middle grades where there is a specific focus on implementing a middle-level program for those students in the higher grade levels (Hough, 1995). For the purpose of this research, an elemiddle school was defined as a school which housed elementary and middle-grade students in the same building or on the same campus. In addition, an elemiddle was defined by a school where the educational program for the middle-grades differed from the traditional self-contained program offered at the elementary.

*Grade configuration.* The grade spans housed in any one-school setting. For example, a middle school could have a 5-8, 6-8, or 7-8 configuration (Calhoun, 1983).

*Intermediate school.* For the purpose of this research, the intermediate school was defined by any school which was organized in one of the following configurations: 5-8, 6-8, 7-8, or 7-9. In addition as defined by the State of Montana Accreditation Manual, the intermediate school must follow either the middle school philosophy or the departmentalized philosophy often seen in junior high settings (Montana Office of Public Instruction [OPI], Administrative Rules of Montana, Section 10.55.902, 2005a).

*K-8 school.* A school is said to have K-8 configuration when it serves students from kindergarten through eighth grade within the same facility or on the same campus

(McEwin et al., 2004). For the purpose of this research a K-8 school was defined as a school which had the characteristics of an elementary school for all grade levels.

*MontCAS.* This acronym was an abbreviation for Montana Comprehensive Assessment System. In the State of Montana, MontCAS was used to identify the NCLB criterion referenced test.

*School size.* The actual number of students housed in one school building. School size can be further be defined by the number of students at a particular grade level. For example, a large school could be defined as any school having at least 100 students at each grade level (Cotton, 1996). For the purpose of this research a large school was identified as one with more than 400 students (Mertens & Anfara, 2006).

*School socioeconomic status (SES).* For the purpose of this research, SES was defined by the percentage of students within a school who participated in the free or reduced lunch program. A school with a high percentage of students on free or reduced lunch was said to have a low SES. A school with a low percentage of students on free or reduced lunch was said to have a high SES (Barth, 2001). For the purpose of this research, a low SES school was defined as a school that had more than 40% of its student population who participated in the free or reduced lunch program. A high SES school was defined as a school that had less than 40% of its student population who participated in the free or reduced lunch program (US Department of Education, 2001, Section 1114, Title I, Part A).

*Student at-risk behavior.* For the purpose of this research, student at-risk behavior was defined as the occurrence of at-risk behaviors as self-reported on the Youth Risk Behavior Survey (YRBS) given to seventh and eighth grade students in the spring of

2005. The YRBS was a survey established by the U.S. Centers for Disease Control and Prevention to help monitor the prevalence of behaviors that not only influence youth health, but also put youth at risk for the most significant health and social problems that can occur during adolescence and adulthood (Montana OPI, 2005b). For the purpose of this research, the indicator of at-risk behavior was reported as a school average based on the percentage of students who indicated participation in the at-risk behavior.

*Student performance.* For the purpose of this research, student performance was measured by individual student scores on the Montana CRT given to eighth grade students in the spring of 2005. The CRT was the main component of the Montana Comprehensive Assessment System (MontCAS). The CRT test measured student performance in the areas of mathematics and reading comprehension. Results from the CRT were used to assess progress towards completing the goals defined by the No Child Left Behind Act (Montana OPI, 2006). Individual student performance was reported as a raw score which was based on the total number of questions answered correctly.

*Student success.* Although there were numerous methods which could have been used to determine student success in school, for the purpose student success will be defined by the following indicators: student performance on the MontCAS (Montana's CRT) and incidence of at-risk behaviors as measured by the YRBS.

### *Summary*

The proper configuration to educate middle-grade students has caused much disagreement among policy makers and professionals in the last 100 years. Configuration decisions have been directed by changes in teaching philosophy, such as the rise of the middle school concept in the 1960's. More recently, middle schools have



been the subject of criticism from parents and community members. The K-8 configuration has regained popularity in several large American cities. In examining the issue of grade configuration, there was a lack of research related to which type of configuration was best for student success. In addition, the middle school concept has been promoted as a way to address the social/emotional needs of the adolescent. However, some believed these needs could have been served better in the K-8 model. There was a lack of research which examined the relationship between grade configuration and student behavior. Therefore this research was timely and important.

## CHAPTER TWO

### Literature Review

Considering the various changes which have occurred in the field of public education, perhaps no other reform effort has ignited quite as much controversy and sparked as much interest as that of early adolescent education (Calhoun, 1983). The history of middle-grades reform has spanned 100 years and experienced two major renovations, starting with the change to a junior high model in the early 1900's and then a switch to middle schools in the 1960's. More recently there has been a movement to revive the original K-8 model. In order to examine the debate regarding middle-level education, the literature review presented here has been organized into four sections. First, the review summarizes the middle-grades reform process from a historical perspective. Second, in an effort to understand the revival of the K-8 movement, an analysis of the most frequent middle school criticisms is presented. Third, school size and socio-economic status, as well as other relevant variables, are considered. Finally, the research is presented which compares middle school student outcomes with those of K-8 students.

#### *History of Grade Span Configurations*

Educators and policy makers have long debated the best way to educate the young adolescent. The K-8 configuration dominated the American educational system for much of our early history. With a mostly rural population dependent on agriculture, the K-8 model was effective in educating the majority. In the 1950's our educational system experienced a shift created by a renewed focus on science and math. This change led to the development of the large comprehensive high school (Calhoun, 1983). With an

increased focus on consolidation, policy makers believed that the K-8 school was no longer an efficient way to prepare students for high school (Cotton, 1996). In 1920, approximately 80% of all high schools students had experienced the K-8 model. By 1960 the percentage had switched. Only 20% of students were still attending K-8 elementary schools, while the other 80% were attending junior high schools (Pardini, 2002).

By the late 1960's there was another shift in the practice of educating middle level students. Educators, dissatisfied with the impersonal feeling of the large junior high model, believed there was a better way to educate the young adolescent. The middle school concept was introduced as a way to address the unique social and emotional needs of the young adolescent (Jackson et al., 2000). The shift to middle school configuration was gradual. In 1971, approximately 69% of schools were organized in the junior high configuration. By 2000 that percentage had dropped to 22% of all schools. In 2000 the majority of schools (59%) were organized in a middle school configuration (National Middle School Association [NMSA], 2001).

### *The Early Years: Pre-1910*

The K-8 model still dominated the educational landscape at the end of the 19<sup>th</sup> Century (Mizell, 2005). The debate over appropriate grade configuration began in 1888, when Harvard President Charles Eliot suggested that America's education system could do a better job of preparing students for college. Eliot, speaking to education officials at a National Education Association meeting, believed that the 8-4 system was inefficient. He suggested that the educational program could be shortened and enriched. He also suggested that age of admission to college be lowered to 18 and college prep courses be introduced at a younger age (Calhoun, 1983). Eliot's theory did not fit with the goals of

the current K-8 system. In the early 1900's, the majority of the population ended education with the eighth grade. It is estimated that only one in every ten students graduated from high school in the early 1900's (Calhoun, 1983). Eliot proposed changing the focus of high school to include more college preparatory work. To achieve this goal, much of the work traditionally left to the high school was being pushed down to the middle-grades.

Encouraged by Eliot's remarks the Committee of Ten, comprised of education policy makers from both university and government, was established in 1893 with the goal of redefining high school education. The Committee agreed that the main purpose of secondary education was to prepare students for college. The Committee believed schools could do a better job in getting students ready for college. It encouraged a downward extension of subjects traditionally reserved for high school. To accomplish this goal, the Committee recommended expanding secondary education into a six year model. This would allow for more high school preparation courses to be introduced in the seventh and eighth grades (Calhoun, 1983).

With similar makeup to the Committee of Ten, in 1895 the Committee of Fifteen supported the recommendation to introduce secondary subjects in the seventh and eighth grades. The Committee also suggested a departmentalized structure for the middle-grades as the best way to teach secondary subjects. Although the Committee still supported the K-8 configuration model, it saw a need to establish different teaching strategies at the upper levels of the K-8 school. The Committee on College Entrance Requirements, a second committee established by the NEA in 1895, was the first to

suggest the 6-6 model as an appropriate grade span configuration for all schools (Calhoun, 1983).

For the next fifteen years several other national education committees were formed to help define the future of secondary education. The issues of grade configuration and middle-level teaching methods were discussed and the 6-6 grade span became the most popular reform idea. In 1918, the Commission on the Reorganization of Secondary Education completed what could be described as a belief statement for secondary schools. The “Cardinal Principles of Secondary Education” established several guidelines for schools. The Commission recommended the establishment of a 6-6 model with the last 6 years divided into a 3-3 configuration. The middle-grade years were to be organized in a junior high school, while the last three years would be reserved for the high school. The purpose of the new school was to introduce academically challenging courses through a departmentalized structure. In addition, the junior high would offer elective courses which would provide an environment where students could explore their interests and abilities (Anfara & Buehler, 2005).

#### *The Junior High: 1910 – 1960*

The first junior high school was established in Richmond, Indiana in 1895. The movement was slow to take hold. In 1910, two large U.S. cities reorganized into a 6-3-3 grade configuration. By 1915, it was estimated that 26 U.S. cities had established the 6-3-3 system. The number of junior high schools grew rapidly to 385 by 1920 and over 5,000 by 1960 (Calhoun, 1983).

Similar to the middle school movement, there is some debate regarding the original purpose of the junior high school. Cuban (1992) suggests that the junior high

school was established for the following reasons: (a) to eliminate time wasted in the later years of the K-8 school, (b) to prepare students for high school, (c) to provide students with prevocational choices, (d) to help students explore aptitudes and interests, (e) to provide conditions for better teaching, and (f) to provide students with socialization opportunities beyond the elementary years. However, the reasons some districts adopted the junior high model had little to do with the reasons outlined in Cuban's list. In many areas of the country, junior high schools were established to help solve overcrowding issues which resulted from population booms in post-war America. The adoption of a junior high school solved two issues related to overcrowding. Seventh and eighth graders were moved up and ninth grade students were moved down. These moves created more space in the elementary and high schools (Calhoun, 1983).

As the growth of our country progressed, the purpose of the junior high evolved. Changes in child labor and mandatory education laws provided a strong incentive for the junior high program to provide vocational education. The new purpose of the junior high focused on providing articulation strategies to support the transition between elementary and high school. In an effort to retain students, the junior high had to alter its program to fit the needs and interests of students (Calhoun, 1983).

The seeds of criticism towards the junior high model started early. A list created in 1945 suggested several concerns with the junior high model. The departmentalized structure was too regimented and too closely matched the high school model. The curriculum overemphasized the importance of subjects rather than teaching methods. Junior high teachers were inadequately trained. There was too much tracking of students

based on ability. Exploration programs were intended to be built on student interest, but had limited offerings (Cuban, 1992).

Research from the early 1960's provided more fuel for critics of the junior high model. Junior high schools were failing to provide education compatible with student interests. This finding was discovered by research that included direct observation. The competency of junior high teachers was also called into question, when comparison research found that elementary and secondary teachers were more effective. The junior high curriculum had not changed significantly since its inception. The junior high concept also suffered from lack of implementation. The stated purpose, a focus on the needs and interests of the early adolescent, was not evident in the 1960's junior high model. Finally, despite 50 years of history, the junior high school had still not solved the articulation problem. Originally designed to provide a transition from elementary to high school, the junior high school had articulation problems of its own. The transition from elementary to junior high school created a difficult adjustment for most students. The problems with articulation were inherent to the system and the junior high model was unresponsive to student needs related to transition (Calhoun, 1983).

The foundation of the middle school concept was born in the early 1960's amidst the increase in criticism of the junior high model, the national expectation to desegregate schools, and conditions of overcrowding. The goals and practices of the middle school were formalized in 1968. The middle school sought to: (a) bridge the gap between the elementary and high school, (b) offer individualized instruction to students with varied physical and mental abilities, (c) offer curriculum which provided for exploration of new experiences, and (d) foster continuous educational progress which included successful

articulation from one school to the next (Calhoun, 1983). Similar to the implementation of junior high schools, factors beyond educational philosophy fueled the adoption of the middle school model. A population boom required more space at the elementary level. Moving sixth grade to the middle school helped the issues with overcrowded facilities. As the new Carnegie unit placed high school-type expectations on ninth grade students, school administrators were encouraged by the reconfiguration recommended in the middle school concept (Cuban, 1992).

The first middle school opened in Michigan in 1950. Much like the junior high movement from the early 1900's, adoption of middle schools was slow to start. By the mid 1960's there were almost 500 middle schools nation-wide. Within six years that number had quadrupled. By 1978 there were more than 4,000 middle schools (Calhoun, 1983). The decline of the junior high school paralleled the rise of the middle school. Across the nation in 1970, junior high schools accounted for three quarters of the total number of schools serving seventh grade students. By 2003, that number had slipped to only one-third of the total schools serving seventh grade students. In 2003, there were more than 7,000 middle schools across the nation (Heller et al., 2003).

#### *The Middle School: 1960 – 1995*

*Philosophy.* In the early 1960s, educators were beginning to understand that the developmental needs of the early adolescents are unique. Their development is defined by rapid physical growth accompanied by extreme changes in social and emotional character. With the onset of puberty, children begin to build skills related to life-long personal development. They begin to form personal identity and develop a self-concept. By interacting with others they acquire social skills. They seek to gain autonomy and



frequently exercise their independence. They develop a personal character and begin to form a set of values (NMSA Research Summary, 1996).

Educators believe that the junior high model, with its content driven curriculum and its departmentalized schedule, was not structured to meet the developmental needs of the young adolescent. Students in this age group require an environment which fosters positive social interactions with adults and peers. They need to feel that there is an opportunity for meaningful participation in families, schools, and communities. They also need an environment with structure and clear limits. The environment should allow for expressions of creativity and physical activity. The middle school seeks to create an environment which focuses on the intellectual, social, emotional, moral, and physical developmental needs of young adolescent (Jackson et al., 2000).

As presented so far, issues of adolescent development are closely related to school environment and support structures. Of particular importance is the understanding of self-esteem and self-competence as viewed by the NMSA. In the opinion of NMSA, the young adolescent may experience varying levels of self-esteem. However, it is also believed that the adolescent possesses an adequate level of self-esteem, which increases over time. Conversely, the levels of self-competence generally decline over time. Self-competence is defined as a student's self-perception of personal academic and performance abilities (*This We Believe*, 2003).

*Middle school components.* Middle schools differ from junior highs not only in organizational structure, but also in instructional practices. In general, a middle school would contain several of the following organizational structures: (a) interdisciplinary teaming, (b) an advisory program, (c) flexible blocks of time, (d) exploratory courses,

and (e) transition programs. Unlike the departmentalized junior high, middle school teachers are divided into interdisciplinary teams with each member on the team serving the same group of students. The middle school schedule allows for common planning times where teachers can meet to discuss individual student progress. The middle school schedule is flexible. Teachers can decide how to structure time and how much to dedicate to each subject. Middle schools also have an advisory program where a small group of students is assigned to one adult for regularly scheduled meetings. The middle school schedule contains exploratory courses which cover subjects outside the traditional core classes. Finally, the middle school model would encourage transition programs where students can visit new schools prior to start of the year (NMSA Research Summary, 1996).

There is evidence that some schools adopted the middle school concept in name only. Because of overcrowding situations in K-6 elementary schools, some school districts transformed their junior high schools rather quickly into middle schools. The grade configuration of most junior high schools was grade seven to nine, while middle school proponents advocated for a six through eighth configuration. Moving sixth grade students to the middle school helped to ease the overcrowding situations in many elementary schools. In a rush to accommodate students, several schools adopted the middle school organizational structure without changing the instructional practices (Beane, 1999a).

*Middle School Instructional Practices.* As defined by the NMSA (1996), there are several key instructional practices evident in exemplary middle schools. First, instructional practices are varied and progressive. Teachers use literature circles to

enhance reading comprehension. They use thematic units to draw connections between subject matter and increase relevance. Students are expected to do less drill and practice and more activities related to developing higher level thinking skills. Second, curriculum in the middle school includes subject matter outside the traditional core areas. The middle school advisory program teaches students how to develop character and social relationships. The exploratory program exposes students to a range of academic, vocational, and recreational subjects. Third, middle school students are exposed to a variety of learning experiences. Classrooms environments contain less direct instruction, and more hands-on and student centered activities. Coursework is structured around group activities which require students to work collaboratively (NMSA Research Summary, 1996).

A final instructional practice encouraged by most middle school research is the concept of heterogeneous grouping. In an effort to foster cooperative learning, the exemplary middle school seeks to eliminate competitive structures. For example, tracking students according to math ability would be evidence of the content driven junior high model. The middle school would encourage teachers to group students with mixed ability. However, parents of advanced students do not favor mixed ability grouping and most middle schools have been pressured to offer advanced math classes (Beane, 1999a). Heterogeneous grouping is just one example of a middle school component that has recently come under the attack of public criticism (Yecke, 2003).

*Revival of the K-8 Configuration: 1995 – 2007*

In 1990, citing the growing number of junior high schools converting to middle schools, author Philip Vassallo predicted that the junior high would be completely extinct

by the year 2000. Using data from the National Center for Education Statistics, Vassallo projected that the number of middle schools would grow to over 20,000 by the turn of the 20<sup>th</sup> Century. As we know today, his prediction did not come true. Of the 14,000 schools which served middle-grades in 2003, one-third still carried the junior high name (Heller et al., 2003). Not only did the junior high survive extinction, but the K-8 school was also revived from the dead. More important in Vassallo's writing is a list of common concerns voiced by those who questioned the middle school concept. In this list was a common parent concern about children leaving the close-knit environment of the elementary school to enter the impersonal world of the middle school. Typically, elementary schools have smaller total population than middle schools. Without knowing it, Vassallo may have been foreshadowing the reform movement that no one expected; the revival of the K-8 school.

One major criticism of the junior high model was an overemphasis on academics while ignoring the developmental needs of the student (Calhoun, 1983). Some have criticized middle schools for the opposite reason. In other words, the middle school concept overemphasizes the social/emotional needs of the student while being lax on student achievement. In 1995, there seemed to be some data which supported the notion of a weak academic focus in U.S. middle schools. The Third International Math and Science Study (TIMSS) showed that U.S. students appeared to lose academic ground by the end of eighth grade. Even though fourth graders in the U.S. scored above the international average, eighth graders scored below the international average (Yecke, 2006).

When examining the recent revival of the K-8 school, the roots of this reform movement might be traced to authors Tucker and Coding (1998). Citing weak academic achievement, the authors labeled the middle school as the “wasteland of our primary and secondary landscape” (p. 153). These authors may have been the first to suggest that districts should abandon the middle school in favor of a K-8 configuration. Tucker and Coding (1998) provide several reasons to support their recommendation. When schools convert to the K-8 model, the feelings of anonymity in the middle school would be replaced by the feelings of community and social support found in elementary schools. The increased focus on academics found in the secondary school would begin to extend into the elementary grades. The older students in the K-8 setting would be able to tutor younger ones, a practice which would benefit both students. Finally, K-8 schools would serve the purpose of keeping children in their neighborhoods at a time when they are vulnerable and need the support of their families.

Since 1998, several large school districts around the country have begun the process of switching their middle schools to K-8 configurations. Among the first to make the switch were Cincinnati and Cleveland. Others, still in the process of switching included: Boston, Baltimore, Milwaukee, Newark, New Orleans, New York City, Oklahoma City, and Philadelphia. Some, like Chicago, resisted the middle school reform movement and had always maintained the K-8 configuration (Gewertz, 2004). With all this switching, one might assume that the number of middle schools would be on the decline. As of 2001, the middle school configuration still represented the majority of schools in America. Fifty-six percent of schools maintained the K-5/6-8 configuration.

Another 34% maintained the K-6/7-8 configuration. Only 10% of schools were using the K-8 configuration model in 2001 (Anfara & Buehler, 2005).

Much like the junior high and middle school, the purpose for districts adopting the K-8 reform model is up for debate. Districts that have made the switch cite a variety of different reasons for converting their middle schools to K-8 schools. Some of the stated reasons are as follows: K-8 schools maintain the community flavor of the elementary school; K-8 schools have shown promising results related to test scores and discipline statistics; and K-8 schools have helped relieve overcrowding in large middle schools. Beyond the stated reasons, there are other possible reasons for making the conversion. Adopting a K-8 configuration eliminates the problems associated with school-to-school transition. Students in K-8 schools stay in their neighborhood schools for a longer period of time, which can have positive effects on parent involvement and participation in after school activities (George, 2005).

The process of reforming the middle-grades, extending back over the past 100 years, demonstrates a classic example of the education reform pendulum. In 1910, under stress related to immigration, industrialization, and the need for a better college preparatory program, the U.S. converted from the K-8/9-12 configuration to the K-6/7-9/10-12 configuration. For the next 50 years, the junior high model became the accepted model for the middle grades. In 1960, dissatisfaction with junior high school's overemphasis on academics led to the development of the middle school philosophy which promoted a more humanized approach. By 1970, tension from overcrowded schools and a national expectation to desegregate the schools encouraged more than 2,000 schools to adopt the middle school model (Calhoun, 1983, Cuban, 1992). In 1989,

the middle school philosophy had a voice and a purpose as outlined in *Turning Points*. In 1995, the NMSA affirmed its goal to create a developmentally responsive school with the publication of *This We Believe*. In the same year scores from the TIMSS assessment were used as an example to criticize middle schools for ignoring student achievement. With the start of the 21<sup>st</sup> Century, under NCLB, the country experienced federal pressure to refocus school goals towards student achievement as measured by test scores. Starting in 1998, while questioning the test performance of their middle schools, many districts began the process of converting all middle schools to a K-8 configuration. From this timeline, it can be concluded that the middle-grades reform model has come full circle.

#### *Middle School Criticism*

The Hatch Amendment was enacted by Congress to support parents who wanted the ability to control content and teaching methods within the public school (Beane, 1999a). In a Congressional hearing related to the Hatch Amendment, supporters presented a list of 75 “satanic” practices present in the public schools. Some items included on the list were: moral education, sex education, and values education. Also present on the list of satanic teaching methods was the middle school concept (Beane, 1999a). Those who supported the Hatch Amendment were dissatisfied with the middle school concept because it encouraged schools to become involved in physical and emotional development of their students. Fundamentalists believe that this type of development is best left to the family.

In 1998, Barbara Byrd-Bennett, Superintendent of Cleveland Schools, declared that the District’s twenty five middle schools were failing. In responding to stakeholders, she cited several issues to rationalize her plan to convert all middle schools to K-8

schools. Middle school students had demonstrated poor academic achievement as evidenced by low sixth grade test scores. Once students entered middle schools, their absence and suspension rates soared. The Cleveland middle schools were too big, creating student feelings of anonymity. In the Superintendent's opinion, the District configuration, K-5/6-8, actually worked against the needs of the young adolescent (Herman, 2004).

When considering a K-8 conversion, educational administrators from around the country voice concerns, similar to those of the Cleveland Superintendent. In making policy decisions, administrators seek advice from educational research. The research used to criticize the middle school concept can be grouped into four general themes. First, the middle school movement suffers from a lack of purpose or clear direction. Second, despite the longevity of the middle school movement, there are still problems related to lack of implementation. Third, several critics have cited research related to the negative effects of school-to-school transition. Finally, most who seek to abandon middle schools believe there is a general lack of academic rigor inherent to the middle school philosophy.

#### *No Clear Purpose*

The middle school concept has been criticized for demonstrating a lack of direction and purpose. Members of the educational community disagree on the main components of a middle school. The level of implementation varies from one middle school to the next. The middle school concept appears to be suffering from an identity crisis, causing the public to ask: "What are middle schools for?" If the reasons to



establish a middle school are unclear, there is no specific reason to preserve the middle school model (Beane, 1999a).

In 1969, middle school advocates defined as the purpose of a middle school to humanize the education of the early adolescent (*This we believe*, 2003). To this end, middle school proponents sought to encourage an educational philosophy that stressed the needs and interests of the students. The focus was on educating the whole child, not merely the intellectual component (Calhoun, 1983). In 1995, the NMSA set forth their vision and purpose in a publication titled, *This We Believe*. The NMSA belief statement defined several guidelines for successful schools for young adolescents. According to these guidelines; successful middle schools provide curriculum that is relevant, learning and teaching strategies that are responsive to the needs of students, and organizational structures that support meaningful relationships.

From 1969 to 1995, the stated purpose for the middle school movement had changed. However, common to both stated purposes is the establishment of a school that is developmentally responsive to the needs of the young adolescent. Contrast this purpose with an alternative view proposed by a middle school critic. According to Yecke (2006), “[t]he middle school concept is the belief that the purpose of these schools is to create students who are...in touch with their political, social, and psychological selves; who eschew competition and individual achievement to focus on identity development...” (p.20). The differences in these three philosophy statements may demonstrate the evolution of a theory that had no clear direction from the start.

Another concern for middle school proponents may lie in the origin of the stated purpose. Most advocates agree that the primary purpose of the middle school concept is

to respond to the developmental characteristics of the young adolescent. These characteristics were founded in the work of G. Stanley Hall (1904). He was the first to describe adolescence as a stage of human development marked by storm and stress. Hall's theory of the early adolescent was generated from work with white, upper-class males. From this group, Hall generalized characteristics to all adolescents. The middle school philosophy of creating developmentally responsive schools is founded on a generalization used to describe the nature of the adolescent. This fact may create a dilemma for those who seek to defend the purpose of the middle school (Beane, 1999b).

Another possible source of confusion arises when the stated purpose for establishing a middle school differs from the actual purpose. In many districts around the country, the conversion from the junior high model to the middle school model was sparked by more than just establishing developmentally responsive programs. The hidden agenda for many school districts was less about student needs and more about the needs of the district (Cuban, 1992, George, 1988).

The civil rights movement of the late 1960's pressured many districts to desegregate their schools. Closing segregated junior high schools and opening new middle schools allowed districts to move sixth grade students out of neighborhood schools. By their nature, these neighborhood schools were often segregated. Rather than focus community attention on the contentious issue of desegregation, districts were able to promote the middle school concept as the latest educational trend. Desegregation was a bonus feature of the middle school package (George, 1988).

Beyond desegregation and perhaps more relevant for many districts across the country was the issue of accommodating shifting enrollments (Calhoun, 1983). In the

late 1970's the number of new students entering schools around the country started to decline (George, 1988). Rather than close schools, districts found that they could fill empty classrooms by moving ninth grade to the high school. In addition, they could add kindergarten to the elementary school by moving sixth grade up. Working in their favor was a convenient coincidence that the new middle school philosophy promoted a K-5/6-8/9-12 arrangement. Faced with the depressing possibility of closing schools, many educational administrators suddenly became middle school advocates. During this time of shifting enrollments, the number of middle schools across the country quadrupled (George, 1988).

#### *Partial Implementation of Middle School Components*

In an educational reform process, implementation problems can result for a variety of different reasons. Perhaps there is a lack of support from the administration, staff, or district. There may be an unclear vision to guide the reform, as might be the case in the middle school movement. Confusion as to why a change is made may also stall implementation of an educational reform. Whatever the reason, partial implementation has sparked criticism on both sides of the middle school debate. When challenged with poor academic achievement, middle school advocates frequently cite partial implementation as a possible cause for the low achievement scores. (Swaim, 2004) From the other side, middle school critics see the lack of implementation as a consequence of a defective philosophy. As the middle school concept has been around for almost 40 years, critics believe middle schools have had plenty of time to prove their worth (Yecke, 2006).

Regardless of the differing opinions concerning the lack of implementation, most agree that sufficient implementation is important to the middle school movement.

Research related to the level of implementation has shown that when schools pursue the middle school practices with strong commitment, students benefit both academically and socially (Beane, 1999b). Researchers Lee and Smith (1993) examined the effects of middle school practices on student engagement. Engagement was assessed by survey data which related to student involvement in academic work and incidence of at-risk behaviors. Looking at data on more than 8,000 students in 377 schools, they discovered that schools which were faithful to the middle school concept demonstrated increases in student engagement. For example, schools saw increases in students completing homework or coming to class prepared. They defined key middle school elements, which related to positive benefits for students. These elements included less departmentalization, more team teaching, and more heterogeneous grouping. The Lee and Smith findings demonstrate a common problem with research related to middle schools: the positive benefits are directly tied to the level of implementation. As such, it is difficult to generalize the success of the middle school concept to all schools when some are middle schools in name only (Beane, 1999b).

The level of implementation, as defined by the types of middle school practices adopted by schools, has been studied several times in the last 20 years. One of the first national studies which analyzed implementation practices in 1,753 schools was the John Hopkins Middle Grades Survey completed in 1988. This Survey identified the following middle school practices as key implementation practices: flexible scheduling, exploratory courses, interdisciplinary teams, common planning periods, and advisory programs. Upon completion of the study the following levels of implementation were reported: 15% of the schools were using some form of flexible scheduling; 35% of the

schools were offering exploratory courses; 37% of the schools had implemented interdisciplinary teams; 10% of the schools gave teachers common planning time; and 39% of the schools used some sort of advisory or homeroom program (Mac Iver & Epstein, 1993).

In 2001, another national research study analyzed the level of implementation regarding middle school recommendations. The study, entitled “America’s Middle Schools: Status and Progress,” surveyed 746 middle schools with the following grade configurations: 5-8, 6-8, and 7-8. The study collected levels of implementation for a variety of middle school practices: interdisciplinary teams, common planning time, exploratory course offerings, advisory programs, and instructional grouping practices. The results show some positive trends relative to levels of implementation: 77% of schools were using some form of team organization; 95% of the schools with interdisciplinary teams provided at least one hour of common planning time; 49% of the schools reported offering exploratory (student-interest driven) courses; and 48% of the schools maintained an advisory program. With the recent criticism related to heterogeneous grouping, it is interesting to note that 78% of the schools reported using ability grouping (non-heterogeneous grouping) in at least one core subject (McEwin, Dickinson, & Jenkins, 2003).

From 1988 to 2001, it appeared that the level of implementation had increased in most areas. However, there were still some inconsistencies related to what is expected and what was actually done in the middle schools. For example, heterogeneous grouping is encouraged as a developmentally responsive practice. However the vast majority (78%) of schools still maintained at least some tracking by ability in their practices. In

addition, practices such as advisories, which have been promoted as a method to strengthen student/teacher relationship, still remained as one of the least implemented practices from the middle school concept (McEwin, Dickinson, & Jenkins, 2003).

An important study to note is an implementation study which used Montana schools. Neill (1999) analyzed implementation practices for the 32 accredited middle schools in Montana. The principal and two teachers from each school were surveyed regarding their perceptions of their schools' level of implementation of middle school practices. The middle school practices were defined as essential elements of developmentally responsive middle schools by the NMSA. Some of these essential elements were: (a) comprehensive guidance and support systems, (b) an adult advocate for every student, (c) flexible organizational structures, and (d) programs and policies that foster wellness and safety.

Results of the survey showed that although the essential elements were perceived to be important to administrators and teachers, the degree to which the elements were actually implemented did not match the perceived importance. The survey allowed respondents to assess the level of implementation with one of the following descriptors: not implemented, partially implemented, moderately implemented, and majorly implemented. For each essential element, respondents described the level of implementation with the not, partial, or moderate descriptor. Given a similar scale, each of the essential elements was rated as very important by the survey respondents.

#### *School-to-School Transition*

The criticism of the middle school movement became part of main stream media in a recent article from *Time* magazine (Wallis, 2005). The article, *Is Middle School Bad*

*for Kids?*, suggests that sixth grade may be the wrong time for children to switch schools. Citing relevant research related to adolescent development, Wallis implies that preteens already experience many changes regarding their physical, emotional, and intellectual development. Expecting students to transition to a new school environment in the midst of all the other changes may be unnecessary and harmful (Juvonen, Le, Kaganoff, Augustine, & Constant, 2004).

In a meta-analysis study, Anfara and Schmid (2007) dissect the possible problems related to school transition. Overall, students exposed to school transitions may experience declines in academic achievement, motivation, and self-esteem (Alspaugh, 1998a; Eccles et al., 1993; Simmons & Blyth, 1987). School transition may also be associated with increased behavior problems (Abella, 2005). Although some believe that many of these problems are associated with puberty, the negative effects related to school transition may also lead to these issues. What creates the negative effects related to school transition? Some researchers have suggested a mismatch between the environment of the new school and the developmental needs of the student (Eccles et al., 1993).

Middle school advocates have recognized the importance of creating environments to help ease the stress related to school transitions. Establishing programs and practices to help students adjust to the new school is a targeted goal outlined in the NMSA policy statement, *This We Believe* (2003). Understanding the difficulties related to the transition from elementary to middle school is an important tool in helping students succeed. Typically, middle schools are larger in size and students become accountable to many teachers, rather than just one as experienced at the elementary level. In addition,

the academic demands are more difficult in middle school. There is an increased emphasis on individual student responsibility with less teacher support than students had experienced in the elementary setting (Herman, 2004). Students entering middle school also face the top-dog / bottom-dog effect. They experience social advantages or disadvantages as they transition from being the oldest students in the elementary school to being the youngest students in the middle school (Simmons & Blyth, 1987).

The environmental challenges experienced by students entering middle school are compounded by the onset of puberty for most students. In a national study which examined middle school research from the past 20 years, researchers proposed that the onset of puberty was a particularly poor time to begin a new phase of schooling (Juvonen et al., 2004). Experts in the area of transition suggest that school-to-school transition at sixth grade may be difficult for the young adolescent for a variety of reasons. First, middle school students experience a culmination of change; and the adjustment to one transition may be affected by the number of other changes experienced at the same time. For example, school transition may be compounded with puberty changes. Second, the transition from elementary to middle school could be described as a discontinuous change (Simmons & Blyth, 1987). Change is considered to be discontinuous when it is sudden or abrupt rather than gradual. Discontinuity is also defined by large differences between pre and post-transition periods. The elementary and middle school environment are generally considered quite different (Eccles et al., 1993).

Students attending K-8 schools may have an easier time adjusting to puberty changes for two specific reasons. First, K-8 students may have the advantage of being able to focus on physical and emotional changes related to puberty without the stress of



school transition. The focal theory of change suggests that it is easier for the adolescent to focus on one major transition at a time (Simmons & Blyth, 1987). Second, the K-8 setting may help to create an arena of comfort for students experiencing changes related to the onset of puberty. The arena of comfort theory is defined by the need for an individual to possess at least one area of life where they can feel relaxed and comfortable (Simmons & Blyth, 1987). Within this arena of comfort the individual can withdraw from life stress and become reinvigorated to help cope with future stress. For most children, this arena of comfort may exist within the home environment. However, for the at-risk youth, the home environment may create more stress than comfort. For these youth, the school environment may become the only arena of comfort. The transition from elementary to middle school may create life stress. The K-8 setting, without transition between fifth and sixth grade, may provide an arena of comfort for those experiencing stress related to the onset of puberty (Yecke, 2006).

Research concerning school-to-school transition has focused on the effects of school transition on student achievement, self-esteem, motivation, and drop-out rate. Researcher John Alspaugh (1998a) observed the negative effects of transition on student achievement. Examining seventh grade achievement for those students who transitioned from the K-6 elementary school to the 7-9 junior high, Alspaugh found a sharp decline in achievement levels during the grade level immediately following the transition. However, when he examined the seventh graders at K-8 schools, he found no drop in achievement from sixth to seventh grade.

In examining Alspaugh's findings, one might assume that the K-8 students would experience a more difficult transition to high school than the middle school students.

Knowing that middle school students had already experienced one transition, it is logical to assume that they would have an easier time with the high school transition, as compared with the K-8 students who did not experience the first transition. Researchers Simmons and Blyth (1987) have defined this theory as the stress inoculation hypothesis. By tracking students' progress through high school, Alspaugh (1998b) did not find evidence to support the stress inoculation theory. The students who had experienced only one school-to-school transition (K-6 then 7-12) had lower high school drop-out rates when compared to those who had experienced multiple transitions (K-6, 7-9, and 10-12). The group with only one school-to-school transition had a mean drop-out rate of 4.37%, while the group with multiple transitions had a mean drop-out rate of 6.68%. Through statistical analysis, Alspaugh determined this difference in the means to be significant at the .01 level. Although Alspaugh acknowledged school size and SES influence high school drop-out rates, he also discovered that larger school districts were likely to have more school-to-school transitions. He concluded that school size and school transitions may be working together to increase drop out rates.

In 2004, Wren examined the effects of school-to-school transition in a study involving 232 schools in Michigan. In looking at student achievement as measured by the percentage of students who passed the state-level achievement test, Wren discovered that the number of transitions had an adverse effect on student achievement. In general, the more grade levels served by a school the better the students performed on the test. Conversely, the more transitions students make throughout their school career, the worse they performed on the achievement test. (Students who had experienced only one transition passed the state test with a mean rate of 36.6%, while students with multiple

transitions had a pass rate of 21.9%) These findings would suggest that students would do better in the K-8 setting when compared with students attending a K-5 and 6-8 configuration. Wren also examined the combination effects of grade span and school-to-school transitions. In doing so she determined that the number of transitions was negatively correlated ( $-0.35, p < .01$ ) to student achievement. As the number of transitions increased, the rate of passing the state test decreased.

Beyond student achievement, other researchers have examined the effects of school transition on issues of motivation, self esteem, and behavior. In a position paper produced by the NMSA, educators working with middle school age students were encouraged to focus on more than just student achievement (*This We Believe*, 2003). According to the NMSA, middle level students possess unique characteristics defined by changes in physical, social, and intellectual development. For example, it is believed that the young adolescent experiences various levels of self-esteem throughout their growth through the middle school years. It is the challenge of the middle school educator to create the environment that is developmentally responsive to the unique needs of the young adolescent (*This We Believe*, 2003). As a result of this focus, researchers have examined variables beyond student achievement.

In a longitudinal study, Simmons and Blyth (1987) compared students attending K-8 schools with those attending K-6/7-9 schools. Using survey data, the researchers examined self esteem issues both before and after the school-to-school transition. In lower grades, students generally experience a growth in self-esteem as they progress through the grade levels. Students experiencing a school transition in the K-6/7-9 setting showed no growth in self esteem upon entering the seventh grade, where as their

counterparts in the K-8 setting showed self-esteem growth when moving from sixth to seventh grade. Upon entering high school, the K-6/7-9 students were still lagging behind in the self-esteem measure. Contrary to the stress inoculation theory, the K-6/7-9 students had a more difficult time with the high school transition than those students coming from the K-8 setting. Overall, Simmons and Blyth (1987) conclude that the K-8 setting, minus the extra school-to-school transition, was more effective in supporting students' growth into adolescence.

In 1993, a team of researchers in Michigan examined the apparent decline in student motivation after the transition into middle school. To start, researchers gathered survey data to examine the nature of the decline in motivation. It was assumed that the motivation drop was a natural characteristic of the early adolescent. Using classroom observation data, researchers discovered that the middle-level classroom was much different than the elementary classroom. In comparing the two settings, the middle-level classroom was defined by an increase of teacher control coupled with a decrease of teacher support. The apparent drop in motivation was not necessarily a consequence of adolescent development, but perhaps a product of the middle school environment (Eccles et al., 1993).

Seidman, Allen, Aber, Mitchell, and Feinmann (1994) examined the effects of school-to-school transition on early adolescent students in Baltimore, Washington D.C., and New York. Two types of grade configurations were analyzed: K-5/6-8 and K-6/7-9. Researchers focused on the students' self-system which was defined by self-esteem, motivation, self efficacy, social skills, and academic performance. Some factors of the

students' self-system were impacted negatively during the school transition. The factors which showed a decline were self-esteem and student achievement.

Levels of self-esteem were assessed by a survey given to students before and after the transition. Survey responses were assigned a point value based on the selections made by the student. Based on a four point scale, the mean response value before the transition was 3.07 and after the transition was 2.86. The mean difference between these values was determined to be consistent at the .001 level.

Student achievement was assessed by self-reported grade point average and showed a decline during the year following the transition. Grade point average was self-reported by the student as the school records for GPA were unavailable to researchers. Grade point average was assessed on a five point scale. Students self-reported a "one" if they believed their class work was generally unsatisfactory and they reported a "five" if they believed their class work was generally excellent. The average self-reported GPA before the transition was 4.14 and the average after the transition was 3.90. This mean difference was consistent at the .001 level.

These findings of self-esteem and self-reported GPA were consistent across ethnicity and gender. The problems associated with transition coupled with a decrease in the students' social support system led to increased feelings of school disengagement. Generally, disengagement from school at a young age is viewed as a precursor for dropping out of high school (Seidman, Allen, Aber, Mitchell, & Feinmann, 1994).

Given all the research related to the negative effects of school transition, one may question the value of maintaining a grade-span configuration which has school-to-school transition built into the system. However, there is some research which addresses the

positive effects of school transition (Anfara & Schmid, 2007). In general, this research could be summed up in following belief statement; transitions provide individuals with a fresh start. Through survey data it has been shown that transitions provide opportunities for students to develop effective coping skills. In addition, a new school can allow students to make life changes related to both academics and social behavior. In addition, transitions can allow students to create a new reputation in an environment free from judgment and preconceived notions. From the perspective of a student, most sixth graders expressed positive comments related to the transition to middle school. Even though most agreed that the transition was stressful, most also agree that the new school environment was more desirable when compared to the previous school.

#### *Lack of Academic Rigor*

Middle schools have also been criticized for being too child centered, focused only on the social and emotional development of the child. Critics believe that the middle school model has failed with regard to academic achievement and behavior control. A large contingency of parents believes that middle schools lack rigor and place little emphasis on the college prep curriculum. Many of these parents experienced the content driven junior high model and they believe their children should have the same experience (Beane, 1999a).

Many also believe that the middle school, which is defined by progressive teaching methods, ignores the basic skills. It is believed that middle school educators are obsessed with creating the correct developmentally responsive program. This obsession manifests itself into an endless search for the ideal self-esteem curriculum or the best arrangement of teams. These searches distract educators from focusing on academic

needs of their students. Closely related to this criticism is the belief that middle school educators have no clear academic expectations. With the depth of middle school training, middle school educators can explain the unique characteristics of the young adolescent in great detail, but many are unable to explain what middle level kids should know and be able to do (Bradley, 1998).

While serving as the Commissioner of Education in the State of Minnesota, Cheri Pierson Yecke has been a vocal critic of the middle school concept. She identifies five core practices that define the middle school philosophy. According to Yecke, these core practices work together to decrease student achievement in the middle school. First, by eliminating tracking and encouraging heterogeneous grouping, middle schools promote an equality of outcomes. Students are expected to be equal rather than some being more successful than others. Second, using strategies such as cooperative grouping, middle schools lessen the value of the individual. Third, middle school strategies support the notion that the group is more important than the individual. This concept has the effect of discouraging individual achievement. Fourth, by encouraging peer tutoring and cooperative learning, middle schools promote the expectation that advanced students will help those students who may struggle with learning. Finally, middle schools advance the notion that competition is negative and should be eliminated. Yecke believes that these practices demonstrate a general lack of academic focus inherent in the middle school philosophy (Yecke, 2003, 2006).

The type of criticism promoted by Yecke and Bradley leads one to question the research related to middle school student achievement. In examining such research, Hough (2004) asserts that there were more than 3,700 middle school studies conducted

between 1991 and 2003. Unfortunately, much of this research has a narrow focus limited to individual practices, such as the effects of teaming or the effects of common planning time. In a research summary prepared for the NMSA, Mertens and Anfara (2006) suggest that there is a need for more research which analyzes the effects of the middle school concept as a whole.

Another problem is that most of middle school research is focused on practices that promote developmentally responsive schools. Because the middle school philosophy is grounded in the social/emotional context, research studies tend to analyze measures of the affective domain. There is very little research related to student achievement and the research that does exist is inconclusive (Heller et al., 2003). In a review of the research, Pamperien (1997) analyzed seventeen studies which focused on the relationship between the middle school concept and academic achievement. Although the studies do not represent all the research related to middle school student achievement they help to describe the inconclusive nature of the existing research. Nine of these studies showed a positive relationship between the implementation of middle school practices and student achievement. The remaining eight studies concluded that there was no relationship between middle school practices and student achievement.

In defense of middle schools, NMSA President Sue Swaim suggests that the middle school movement cannot be blamed for problems related to poor implementation. In her opinion, the reality is that most schools which call themselves “middle schools” do so without implementing several of the recommended practices (Swaim, 2004). The level of implementation seems to be important when one considers the relationship to student achievement. In reviewing the literature, Pamperien (1997) found that the level



of implementation of middle school practices was positively related to student achievement.

A research summary developed for the NMSA, found similar promising results related to student achievement. Researchers Mertens and Anfara (2006) analyzed the results of several large scale studies. In general, they report there is a foundation of research which ties the implementation of the middle school philosophy to improved student academic and social development. In arriving at this conclusion, they cite five middle school studies completed between 1993 and 2003. Most of these studies involved several middle schools with varying levels of implementation. Although the details of each study are slightly different, the findings are similar. Implementation of middle school reform elements is positively correlated with student learning and achievement.

One particular study analyzed the relationship between the level of implementation and student achievement (Felner et al.,1997). The study examined over 15,000 students enrolled in 31 schools in the state of Illinois. To be part of the study, schools had to be at least partially committed to the “Turning Points” reform model. The “Turning Points” document was a formal list of recommendations published by the NMSA (Jackson et al., 2000). The Turning Points recommendations included: cross disciplinary student teams, student/teacher advisories, interdisciplinary instruction, and common planning time for teachers.

Schools were split into three groups depending on level of implementation of recommended middle school practices. Student achievement was defined by scaled scores on a math, reading, and language assessment. The mean difference between schools of high implementation when compared with schools of low level of

implementation was 50 points in mathematics, 61 points in language, and 28 points in reading. In general, schools with a higher level of implementation showed higher test scores. When considering longitudinal data across grades, those schools with little or no implementation showed an overall drop in test scores from grade six to eight. Schools considered at the highest level of implementation showed test score gains of 20 points from grade six to grade eight. Researchers also discovered a positive correlation between increases in the level of implementation and test scores ( $r = .51$  with  $p < .001$ ).

An assumption may be that this finding was limited by the pre-existing differences between schools. For example, it may be logical to assume that the high implementation schools all came from the more affluent areas, which would account for the academic gains. However, among the three groups -- high, middle, and low level of implementation -- each included schools of various size and demographic makeup. There were no dramatic demographic differences among the three groups (Felner et al., 1997).

Despite the promising findings of Felner (1997) and Mertens and Anfara (2006), middle school critics continue to find assessment data which may support the notion of a weak academic focus in the middle school model. In 1998, U.S. Secretary of Education Richard Riley shared the following observation: "We do a good job of teaching math and science in the early years, but we drift in the middle years and eventually fall behind the international standard in high school" (Herman, 2004, p. 3). In his statement, Secretary Riley was referring to the results of the Third International Math and Science Study (TIMSS) released in 1995.

In addition to the international comparisons, middle schools students were also showing declines in national comparisons with other grade levels. In looking at 1996

data from the National Assessment of Educational Progress (NAEP), researchers discovered that only 23% of eighth grade students were proficient in mathematics. Proficiency was defined by NAEP as possessing the fundamental skills most would agree are needed to be successful in high school (Cooney, 1998). Even worse for middle school proponents, the percentage of eighth graders scoring below the basic level was higher than the percentage of fourth graders scoring below the basic level (Cooney, 1998). Using similar NAEP data, researchers with the Southern Regional Education Board noted that eighth grade scores had increased slightly from 1996 to 2000. An assumption may be that the NAEP increase was a result of more schools implementing middle school recommendations, but the increase was more likely attributed to better performance at the fourth grade level (Heller et al., 2003).

Although NAEP scores do not look encouraging for eighth graders, the roots of the criticism of the middle school movement may actually be traced to the release of the TIMSS results in 1995. The TIMSS included one-half million students in 41 nations. In the U.S. 33,000 students in over 500 schools were assessed in the areas of math and science. Overall, the U.S. students' international standing was stronger in fourth grade than in eighth grade in both math and science (Calsyn, Gonzales, & Frase, 1999). In reaching this conclusion, student scores were averaged and then compared among the various countries participating in the study. Specifically, the average score for U.S. fourth graders was 16 points above the international average in math, while U.S. eighth grader average score was 13 points below the international math average. Of the 41 countries that participated, 27 countries scored higher than U.S. eighth graders in math and 16 countries scored higher than U.S. eighth graders in science (Yecke, 2003).

These results partially inspired a follow up study in 1999, the TIMSS-R (R for repeat). This study focused only on eighth grade achievement. Officials believed that the impressive U.S. fourth grade scores in 1995 were a result of curriculum changes which occurred in the early 90's. Assuming that the academic gains would carry from fourth to eighth grade, it was believed the TIMSS-R in 1999 would show eighth graders above the international average. The results were just the opposite.

In 1995, U.S. fourth graders were above the international average in math. In 1999, the U.S. eighth grade average was 22 points below the international average. There was also a drop in the average U.S. Science scores. In 1995, U.S. fourth grade average was 28 points above the international average. By 1999, the U.S. eighth grade average was nine points below the international average. These results led TIMSS officials to suggest that the pace of learning between fourth and eighth grade was slower in the U.S. than in other countries. Officials at the U.S. Department of Education concluded that when U.S. students are compared with other international students, the U.S. students do not necessarily start behind, but they fall behind during the middle school years (Calsyn et al., 1999; Yecke, 2003).

When considering international comparisons, several limitations are typically assumed. For example, an assumption might be that the comparison is flawed because countries outside the U.S. only test their brightest students. The officials at TIMSS conducted additional research to investigate these common assumptions. In the original TIMSS, an advanced math and science test was given to a sample of the top 10-20% of all students. Sampling procedures included selecting students who taking or had taken

advanced math and science classes. In this comparison, the average of U.S. scores was 59 points below the international average (Calsyn et al., 1999).

With many large districts around the country abandoning the middle school concept in favor of the K-8 configuration, the middle school philosophy has clearly been called into question. In an effort to justify reconfiguration plans, school administrators have provided the public with several criticisms of the middle school concept. First, there is an apparent lack of purpose related to the middle school concept. In many areas around the country, school districts adopted a middle school model to meet the needs of the district. This mismatch between intended purpose and actual purpose has led to implementation problems. Second, partial implementation of key middle school components has created problems with student achievement and staff buy-in. Third, the negative effect of school-to-school transitions conflicts with the middle school model, which promotes separate schools for the early adolescent. Finally, data from national and international assessments demonstrate a drop in academic achievement during the middle school years. These scores may add credence to the perception that there is a lack of academic rigor inherent in the middle school philosophy.

#### *Variables for Consideration*

The potential relationship between school grade configuration and student success is dependent on several possible variables. Beyond the independent variable of grade configuration, student success can be influenced by school size and school SES. In this study, student success was defined as student performance on CRT as well as incidence of at-risk behavior. All of these variables will be explored in this section of the literature review.

### *Primary Independent Variable*

*Grade span and school configuration.* Given the many possible school configuration models, this research focuses on three types; intermediate level, elemiddle, and K-8 configuration. The intermediate level school typically serves as few as two grade levels (7-8) and as many as four grade levels (5-8). In a 2000 analysis, the NMSA found that the majority of intermediate schools had one of the following configurations: 5-8, 6-8, 7-8, and 7-9. Of these four configurations, the 6-8 model was the most popular with 59% of the schools reporting this configuration. The 7-8 model was a distant second, used in 17% of the middle grade schools. (NMSA Research Summary, 2001) The NMSA did not report the number of K-8 schools as part of its sample. Perhaps this type of school was not recognized as a legitimate middle grades school. However, some K-8 schools strive to offer a middle school program for their upper grade levels (Hough, 2005). These schools are especially prevalent in rural states like Montana or North Dakota.

David Hough, editor of the National Middle School Research Journal, has coined the name “Elemiddles” to describe those schools that include both primary and middle grades. Different from the standard K-8, the elemiddle has a specific focus on implementing effective middle school components. In a 1995 study, Hough found that K-8 schools were able to implement the components of an exemplary middle school better than any other configuration. Hough explained that the elemiddle configuration was able to provide the best of both worlds; the key middle school components offered within the supportive and nurturing environment of the elementary.

Recently districts around the country, such as Cleveland, Philadelphia, and Oklahoma City are abandoning the middle school model in favor of the K-8 school configuration. William Moloney, who serves as the Commissioner of Education for Colorado, believes that the resurgence of the K-8 school is the next big idea in school reform (Pardini, 2002). However, the K-8 reform movement has been slow to take hold. According to the National Center for Educational Statistics, 11% of all elementary schools serve students in a K-8 configuration. However, in some states the percentage of K-8 schools is higher than the national average. In Montana, approximately 30% of all schools serving eighth graders have a K-8 configuration. In South Dakota the percentage of K-8 schools is 35% (<http://www.nces.ed.gov/ccd/schoolsearch/>).

#### *Interacting Independent Variables*

Offenberg (2001) conducted a three-year study in Philadelphia where he examined school configuration and eighth grade student achievement. Using standardized test scores, he found that eighth grade students attending K-8 schools outperformed eighth graders attending other types of schools. Some middle school advocates believed that SES could explain Offenberg's finding. In Philadelphia, the K-8 schools served students from mostly middle class families. However, when Offenberg controlled for SES level, he found that SES was only partly responsible for the differences in the test scores. In his analysis he discovered that K-8 schools typically had a smaller number of students per grade level. He believed that this fact might have contributed to better test scores (Offenberg, 2001). The Philadelphia study demonstrates the importance of two independent variables that need to be considered when researching student achievement; school size and SES.

*School size.* When researching a relationship between school configuration and student achievement, it is difficult to control for other variables, which may influence achievement. For example, most K-8 schools exist in rural settings with a small school size (Coladarci & Handcock, 2002). Could differences in achievement be attributed simply to the K-8 configuration, or does the smallness factor have anything to do with higher achievement?

Appropriate school size has long been the subject of debate among educational leaders and policy makers. School consolidation began in the 1950s, when the public believed that large schools were necessary for a strong curriculum to educate future scientists. In 1959, James Conant was the first to quantify what he believed to be an appropriate school size. Conant said that schools must have at least 100 students per grade level in order to be cost effective and sufficiently large enough to offer a varied curriculum (Cotton, 1996). Others have recommended that appropriate school size depends on the type of school. Sergiovanni (1995) recommended 200-350 students for a lower-grade school and 400-500 for a high school as the optimum size.

Cotton (1996) conducted a meta-analysis of research related to the relationship of school size and student achievement. In reviewing 31 research studies, she discovered about half of the research showed no difference in achievement for students in large schools as compared to students in small schools. The other half of the research showed that students in small schools have better achievement than students in large schools. In looking at the research, Cotton raises an important issue related to an extraneous variable. Assuming that small schools typically exist in rural environments, researchers were unable to control for the “ruralness” factor. Do students in small schools perform better



because of the school size, or is the performance related to the rural setting (Cotton, 1996)?

Research on school size generally favors smaller rather than larger schools. Sergiovanni (1995) explains that small schools are better at solving internal problems and more focused on academics. Small schools provide for more parent involvement which may explain why small schools have a higher degree of parent satisfaction. Although Sergiovanni admits that some of these same qualities can exist in large schools, he concludes that it is more difficult to foster this type of environment in a large school where students, parents, and teachers have fewer opportunities to build relationships which foster a positive, more personal environment.

Most education policy makers will admit that smaller is better, but it is difficult to justify maintaining several small schools in times of both shrinking enrollment and dwindling budgets. In terms of school facilities, the American education model has been built on the notion of economy of scale (Bickel, Howley, Williams, & Glascock, 2000). Education officials, from the state level down to the local school boards, believe that it is cheaper to educate more students in one large school rather than several small ones. This philosophy has been translated to mean consolidation is better. After years of operating under the economy of scale philosophy, schools nationwide have become larger. For example, more than 70% of high school students currently attend schools with more than 1,000 students (Allen, 2002).

However, larger does not always translate into better or cheaper. There is a possibility that a school can become too large and fall victim to a diseconomy of scale. Large schools can be more expensive to run if they include multiple levels of

administrators, supervisors, and support staff (Sergiovanni, 1995). A study in Texas concluded that small K-12 unit schools not only had better student achievement, but also were cheaper to run than large comprehensive high schools. Researchers in Texas discovered that it was cheaper to educate students in the small, K-12 unit school; over \$1,000 per student cheaper (Bickel, Howley, Williams, & Glascock, 2000). A study in Mississippi found that higher per pupil expenditure did not always translate to higher achievement. In comparing schools with various grade spans, researchers discovered that the schools spending less money per pupil had a higher average on a state achievement test when compared to schools spending more money per pupil (Cox, 1996).

A middle school study which focused on school size was conducted by Peterson in 2002. The schools used for this study had similar socioeconomic (SES) levels. Using achievement data in North Carolina, Peterson compared small schools (<350 students) with large schools (>350) to see if there were any differences in student success factors such as achievement and participation in extra-curricular activities. Peterson found that students in larger schools were more likely to participate in activities and performed better on a math achievement test than their counterparts in small schools.

Achievement was measured by North Carolina – End of Grade State test. Students in larger middle schools had an average math scaled score of 166.40 compared with an average scaled score of 163.76 for students in small schools. This mean difference was determined to be significant at the .0083 level. (The independent t-test for this comparison yielded a p-value of 0.003.) When he examined reading achievement Peterson found no significant difference between the small and large schools. The average reading scaled score for students from large middle schools was 156.22,

compared with the average score of 155.38 for students from small middle schools. This mean difference was determined to be not significant ( $p < .083$ ). The p-value for this comparison was calculated to be 0.269 (Peterson, 2002).

The debate on school size has been reborn with the recent focus on high school reform. Believing that small schools academically outperform large schools, many high schools have received grants to divide their large high schools into smaller learning communities. However, the belief that small schools are better may be based on an incorrect statistical interpretation. The assumption is based on data which contain a higher than expected proportion of small schools in the group of high performing schools. Researchers Wainer and Zwerling (2006) argue that educators and policy makers have ignored 50% of the statistical data related to student achievement. In looking at Pennsylvania test data it appears that there were several small schools among the top scoring schools in the state. However, Wainer and Zwerling found just as many small schools among the lowest performing schools in the state. A correlation test showed no relationship between school size and student performance. (Although the researchers did not report an r-value for this correlation, the graph of the data shows a flat regression line which would translate to an r-value close to zero.) From this, the assumption may be the data related to school size and student achievement are inconclusive.

*Socio-economic status.* The influence of SES cannot be ignored. There are an increasing number of students showing up at school with problems related to poverty. These students typically experience family structure problems, poor medical care, substance abuse, physical abuse, emotional handicaps, neglect, and the hopelessness that comes from poverty (Nelson, 2006). In general, the influence of poverty outweighs other

variables in terms of effect on student achievement. The poverty influence is strengthened in situations where poor children attend school with other poor children. Because U.S. schools are generally segregated by social class, this situation occurs frequently (Howley, 1999). Research has shown the negative influence of poverty on student academic performance (Barth, 2001; Howley, 1996; Roberts, 2002).

In education, the gap between high and low performing students is known as the achievement gap (Nelson, 2006). Typically, the achievement gap is influenced by school factors such as environment, curriculum, instructional practices, school size, and configuration. Socio-economic status has the greatest impact on the achievement gap (Howley, 1999). An example of the impact of SES can be seen in high school dropout rates. Nationwide, students from high SES families dropout of high school at a rate of 3% as compared with a dropout rate of 27% for students coming from low SES families (Barton, 2005). Frustrating for many educators; the level of SES is one factor that is not easily influenced by the actions of school staff. Understanding how poverty affects the achievement gap can help schools to minimize the influence of SES.

Factors that are linked to poverty include; low birth weight, lack of access to health care, hunger, and poor nutrition. These factors can influence educational achievement. The lack of parent availability and family mobility are other characteristics of poverty that can influence student achievement (Nelson, 2006). Parents who struggle with poverty typically spend more hours working, which means they have less time to read to their children or be involved in school activities. Family mobility can translate to students transitioning between schools frequently. Several research studies have shown

the negative effects of school transition with regard to student achievement (Simmons & Blyth, 1987; Alspaugh, 1998b; Wren, 2004).

The reading gap, conversation gap, and role model gap, are three other characteristics of poverty which have been shown to impact the achievement gap (Nelson, 2006). For a variety of reasons, children from low income families are typically not exposed to the types of activities which help higher income students prepare for school. For example, exposing children to meaningful reading materials is an activity which can greatly help to close the achievement gap.

The conversation gap has been defined as the difference between the conversations which exist in middle-class families and those which exist in low-income families. Similar to what they experience on the job, parents from low-income families tend to converse with their children through orders and directions. Conversely, parents from middle-income families tend to engage their children in conversations which build negotiation and reasoning skills.

Finally, the role model gap refers to the types of adults that children are exposed to in their daily lives. Outside of school, children of poverty are less likely to interact with adults that have education beyond high school. These three problems of the reading gap, the conversation gap, and the role model gap, help define a cycle of poverty that perpetuates from parent to child. Adults who work multiple jobs to support a family will have little time to read to children. As a result, their children will enter school already behind their peers. The achievement gap widens as the child progresses through school. Upon reaching adulthood, the child leaves school with very few high paying job

opportunities. This cycle of generational poverty creates a situation where it is almost impossible for students to transcend the achievement gap (Nelson, 2006).

*School size and SES.* Because school size and SES both influence student achievement, several researchers have examined these two factors in tandem. In 1996, Howley examined how school size effected the achievement of students with varying levels of SES in West Virginia. He discovered that small schools enhanced the achievement of poor students, while large schools enhanced the achievement of affluent students. Conversely, Howley discovered that large schools negatively influenced the achievement of poor students. Cotton (1996) reported a similar finding in her meta-analysis. She discovered that the negative impact on minority and low-SES students was greater in large schools. In examining high school drop out rates, Alspaugh (1998b) reaffirmed findings of previous researchers that SES level and school size negatively influenced drop out rates. However, when Alspaugh examined these two variables in tandem, he discovered that SES levels had less of an impact in small, rural schools.

Research related to school size combined with SES has been completed in a variety of different settings across the country. The “Mathew Project” was a school size and SES study completed in California, Alaska, and West Virginia. Assuming that poverty has the strongest influence over student achievement, researchers sought to investigate how school size interacts with poverty to influence achievement. In these states, small schools were found to mitigate the negative impact of poverty on achievement. The level of SES was directly correlated with the school size needed to maximize achievement. A community which was severely impoverished would require a

small school to contrast the influence of poverty (Howley, 1999). After discovering this correlation, the Mathew Project hoped to replicate the study in other areas of the country.

For two reasons, researchers from the Mathew Project discovered that Montana's unique school structure presented a perfect environment for replication. First, despite the many years of school consolidation in other parts of the country, Montana had maintained several small schools and districts. Second, despite the fact that Montana was mostly a rural state there was a high percentage of minority students (13%). This percentage was high when compared to an average of only 4% in other rural states. Using Montana schools, researchers determined that the small school factor helped to enhance student achievement in poor communities (Howley, 1999). Montana represented the model for equity. In the impoverished communities, student achievement had equaled or even outperformed students from more affluent communities. As a result, researchers recommended that Montana maintain the commitment to small schools and districts (Howley, 1999).

Other studies have investigated the effect of school size and SES on student achievement at the middle school level. In 2002, John Roberts studied 156 middle schools in South Carolina. Specifically, Roberts studied the relationship between the size of the middle school and student achievement, while controlling for SES. Student achievement was defined by the percentage of students scoring above the "basic" level on South Carolina Achievement Test. Roberts used a "partial correlation" statistical technique to control for the influence of SES. In examining the relationship between school size and test scores, Roberts found a negative correlation. As the size of the middle school became larger, the percentage of students passing the test grew smaller.

The correlation value for the language arts test was  $-.2089$  ( $p=0.009$ ) and the correlation value for the math test was  $-.1655$  ( $p=0.04$ ). When Roberts ran a standard correlation test, without controlling for SES, he found a positive correlation value ( $0.243$ ) for both math and language arts.

### *Potential Dependent Variables*

When comparing two types of grade configurations, the educational effectiveness of a particular configuration has been evaluated using a variety of different data. In looking at student data, there are many of factors which have a potential impact on student achievement (Guilfoyle, 2006). For example, when comparing middle schools with K-8 schools, researchers have analyzed student behavior data such as attendance and discipline (Cook, 2005). In addition, researchers have examined student performance data such as grade point average and drop-out rates (Alspaugh, 1998b; Weiss & Kipnes, 2006). Several researchers have also examined student perceptions related to school environment (Freeman, 2005; Moore, 1984; Poncelet, 2004). Each of these studies will be examined in more detail in the final section of this literature review.

Student performance on standardized tests has been frequently cited by school administrators as a reason for abandoning middle schools in favor of adopting a K-8 configuration (George, 2005; Pardini, 2002). In addition, scores from national assessments (NAEP) and international assessments (TIMSS) have frequently been used to criticize the effectiveness of the middle school concept (Beane, 1999a; Tucker & Coddling, 1998; Yecke, 2003). Therefore, the focus of this section will be to consider the history and importance of student test scores.



*Student test scores.* Although the possible measures to evaluate student success are abundant, the public pressure created by the NCLB act has focused attention on standardized tests (Guilfoyle, 2006). The goal of standardized testing and NCLB assumes that schools (teachers and students) will take their jobs more seriously when faced with incentives and threatening punishments (Nichols, Glass, & Berliner, 2005). The increased focus on testing has influenced educational research. Many of the recent studies, which compare schools based on configuration, have used student test data as the primary dependent variable (Comer, 2006; Cook, 2005; Freeman, 2005; Peterson, 2002; Poncelet, 2004; Weiss & Kipnes, 2006).

The concept of high-stakes testing has not come directly from NCLB. The public pressure related to testing can be traced to a report released in the mid-1980's called *A Nation At Risk* (Yecke, 2003). This report was one of the first to label the American public education system a failure. The public concern generated from this report led to an increased focus on standards-based education (Tucker & Coddling, 1998). "What students should know and be able to do," became the new marching order for educators. This objective led to high school exit exams, perhaps the first high-stakes test. In 2001, the NCLB Act was launched with a single goal, all students can succeed at a proficient level; and a single system of implementation, standards based assessment and accountability (Nichols, Glass, & Berliner, 2005).

The NCLB system of accountability can be summarized in six steps. First, states are required to identify curriculum standards. Next from these standards, they should create a state-wide assessment system to monitor student progress towards achieving the standards. Then, schools and districts must publish report cards showing performance on

the state assessment, disaggregated by student subgroups. In addition, as part of this reporting system, each state is required to create a system of labels to publicly communicate to the success or failure of the school. All schools should also develop a plan to make sure all students are making adequate yearly progress towards achieving the level of proficiency by the year 2014. Finally, the system of accountability should include sanctions for those schools that do not make progress towards 100% proficiency (U.S. Department of Education, 2002).

Parts of the NCLB accountability system have created some positive accomplishments, although other parts have created problems (Wallis & Steptoe, 2007). From the perspective of many, perhaps NCLB greatest impact has been an accountability system which has shed a public light on the learning needs of our less than privileged students (Jennings & Rentner, 2006; Wallis & Steptoe, 2007). The NCLB requirement of disaggregating the test scores by student subgroups has encouraged educators and communities to focus more closely on the achievement gap (Guilfoyle, 2006).

However, the accountability system has created “unintended consequences,” such as narrowing the curriculum towards a focus on reading and math. In a survey of school administrators, from 1999 to 2004, the time spent on reading comprehension instruction had increased by 40 minutes per week, while the time spent on other subjects had decreased. Over the same year span, the instructional time spent on social studies had decreased by 17 minutes and science by 23 minutes per week (Wallis & Steptoe, 2007).

In analyzing the focus of NCLB, the presumption would be that assessing math and reading was the only method to evaluate the effectiveness of a school. However, in measuring student success, NCLB leaves many other areas off the list. For example,

student aptitude in subjects such as history, art, and physical education is not measured by the NCLB accountability system. The success of a school can also be evaluated by examining school culture and student well-being, which can impact student achievement and learning (Guilfoyle, 2006). Therefore, going beyond the NCLB accountability model would be important in designing a comparison research study.

*Student risk behavior.* As stated earlier, several studies have used student test scores to assess the effectiveness of a school configuration model. However, most have looked at other factors beyond test scores. For example, Abella (2005), Comer (2006), and Freeman (2005), all examined student behavior data as it relates to attendance and discipline. (The details of their findings are discussed in the last section of this literature review.) As the intention of the present study to focus middle-level students, it would be important to examine factors of success for students of this age level.

The NMSA has prescribed the method of success for middle-level educators; a focus on creating a developmentally responsive school environment (*This We Believe*, 2003). To accomplish this task, the NMSA has outlined the following developmental needs which should define the environment: (a) positive social interactions with adults and peers, (b) structure and clear limits, (c) physical activity, (d) creative expression, (e) meaningful participation in school community, and (f) opportunities for self-definition (NMSA, 1996).

According to the NMSA experts, schools that have an environment with a focus on meeting these developmental needs will create a successful learning experience for all students. Other researchers have shown the positive correlation that exists between implementing the NMSA recommendations and increased student learning (Mertens &

Anfara, 2006). When the school environment ignores the needs of the young adolescent, a lack of success can occur. Not meeting the developmental needs can lead to feelings of alienation and loss of self-esteem. As a result, in order to cope, students may turn to destructive methods such as engaging in at-risk behavior involving delinquency and drugs (NMSA, 1996).

With the knowledge of the importance of school environment and its relationship to the success of the middle-level student, it would be important to examine student behavior data in schools serving middle-level students. In a nation-wide poll, targeting seventh and eighth grade students, researchers attempted to assess school environment issues for the middle-level student (Bushaw, 2007). The poll was conducted by the National Association of Secondary Principals in February/March of 2007. There were 1,814 (seventh and eighth grade) students responding to the survey, with the majority attending middle schools, 13% attending K-8 schools, and 6% attending high schools. When asked about the most common student behaviors in their schools they cited a variety of delinquency issues. The most commonly problem was bullies and gangs, cited by 15% of the respondents. The next most common problem was inappropriate behavior by other students (14%), closely related to social pressures to engage in inappropriate behavior (12%). Other significant issues for the middle-level students were fights (7%) and illegal activities (7%).

Another national survey conducted in 2005 assessed student at-risk behavior for the middle-grades. The 2005 Youth Risk Behavior Survey (YRBS) was given to sixth, seventh, and eighth graders in eleven different states across the country. Created by the U.S. Centers for Disease Control, the YRBS focuses on risk behaviors that contribute to

unintentional injury and violence, tobacco use, and alcohol use. Because several of the YRBS questions relate directly to delinquency behaviors, the results are relevant to this research. With regard to behaviors that contribute to violence: the percentage of students who had ever carried a weapon, defined as a gun, knife, or club, was 42.6% and the percentage of students who had ever engaged in a physical fight was 60.7%. With regard to behaviors that involve drugs or alcohol: the percentage of students who had ever drunk alcohol, other than a few sips, was 48.7% and the percentage of students who had ever used marijuana was 17.3% (Montana Office of Public Instruction, 2005b).

Assessing the relationship between grade configuration and potential variables of student success can be challenging. The pressured related to NCLB has established the importance of examining student test scores. However, student scores are not the only measure of student success. The recommendations from the NMSA emphasize the importance of assessing the student behaviors to evaluate the effectiveness of a school.

#### *Relevant Comparison Studies*

Considering the relationship between grade span configuration and student success in the middle grades, the research is limited (Coladarci & Hancock, 2002; Yakimowsky & Connolly, 2001). Over the years, limited research has been conducted comparing K-8 schools to junior high schools. More recently there has been some research comparing K-8 schools to middle schools. In a summary of research conducted in 2004, Herman found no conclusive evidence which would show that middle level students attending K-8 schools do better than similar students in middle schools. When considering almost 4000 studies conducted between 1991 and 2002, Hough (2005)

concluded that middle school researchers had avoided the topic of comparing student success based on grade configuration.

Historically, it would seem that for quite some time administrators and policy makers have been questioning which grade span is best for the education of the young adolescent (Calhoun, 1983). Evidence of this dilemma can be seen by the progression of configuration changes. The practice of testing various grade spans has resulted in a few relevant comparison studies which have examined the differences in student outcomes between those attending K-8 schools and those attending junior high and middle schools. The research presented here represents all of the relevant comparison studies currently available. Considering the current research related to grade span, many researchers have concluded it would be difficult to find that one configuration better than another (Epstein, 1990; McEwin et al., 2005; Paglin & Fager, 1997; Swaim, 2004).

#### *Large Scale and Meta-Analysis Studies*

One of the earliest large scale comparison studies was conducted in 1989 by the Center for Research on Elementary and Middle Schools at John Hopkins University. This national survey helped researchers to compare practices and goals among schools that served middle-level students. The study included a variety of different grade spans, all of which served seventh grade students: K-8, K-12, Middle Schools, Junior Highs, and Junior/Senior Highs. In terms of enrollment, researchers discovered that K-8 schools generally served a smaller number of students than middle schools. The K-8 school had an average enrollment of 42 students per grade level and the middle school had an average of 180 students per grade level.

The John Hopkins study also examined the differences in core goals and practices. Principals were asked to rank seven goals in order of importance for their schools. All grade configurations ranked “basic skills or mastery of subject matter” as the most important goal for students. The “citizenship” goal was ranked last by all schools. Other goals important to middle school principals were: personal growth and work habits. The principals of K-8 schools placed less importance on personal growth and more importance on higher level skills (Epstein, 1990). From these findings it was concluded that middle school principals were more interested in developing practices to strengthen personal and social development, whereas K-8 principals placed more emphasis on academic skill development. In the current state of accountability, middle schools have been criticized for overemphasizing personal development and a lack of academic rigor (Yecke, 2003, 2006).

Using the data from the John Hopkins study, in 1990 researcher Joyce Epstein reached the following conclusions regarding the importance of grade span configurations. First, it would be difficult to conclude that one grade span or program is best for all middle-level students. Second, most middle schools surveyed had yet to implement all of the practices recommended for the social and emotional development of the young adolescent. All the principals surveyed had plans to implement the recommended practices within the next five years. Despite good intentions, the lack of implementation is seen today as a major flaw in the middle school movement (Look, 2001). Finally, based on the survey data, Epstein proposed that debate over appropriate grade span was misguided. She encouraged middle-level educators to focus on practices rather than determining appropriate grade span.

In conjunction with the Northwest Regional Education Laboratory, Paglin and Fager (1997) conducted a meta-analysis study related to middle-level grade configuration. In analyzing various school districts that had reconfigured grades, they discovered the following as common reasons for changing: school reform, increasing or decreasing enrollment, financial concerns, consolidations, community values, and geographical considerations. They also identified several factors that one must consider when making a grade reconfiguration decision. Such factors include: transportation, parent involvement, course offerings, effects on student performance, loss of neighborhood school, and school redesign. Another important concern raised by Paglin and Fager was the increase of school-to-school transitions.

In a careful examination of all available research related to grade span, Paglin and Fager (1997) reached the following conclusions: There was no research available which tested the cause and effect relationship between grade configuration and academic achievement. Few studies controlled for factors such as school size, socio-economic status and teacher experience. The current research was not generalizable and the results of several studies were varied and did not translate into clear guidelines. Overall Paglin and Fager proposed a familiar conclusion: practices and programs were more important than grade span.

In 2004, the Rand Corporation published a summary of middle school research in an effort provide a “state of America’s middle school” and identify challenges facing the middle school concept (Juvonen et al., 2004). The Rand report identified student achievement as one of the largest challenges facing the middle school movement and the Third International Math and Science Study (TIMSS) as the piece of evidence most cited



in the criticism of middle schools. In addition, Juvonen acknowledges that middle schools are often blamed for student disengagement from school which is seen as a precursor to dropping out of high school.

In a review of 20 years of middle school literature, Juvonen acknowledges that the existing research is limited and there are gaps of information in the research. Having acknowledged the limitations, Juvonen outlines several conclusions which can be drawn from a review of all literature. Changes in school configuration typically have very little to do with education or development concerns. Typically schools are reconfigured because of changes in the labor market and building capacity. With regard to conditions for learning, Juvonen found that middle school students in America had a negative perception of their school when compared to other students on an international level. Finally, based on research related to adolescent social development, Juvonen concludes that the onset of puberty is not an ideal time to begin a new phase of schooling.

In an effort to improve the state of America's middle-level education, the report from the Rand Corporation provides the following overall recommendation: schools should explore alternative structures that allow for a reduction in the number of transitions across grades K-12. Moreover, school districts should develop programs to facilitate the alignment of curriculum, goals, and instructional practices across grade levels. Because of the depth and breadth of this particular study, the alternative configuration recommendation is not taken lightly by education policy makers. Many have translated this to mean that middle schools should consider converting to a K-8 model.

In a 2004 NMSA study, McEwin, Dickinson, and Jacobson examined the programs and practices which exist in K-8 schools. Having just completed a similar study of middle schools in 2003, the researchers were in a unique position to compare the two configurations based on survey data. The middle school study included 746 schools, while the K-8 sample included 101 schools. The comparison study hoped to answer the following questions: Do middle level students in K-8 settings experience school differently? Are middle level programs in K-8 schools effective? What are the differences in basic and elective course offerings in K-8 schools versus those offered in middle schools?

In comparing the survey data the researchers arrived at the following conclusions. Middle school practices were not as prevalent in the K-8 setting. For example, interdisciplinary teaming was present in only one-third of the K-8 schools. By comparison, 77 percent of the middle schools used interdisciplinary teaming. In the K-8 school more time was spent on the core subjects and less time on elective classes. Students in K-8 schools spent up to 90 extra minutes per day in subjects such as math and language arts. Finally, the researchers discovered an interesting finding related to the administrative perception of middle-level students. When asked to consider the best educational environment for middle-level students, 84 percent of the K-8 principals believed that the middle grade students should be placed in a separate middle school.

Although NMSA study examined specific programs and practices, it is limited to a comparison of qualitative data. The researchers admit that there is still a need for an empirical, large scale study which examines the relationship between grade span and student achievement. Without a large scale achievement study, they conclude there is no

definitive answer as to which configuration is better for students. However, based on their comparison of survey data coupled with their understanding of the middle school concept, the authors conclude that a separately organized middle school is better for students. Others would describe this assumption as a leap of faith without the presence of a research base to support middle level practices (Heller et al., 2003).

In the absence of a large scale study comparing grade span with respect to student achievement, the next three sections of this literature review will examine small scale student achievement studies. The studies have been divided by grade level. The first section will focus on research which is related to the transition grades from elementary to middle school. The second section focuses on studies concerning the eighth grade, another important transition year for both K-8 and middle school. The third section focuses on longitudinal studies which examine the long term effects of grade span configuration across multiple grade levels.

#### *Transitional Grades: Fifth and Sixth*

Although school grade configuration can change for a variety of reasons, many times the change is based on issues with enrollment, finances, or facility. Many school districts have added or removed fifth and sixth grade from the middle school depending on the needs of the district. In the process of converting junior highs to middle schools, many districts began by moving sixth graders out of the elementary schools. As justification for the reconfiguration, middle school experts would argue that sixth graders were physically and psychologically closer to seventh and eighth graders (Cromwell, 1998). At times these placement decisions are made without consideration for the educational and developmental needs of the students. This age group represents a unique

subgroup within the larger system. Sometimes their group's needs are overlooked as they are at disadvantage by being in the middle of two different developmental age groups, the primary grades and the middle grades (Jenkins & McEwin, 1992). The unique nature of the age group and the constant reconfiguring of grades leave many parents and educators wondering about the best place for the fifth and sixth grader.

In 1987, Becker examined the effects of grade span on sixth grade achievement as measured by a standardized state assessment. In his study, Becker analyzed 8000 Pennsylvania students in 330 different schools. The sixth grade students were split almost evenly between K-6 elementary schools and 6-8 middle schools. The sample was drawn from rural areas of Pennsylvania. In the elementary setting, students experienced instruction in a self-contained model. In middle school sixth grade students experienced departmentalization and interdisciplinary teaming. In addition to grade span, Becker examined socio-economic status (SES) as an interacting variable and school size as a control variable.

In considering instructional organization, those students experiencing self-contained situations had higher test scores. This factor was more prevalent for students from low SES backgrounds. Sixth grade students in the low SES group scored better when they were educated in elementary settings (one-fourth a standard deviation higher). Even though several subject areas were tested, the correlation results for the low SES group were highest in reading and math scores (correlation values of .24 and .22 respectively). When Becker examined the relationship between school configuration and student achievement for the high SES group, he found the opposite results. Generally,

students from this group scored higher in middle school settings represented by a negative correlation ( $r = -.07$ ).

In 1997, researchers Tucker and Andrada examined the effects of grade span on student achievement as students transitioned from elementary to middle school. Specifically, they examined the differences between standardized tests given in the fourth and sixth grade. The sample included 500 students in Connecticut who were split into two different grade configurations, K-5/6-8 and K-8. Growth in achievement was defined by changes in the average scores on the Connecticut mastery test in reading, writing, and mathematics.

In comparing scores, Tucker and Andrada discovered an accountability issue related to transition between K-5 and 6-8. Because the students in this group took the fourth and sixth grade test in two different schools, it seems neither school was perceived as “accountable” for the scores. In the K-8 setting, however, the students took the fourth and sixth grade test in the same school and there was a stronger sense of public accountability for the both the test scores. In looking at growth between the two tests, the K-8 schools showed more progress in all subject areas.

In 1998, Franklin and Glascock examined the relationship among grade configuration, student performance, and behavior in sixth graders attending public schools in rural Louisiana. The grade spans included K-6, MS (6/7-8/9), and Unit Schools (K-12). The sample size was approximately 215 schools. Student performance was measured by a CRT. Behavior performance was measured by attendance and suspension data.

Sixth graders attending elementary and unit schools scored significantly higher on the CRT than their peers attending middle schools. The K-6 students had behavior scores higher than those attending middle school. This was especially true for boys who experienced more suspensions in the middle school setting, in contrast to their counterparts in the elementary or unit school setting. The researchers attribute the differences in performance to school transition and school size.

In 2000, Vaccaro examined the potential relationship between configuration and student achievement in several content areas. Vaccaro researched test scores for students attending K-8 and middle schools in East Tennessee. In all, there were 119 schools; 57 were middle schools and 62 were K-8 schools. Urban, suburban, small town, and rural communities were represented in the sample of schools chosen for this study. To gauge student achievement, Vaccaro used student scores on a CRT between the years of 1997 to 1999. To evaluate and assess schools the State of Tennessee reports academic gains on the CRT test. This growth model helps to minimize the effects of the individual differences that exist among students.

To test for significant differences between the academic gains of those students attending K-8 schools and those attending middle schools, Vaccaro used an effect size calculation. The effect size was defined as the difference between the average K-8 mean gain and the average middle school mean gain divided by the standard deviation of the middle school group. If the effect size was at least 0.5, then the comparison between the two groups was determined to have large educational significance. Vaccaro found a significant effect size (0.72) for the comparison between seventh grade math scores favoring those students attending middle schools over their counterparts in the K-8. An

educationally significant effect size (0.61) was also found for the seventh grade science scores favoring those students attending middle schools. The academic gains in reading favored sixth grade students attending K-8 schools (effect size of 0.71). From these significant effect sizes, Vaccaro concluded that the K-8 setting was more appropriate for sixth grade while the middle school setting was more appropriate for seventh grade. However, because of the limited number of comparisons which reached the level of educational significance, Vaccaro stated that he could not conclude one configuration was better than the other.

Since the late 1990's, several large school districts have begun the process of converting their middle schools into K-8 elementary schools. One such district is Cleveland Municipal School District, which began the process of converting schools in 1999. In 2004, Poncelet used a quantitative approach to compare sixth grade student achievement between several K-8 and middle schools in Cleveland. The sample size was approximately 4000 students with 25% attending K-8 schools and 75% attending middle schools. Achievement was defined by student performance on the Ohio Proficiency Test. In both reading and math, sixth graders attending K-8 schools scored higher than students attending middle schools. The results of the achievement test were reported by a scaled score. The scores were averaged for each group. Then scale scored difference was calculated by subtracting the mean score of the middle school group from the mean score of the K-8 group. The scale score difference in reading was 7.43 points, which translated to a small to medium effect size (Cohen's  $d = 0.29$ ). The scale score difference in math was 9.41 points, which translated to a medium effect size (Cohen's  $d = 0.38$ ).

Although grade configuration is not a new topic, the revival of the K-8 school has re-opened interest in the topic. Recently there have been a few doctoral dissertations related to the topic. In 2005, Freeman examined grade span effects on sixth grade achievement in North Carolina. In 2006, Comer focused on grade span and student achievement for fifth graders in Tennessee. Both studies found advantages for fifth and sixth grade in the elementary setting, but they differed in terms of variables.

The Freeman study examined academic achievement and attendance, while controlling for socio-economic status, school size, and previous achievement. Freeman discovered a significant difference in achievement favoring K-8 setting over the middle school setting. Student achievement was defined by student scores on the End-of-Grade math and reading test. On the reading test, the K-8 average score was .87 points higher than the middle school average. The mean difference for the K-8/middle school comparison was statistically significant at the .05 level ( $p=.00$ ). On the math test, the K-8 average score was .54 points higher than the middle school average. The mean difference for the K-8/middle school comparison was statistically significant at the .05 level ( $p=.048$ ). As a follow up, Freeman conducted interviews with educational professionals at each setting in an effort to comprehend the success of the K-8 setting. Results from the interviews are reported later in the section on social/emotional context.

The Tennessee study (Comer, 2006) compared various grade span configurations with regards to student achievement, promotion rates, and attendance. Student achievement was measured by the students' ability to reach NCLB benchmarks set by the state. For fifth grade students, the elementary setting was better in terms of attendance and likelihood of achieving the NCLB benchmarks. The attendance rate for students



attending K-8 schools was .35 points higher than those attending middle schools. This mean difference met the alpha level ( $p < .05$ ) with a p-value of .001. Comer discovered no significant difference in promotion/retention rates between the various grade span configurations. The promotion rates for K-8 students were slightly lower than the rates of middle school students (97.92 and 98.26 respectively). This mean difference did not meet the alpha level ( $p < .05$ ) with a p-value of 0.206.

The research available seems to demonstrate an advantage for fifth and sixth grade students attending elementary and K-8 settings when compared to their counterparts attending middle schools. This finding is consistent when extraneous variables, such as school size and SES are controlled. At least one researcher concluded that the advantage of the K-8 setting decreased as grade level increased (Franklin & Glascock, 1998). In addition, one study (Vaccaro, 2000) found that seventh grade students performed better in a middle school setting. Therefore, it would be important to examine comparisons at higher grade levels.

### *Eighth Grade Research*

One of the earliest studies which focused on eighth grade achievement was completed by a researcher in New York, who examined the differences in students attending K-8 and those attending junior high schools. Using a sample of approximately 2000 seventh and eighth grade students in 18 different schools, this research compared the two grade span configurations looking for differences in achievement, attendance, self-esteem, attitude towards school, and perception of teacher's behavior with regard to how they handled pupil control (Moore, 1984).

Extraneous variables were controlled in the process of selecting the sample. The schools chosen for the New York study were similar in demographics, in that there were no important differences in ethnicity, size, and socio-economic status among the schools. Results showed advantages for the K-8 setting. Students enrolled in K-8 schools were more likely to exhibit a positive attitude towards school and miss fewer days. Regarding attendance, the mean difference between the two groups was consistent at the .01 level. Concerning reading achievement, students in K-8 consistently scored higher than their junior high counterparts. There was a significant difference (at the .01 level) in reading achievement as students transitioned from the elementary setting to the junior high. Because K-8 students experienced no elementary to junior high transition, their scores improved from sixth to seventh grade (Moore, 1984).

Another study which focused on eighth grade achievement was conducted by Wihry, Coladarci, and Meadow (1992). Similar to previous studies, researchers examined the achievement scores of eighth grade students from various grade span configurations (K-8, 5/6-8, 7-8/9, and 7-12). The sample size included 163 schools in the State of Maine. Student achievement was measured by scores in the areas of math, science, reading, writing, and social studies. Using a multiple regression approach, researchers included other school characteristic variables such as school size, SES, instructional expenditures, student-teacher ratio, and teacher qualifications. Overall, researchers determined that the regression model accounted for the majority of variance when predicting the full-scale test scores ( $R^2 = .50$ ).

Using the regression method allowed researchers to examine other relationships closely tied to student achievement. As expected, SES was found to be a predictor of

student performance. Community socio-economic status was assessed by calculating the percentage of adults in the community who had completed four or more years of college. For each percentage of increase in the SES measure, the full scale student score increased by 32 points, which translated to two thirds of a standard deviation. Teacher experience, as measured by education and years of experience, also influenced student achievement. For each additional year of teacher experience, full scale student scores increased by 12.5 points, which translated to one-fifth of a standard deviation.

Researchers also discovered that grade span also had an influence on student achievement. Specifically, grade span had an influence on the full-scale test score (math and reading combined). For eighth grade students, the K-8 grade span configuration was more advantageous than other grade span settings. The full-scale scores for eighth grade students attending middle schools or junior high schools (serving grades 7-9) was 15 points below eighth graders in the K-8 schools (one-third of a standard deviation). The grade span which fared the worst when compared with the K-8 setting was the junior/senior high school, serving grades seven through twelve. The full-scale test score of eighth grade students in the 7-12 span was 38 points below that of the eighth graders in the K-8 school. This scale score difference was .79 of a standard deviation.

In Mississippi, Cox (1996) examined the performance of eighth grade students in four different grade spans: K-8, 5-8, K-12, and 7-12. Other variables used in the research were per pupil expenditure, racial composition of the school, and SES levels. Student achievement was measured using the math portion of the Stanford Achievement Test. In analyzing the means of each group, Cox found very little variation in the means. For example, the average test score of the 5-8 configuration was only 0.03 points higher than

the K-8 mean score. However, an ANOVA test showed a significant difference, at the .01 level ( $p = .0056$ ), among the means. A Scheffe' post-hoc comparison test was run to determine where the significant difference existed. This test showed a significant difference between the average score of the 5-8 group (49.57) when compared with the average score of the 7-12 group (47.55).

One interesting finding was discovered in looking at per pupil expenditures. The schools which had the lowest per pupil expenditure actually scored higher than the schools which had the highest per pupil expenditures. For example, in the K-12 group, the mean score for schools with a low per pupil expenditure was 1.35 points higher than the mean score for the schools with a high per pupil expenditure (Cox, 1996).

A study which focused on school configuration and school size for three counties in North Carolina was conducted by Peterson in 2002. This study focused on student achievement, participation in activities, perception of teacher support, and level of self-concept. The configurations used for the research were K-8 and 5/6-8 middle schools. School size was also used as an independent variable. Schools with greater than 350 students were considered large schools and less than 350 were small schools. The sample included 350 students from the K-8 population and 470 students from the middle school population. The schools were located in relatively small communities with low socio-economic levels.

In Peterson's study, he found no significant differences in math and reading achievement between the two grade configurations. The differences between the means were small. For example, the mean reading scaled score for the K-8 schools was 155.92, while the mean reading scaled score for the 5/6-8 schools was 155.66. Significance was

set at the 0.0083 level. The p-value for the comparison between the reading scores was 0.744, which did not meet the  $p < .0083$  requirement for significance. The mean math scaled score for the K-8 schools was 165.48, while the mean math scaled score for the 5/6-8 schools was 164.97. Again, the level of significance was not met. The p-value for the math comparison was .585 which did not meet the  $p < .0083$  level (Peterson, 2002).

Another researcher (Cook, 2005) looked at student achievement data using Milwaukee students attending various grade span configurations. This particular study was commissioned by the Milwaukee Public School system and conducted by a researcher working in the educational research division of the school district. The research looked at data from 85 different schools, divided into four subgroups: Old K-8, New K-8, Transitional K-8, and Middle Schools. The researcher suggested the use of effect size to measure practical significance. Student performance was measured in two areas: First was student engagement which included attendance, suspensions, tardiness, and school perception. Second was achievement growth as measured by a nationally standardized test (CTB – Terra nova).

Cook discovered that the K-8 students outscored their middle school counterparts with the “Old K-8s” scoring the highest. Even after controlling for factors such as poverty, ethnicity, special education, and English language barriers, the K-8 students did better. Concerning engagement, across all indicators, students in K-8 schools tended to have a higher degree of engagement as evidenced by a moderate effect size. Cook noted that student achievement seemed to be similar between schools up to grade five. In grades six and seven, K-8 students had an average scaled score of four points higher. By eighth grade the scaled score for K-8 students was an average of ten points higher.

One final study related to eighth grade performance was conducted by Weiss and Kipnes (2006) using students from 45 schools in Philadelphia. The researchers examined differences in student performance indicators between grade configurations. Their initial results, using a linear regression model of analysis, showed an advantage for students attending K-8 schools. The average of all course grades for students at K-8 schools was 78.7% as compared with 77.2% for students at middle schools. This difference was significant at the .001 level. The percentage of students who failed at least one class at the K-8 schools was 26.5% as compared with 37.5% at the middle school. This difference was significant at the .001 level. The percentage of students who had missed more than 20 schools days was 16% at K-8 schools as compared with 20% at middle schools. This difference was significant at the .05 level.

However, when Weiss and Kipnes examined the data using a multilevel regression model, which controlled for extraneous variables, their final result was quite different. Between the two grade configurations, they found that variance in student performance was not associated with the difference in grade span configuration. The characteristics of the individual schools had a strong influence on student outcomes. For example, the population of students attending the district middle schools was defined by families with parents who had less education and relied on more public assistance than the families attending the K-8 schools.

The results of this last study are noteworthy. In their own words, Weiss and Kipnes (2006) realize the importance of their findings, "...our findings can speak to the value of the current reforms to eliminate middle schools. Our results suggest that such reform efforts are likely to have little effect on students' performance..." (p. 267).

*Longitudinal Studies and Influences on High School*

Up to this point the majority of studies reported in this review could be described as cross-sectional research. All the studies described represent a comparison between grade spans using a single moment in time. The next section will examine studies which will show grade span effects across grade levels over a longer period of time. Several of the studies follow a cohort group of students as they move from elementary through middle level and onto high school. Up to this point, it would appear that most of the research shows advantage towards K-8 schools. It would be important to know if these advantages translate to better achievement in high school.

Perhaps one of the very earliest longitudinal grade configuration comparisons was done in Grand Rapids, Michigan after the school district had recently converted to a K-6 / 7-9 (junior high) model. Stetson (1917) wanted to know if the new junior high was more or less cost efficient than the traditional K-8 model. He followed 404 students as they progressed from sixth grade through high school graduation. Half the students had attended a traditional elementary school and the other half had gone through the junior high. Stetson estimated that it cost more (73% higher) to educate a student under the junior high model. After examining school achievement grades, Stetson determined there was no significant difference in academic marks between the two configurations. Therefore, he concluded that the higher cost related to implementing a junior high did not translate to better scholastic work in high school.

A Missouri study examined student achievement as students moved from elementary to middle and high school (Alspaugh, 1998a). Three configuration groups were compared in the study. The first group experienced one transition as its students

moved from K-8 to high school. The second group experienced two transitions, one between fifth and sixth, the second between eighth and ninth. The third group was similar to the second in terms of transition, but was unique in that multiple elementary schools converged into one middle school during the fifth to sixth transition. The three groups covered 16 districts across the state. Alspaugh controlled for school size and SES.

The findings suggest that students experience an achievement loss when they transition between schools. This loss was seen in the fifth/sixth transition. Achievement was defined by the average student scaled score on the Missouri Mastery and Achievement Tests. The fifth/sixth transition loss was greatest for the group which transitioned from multiple elementary schools into one middle school. As students moved from fifth to sixth grade in the K-8 school, on average they showed a gain of 7.40 points on the scaled score. The students who transitioned from the elementary school to the middle school showed an achievement drop of 7.06 points on the scaled score. This mean difference was significant at the .05 level ( $p = .039$ ). As stated earlier, all students experienced an achievement loss when they transitioned to high school; however Alspaugh discovered that the high school transition loss was greater for those students coming from the middle school. This finding is in contrast with the translation inoculation theory which assumes students will have an easier time with the second transition after recovering from the first transition (Simmons & Blyth, 1987).

As a follow up to his original study, Alspaugh examined drop out rates between the various grade configurations (Alspaugh, 2000). He discovered high school drop out rates were higher for those districts which contained middle schools. The average high



school drop out rate for students who had attended K-8 schools was 3.28% as compared to 5.49% for the students who had attended a middle school.

Gronna (1998) studied the effects of grade of school transition on student achievement in the State of Hawaii. Specifically, she sought to determine if student achievement was attributable to the transition (or non-transition) of a student. With a middle school configuration, students experience at least one transition between the fifth and sixth or the sixth and the seventh grade. Conversely, students attending a K-8 school experience no transition. As discovered by Alspaugh (1998a), transition between schools typically results in an achievement loss. Gronna sought to verify this fact for students attending schools of various configurations in Hawaii.

This study examined three types of configurations: K-8 (n=554 students), K-5/6-8 (n=2086), and K-6/7-8 (n=4523). Gronna found that students who experienced a transition had lower achievement scores than those who did not transition before the eighth grade. Specifically, students who did not experience a transition could expect a higher score on a reading achievement test (a difference of 0.2 standard deviations). In addition, Gronna discovered that students who transitioned after sixth grade scored significantly higher (a difference of 0.13 standard deviations) than those who transitioned after fifth grade. These findings led Gronna to conclude that not transitioning prior to eighth grade was a better configuration model for student achievement. In addition, from the perspective of longitudinal effects on student achievement, if students had to transition into middle school, they were better off making the transition after sixth grade.

Offenberg (2001), in conjunction with the Office of Accountability and Assessment for the School District of Philadelphia, completed a comparison study

between approximately 40 middle schools and 40 K-8 schools. Philadelphia had just started the process of converting its middle schools into K-8 schools. What makes this study different than the previous ones is the sampling method. Offenbergh used individual schools as the unit of analysis rather than using individual students.

Schools were measured on the following performance standards: the average test scores in reading, math, and science, the probability of students attending high schools with a college prep curriculum, and the average grade point average of their students at the end of the ninth grade year. Although it may seem unfair to use high school enrollment patterns as a school performance measure, it helps to understand the makeup of Philadelphia's high school system. As students in Philadelphia enter high school they reserve options to apply for several spots in magnet high schools around the city. These high schools typically offer a more challenging "college prep" curriculum. The number of students applying to these magnet high schools is presumed to be a measure of success for the middle or K-8 school.

In analyzing the data, Offenbergh controlled for SES status and used school size as an interacting variable. Offenbergh's study had four relevant results: First, K-8 schools had average test scores 3.5 to 8.5 points higher than middle school test scores. Second, the percentage of K-8 students enrolling in the more competitive magnet high schools was 11 points higher than the percentage from middle schools. Third, the K-8 students demonstrated a higher level of school engagement as evidenced by higher grade point averages, better attendance, and less loss of credits after their first year of high school. Finally, the achievement gap between K-8 and middle schools was dependent on the size of the school. As the eighth grade class became larger, the achievement gap between the

two grade configurations was smaller. This finding is of particular importance when one considers that the majority of K-8 schools have a smaller number of students per grade level when compared to middle schools (Epstein, 1990 and McEwin et al., 2004).

Under the direction of the Board of Education for the Baltimore City Schools, Yakimowski (2001) conducted a study comparing the district's K-8 and K-5/6-8 schools. In the Baltimore study it is interesting to note that the Superintendent requested the research to provide information to the "Space Utilization Committee." This fact reminds us that school configurations decisions are sometimes driven more by school facility issues and less by what is educationally best for students (Paglin & Fager, 1997).

In looking at 2500 middle school students and 400 K-8 students, Yakimowski discovered that K-8 students scored better on standardized tests, but had less opportunity to take college prep courses. Achievement was measured using the CTB-Terra Nova test in the areas of reading and math. After controlling for previous year's achievement, the researchers found that K-8 students were significantly more likely to pass certain sections of the test. On average K-8 students scored higher on the reading and math portion of the test. Other subject areas were not significant. In looking at attendance, the researchers found no significant difference between the two grade configurations.

In secondary analysis, Yakimowski examined differences in course offerings between the two configurations. In general, there was a lack of "gate keeper" courses available in the K-8 setting. Atanda (1999) found that certain courses taken in the eighth grade increased the likelihood that a student would attend college after high school. He identified Algebra I and a foreign language as gate keeper courses. In other words, opportunities to take these courses in the eighth grade would "open the gate" to more

educational options after high school. The researchers in Baltimore discovered that a smaller percentage of K-8 schools offered “gate keeper” courses as compared to middle schools.

In 2005, while working with the Miami-Dade school district, Abella examined differences between students attending K-8 and middle schools. Abella not only compared achievement levels during the middle level years, but also tracked the students as they left the eighth grade to determine if the achievement differences still existed in high school. The sample included over 4000 students attending middle schools and 360 students attending K-8 schools. Abella selected schools which were similar in demographics with regard to SES and ethnicity. There was a difference in school size. The middle school sample had an average of 534 students per grade while the K-8 sample had an average of 192 students per grade. In addition to academic performance, Abella also compared the two groups on the basis of attendance and disciplinary infractions.

Similar to previous studies, when comparing reading and math scores Abella discovered that K-8 students showed more improvement as they progressed from sixth to eighth grade. However, when the two groups transitioned to high school the difference in scores was no longer statistically significant. By the end of the ninth grade year, the two groups had similar math and reading scores.

In considering rate of attendance, most students increased their level of absenteeism as they progressed into higher grade levels. However, in Miami, the K-8 students were less likely to increase their level of absenteeism. This finding carried onto high school. The K-8 alumni had a lower level of absenteeism at the end of the ninth grade when compared to the middle school alumni.

In looking at suspension rates, Abella noted that the number of out of school suspensions increased as the student groups progressed through the middle level grades. However, the K-8 group progressed at a slower rate. By the time the two groups reached high school, there was no significant difference in suspension rates. In one interesting finding related to suspensions, Abella discovered a sharp increase in out of school suspensions at the eighth grade level for those students attending K-8 schools. At the eighth grade level there was no significant difference in suspension rates between the K-8 students and their middle school counterparts.

Using longitudinal data, Blair (2007) examined the impact of grade configuration on academic achievement in the Middle Tennessee Area. Using a growth model, he studied academic gains on a CRT for students attending K-8 and middle schools from 2003 to 2005. The use of a growth model allowed for individual student longitudinal data. This type of data helped control for difference among students by using yearly gains where each student was compared to their own past performance. Then these yearly gains are compared between groups of students to determine if there are any significant differences. Blair defined a significant difference to be an effect size of at least 0.5 between K-8 and middle school groups.

Blair used scores from 160 schools (100 middle schools and 60 K-8 schools) in both rural and urban areas of Tennessee. In comparing scores from sixth, seventh and eighth grade students in the areas of language arts, math, science, and social studies, Blair found very few comparisons which reached the required effect size of 0.5. For the majority of comparisons, the effect size was calculated below the 0.5 level. (Effect size ranged from 0.01 to 0.41). In the review of literature, Blair discussed the issues related to

middle school criticism and advantages of the K-8 configuration. However, Blair countered previous literature when he concluded that his results did not show one grade configuration was more appropriate than another. To support this conclusion and refute the research in his review of literature, in his final conclusions Blair cites several researchers who found that there were factors, other than configuration, that have an influence on achievement. For those school districts who were considering a switch of configuration, Blair suggests that it would be important to consider other community factors when deciding on an appropriate school grade organization.

Nobles (2008) studied grade configuration and the achievement of eighth grade students on the end of grade reading and math test in the State of North Carolina. The data from the end of grade test were collected over the course of three school years (2003, 2004, and 2005) and scores reported as a growth index from one year to the next. Nobles gathered data from 107 schools and grouped the scores based on configuration; 50 K-8 schools and 57 middle schools. In addition, the configuration groups were further divided based on gender and ethnicity. This additional sorting allowed for the researcher to control for the potential effects of gender and ethnicity.

In North Carolina the results of the end of grade test were reported as a developmental scaled score. Nobles analyzed ten comparisons between groups based on configuration. Among these ten comparisons, only two were found to be statistically significant at the 0.05 level. First, the scaled math score for white students attending middle schools was found to be higher than the scaled score for white students attending K-8 schools (mean difference of 1.24 points on the scaled score). The t-test analysis calculated a p-value of  $<.0001$ . Second, the scaled reading score for white students

attending middle schools was found to be higher than the score for white students attending K-8 schools (mean difference of 0.74 points on the scaled score). The t-test calculated a p-value of 0.02. Nobles did not find any other comparisons which reached the level of significance at the 0.05 level. Even though she cited several studies in her review of literature that showed support for a particular grade configuration, Nobles concluded that her research would not support one configuration over another. From these findings, she concluded that there may be factors other than configuration which impact student achievement. She suggested that the decision to change a grade configuration should focus on what is best for the community rather than relying on potential impacts on student achievement.

#### *Social and Emotional Research*

Since the establishment of the NMSA, middle school experts have professed the importance of creating an environment which is developmentally responsive to the needs of the “young adolescent.” In a position paper published, the NMSA outlines its beliefs regarding the process of designing successful schools for the young adolescent (*This We Believe*, 2003). This philosophy statement is built on the understanding that young adolescents possess an identifiable list of characteristics with regards to physical, intellectual, moral, psychological, and social-emotional development. Based on these characteristics and the unique nature of the young adolescent, the NMSA would encourage a separate school which could focus on the developmental needs of the young adolescent.

The previous three sections presented a summary of research from seventeen studies which were designed to examine the impact of grade configuration on student

performance, specifically at the middle-grades level. In all the studies, student performance was measured by not only academic achievement, but also relevant measures of student behavior such as attendance and discipline. With the understanding that the NMSA has placed great importance on the social and emotional development of the young adolescent, it is interesting to note that only five of the seventeen comparison studies have examined issues related to personal development (Freeman, 2005; Peterson, 2002; Poncelet, 2004; Moore, 1984; Weiss & Kipnes, 2006).

The earliest comparison study which, in part, addressed the topic of personal development was completed by Moore (1984) using K-8 and junior high schools in New York City. Even though comparing student achievement was the main focus of the study, Moore also examined differences in teacher control behavior and student levels of self-esteem. He discovered that K-8 schools had developmental and social advantages over junior high schools. From the perspective of the student, teachers in K-8 schools used a more humanistic approach to classroom management. Students in K-8 schools had higher levels of self-esteem when compared with students attending middle schools. Both of these findings help to define the characteristics of the K-8 environment.

Peterson (2002) examined self-concept as one of the dependent variables in his comparison study between K-8 and 5/6-8 middle schools. He defined self-concept to be an individual's feeling of self worth. This variable was evaluated with student surveys administered to both configurations. Peterson's analysis showed no difference in the survey responses as reported by students in K-8 and middle schools. In a second comparison, he also noted school size had no influence on survey responses.



In a qualitative analysis, Poncelet (2004) examined the K-8 environment as compared to the middle school environment. Given the NMSA recommendations for creating a developmentally responsive middle school, Poncelet found that the K-8 setting matched the NMSA recommendation. She described K-8 classrooms that were student centered, where students engaged in advancing levels of cognitive work. The K-8 environment provided students with opportunities of self expression and participation in multiple activities. Upper level K-8 students acted as not only learners, but also mentors. The K-8 setting fostered student connections to caring adults. Poncelet concluded that many of the characteristics of the K-8 school aligned with the NMSA philosophy of creating a developmentally responsive school.

Freeman (2005) reached similar findings when he examined the K-8 environment using a qualitative method. After conducting interviews with educational professionals working in both K-8 and middle school settings, Freeman concluded that students needed personal competence to navigate adolescent transitions successfully. Freeman described personal competence with the following student characteristics: level of maturity, self-esteem, and motivation. With the absence of personal competence, the students required personal support to aide with the transition. Personal support was defined by student-teacher and student-student relationships. Using data from the interviews, Freeman characterized the K-8 setting as fostering an environment which offered more support for the individual student.

The most recent comparison study which included student personal development was conducted by Weiss and Kipnes in 2006. Their initial analysis showed differences in self-esteem between the two grade configurations, but their final analysis showed no

significant difference. After controlling for differences in school characteristics, these researchers discovered far fewer differences in self-esteem than previous studies might suggest.

In 1987, Simmons and Blyth published their work which examined the effects of school transition on student self-esteem. This study was not included as a comparison study in the previous section because it did not include student achievement as a dependent variable. However, the importance of the study is relevant here in the discussion regarding social and emotional development. Findings from this research helped guide middle school policy and practices regarding support for students during the transition years.

The purpose of the Simmons and Blyth study was to compare the environmental characteristics in the elementary and junior high school to determine the best setting for the adolescent learner. The study sampled over 700 sixth grade students from the Milwaukee school system who were arranged in one of two grade configurations: K-6/7-9, and K-8. The research spanned multiple years and followed the sixth grade students as they progressed through the middle-level years and onto high school. Simmons and Blyth investigated effects on self-esteem, victimization, maturity, leadership, and participation in extra-curricular activities. Overall, the researchers concluded that the K-8 grade span more effectively supported students' personal growth into adolescence, particularly for female students.

The Simmons and Blyth research showed that school transition was associated with an overall drop in self-esteem levels. This finding was most noticeable for sixth grade girls who transitioned into junior high. Their sixth grade counterparts in the K-8

schools did not experience the drop in self-esteem as there was no school transition. The drop in self-esteem was associated with less participation in extra-curricular activities, less opportunities for leadership, and higher levels of victimization. As the sixth grade girls progressed through seventh and eighth grade in the junior high, they never recovered from the drop in self-esteem associated with the transition from elementary. One might expect that the junior high girls would have an easier time with the high school transition, having just experienced a transition into junior high. However, the data collected by Simmons and Blyth showed that the drop in self-esteem carried onto high school. It seems that the female students from the junior high were never able to overcome the drop in self-esteem they had experienced in the transition from elementary.

#### *Summary*

The research presented here is a summary of numerous research studies which have compared K-8 schools to other configurations. Much of the early research, prior to 1990, has compared K-8s to junior high schools. More recent research has compared K-8s to middle schools. The research is not isolated to one geographic region. A variety of cities and states are represented. Fourteen studies drew samples from schools across a particular state, while the others focused on individual school districts. Many of the studies examined not only student achievement, but also other measures of student performance. Several compared attendance data, behavior data, and school environment factors. Most of the researchers recognized the influence of multiple variables on student achievement. Controls for socio-economic status, school size, and previous achievement were taken into consideration using appropriate statistical methods.

In comparing middle level students attending K-8 schools to those attending other grade configurations, most of the research favors the K-8 setting. Students attending K-8s were more likely to have higher achievement, fewer behavior problems, and better attendance. Many of these advantages carried into high school and resulted in lower drop-out rates for the K-8 students (Alspaugh, 1998b). Beyond the achievement measures, students attending K-8 schools experienced higher levels of self-esteem and maintained better attitudes towards school than their middle school peers. Much of the students' success was attributed to the K-8 environment (Freeman, 2005).

Five studies showed no significant differences between the various grade configurations after accounting for control variables (Cox, 1996; Peterson, 2002, Weiss & Kipnes, 2006; Blair, 2007; Nobles, 2008). With regard to control variables, Offenber (2001) demonstrated the importance of school size. He noted that the achievement differences between the two configurations became smaller as the size of the schools became larger. There was also one study which showed no significant difference between the two groups once they had reached high school (Abella, 2005).

The comparison research showed very few disadvantages to the K-8 setting. Becker (1987) noted that sixth grade students who came from families with a high SES level performed better in non-elementary settings. In other words, the K-8 setting was unfavorable towards student achievement for students from affluent families. Yakimowski (2001), found K-8 schools less likely to offer Algebra I and foreign language, which are considered to increase the likelihood that a student will attend college. Abella (2005), found a sharp increase in eighth grade discipline incidents which resulted in suspensions for those students attending K-8 schools.

Finally, it would be important to note that four of the comparison studies were conducted by individual school districts and the individuals performing the research were employees of the district. Abella (2005), Cook (2005), Offenber (2001), and Yakimowsky (2001), were all employed within the research division of their individual school districts. In 2004, Hough noted that the majority (95%) of the middle school studies have been conducted independently by schools and school districts, which has created some limitations in the results. In describing the role of the researcher, Creswell (2003) defines “backyard” research as that which involves conducting research within the researcher’s own organization. This type of research can lead to limitations regarding the disclosure of information. In addition, this type of research in a school district can lead to issues of power related to expectations from the school board or superintendent. The goal of this research was to examine differences in student performance in various grade configurations in the State of Montana. Chapter Three will discuss methodology and research design to achieve this goal

## CHAPTER THREE

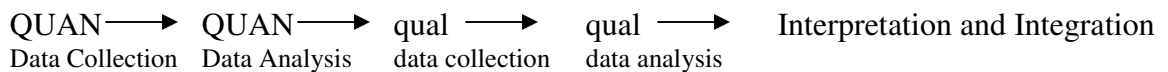
### Methodology

#### *Mixed Method Research Design*

The primary purpose of this study was to explore, through quantitative methods, the potential relationship between grade configuration and student success. The secondary purpose was to use qualitative methods to assist in explaining and interpreting the findings from the quantitative study. The implementation sequence would have been described as sequential, with the quantitative study conducted first and followed by the qualitative study. With regard to data collection and analysis, the priority was given to the quantitative data. The quantitative and qualitative data was integrated at the data interpretation phase. Therefore, this mixed method design was described as a sequential explanatory design (Creswell, 2003).

#### *Visual Model and Procedures*

As defined by Creswell (2003), the following visual model was provided to clarify the process of a sequential explanatory design.



#### *Quantitative Methodology*

The quantitative phase of this study explored the potential relationship between school grade configuration and student performance on a CRT, for eighth grade students attending intermediate level schools and those attending K-8 schools located in the State of Montana. In addition, the quantitative phase of this study explored the potential relationship between school grade configuration and incident of student at-risk behavior for seventh and eighth grade students attending intermediate level schools and those

attending K-8 schools in the State of Montana. In combination with grade configuration, the interacting variables of school size and school SES, were used to compare intermediate level schools with K-8 schools.

*Quantitative Research Questions*

What was the relationship between the grade configuration of a school and student performance on a CRT for eighth grade students attending an intermediate school as compared with those attending a K-8 school in the State of Montana?

What was the relationship between the size of a school and student performance on a CRT for eighth grade students attending an intermediate school as compared with those attending a K-8 school in the State of Montana?

What was the relationship between school SES and student performance on the CRT for eighth grade students attending an intermediate school as compared with those attending a K-8 school in the State of Montana?

What was the relationship between school grade configuration and occurrence of student at-risk behaviors for seventh and eighth grade students attending an intermediate school as compared with those attending a K-8 school in the State of Montana?

What was the relationship between school size and occurrence of student at-risk behaviors for seventh and eighth grade students attending an intermediate school as compared with those attending a K-8 school in the State of Montana?

What was the relationship between school SES and occurrence of student at-risk behaviors for seventh and eighth grade students attending an intermediate school as compared with those attending a K-8 school in the State of Montana?

### *Quantitative Research Design*

This research analyzed existing CRT data for adolescent students attending public schools in the State of Montana. In addition, this research examined existing Youth Risk Behavior Survey (YRBS) results for adolescent students attending public schools in the State of Montana. As students were already grouped into existing schools, the researcher had no control over school assignment. Schools were grouped into categories based on grade configuration. The three configuration categories were schools containing only intermediate grades, schools containing grades K-8, and schools defined as elemiddle schools. The variance among groups was defined by school configuration, school size, and socioeconomic level. School demographic data was collected from the 2005/2006 school year. Student performance data was collected from the Montana CRT given to eighth grade students in Spring 2005. Incidence of student at-risk behavior was collected from the YRBS given to seventh and eighth grade students in Spring 2005. Therefore, the quantitative phase of this study was described as an ex-post facto study consisting of a single observation in time.

### *Quantitative Population*

The population for the quantitative research was any public school within the State of Montana which served students in the eighth grade. Assignment to subgroups was determined by school configuration. The three comparison subgroups were the intermediate level schools, the K-8 schools, and those which fit the characteristics of elemiddle schools. Within Montana there were 326 schools which served 12,050 eighth grade students. Of these schools, 44 schools were organized in an intermediate level configuration, 112 schools were K-8 schools, and 170 were elemiddle schools. The



enrollment of some of the K-8 schools limited the ability to implement middle school or junior high practices (<http://www.nces.ed.gov/ccd/schoolsearch/>).

Concerning school size, total student enrollment data varied greatly between schools and configurations. Within the intermediate schools subgroup, the enrollments ranged from a low of 13 students to a high of 1,013 students with an average of 370 students per school. Of the 112 K-8 schools, the enrollments ranged from a low of one student to a high of 180 students with an average of 22 students per school. Among the 170 elemiddle schools, the enrollments ranged from a low of fourteen students to a high of 1,203 students with an average of 202 students per school (<http://www.nces.ed.gov/ccd/schoolsearch/>).

In all three categories, the percentage of students who qualified for free or reduced lunch varied from a low of 0% to a high of 99%. Overall the elemiddle schools had the highest average of students who qualified for free and reduced lunch, 44%. Within the intermediate schools subgroup, the average percentage of students who qualified for free or reduced lunch was 39%. Among the K-8 schools the average percentage of students who qualified for free or reduced lunch was 17%. However, the K-8 percentage was influenced by the small enrollments of the K-8 schools. Because of the low enrollment several of the K-8 schools did not offer a lunch program. As a result, many of these schools had a free or reduced lunch percentage of zero (<http://www.nces.ed.gov/ccd/schoolsearch/>).

The population for the data regarding at-risk behavior was any school serving seventh and eighth grade students in the State of Montana. In 2005, there were 9,178 seventh and eighth grade students who participated in the YRBS. The population was

limited to Montana for several reasons. First, the YRBS was an optional survey. Not every state chose to administer the YRBS. Some states have limited the years in which the YRBS was administered. Montana was one of the few states in the Northwest which had consistently administered the survey, state-wide, each year it has been offered (<http://www.cdc.gov/HealthyYouth/YRBS/index.htm>). Second, the number of K-8 schools in Montana was among the highest in the Northwest (<http://www.nces.ed.gov/ccd/schoolsearch/>). Finally, Montana was one of the only states in the Northwest which maintained individual school data (P.L. Stewart, personal communication, September 11, 2007).

#### *Quantitative Sample*

A sample of schools was chosen for inclusion in this research. The schools contained in the sample were located in close proximity of the seven large population centers in the State of Montana: Billings, Bozeman, Butte, Great Falls, Helena, Kalispell, and Missoula. Each of these communities contains both intermediate level and K-8 schools. (See Appendix B) With regard to student performance as measured by the CRT, all eighth grade student scores from all schools in the State were included as part of the analysis. Conversely, in analyzing the occurrence of at-risk behavior, not all students from the schools in the sample were surveyed. By design, the YRBS was given to a random sample of students from each school. Depending on the size of the school, the YRBS typically surveyed one hundred students from each school (Montana OPI, 2005b).

#### *Quantitative Data Collection*

*CRT data.* Student performance was measured by examining individual student raw scores on the MontCAS, Phase 2, CRT. As part of the data collection process, the

Montana Office of Public Instruction was contacted and asked to submit CRT scores for this research. Scores were typically released in the form of a scaled score. The range of the scaled score was between 200 and 300. The scaled score was calculated through a simple linear transformation of the raw score (Montana OPI, 2006). Given the scaled score a raw score was calculated for each student by the researcher.

*YRBS data.* The occurrence of student at-risk behaviors was reported as the percentage of students identifying their own participation in at-risk behaviors on the YRBS. The results of the YRBS were given as a percentage of students within each school who reported engagement in a particular at-risk behavior. For example, the YRBS gave the percentage of students within an individual school that carried a weapon on school property in the past 30 days. Schools that are part of the sample were asked to submit their school-level YRBS data from the 2005 survey. This report was accessed electronically through the Montana Office of Public Instruction data warehouse. Given that each school surveyed approximately 100 students, the school-level report gave the overall percentage of students who identified participation in a specific in at-risk behavior.

#### *Quantitative Variables and Level of Data*

The independent variables were school configuration, school size, and school SES. The dependent variables were student performance on the CRT and incidence of at-risk behavior. The level of data was interval and ratio. The incidence of student at-risk behavior was interval data. The school size, school SES, and student performance on the CRT was ratio data. Schools were identified with one of the following configurations: intermediate, K-8, or elemiddle. School size was measured by the total number of

students enrolled in the school during the 2005/2006 school year. School socioeconomic level was measured by the percentage of students participating in the free or reduced lunch program during the 2005/2006 school year (<http://www.nces.ed.gov/ccd/schoolsearch/>).

Student performance was measured by the individual raw scores on the mathematics and reading portion of the 2005 CRT. The Montana CRT scores are generally reported as a scaled score which would be considered ordinal level data. The researcher requested and received the CRT data as raw scores which would be considered ratio level data. A raw score was reported as an integral number and would represent the sum of all points achieved on the 2005 CRT (Montana OPI, 2006). The means and mean differences were reported consistent with the conventional rules of significant figures or digits, which may be found in detail at (<http://www.chem.sc.edu/faculty/morgan/resources/sigfigs/index.html>, and/or <http://chemed.chem.purdue.edu/genchem/topicreview/bp/ch1/sigfigs.html>). Consequently, the means and mean differences were reported as integers, the same level of precision as the data from which they were computed.

The incidence of student at-risk behavior was reported as the percentage of students engaged or not engaged in a particular at-risk behavior as measured by the YRBS administered in the spring of 2005 (Montana OPI, 2005b). The YRBS was given every other year to a sample of seventh and eighth grade students in the State of Montana. The data for this research was gathered from the 2005 YRBS. There was a survey given in 2007, although the number of questions was reduced from 90 questions on the 2005 survey to only 49 questions on the 2007 survey. Many of the questions

which related to at-risk behaviors at school were eliminated. Because this research was focused on the influence of school configuration, it would be appropriate to examine survey data which related to school behaviors (Montana OPI, 2005b). Not all questions on the YRBS were used in the analysis (See Appendix E).

*Null Hypothesis for Student Performance on CRT*

There were no experimentally important or consistent mean differences in student reading and math scores on the MontCAS (Montana's CRT) among the three grade configurations (intermediate, K-8, and elemiddle).

There were no experimentally important or consistent mean differences in student reading and math scores on the MontCAS between the two school size categories (Large >400 and Small <400).

There were no experimentally important or consistent mean differences in student reading and math scores on the MontCAS based on school configuration and size.

There were no experimentally important or consistent mean differences in student reading and math scores on the MontCAS between the two socioeconomic categories (Low SES >40% and High SES <40%).

There were no experimentally important or consistent mean differences in student reading and math scores on the MontCAS based on school configuration and SES level.

*Hypothesis for Incidence of Student At-Risk behavior*

There was an important difference in the incidence of student at-risk behavior, as determined the YRBS results, among schools with an intermediate configuration, a K-8 configuration and an elemiddle configuration.

There was an important difference in the incidence of student at-risk behavior, as determined the YRBS results, among large schools with an intermediate configuration, a K-8 configuration and an elemiddle configuration.

There was an important difference in the incidence of student at-risk behavior, as determined the YRBS results, among small schools with an intermediate configuration, a K-8 configuration and an elemiddle configuration.

There was an important difference in the incidence of student at-risk behavior, as determined the YRBS results, among high SES schools with an intermediate configuration, a K-8 configuration and an elemiddle configuration.

There was an important difference in the incidence of student at-risk behavior, as determined the YRBS results, among low SES schools with an intermediate configuration, a K-8 configuration and an elemiddle configuration.

### *Hypothesis Definitions*

*Experimental importance.* For the purpose of this research, experimental importance, as it related to student performance on the CRT, was defined to be a mean difference of 10 points on the CRT scaled score among any groups. The lowest possible scaled score for the CRT was 200 while the highest scaled score was 300. For the math CRT, a 10 point difference on the scaled score translated to a difference of three points on the raw score. For the reading CRT, a 10 point difference on the scaled score translated to a difference of three points on the raw score. The highest possible raw score on the mathematics portion of the 2005 CRT was 66 points. The highest possible raw score on the reading portion was 60 points (Montana OPI, 2006).

A ten point scaled score difference has been supported by recent comparison studies. While comparing eighth grade performance data for intermediate and K-8 schools in Baltimore, Yakimowski (2001) found a scaled score mean difference of seven points on the reading test and nine points on the math test. In comparing scores in Cleveland, Poncelet (2004) found a scaled score mean difference of 7.43 points in reading and 9.41 points in math. In Milwaukee, Cook (2005) discovered a twelve point scaled score mean difference on a reading test and an eighteen point scaled score mean difference on a math test. All comparisons favored student performance for students attending K-8 schools rather than those attending intermediate schools.

*Experimental consistency.* An experimental consistency was set at the 0.05 level.

*Important difference.* For the purpose of this research, an important difference as it related to student at-risk behaviors was defined to be a mean difference of 5% of students reporting engagement in at-risk behavior between any two groups.

A five percent difference was supported by comparison research as it related to the YRBS. The State of Montana has performed two comparison studies using the YRBS which are relevant to this research. In 2001, a YRBS comparison was made between low and high-performing schools in Montana. The schools were classified as low or high based on performance on a standardized test. Researchers discovered a mean difference of 7% between low and high-performing schools with the low-performing schools reporting a higher rate of participation in at-risk behaviors. In 2005, a mean difference of 9% was discovered when comparing a sample of students attending rural schools versus a sample attending urban schools. The students attending rural schools reported a higher level of participation in at-risk behaviors (Montana OPI, 2005b).

Another indicator which was tied closely to at-risk behaviors was the level of school engagement. School engagement can be associated with regular attendance and a positive attitude towards school. Cook (2005) compared intermediate school students with K-8 students who were attending school in Milwaukee. He discovered a 4% difference in attendance rates with the K-8 students showing a higher rate of attendance. He also found a 7% difference in suspension rates with the intermediate school students showing a higher rate of suspension. In a similar study, Abella (2005) found a 5% difference in suspension rates when he compared students attending K-8 schools with those attending middle schools in Miami.

#### *Quantitative Statistical Procedure*

For the purpose of analyzing the data related to student performance on the CRT, a factorial ANOVA was used to calculate the experimental importance and consistency among school configuration, size, and SES. The factorial ANOVA was used to create multiple comparisons in one statistical procedure, rather than performing multiple separate ANOVAs for each variable. The use of the factorial ANOVA allows for multiple null hypotheses testing within one statistical procedure, which avoids the chance of encountering test-wise errors (Howell, 2002). GBSTAT 6.5 was used for computation. Descriptive statistics compared the relationship between grade configuration, school size, SES, and incidence of student at-risk behavior. Specific YRBS questions used in the descriptive statistics analysis are presented in Appendix E.

#### *A Priori Considerations*

For all tests, the assumption of normality was met by sufficient sample size. The assumption of homogeneity of variance was met using a statistical test explained in



Chapter Four. All measurements were independent. If the F-value calculated by the factorial ANOVA test was less than 0.05 for any two groups, a Tukey HSD (honest significant difference) post-hoc test was conducted to determine the level of experimental consistency. The research design contains multiple comparisons among various means. The Tukey HSD post-hoc is generally used when analyzing multiple comparisons, while limiting the error rate for all possible null hypotheses (Howell, 2002).

### *Threats to Validity*

#### *Internal Threats*

By choosing a non-experimental design the researcher accepted that there was no cause-effect relationship between the program and the outcome. One area of concern was instrumentation. For example, it was important to consider if the CRT was an appropriate instrument to measure student achievement. Typically CRTs were designed to measure student understanding of knowledge and skills of specific curriculum standards. These types of tests were valid in terms of providing assessment information for students, schools, and communities within a given state for which the test was developed. At least within the State of Montana, statistical analysis has been conducted to demonstrate reliability and validity of the CRT. The reliability of the CRT has been verified through a repeated measures process where specific test questions are checked for consistency. The validity of the CRT has been verified through an item analysis where educators from the State of Montana examine test questions to determine the connection to the State curriculum standards (Montana OPI, 2006).

Another instrument used in this study was the YRBS. When considering threats to validity, it was important to consider the reliability and validity of this instrument.

There was some research related to the reliability of the YRBS. The YRBS was designed and established by the Centers for Disease Control (CDC). In 1991, the CDC analyzed the reliability of the YRBS. To test reliability the questionnaire was given twice to the same group of students. Approximately three-fourths of the questions were rated as having substantial or higher reliability (<http://www.cdc.gov/HealthyYouth/YRBS/index.htm>).

With regard to validity of the YRBS, researchers generally agree that the survey has limited validity due to selection and administration process. The selection of subjects was not considered a random sample; therefore it was only valid to generalize the results to the individual students who responded to the survey. In addition, the validity was limited by the self-reporting nature of the survey. In self-reported surveys, students may under-report negative behaviors and over-report those behaviors which were more socially acceptable. Steps were taken during administration to minimize the effects of false self-reporting. For example, no individual identification data was collected (Montana OPI, 2005b).

Another threat to internal validity occurred in the selection process. Because the research examined pre-existing schools, two groups that appeared similar could have been in reality non-equivalent. For example, when selecting two schools of equal size and grade configuration, there was a chance that a difference in SES could have made the comparison less valid (Howell, 2002).

A final threat to internal validity occurred in the process of group assignment. Random assignment of subjects to various groups is important in protecting the internal validity of our comparison (Howell, 2002). Schools were not randomly assigned to

subgroups. Schools were assigned based on configuration, size, and SES. This fact presented a potential threat to internal validity.

### *External Threats*

One potential external threat to validity was related to how the subjects were selected. To protect external validity, subjects should be selected randomly (Howell, 2002). The schools in this study were not selected randomly from the larger population. Schools were selected on the basis of grade configuration and location. In order to assess external threats to validity it is important to consider if the characteristics of the sample could be generalized to a larger population. If the sample consisted of a selection of Montana schools of various size and configuration, then the research may have been generalizable to other schools within Montana. The sample of schools included in the CRT research was selected from all schools in the State of Montana serving eighth grade students. This sample consisted of schools of various size and configuration and could be generalizable to other schools in the State. With regard to student at-risk behaviors, the research may also have been generalizable to other schools within the State of Montana. However the sample of school included in the YRBS research was much smaller and contained only those schools that chose to participate in the research. Therefore the YRBS results were not a true reflection of all schools in the State. In addition, the types of at-risk behaviors were limited to the only those questions asked on the YRBS. Therefore the conclusions were only related to those at-risk behaviors addressed on the YRBS. There are other at-risk behaviors not part of the YRBS and therefore not reported in this research.

### *Qualitative Methodology*

The qualitative phase of this study examined the potential similarities and differences in school characteristics, which existed between intermediate level and K-8 schools. The qualitative data was gathered through interviews with educators from both types of configurations. The interviews were used to explore specific school characteristics related to configuration, which helped to explain the relationship between grade configuration and student success.

### *Qualitative Research Questions*

The central question of this qualitative study was: How was the grade configuration of a school related to student success as perceived by educators working in various grade configurations in one Montana community?

*Sub-questions.* The central question is supported by the following sub-questions: (a) How did the characteristics of the school grade configuration help to address students' developmental and academic needs? (b) What were the advantages associated with the grade configuration of the school? (c) What were the disadvantages associated with the grade configuration of the school? (d) How were student-teacher or student-student relationships defined by characteristics related to school grade configuration? (Freeman, 2005).

### *Qualitative Population*

The primary goal of the qualitative research was to examine similarities and differences in school characteristics of the two basic grade configurations, intermediate level and K-8 schools. To achieve this objective, interviews were conducted with educators from intermediate schools as well as those from K-8 schools. Researchers

Coladarci and Hancock (2002) discovered the importance of a school's geographic location when analyzing student performance. They found that the performance of a school may be less dependent on configuration and more dependent on the size of the school and community. In an effort to minimize the effects of differences between rural and urban communities, comparison schools should be chosen from one community. Several of Montana's urban communities provided a sufficient population for schools of various configurations to exist within the same community. Therefore, the participants for the qualitative comparison interviews were drawn from a population of schools within the community of Missoula, Montana.

#### *Qualitative Participants*

In selecting the qualitative participants, consideration was given not only to which community would be selected, but also the individual educators who were selected for an interview. Creswell (1998) proposed sixteen strategies which could be used in the process of purposeful sampling in qualitative research. In developing the purposeful sample for this research, a few of these strategies were relevant. First, a homogeneous sampling strategy encouraged the researcher to focus on subjects which were similar in nature. This strategy reduced the opportunity for pre-existing differences to influence the outcome. Second, a criterion sampling strategy supported the selection of subjects which met a specific set of criteria. Third, combination sampling allowed for flexibility in the research by selecting subjects which met multiple interests and needs (Creswell, 1998).

In selecting a community for participation in this qualitative research, it was important to consider a location which had a sufficient number of schools with various configurations. Selecting such a community ensured that there were enough educators

who had experience with different grade configurations. At the same time, the schools in the community were similar in size and SES status. Keeping Creswell's (1998) purposeful sampling strategies in mind, the community of Missoula was selected for participation for the following reasons: (a) Missoula contained a variety of schools whose configurations met the definitions generally associated with middle-grades and K-8 schools, (b) all schools were in close proximity to a large population center which helped to ensure a homogeneous sample, (c) schools in the community represented a combination of needs and interests with regard to configuration and size, (d) some schools in the community had experienced a switch in configuration from a K-8 school to an intermediate school.

Educators selected for participation in qualitative interviews met a specific set of criteria. First, consideration was given to length and type of experience. In order to achieve a homogeneous sample, educators with similar experience were selected. The sample included educators with sufficient experience, more than five years in the field. In addition, the sample included educators with similar types of experience, teachers and administrators. Second, in choosing the sample, consideration was given to educators which met a combination of needs and interests. Educators who can speak from both professional and personal experience were considered. For example, an educator who had children of their own, were able to speak from the perspective of the teacher as well as the parent. In addition, educators who had experienced both grade configurations were considered. As Missoula had schools of various grade configurations, several educators have had the opportunity to work in both types of configurations. These educators

provided a unique perspective with regard to the advantages and disadvantages of both types of configurations.

*Qualitative sample size.* A representative sample of educators was interviewed in this study. In the Missoula sample, there were ten schools. (See Appendix B) Seven of these were K-8 schools and three were intermediate schools. Choosing two or three educators from each school resulted in 22 interviews. This sample size was supported as being sufficient for a qualitative study (Creswell, 1998).

#### *Qualitative Data Collection*

*Interview process.* For the qualitative section of this research, data was collected through face-to-face interviews with educators selected through the purposeful sampling process described previously. Prior to starting the interview, consent was obtained from the interviewee using a written consent form which had been approved by The University of Montana - Institutional Review Board (See Appendix C). As part of the consent process, purpose of the study and plans for using the results were discussed with the interviewee. Interviews were conducted using an interview protocol (See Appendix D). The protocol contained open-ended questions designed to elicit thoughtful responses from the individuals being interviewed. In addition, consideration was given to the validity of the questions. These questions have been used in a similar recent doctoral dissertation study (Freeman, 2005). In conducting interviews, consideration was given to the time and setting. In most cases, interviews were conducted on site within the school setting. Participants were offered an off-campus site if they felt uncomfortable conducting the interview at their school. With regard to interview time, participants were informed of the length of time needed to complete the interview. If the interview

occurred within the work day, consent from a building administrator was obtained prior to the interview.

*Accuracy and verification.* To ensure accuracy in the interview data collection process, question responses were recorded by the interviewer in both written and audio format. During the interview, the interviewer took notes on the protocol form. After the interview, the interviewer used the audio recording to verify the written notes.

To ensure appropriate level of verification standards, qualitative researcher Creswell (1998) encourages use of a peer review process. Throughout the data collection process, the results of the individual interviews were to be presented in a debriefing session to the dissertation committee chairperson. This individual was a university professor and served as the primary advisor for this research. In the debriefing session, responses to interview questions and interpretation of results were to be discussed.

In addition, Creswell (1998) encourages use of a triangulation procedure to verify the findings of qualitative analysis. Triangulation occurred between the results of the qualitative findings and two external pieces of data. First, educators who had worked in both grade configurations were selected as secondary participants in the research. After the initial comparison interviews were completed and the responses were categorized, then the secondary participants were interviewed. Their responses helped to verify and support the summary results of the one-on-one comparison interviews. Second, the quantitative results were used to support and verify the results of the qualitative analysis. In the integration phase, the results from the quantitative analysis were compared with the results from the qualitative analysis. This comparison resulted in similar conclusions and will be discussed in Chapter Five.



### *Qualitative Data Analysis*

The process of data analysis in qualitative research involves examining text rather than analyzing measures or indicators. As such, the process of qualitative analysis was much different than the quantitative process presented in the previous section. Creswell (1998) promotes a multi-step data analysis process to be used when examining and interpreting qualitative data. These steps were followed to analyze the data gathered during the interviews with educators.

After data collection, the first analysis stage involves data management. In this stage, the researcher organizes data into usable format which may include files or units. Interview data from the present study was initially organized by grade configuration and further stratified by the role of the educator. For example, data could be organized into a major unit of intermediate schools and a minor unit of intermediate school teacher. Also included in this step was a process of data conversion. For example, data such as grade configuration advantages or disadvantages were separated and converted from interview response to a list of advantages or disadvantages.

The second stage of the data analysis process is described as reading and memoing (Creswell, 1998). The goal of this step was to obtain a global view of the entire data set. This step involved reading through all interview notes and clarifying meaning through the audio recording. While reading through interview notes, memos were written in the side margins of the interview protocol. Memos related to general ideas or key concepts which were used to summarize responses. In addition, these memos helped the researcher to begin the process of identifying similarities and themes among the responses.

The third stage of the data analysis process is described as classifying and interpreting the data (Creswell, 1998). In this stage, qualitative researchers begin to develop a classification system. Generally, the system is dependent on themes or categories which are discovered in the process of creating memos in the side margins of the interview notes. In the present study, the researcher developed categories related to advantages and disadvantages of the various grade configurations. In addition, there were some general themes which emerged related to the particular characteristics of the various grade configurations and how these characteristics supported student success. After developing themes from the data, the qualitative researcher generally begins to form interpretations. With regard to this study, the interpretation sought to form global meanings from the data which explained the potential differences in the various school configurations.

The fourth and final stage of qualitative data analysis is representation (Creswell, 1998). In this stage, the qualitative researcher seeks to present the data in both a narrative and visual form. The goal of the narrative report would be to summarize data and present themes or categories. The visual map supports the summary by displaying the data in a format which allows the researcher to show potential connections or relationships between themes or categories.

#### *Role of the Researcher*

The researcher in this study was a doctoral candidate in the Educational Leadership Program at the University of Montana. The researcher had a professional background in the K-12 public education system for sixteen years at the time of publication. The researcher was employed as an administrator at the high school level in

Missoula, Montana. Previously, the researcher was employed as a middle school principal in the same district. The employment experience of the researcher may have created bias for two specific reasons. First, prior beginning this research, the researcher had working experience related to implementing the middle school concept and the tenants of the middle school philosophy as defined by the NMSA. This fact may have created bias in support of the middle school concept when comparing schools based on configuration. Second, the researcher was employed in Missoula, Montana. Creswell (1998) warns of the danger of conducting research in “one’s own backyard.” Because the qualitative research focused on school in Missoula and the researcher was employed by the Missoula School District, there may have been bias towards the middle school grade configuration.

The researcher has completed The University of Montana Institutional Review Board’s on-line tutorial concerning graduate research involving human subjects. The researcher received approval of the research plan through the application procedure set forth by the Institutional Review Board. For participation in the qualitative portion of this study, human subjects were required to sign a consent form (See Appendix C).

#### *Integration of Qualitative and Quantitative Data*

The sequential mixed method design encourages the researcher to integrate the data as a final stage in the analysis process (Creswell, 2003). In this study, the quantitative research took priority over the qualitative research. Upon completion of both phases, the qualitative data helped to explain and support the quantitative findings. Therefore, the final analysis of this mixed method design sought to integrate the themes

and general ideas which arose from the qualitative data with the results of the hypothesis testing from the quantitative research.

### *Delimitations*

#### *Selection of Sample*

To overcome the limitations related to rural versus urban schools, comparison schools were selected that were relatively close in size and geographic location. Some locations in Montana had examples of both intermediate schools and K-8 schools within close proximity. For example, within the community of Missoula, Montana there were three intermediate schools and three schools which possessed the characteristics of a K-8 school. Using Mertens and Anfara's (2006) school size definition, all of these schools were defined as large with enrollments of more than 400 students. Because of their proximity to a relatively large population center, all of these schools were considered more urban than rural.

The quantitative sample (CRT scores) consisted of all schools in the state of Montana. Therefore the results of the CRT analysis could be generalized to all middle-level schools in the State of Montana, assuming schools are matched with the school configuration definitions outlined in Chapter One. The YRBS sample of this study was confined to specific schools selected by the researcher. To minimize the effects of school location the YRBS sample consisted of schools in and around the seven large population centers in the State of Montana. Therefore the results of the YRBS sample cannot be generalized to all schools in Montana. The qualitative (1-1 comparison interviews) sample was a purposeful sample which focused on schools of various configurations within one specific community in Montana (the community of Missoula). Therefore the

results of the qualitative analysis could only be assumed to be true for the community of Missoula; however the results could be transferable to similar communities in Montana.

### *Selection of Variables*

*Independent variables.* Although there were many independent school variables which may have had an influence on student success, this study only considered school configuration, size, and SES level. Socioeconomic status was used because of the strong correlation between a student's SES level and performance in school (Nelson, 2006). Although SES level is typically tied to an individual student, the methodology of this research defined the SES of an entire school by the number of individual students who had signed up for the free or reduced lunch program. This fact may limit the results of this study. It may be inappropriate to assume that all students in a particular school are influenced by the SES level of only a few students. In addition, the schools were separated into various configuration groups as defined in Chapter One, definition of terms. Although these configuration designations are supported by research, the results of this study are limited by the definitions of those configurations.

*Dependent variables.* There were many dependent variables which could be used to describe a student's success in school. For the purpose of this study, student success was measured by performance on the state level assessment, a CRT. In addition, student success was described by occurrence of at-risk behaviors as reported on the YRBS.

When comparing two types of grade configurations, the educational effectiveness of a particular configuration could have been evaluated using a variety of different data. For example, when comparing middle schools with K-8 schools, researchers have analyzed student behavior data such as attendance and discipline (Cook, 2005). In

addition, researchers have examined student performance data such as grade point average and drop-out rates (Alspaugh, 1998b; Weiss & Kipnes, 2006). Also, several researchers have examined student perceptions related to school environment (Freeman, 2005; Moore, 1984; Poncelet, 2004).

Even though the possible measures of success were abundant, the public pressure created by the NCLB, has limited the measures of school success to student performance on the standardized test (Nichols, Glass, & Berliner, 2005). Under NCLB, each state was required to implement an accountability system for schools, based on test scores. By 2014, all students within any particular school should have demonstrated proficiency on the CRT. Built into the accountability system were consequences for schools which fail to show progress towards getting all students to proficiency (U.S. Department of Education, 2002). Schools which fail to show progress were given the label of “needs improvement.” Under NCLB, states required schools to publish annual report cards for the public, which identified student performance as a group as well as subgroups (Nichols, Glass, & Berliner, 2005).

The public focus on NCLB was another limitation. A review of comparison studies has shown that a focus on student test data was a common theme among the research (Cook, 2005; Offenber, 2001; Yakimowsky & Connolly, 2001). In addition, student performance on standardized tests has been frequently cited by school administrators as a reason for abandoning middle schools in favor of adopting a K-8 configuration (George, 2005; Pardini, 2002). Therefore, it was necessary to consider student test data as a potential measure of school effectiveness. However, it was

important to acknowledge that a focus on student test data may have limited the results of this research.

### *Limitations*

A first limitation to this study was related to the possible relationship between school size and school configuration. In some states, such as Montana and North Dakota, rural schools generally have a smaller population of students than urban schools. In schools with smaller populations, it is difficult to implement all components of a middle school (Deyoung, Howley, & Theobald, 1995). Some small schools were required to follow the K-8 model simply by the nature of their population. Therefore a comparison between K-8 schools and intermediate schools could have resulted in a comparison between rural schools and urban schools. In other words, in a rural state like Montana, the variance between schools may be more related to size and location, rather than configuration.

A second limitation to this study was related to the configuration of schools in Montana. Some states have provided a monetary incentive for schools to adopt a grade structure other than the self-contained classroom model for the middle-grades. For example, in Montana, schools received extra funding when middle-grades are organized into a middle school or junior high model. Funding given to schools was based on a per pupil formula. High school students were paid at a higher rate than elementary students. Seventh and eighth grade students could have been funded at the high school rate for schools that demonstrated a middle school or junior high program (Montana OPI, Section 10.55.902, 2005a). This incentive created a unique situation, which may have encouraged a middle-grades organization, while discouraging a K-8 structure. Schools

that might have chosen to adopt a K-8 configuration may have been discouraged by the monetary incentive.

A third limitation to this study was related to the lack of connection between the school name and instructional practices. For the purpose of this research, designation of school configuration was based on data from the National Center of Educational Statistics (NCES) (<http://www.nces.ed.gov/ccd/schoolsearch/>). In the NCES data, some schools may have carried the name middle school in their title. However, these schools may have had various configurations, 5-8, 6-8, or 7-8. In addition, the NCES data did not specify specific pedagogical practices within each school. Therefore, it would be inappropriate to assume these schools had implemented all the practices recommended by the NMSA. Therefore, the results of this study were not generalized to all schools which carry the words middle school in their name.

A fourth limitation to this study was related to the use of the YRBS. To measure incidence of at-risk student behavior, results from the YRBS were analyzed. This survey was given nationally to seventh and eighth grade students every other year. The nature of how this survey was administered creates some limitations. The results of the survey were dependent on the accuracy of the individual student answers. The YRBS asked students to report the occurrence of specific at-risk behaviors within their own lives. Self-reported surveys can cause some students to over-report behaviors which are socially desirable and under-report behaviors which would seem risky or unhealthy (Montana OPI, 2005b). However, the anonymity of the administration process may have helped to minimize the effects of self-reporting (Montana OPI, 2005b). The data gathered from this instrument was dependent on the integrity of the responses from the



participating students. Therefore, the results of this study were only generalized to the specific population of students who took the survey.

### *Summary*

With the increased focus on student achievement and school accountability, schools across the country are converting from the middle school (intermediate) model to the K-8 model. School officials are making these changes in the absence of any definitive research which shows the potential relation between school configuration and student performance. Other studies, presented in Chapter Two, have shown the powerful impact of interacting variables such as school size and SES. With these thoughts in mind, this present study was designed to explore the potential relationship between grade configuration and student success. To compare schools of various configurations, the research design focused on three pieces of data: (a) student performance on the MontCAS, (b) incidence of at-risk behavior as reported on the YRBS and (c) one-on-one comparison interviews with educators currently working in schools of various configurations. The results of the quantitative and qualitative analysis of the three data sets were presented in Chapter Four.

## CHAPTER FOUR

### Results

#### *Introduction*

This research focused on the potential relationship between grade configuration and student performance of intermediate level students in the State of Montana. Performance on a CRT and incidence of at-risk behavior were indicators of student success. In addition, this research explored perspectives of educators working in a variety of configurations. Interviews were conducted with educators to determine school characteristics which impacted student success.

In this chapter, the results have been organized and presented in three sections, each with a distinct process of analysis. The results have been gathered from three data sources: (a) CRT scores in the areas of reading and mathematics, (b) at-risk behaviors as reported on the YRBS, and (c) one-on-one comparison interviews with educators experienced in the education of intermediate level students. The CRT scores have been collected from the 2005 math and reading test given to all eighth graders in the State of Montana. The YRBS results were gathered from a selection of schools from across the State of Montana who agreed to participate in the research. The comparison interviews were conducted with educators from one Montana community who have experience in the area of intermediate level education. Table 4.1 outlines the three sources of data, method of analysis, the number of schools, and the general location of schools who participated in this research.

Table 4.1 *Overview of Data and Analysis*

Source of Data	Year of Data	Method of Analysis	Number of Schools	Location of Schools
Criterion Referenced Test Scores	2005	Statistical Analysis (ANOVA)	271	All schools in State of Montana
Youth Risk Behavior Survey Results	2005	Descriptive Analysis	24	Sample of Montana schools
One-on-one Comparison Interviews	2008	Qualitative Analysis	8	From Missoula County, State of Montana

### *Analysis of Criterion referenced Test (CRT) Data*

#### *Introduction*

The CRT data was collected from test results from the 2005 Montana Comprehensive Assessment System (MontCAS) in the areas of reading and mathematics. The MontCAS was given to all eighth grade students in March, 2005. The scores were reported as raw scores for each student participating in the test. The raw score was defined by the total number of questions answered correctly. A raw score of 60 represents a perfect score on the reading test and a raw score of 66 represents a perfect score on the math test (Montana Office of Public Instruction, 2006).

#### *Data Categories and Sub-groups*

The raw scores were separated into categories based on the following criteria. First, scores were divided based on type of test, math or reading. Next, scores were divided into three categories based on the grade configuration of the school; K-8 (K8), Elemiddle (EM), and Intermediate School (IS). Each configuration group was further

defined by school size and school socio-economic status. A score from a school with an enrollment of greater than or equal to 400 students was assigned to a large (LRG) category group. A score from a school with less than 400 students was assigned to a small (SML) category group. In addition, if a score came from a school that had less than 40% of the total enrollment participating in the free and reduced lunch program, then that score was placed in the high SES (HSES) category group. If a score came for a school that had a free and reduced lunch population greater than or equal to 40% of the total enrollment, then that score was placed in the low SES (LSES) category group. Then category groups were further divided into subgroups. For example, a score which came from an intermediate school with an enrollment of less than 400 was placed in the small intermediate school (SMLIS) subgroup. Table 4.3 and 4.4 have descriptive data related to the subgroups.

### *Sample Size*

In 2005/2006 there were 326 schools serving intermediate grade students in the State of Montana (<http://www.nces.ed.gov/ccd/schoolsearch/>). In 2005, the MontCAS only tested eighth grade students. There were 271 schools selected for this research. Table 4.2 provides some descriptive data for sample schools by category. There were 11,690 individual scores from the 271 participating schools utilized in this study. In 2005 there were 12,050 eighth grade students attending Montana schools. The sample of 11,690 represents 97% of the total population. Because of the size of this sample results can be verified with a 99.9% confidence interval and a 0.5% margin of error. Schools and scores were excluded from participation in this study for two reasons. First, some schools in the population did not serve eighth grade students in 2005 and no scores were

reported for those schools. Second, some schools had fewer than 10 scores reported. To protect the confidentiality of those students, the Office of Public Instruction does not report scores for schools with fewer than 10 scores.

Table 4.2 *Descriptive Statistics for Categories of Sample Schools*

Category	Number of Schools	Number of Scores	Percentage participating in free/reduced lunch program	Average Enrollment
K-8 (K8)	58	219	17%	28 students
Elemiddle (EM)	156	3,780	44%	202 students
Intermediate School (IS)	57	7,691	39%	370 students
SML School (< 400)	229	4,892	-	110 students
LRG School ( $\geq$ 400)	42	6,798	-	606 students
HGH SES (< 40% F/RL)	141	7,151	19%	-
LOW SES ( $\geq$ 40% F/RL)	130	4,539	58%	-

Table 4.3 *Descriptive Statistics for Subgroups Defined by SES Level*

Subgroup		Number of Schools	Number of Scores	Free/Reduced Lunch % Range	Free/Reduced Lunch % Average
HIGH SES (<40% F/RL)	K8	42	135	0 – 35%	3%
	EM	69	1,792	0 – 39%	25%
	IS	30	5,224	13 – 38%	28%
LOW SES (>40% F/RL)	K8	17	84	43 – 90%	61%
	EM	87	1,988	40 – 92%	58%
	IS	26	2,467	40 – 94%	57%

Table 4.4 *Descriptive Statistics for Subgroups Defined by Size*

Subgroup		Number of Schools	Number of Scores	Enrollment Range	Enrollment Average
Small (< 400)	K8	59	219	10 – 180	28 students
	EM	136	2,333	17 – 378	131
	IS	34	2,340	30 – 389	172
Large (>400)	K8	0	n/a	n/a	n/a
	EM	19	1,447	400 – 1,203	604
	IS	23	5,351	400 – 1,013	607

### *Test for Homogeneity of Variance*

The sample sizes of the various subgroups were not equal, therefore, the Hartley's F-max test was utilized to test homogeneity of variance. Among the socio-economic subgroups, the sample sizes range from a low of 84 scores in the LSES K8 to a high of 5,224 scores in the HSES IS. Among the SES subgroups, the Hartley's F-max test calculated a p-value greater than the 0.05 level. Also, the sample sizes among the size subgroups were not equal. The sizes range from a low of 219 scores in the small K8 group to a high of 5,351 in the large IS group. Among the size subgroups, the Hartley's F-max test calculated a p-value greater than the 0.05 level. As a result of the Harley F-max test for homogeneity of variance, the variances among the subgroups were found to be statistically equal or pooled.

### *CRT Reading Data as Defined by Configuration and SES Level*

*Means.* Considering all schools in the sample, the schools in the HSES group had a mean raw score of 41 on the reading test. The LSES group had a mean raw score of 38. The schools in the K8 category had a mean reading score of 39. The schools in the EM category had a mean reading score of 41 and the schools in the IS category had a mean reading score of 40. The means ( $\bar{x}$ ) and standard deviations (s) for the subgroups are reported in Table 4.5.

Table 4.5 *Descriptive Statistics for CRT Reading Scores (SES sub-groups)*

Subgroups	K8			EM			IS		
	n	$\bar{x}$	s	n	$\bar{x}$	s	n	$\bar{x}$	s
HSES	135	39	10	1,792	42	9	5,224	42	10
LSES	84	38	10	1,988	39	10	2,467	39	10

*Mean differences.* A difference of at least three points was defined as an important difference in the methodology of this research. There was not an important difference among the means based on only school configuration. There was an important difference among means based only on SES level. The mean raw score for all students in the HSES group was three points higher than the mean raw score for all students in the LSES group. Considering the subgroup means, there was an important mean difference between several groups. The mean score of the HSES EM group was three points higher than the mean score of the high SES K8 group. The mean score of the HSES IS group was three points higher than the mean score of the HSES K8 group. The mean score of the HSES EM group was three points higher than the mean score of the LSES EM group. The mean score of the HSES IS group was three points higher than the mean score of the LSES IS group.

*F-value.* There were some important differences among the means of the subgroups. A 2x3 factorial analysis of variance (ANOVA) was used to calculate the F-value to determine experimental consistency among subgroups. In comparing the means between the LSES and HSES groups, the factorial ANOVA test calculated an F-value of 31.06 resulting in a p-value of <.0001. Comparing the means among the three grade



configurations, the factorial ANOVA test calculated an F-value of 7.56 and a resulting p-value of 0.0005 ( $p < .01$ ). In comparing the means among the subgroups, the factorial ANOVA test calculated an F-value of 3.50 and a p-value of 0.0303 ( $p < .05$ ). Considering that the factorial ANOVA resulted in identifying consistent differences among the means, a Tukey/Kramer post-hoc test was used to determine which subgroups had consistent mean differences. The results of the Tukey/Kramer test for the reading scores defined by SES are presented in Table 4.6.

Table 4.6 *Tukey/Kramer's Post Hoc Test for CRT Reading Scores Defined by SES*

	HSES K8	LSES K8	HSES EM	LSES EM	HSES IS	LSES IS
HSES K8	n/a	0.42	13.49*	0.01	14.99*	0.25
LSES K8		n/a	14.04*	0.56	15.14*	0.17
HSES EM			n/a	105.83**	0.12	143.53**
LSES EM				n/a	171.15**	1.49
HSES IS					n/a	243.89**
LSES IS						n/a

\*  $p < .05$       \*\* $p < .01$

In Table 4.6, there are several comparisons which reached the *a priori* condition of  $p < .05$  for experimental consistency, however it is important to highlight those comparisons that also showed experimental importance. The Tukey/Kramer post-hoc test shows experimental consistency at the .05 level between the HSES EM and the HSES K8. Experimental consistency at the .05 level was also found between the HSES IS and the HSES K8. Experimental consistency at the .01 level was found between the HSES

EM and the LSES EM. In addition, experimental consistency at the .01 level was found between the HSES IS and the LSES IS.

*CRT Math Data as Defined by Configuration and SES Level*

*Means.* Considering all schools in the sample, the schools in the HSES group had a mean raw score of 31 on the math test. The LSES group had a mean raw score of 28. The schools in the K8 category had a mean math score of 29. The schools in the EM category had a mean math score of 30 and the schools in the IS category had a mean math score of 30. The means ( $\bar{x}$ ) and standard deviations (s) for the subgroups are reported in Table 4.7.

Table 4.7 *Descriptive Statistics for Math CRT Scores (SES sub-groups)*

Subgroups	K8			EM			IS		
	n	$\bar{x}$	s	n	$\bar{x}$	s	n	$\bar{x}$	s
High SES	134	28	10	1,794	32	10	5,219	32	11
Low SES	84	29	10	1,985	28	9	2,467	27	9

*Mean differences.* A difference in at least three points was defined as an important difference in the methodology of this research. There was not an important difference among the means based on only school configuration. The mean difference between the EM and K8 groups and between the IS and K8 groups was one point. This difference does not meet the three point difference required to be an important difference. There was an important difference between the means based only on SES level (mean difference was three points). Among the subgroup means, there was an important mean difference between several groups. The mean score of the HSES EM group was four

points higher than the mean score of the HSES K8 group. The mean score of the HSES IS group was four points higher than the mean score of the HSES K8 group. The mean score of the HSES EM group was four points higher than the mean score of the LSES EM group. The mean score of the HSES IS group was five points higher than the mean score of the LSES IS group.

*F-value.* A 2x3 factorial analysis of variance (ANOVA) was used to calculate the F-value to determine experimental consistency among subgroups. In comparing the means between the LSES and HSES group, the factorial ANOVA test calculated an F-value of 23.65 resulting in a p-value of <.0001. Comparing the means among the three grade configurations, the factorial ANOVA test calculated an F-value of 3.21 and a resulting p-value of 0.04 (p<.05). In comparing the means among the subgroups, the factorial ANOVA test calculated an F-value of 11.71 and a p-value of <.0001. Considering that the factorial ANOVA resulted in identifying consistent differences among the means, a Tukey/Kramer post-hoc test was used to determine which subgroups had consistent mean differences. The results of the Tukey/Kramer test for the math scores defined by SES and configuration are presented in Table 4.8.

Table 4.8 *Tukey/Kramer's Post Hoc Test for CRT Math Scores Defined by SES*

	HSES K8	LSES K8	HSES EM	LSES EM	HSES IS	LSES IS
HSES K8	n/a	0.44	17.06**	0.02	18.18**	0.56
LSES K8		n/a	6.17	0.53	6.51	2.04
HSES EM			n/a	121.11**	0.01	197.48**
LSES EM				n/a	188.06**	6.61
HSES IS					n/a	323.08**
LSES IS						n/a

\* p&lt;.05

\*\*p&lt;.01

In Table 4.7, there are several comparisons which not only reached the *a priori* condition of p<.05 for experimental consistency but also met experimental consistency at the p<.01 level. The Tukey/Kramer post-hoc test shows experimental consistency at the .01 level between the HSES EM and the HSES K8. Experimental consistency at the .01 level was also found between the HSES IS and the HSES K8. Experimental consistency at the .01 level was found between the HSES EM and the LSES EM. Experimental consistency at the .01 level was found between the HSES IS and the LSES IS.

#### *CRT Reading Data as Defined by Configuration and School Size*

*Means.* Considering all schools in the sample, the schools in the large school (LRG) group had a mean raw score of 41 on the reading test. The small school (SML) group had a mean raw score of 40. The schools in the K8 category had a mean reading score of 39. The schools in the EM category had a mean reading score of 41 and the schools in the IS category had a mean reading score of 40. The reading scores were further divided into subgroups based on the configuration and the school size. There

were no large K8 schools. The means ( $\bar{x}$ ) and standard deviations (s) for the subgroups are reported in Table 4.9.

Table 4.9 *Descriptive Statistics for CRT Reading Scores (school size sub-groups)*

Subgroups	K8			EM			IS		
	n	$\bar{x}$	s	n	$\bar{x}$	s	n	$\bar{x}$	s
LRG (>400)	0	n/a	n/a	1,447	41	10	5,351	42	10
SML (<400)	219	39	10	2,333	40	10	2,340	39	10

*Mean differences.* A difference in at least three points was defined as an important difference in the methodology of this research. There was not an important difference among the means based on only school configuration. The largest mean difference was between the EM and K8 category (two points). This difference does not meet the level of difference defined as an important difference. Also, there was not an important difference among means based only on size level (mean difference of one point). Considering the subgroup means, there was only one mean difference which met the minimum important difference of three points. The mean reading score of the LRG IS group was three points higher than the mean score of the SML IS group.

*F-value.* As there was no large K8 group, a 2x2 factorial analysis of variance (ANOVA) was used to calculate the F-value to determine experimental consistency among the EM and IS subgroups. A one-way ANOVA was used to calculate the F-value among the small school subgroups. In comparing the means between the large and the small schools, the factorial ANOVA test calculated an F-value of 74.72 resulting in a p-

value of  $<.0001$ . Comparing the means between the two grade configurations, EM and IS, the factorial ANOVA test calculated an F-value of 0.01 and a resulting p-value of 0.9126. In comparing the means among the subgroups, LRGEM, LRGIS, SMLEM, and SMLIS, the factorial ANOVA test calculated and F-value of 30.04 and a p-value of  $<.0001$ . Among the means of the small school subgroups, the one-way ANOVA test calculated an F-value of 8.74 with a resulting p-value of 0.0002 ( $p<.01$ ). Considering that the ANOVA tests resulted in identifying consistent differences between the means, Tukey/Kramer post-hoc tests were used to determine which subgroups had consistent mean differences. The results of the Tukey/Kramer tests for the reading scores defined by size are presented in Table 4.10 and Table 4.11.

Table 4.10 *Tukey/Kramer's Post Hoc Test for CRT Reading Scores (size sub-groups)*

	LRG EM	LRG IS	SML EM	SML IS
LRG EM	n/a	14.25**	3.87	29.70**
LRG IS		n/a	51.30**	140.82**
SML EM			n/a	15.83**
SML IS				n/a

\*  $p<.05$       \*\* $p<.01$

Although table 4.10 shows several comparisons that met the *a priori* condition of experimental consistency at the .05 level, there was only one comparison that met the *a priori* condition for an important mean difference (LRG IS and SML IS). The Tukey/Kramer post-hoc test showed experimental consistency at the .01 level between the means for the LRG IS and the SML IS subgroup.

Table 4.11 *Tukey/Kramer's Post-Hoc Test for CRT Reading Scores (size sub-groups)*

	SML K8	SML EM	SML IS
SML K8	n/a	4.96	0.36
SML EM		n/a	15.40**
SML IS			n/a

\*p<.05      \*\*p<.01

Table 4.11 shows that one comparison met the requirement for experimental consistency at the .01 level (SML EM and SML IS). However, among the small configuration subgroups, there were no comparisons that met the requirement for experimental importance. The mean difference between the SML EM and the SML IS subgroups was only one point.

*CRT Math Data as Defined by Configuration and School Size*

*Means.* Considering all schools in the sample, the schools in the large school group had a mean raw score of 31 on the math test. The small school group had a mean raw score of 29. The schools in the K8 category had a mean math score of 29. The schools in the EM category had a mean math score of 30 and the schools in the IS category had a mean math score of 30. The math scores were further divided into subgroups based on the configuration and the school size. There were no large K8 schools. The means and standard deviations for the subgroups are reported in Table 4.12.

Table 4.12 *Descriptive Statistics for CRT Math Scores (size sub-groups)*

Subgroups	K8			EM			IS		
	n	$\bar{x}$	s	n	$\bar{x}$	s	n	$\bar{x}$	s
LRG (>400)	0	n/a	n/a	1,444	30	10	5,346	31	11
SML (<400)	219	28	10	2,335	30	10	2,340	29	10

*Mean differences.* A difference of at least three points was defined as an important difference in the methodology of this research. There was not an important difference among the means based on only school configuration. The mean difference between the EM and K8 subgroups as well as the IS and K8 subgroups was one point. This difference does not meet the level of difference defined as an important difference. Also, there was not an important difference among means based only on school size (mean difference between the large and small school groups was two points). Considering the subgroup means, there were no mean differences which met the level of important difference. The largest mean difference exists between the LRG IS group and the SML IS group (mean difference of two points).

*F-value.* As there was no large K8 group, a 2x2 factorial analysis of variance (ANOVA) was used to calculate the F-value to determine experimental consistency among EM and IS subgroups. A one-way ANOVA was used to calculate the F-value among the small school subgroups. In comparing the means between the large schools and the small schools, the factorial ANOVA test calculated an F-value of 38.91 resulting in a p-value of <.0001. Comparing the means between the two grade configurations, EM



and IS, the factorial ANOVA test calculated an F-value of 0.00017 and a resulting p-value of 0.9897. In comparing the means among the subgroups, LRGEM, LRGIS, SMLEM, and SMLIS, the factorial ANOVA test calculated and F-value of 33.09 and a p-value of <.0001. Between the means of the small school subgroups, the one-way ANOVA test calculated an F-value of 10.11 with a resulting p-value of <.0001. Considering that the ANOVA tests resulted in identifying consistent differences between the means, Tukey/Kramer post-hoc tests were used to determine which subgroups had consistent mean differences. The results of the Tukey/Kramer tests for the math scores as defined by size are presented in Table 4.13 and Table 4.14.

Table 4.13 *Tukey/Kramer's Post Hoc Test for CRT Math Scores (size sub-groups)*

	LRG EM	LRG IS	SML EM	SML IS
LRG EM	n/a	16.25**	0.09	15.13**
LRG IS		n/a	27.32**	101.51**
SML EM			n/a	16.85**
SML IS				n/a

\* p<.05      \*\*p<.01

Although table 4.13 shows many comparisons which met the *a priori* condition for experimental consistency, there were no comparisons among the subgroups which met the requirement for experimental importance. Considering the comparison of test scores based on configuration and size, it would be appropriate to conclude that all the comparisons among the subgroups were consistently unimportant.

Table 4.14 *Tukey/Kramer's Post-Hoc Test for CRT Math Scores (size sub-groups)*

	SML K8	SML EM	SML IS
SML K8	n/a	4.35	0.09
SML EM		n/a	18.75**
SML IS			n/a

\*p<.05      \*\*p<.01

In table 4.14 there was one comparison (SML EM and SML IS) which met the *a priori* condition for experimental consistency at the .01 level. However, this comparison did not meet the requirement for experimental importance. The mean difference between these two groups (SML EM and SML IS) was only one point.

#### *Null Hypothesis Analysis for CRT Scores*

In the methodology section of this study, there were five null hypotheses developed to examine the potential relationship between grade configuration and student performance on the MontCAS test. Student scores were reported in reading and mathematics, therefore each null hypothesis will have conclusions for both subject areas. For each null hypothesis, statistical analysis was conducted to determine if there was an experimentally important and consistent difference among the means. This information will be used to reject or fail to reject the null hypothesis. For the purpose of this research, an experimentally important difference was defined to be a difference in the mean score of at least three points on the MontCAS test. In addition, experimental consistency was set at the .05 level. In order to reject the null hypothesis, both *a priori* conditions must be met.

### *Null Hypothesis #1*

There will be no experimentally important or consistent mean difference in student math and reading scores on the MontCAS (Montana's CRT) among the three grade configuration groups (K-8, elemiddle, and intermediate school).

*Reading scores.* A factorial ANOVA statistical analysis was used to calculate the means and F-value for the reading scores of the three configuration groups. Among the three configuration groups, the F-value was 7.56 resulting in a p-value of 0.0005. This value met the *a priori* condition for experimental consistency ( $p < .05$ ). However, the difference in the means among the three configuration groups did not meet the *a priori* condition for experimental importance. The means for each configuration group (K-8, elemiddle, and intermediate school) were 39, 41, and 40, respectively. The greatest mean difference was two points which did not meet the *a priori* level of a three point mean difference. Because the level of experimental importance was not met, the appropriate conclusion would be to fail to reject null hypothesis #1 as it related to MontCAS reading scores. In addition, because the p-value was extremely low, it could be concluded that there was a consistently unimportant mean difference in the MontCAS reading scores among the three grade configurations (K8, EM, IS).

*Math scores.* A factorial ANOVA statistical analysis was used to calculate the means and F-value for the math scores of the three configuration groups. Among the three configuration groups, the F-value was 3.2 resulting in a p-value of 0.04. This value met the *a priori* condition for experimental consistency ( $p < .05$ ). However, the difference in the means among the three configuration groups did not meet the *a priori* condition for experimental importance. The means for each configuration group (K-8, elemiddle, and

intermediate school) were 29, 30, and 30, respectively. The greatest mean difference was one point which did not meet the *a priori* level of a three point mean difference. Because the level of experimental importance was not met, the appropriate conclusion would be to fail to reject null hypothesis #1 as it related to MontCAS math scores.

#### *Null Hypothesis #2*

There will be no experimentally important or consistent mean difference in student math and reading scores on the MontCAS (Montana's CRT) between the two school size categories (Large >400 and Small<400).

*Reading scores.* A factorial ANOVA statistical analysis was used to calculate the means and F-value for the reading scores of the two school size groups. Between the two size groups, the F-value was 74.72 resulting in a p-value of <.0001. This value met the *a priori* condition for experimental consistency ( $p < .05$ ). However, the difference in the means between the two school size groups did not meet the *a priori* condition for experimental importance. The means for each school size group (large and small) were 41 and 40, respectively. The mean difference was one point which did not meet the *a priori* level of a mean difference of three points. Because the level of experimental importance was not met, the appropriate conclusion would be to fail to reject null hypothesis #2 as it related to MontCAS reading scores between large and small schools. In addition, because the p-value was extremely low, it could be concluded that there was a consistently unimportant mean difference in the MontCAS reading scores between the two school groups defined by school size (LRG and SML).

*Math scores.* A factorial ANOVA statistical analysis was used to calculate the means and F-value for the math scores of the two school size groups. Between the two

size groups, the F-value was 38.91 resulting in a p-value of  $<.0001$ . This value met the *a priori* condition for experimental consistency ( $p<.05$ ). However, the difference in the means between the two school size groups did not meet the *a priori* condition for experimental importance. The means for each school size group (large and small) were 31 and 29, respectively. The mean difference was two points which did not meet the *a priori* level of a mean difference of three points. Because the level of experimental importance was not met, the appropriate conclusion would be to fail to reject null hypothesis #2 as it related to MontCAS math scores between large and small schools. In addition, because the p-value was extremely low, it could be concluded that there was a consistently unimportant mean difference in the MontCAS math scores between the two school groups defined by school size (LRG and SML).

#### *Null Hypothesis #3a*

There will be no experimentally important or consistent mean difference in student math and reading scores on the MontCAS (Montana's CRT) for the small school configuration groups (SML K8, SML EM, and SML IS).

*Reading scores.* A statistical analysis, one-way ANOVA, was used to calculate the means and F-value for the reading scores of the small schools grouped by configuration. Among the three configurations for schools less than 400 students, the F-value was 8.74 resulting in a p-value of 0.0002. This value met the *a priori* condition for experimental consistency ( $p<.05$ ). However, the difference in the means among the three configurations for small schools did not meet the *a priori* condition for experimental importance. The means for each group (SML K8, SML EM, and SML IS) were 39, 40, and 39 respectively. The largest mean difference was one point which did not meet the *a*

*priori* level of a mean difference of three points. Because the level of experimental importance was not met, the appropriate conclusion would be to fail to reject null hypothesis #3 as it related to MontCAS reading scores for small schools grouped by configuration. In addition, because the p-value was extremely low, it could be concluded that there was a consistently unimportant mean difference in the MontCAS reading scores among the three small configuration groups (SML K8, EM, and IS).

*Math scores.* A one-way ANOVA was used to calculate the means and F-value for the math scores of the small schools grouped by configuration. The three groups were: SML K8, SML EM, SML IS. Among the three configurations for schools with less than 400 students, the F-value was 10.11 resulting in a p-value of  $<.0001$ . This value met the *a priori* condition for experimental consistency ( $p<.05$ ). However, the difference in the means among the three configurations for small schools did not meet the *a priori* condition for experimental importance. The means for each group (SML K8, SML EM, SML IS) were 28, 30, and 29 respectively. The largest mean difference was two points which did not meet the *a priori* level of a mean difference of three points. Because the level of experimental importance was not met, the appropriate conclusion would be to fail to reject null hypothesis #3 as it related to MontCAS math scores for small schools grouped by configuration. In addition, because the p-value was extremely low, it could be concluded that there was a consistently unimportant mean difference in the MontCAS math scores among the three small configuration groups (SML K8, EM, and IS).

*Null Hypothesis #3b*

There will be no experimentally important or consistent mean difference in student math and reading scores on the MontCAS (Montana's CRT) among the four groups defined by size and configuration (LRG EM, LRG IS, SML EM, and SML IS).

*Reading scores.* A 2x2 factorial ANOVA was used to calculate the means and F-value for the reading scores of large and small, elemiddle and intermediate schools. Among the four groups, LRG EM, LRG IS, SML EM, and SML IS, the F-value was 30.05 resulting in a p-value of <.0001. This value met the *a priori* condition for experimental consistency (p<.05). A Tukey-Kramer post hoc procedure was used to determine if there was experimental consistency between any two of the four configuration groups. When comparing the means among the four groups, the post-hoc procedure showed experimental consistency at the 0.01 level for several of the comparisons. However, for the majority of the comparisons, the difference in the means among the four groups did not meet the *a priori* condition for experimental importance. The means for each group (LRG EM, LRG IS, SML EM, SML IS) were 41, 42, 40, and 39 respectively.

One comparison met the *a priori* conditions for experimental consistency and importance. Between the large and small intermediate school group, the post-hoc procedure calculated an F-value of 140.82 which was consistent at the <.01 level. This p-value met the *a priori* level of <.05. The mean difference between the large and small intermediate school group was three points. This mean difference met the *a priori* condition for experimental importance set at three points. Because the level of experimental importance and consistency were met, the appropriate conclusion would be

to reject null hypothesis #3 as it related to MontCAS reading scores for large and small intermediate schools.

*Math scores.* A 2x2 factorial ANOVA was used to calculate the means and F-value for the math scores of large and small, elemiddle and intermediate schools. Among the four groups, LRG EM, LRG IS, SML EM, and SML IS, the F-value was 33.09 resulting in a p-value of  $<.0001$ . This value met the *a priori* condition for experimental consistency ( $p<.05$ ). A Tukey-Kramer post hoc procedure was used to determine if there was any experimental consistency between any two of the four configuration groups.

When comparing the means among the four groups, the post-hoc procedure showed experimental consistency at the 0.01 level for several of the comparisons. However, for all of the comparisons, the difference in the means among did not meet the *a priori* condition for experimental importance. The means for each group (LRG EM, LRG IS, SML EM, SML IS) were 30, 31, 30, and 29 respectively. The largest mean difference was between the large and small intermediate school groups (a difference of two points). This did not meet the level of an important difference. Because the level of experimental importance was not met for any of the comparisons, the appropriate conclusion would be to fail to reject null hypothesis #3 as it related to MontCAS math scores based on configuration and size. In addition, because the p-value was extremely low, it could be concluded that there was a consistently unimportant mean difference in the MontCAS math scores among the four groups defined by size and configuration.



*Null Hypothesis #4*

There will be no experimentally important or consistent mean difference in student math and reading scores on the MontCAS (Montana's CRT) between the two socioeconomic (SES) categories, HSES (<40% F/RL) and LSES (>40% F/RL).

*Reading scores.* A factorial ANOVA statistical analysis was used to calculate the means and F-value for the reading scores of the two SES categories (HSES and LSES). Between the two SES groups, the F-value was 31.06 resulting in a p-value of <.0001. This value met the *a priori* condition for experimental consistency (p<.05). The difference in the means between the two SES groups met the *a priori* condition for experimental importance. The means for each school size group (HSES and LSES) were 41 and 38, respectively. The mean difference was three points. Because the level of experimental importance and consistency were met, the appropriate conclusion would be to reject null hypothesis #4 as it related to MontCAS reading scores between HSES and LSES schools.

*Math scores.* A factorial ANOVA statistical analysis was used to calculate the means and F-value for the math scores of the two SES categories (HSES and LSES). Between the two SES groups, the F-value was 23.65 resulting in a p-value of <.0001. This value met the *a priori* condition for experimental consistency (p<.05). The difference in the means between the two SES groups met the *a priori* condition for experimental importance. The means for each school size group (HSES and LSES) were 31 and 28, respectively. The mean difference was three points. Because the level of experimental importance and consistency were met, the appropriate conclusion would be

to reject null hypothesis #4 as it related to MontCAS math scores between low SES schools and high SES schools.

*Null Hypothesis #5*

There will be no experimentally important or consistent mean difference in student math and reading scores on the MontCAS (Montana's CRT) based on configuration and SES.

There were six sub groups defined by configuration and SES. Three of the configuration subgroups had less than 40% free or reduced lunch: HSES K8, HSES EM, and HSES IS. Three of the configuration subgroups had more than 40% free or reduced lunch: LSES K8, LSES EM, and LSES IS. Among these six subgroups there were 15 possible comparisons (See Table 4.8). For the purpose of testing null hypothesis #5 this research focused on 9 out of the 15 comparisons. Using ANOVA statistical analysis, the means from HSES K8, HSES EM, and HSES IS were compared to determine if there was an experimental important and consistent difference. The means from LSES K8, LSES EM, and LSES IS, were also compared to determine if there was an experimental important and consistent difference. The means from the HSES K8 subgroup and the LSES K8 subgroup were compared. The means from the HSES EM and the LSES EM subgroup were compared. The means from HSES IS and the LSES IS subgroup were compared to determine if there was an experimental important and consistent difference. Each of these comparisons was tested with reading and math scores.

*Null Hypothesis #5a*

There will be no experimentally important or consistent mean difference in student test math and reading scores on the MontCAS (Montana's CRT) among the three configuration groups defined as HSES groups.

*Reading scores for the three high SES configuration groups.* A factorial ANOVA was used to calculate the means and F-value for the reading scores of the three high SES configuration groups (HSES K8, HSES EM, HSES IS). Among the high SES configurations groups, the F-value was 3.50 resulting in a p-value of 0.03. This value met the *a priori* condition for experimental consistency ( $p < .05$ ). Because the *a priori* condition was met for experimental consistency, a Tukey-Kramer post hoc test was conducted to determine if there was consistency among of the high SES configuration groups. The post hoc test showed consistency at the 0.05 level for two comparisons. First there was experimental consistency between the HSES EM group and the HSES K8 group. There was also experimental consistency between the HSES IS group and the HSES K8 group.

The means for each high SES configuration group (HSES K8, HSES EM, and HSES IS) were 39, 42, and 42 respectively. The mean difference between the HSES EM and HSES K8 was three points, which met the *a priori* level of an important mean difference. The mean difference between the HSES IS and HSES K8 was three points which met the *a priori* level of an important mean difference. Because the post hoc test showed experimental consistency among these groups and the level of experimental importance was met, the appropriate conclusion would be to reject null hypothesis #5 as it related to the comparison between the MontCAS reading scores of the HSES EM and

the HSES K8 group and the comparison between the MontCAS reading scores of the HSES IS and the HSES K8 group.

*Math scores for the three high SES configuration groups.* A factorial ANOVA was used to calculate the means and F-value for the math scores of the three high SES configuration groups (HSES K8, HSES EM, HSES IS). Among the high SES configurations groups, the F-value was 11.71 resulting in a p-value of  $< 0.0001$ . This value met the *a priori* condition for experimental consistency ( $p < .05$ ). Because the *a priori* condition was met for experimental consistency, a Tukey-Kramer post hoc test was conducted to determine if there was consistency among of the high SES configuration groups. The post hoc test showed consistency at the 0.01 level for two comparisons. First there was experimental consistency between the HSES EM group and the HSES K8 group. Also there was experimental consistency between the HSES IS group and the HSES K8 group.

The means for each high SES configuration group (HSES K8, HSES EM, and HSES IS) were 28, 32, and 32 respectively. The mean difference between the HSES EM and HSES K8 was four points which met the *a priori* level of an important mean difference. The mean difference between the HSES IS and HSES K8 was four points, which met the *a priori* level of an important mean difference. Because the post hoc test showed experimental consistency among these groups and the level of experimental importance was met, the appropriate conclusion would be to reject null hypothesis #5 as it related to the comparison between the MontCAS math scores of the HSES EM and the HSES K8 group and the comparison between the MontCAS math scores of the HSES IS and the HSES K8 group.

*Null Hypothesis #5b*

There will be no experimentally important or consistent mean difference in student math and reading scores on the MontCAS (Montana's CRT) among the three configuration groups defined as low SES groups.

*Reading scores for the three low SES configuration groups.* A factorial ANOVA was used to calculate the means and F-value for the reading scores of the three low SES configuration groups (LSES K8, LSES EM, LSES IS). Among the low SES configurations groups, the F-value was 3.50 resulting in a p-value of 0.03. This value met the *a priori* condition for experimental consistency ( $p < .05$ ). Because the *a priori* condition was met, a Tukey-Kramer post hoc test was conducted to determine if there was experimental consistency among the mean differences of the low SES configuration groups. The post hoc test did not show an experimental consistent difference between the means of the three low SES configuration groups.

The means for each low SES configuration group (LSES K8, LSES EM, and LSES IS) were 38, 39, and 39 respectively. The mean difference among the three low SES configuration groups did not meet the *a priori* level of an important mean difference of three points. Because the post hoc test did not show experimental consistency and the mean differences did not meet the level of experimental importance, the appropriate conclusion would be to fail to reject null hypothesis #5 as it related to the comparison of MontCAS reading scores among the three low SES configuration groups.

*Math scores for the three low SES configuration groups.* A factorial ANOVA was used to calculate the means and F-value for the math scores of the three low SES configuration groups (LSES K8, LSES EM, LSES IS). Among the low SES

configurations groups, the F-value was 11.71 resulting in a p-value of  $<.0001$ . This value met the *a priori* condition for experimental consistency ( $p<.05$ ). Because the *a priori* condition was met, a Tukey-Kramer post hoc test was conducted to determine if there was experimental consistency among the mean differences of the low SES configuration groups. The post hoc test did not show an experimental consistent difference between the means of the three low SES configuration groups.

The means for each low SES configuration group (LSES K8, LSES EM, and LSES IS) were 29, 28, and 27 respectively. The mean difference among the three low SES configuration groups did not meet the *a priori* level of an important mean difference of three points. Because the post hoc test did not show experimental consistency and the mean differences did not meet the level of experimental importance, the appropriate conclusion would be to fail to reject null hypothesis #5 as it related to the comparison of the MontCAS math scores among the three low SES configuration groups. In addition, because the p-value was extremely low, it could be concluded that there was a consistently unimportant mean difference in the MontCAS math scores among the three low SES configuration groups.

#### *Null Hypothesis #5c*

There will be no experimentally important or consistent mean difference in student math and reading scores on the MontCAS (Montana's CRT) between the two K-8 configuration groups defined by SES level.

*Reading scores.* A factorial ANOVA was used to calculate the means and F-value for the reading scores of the two K-8 configuration groups defined by SES (HSES K8 and LSES K8). Between these two groups, a Tukey-Kramer post hoc test calculated

an F-value of 0.42 resulting in a p-value of  $>0.05$ . This value did not meet the *a priori* condition for experimental consistency ( $p < .05$ ).

The means for each K-8 configuration group (HSES K8, LSES K8) were 39 and 38 respectively. The mean difference between the two K-8 configuration groups did not meet the *a priori* level of an important mean difference of three points. Because the post hoc test did not show experimental consistency and the mean differences did not meet the level of experimental importance, the appropriate conclusion would be to fail to reject null hypothesis #5 as it related to the comparison of MontCAS reading scores between the two K-8 configuration groups (HSES K8, LSES K8).

*Math scores.* A factorial ANOVA was used to calculate the means and F-value for the math scores of the two K-8 configuration groups defined by SES ( HSES K8 and LSES K8). Between these two groups, a Tukey-Kramer post hoc test calculated an F-value of 0.44 resulting in a p-value of  $>0.05$ . This value did not meet the *a priori* condition for experimental consistency ( $p < .05$ ).

The means for each K-8 configuration group (HSES K8, LSES K8) were 28 and 29 respectively. The mean difference between the two K-8 configuration groups did not meet the *a priori* level of an important mean difference of three points. Because the post hoc test did not show experimental consistency and the mean differences did not meet the level of experimental importance, the appropriate conclusion would be to fail to reject null hypothesis #5 as it related to the comparison of MontCAS math scores between the two K-8 configuration groups (HSES K8, LSES K8).

*Null Hypothesis #5d*

There will be no experimentally important or consistent mean difference in student math and reading scores on the MontCAS (Montana's CRT) between the two elemiddle configuration groups defined by SES level.

*Reading scores.* A factorial ANOVA was used to calculate the means and F-value for the reading scores of the two elemiddle configuration groups defined by SES (HSES EM, LSES EM). Between these two groups, a Tukey-Kramer post hoc test calculated an F-value of 105.83 resulting in a p-value of  $<.01$ . This value did meet the *a priori* condition for experimental consistency ( $p<.05$ ).

The means for each elemiddle configuration group (HSES EM, LSES EM) were 42 and 39 respectively, resulting in a mean difference of three points. The mean difference between the two elemiddle configuration groups did meet the *a priori* level of an important mean difference. Because the post hoc test showed experimental consistency and the mean differences did meet the level of experimental importance, the appropriate conclusion would be to reject null hypothesis #5 as it related to the comparison of MontCAS reading scores between the two elemiddle configuration groups defined by SES. (HSES EM, LSES EM).

*Math scores.* A factorial ANOVA was used to calculate the means and F-value for the math scores of the two elemiddle configuration groups defined by SES (HSES EM, LSES EM). Between these two groups, a Tukey-Kramer post hoc test calculated an F-value of 121.11 resulting in a p-value of  $<.01$ . This value did meet the *a priori* condition for experimental consistency ( $p<.05$ ).



The means for each elemiddle configuration group (HSES EM, LSES EM) were 32 and 28 respectively, resulting in a mean difference of four points. The mean difference between the two elemiddle configuration groups did meet the *a priori* level of an important mean difference. Because the post hoc test showed experimental consistency and the mean differences did meet the level of experimental importance, the appropriate conclusion would be to reject null hypothesis #5 as it related to the comparison of MontCAS math scores between the two elemiddle configuration groups defined by SES (HSES EM, LSES EM).

*Null Hypothesis #5e*

There will be no experimentally important or consistent mean difference in student math and reading scores on the MontCAS (Montana's CRT) between the two intermediate school configuration groups defined by SES level.

*Reading scores.* A factorial ANOVA was used to calculate the means and F-value for the reading scores of the two intermediate school groups defined by SES (HSES IS, LSES IS). Between these two groups, a Tukey-Kramer post hoc test calculated an F-value of 171.15 resulting in a p-value of <.01. This value met the *a priori* condition for experimental consistency ( $p < .05$ ).

The means for each intermediate school group (HSES IS, LSES IS) were 42 and 39 respectively, resulting in a mean difference of three points. The mean difference between the two intermediate school groups met the *a priori* level of an important mean difference. Because the post hoc test showed experimental consistency and the mean differences did meet the level of experimental importance, the appropriate conclusion would be to reject null hypothesis #5 as it related to the comparison of MontCAS reading

scores between the two intermediate school configurations groups defined by SES level (HSES IS, LSES IS).

*Math scores.* A factorial ANOVA was used to calculate the means and F-value for the math scores of the two intermediate school groups defined by SES (HSES IS, LSES IS). Between these two groups, a Tukey-Kramer post hoc test calculated an F-value of 323.08 resulting in a p-value of  $<.01$ . This value met the *a priori* condition for experimental consistency ( $p<.05$ ).

The means for each intermediate school group (HSES IS, LSES IS) were 32 and 27 respectively, resulting in a mean difference of five points. The mean difference between the two intermediate school groups met the *a priori* level of an important mean difference. Because the post hoc test showed experimental consistency and the mean differences met the level of experimental importance, the appropriate conclusion would be to reject null hypothesis #5 as it related to the comparison of MontCAS math scores between the two intermediate school configuration groups (HSES IS, LSES IS).

The results of the quantitative analysis of student performance on the MontCAS showed very few important and consistent mean differences among the groups defined by configuration and school size. There were a few differences among the groups defined by configuration and SES, however most of these differences were related to SES more than configuration. Finding few differences was contrary to the results of other comparison research outlined in Chapter Two. The analysis of the YRBS results in the next section may have helped to support the findings from the CRT analysis.

*Analysis of the Montana Youth Risk Behavior Survey Results**Introduction*

For the purpose of this research, student at-risk behavior has been defined as the occurrence of at-risk behaviors as self-reported on the YRBS given to seventh and eighth grade students in the spring of 2005. The YRBS is a survey developed by the U.S. Centers for Disease Control and Prevention. The purpose of the survey is to help monitor the prevalence of behaviors that put youth at risk for the most significant health and social problems, which can occur during adolescence and adulthood (Montana OPI, 2005b). The 2005 survey results were used because the survey changed in 2007. The 2007 survey eliminated several questions related to in school behavior. Not every survey question was used in this research. Appendix E has the list of questions used in this research.

The survey is optional. It is offered to every school in the State of Montana. Some schools/districts have made the decision not to administer the survey. In 2005, there were 149 schools who administered the survey to their seventh and eighth grade students (Montana OPI, 2005b). In 2005 there were 326 schools in the State of Montana who had the potential to serve seventh and eighth grade students (<http://www.nces.ed.gov/ccd/schoolsearch/>). If a school chose to participate in the survey, and if the school had larger than 100 students in the seventh and eighth grade combined, then a 100 random students were chosen to participate. If a school had less than 100 combined seventh and eighth graders, then all students in those two grades participated in the survey (Montana OPI, 2005b).

*Sample*

In 2005 there were 9,178 seventh and eighth grade students who participated in the Montana YRBS. That same year, there were a total of 23,971 seventh and eighth grade students attending schools in Montana (Montana OPI, 2006). Survey results are not readily available from the Office of Public Instruction. Each individual school district must be contacted to obtain survey results (S. Court, Director of Montana YRBS, personal communication, September 20, 2007). To minimize the effects of school size, location, and socio-economic status, YRBS results were requested from a select group of the total schools who participated in the 2005 survey. Appendix B shows the sample of schools who received a request to release their YRBS results. Results were requested from 69 different schools from across the State of Montana. Of these schools, 28 did not participate in the YRBS in 2005. From the 41 remaining schools, 24 schools submitted their individual school results for this research and 17 schools did not submit their results. The sample size was 1,971 students out of a possible 9,178 students who took the YRBS in 2005. For a population of 9,178 an appropriate sample size would need to be at least 369 students. However, the sample for this research was 1,971 students. This sample size was sufficiently large enough to produce a 99.9% confidence interval with a 0.05% margin of error. Table 4.15 presents demographic data related to the population and sample.

Table 4.15 *Demographic Data for Population and Sample of YRBS Schools*

	State of Montana	MT Schools who participated in YRBS	Sample Schools
Number of Schools serving 7 <sup>th</sup> & 8 <sup>th</sup>	326	149	24
Number of 7 <sup>th</sup> and 8 <sup>th</sup> grade students	23,971	9,178	1,971
Average Enrollment	172 students*	Unknown**	483 students
Percentage of Free or Reduced Lunch	32%*	Unknown**	36%

\* Represents all schools in State

\*\*Information not reported by OPI

*Categories.* The schools in the sample were separated into categories based on demographic information about the school. There were two categories based on school configuration. The K-8 schools (K8) had a one of the following grade configurations: K-8, K-6/7-8, K-2/3-5/6-8, or K-5/6-8. The intermediate schools (IS) had one of the following grade configurations: 5-8, 6-8, 7-8, or 7-9. There were two categories based on school SES. The SES was assessed by the school's participation rate in the free or reduced lunch program. Schools that had a participation rate of greater than or equal to 40% were classified as low SES schools (LSES). Conversely, schools with less than a 40% participation rate in the free/reduced lunch program were classified as high SES schools (HSES). Finally, there were two categories defined by school size. Schools with greater than or equal to 400 students were classified as large schools (LRG). Schools with less than 400 students were classified as small schools (SML). Table 4.16 provides demographic data for all categories in the sample.

Table 4.16 *Demographic Data for Categories of YRBS Sample Schools*

	<b>State of MT</b>	<b>Sample</b>	<b>ALL IS</b>	<b>ALL K8</b>	<b>HGH SES</b>	<b>LOW SES</b>	<b>LRG SCH</b>	<b>SML SCH</b>
Number of Students Surveyed	9,178	1,971	1,147	824	1,284	687	1,281	690
Number of Schools	149	24	11	13	13	11	14	10
Average Total Enrollment	Unknown	483	590	392	516	444	657	239
Average F/RL%	32.0%	36.0%	31.0%	40.0%	27.0%	46.0%	33.0%	39.0%
Females	49.9%	51.0%	53.0%	48.5%	52.2%	49.8%	51.6%	51.0%
Males	51.1%	49.0%	47.0%	51.5%	47.8%	50.2%	48.4%	49.0%
7 <sup>th</sup> Grade	47.7%	47.0%	47.0%	47.0%	48.7%	44.2%	46.0%	49.7%
8 <sup>th</sup> Grade	52.2%	53.0%	53.0%	53.0%	51.3%	55.8%	54.0%	50.3%

#### *Data Handling Procedure*

The YRBS asks students to report the occurrence of participation in at-risk behavior. For example, item number 14 on the survey asks students to report how many days (in the past 30) they carried a weapon such as a gun, knife, or club on school property. Students are given five choices to answer this question: (a) 0 days, (b) 1 day, (c) 2 or 3 days, (d) 4 or 5 days, (e) 6 or more days. For the purpose of this research, the reported percentage will represent the total number of students who answered the question with at least one occurrence of the at-risk behavior, divided by the total number of students who answered the question. Using the research sample, 94 out of 1,971 students reported at least one occurrence of bringing a weapon on school property in the past 30 days. So the reported percentage for this question would be 4.8% (94/1971).

*Comparison Tables and Percent Differences*

The following set of tables will compare YRBS answers for categories and sub-categories. In addition, each table contains a column of data for percent difference. For the purpose of this research, the percent difference will be the difference between the two categories being compared. A negative percent difference would demonstrate that the second category had a higher percent than the first category. The research methodology defined an important percent difference to be a difference of at least 5%

Table 4.17 *Comparison Table: State of Montana and Research Sample*

Descriptor	State of Montana	Sample	%Diff *
Number of Schools	149	24	
Number of Students Surveyed	9,178	1,971	
<b>At-Risk Behaviors at School</b>			
Weapon on school property; past 30 days	5.4%	4.8%	0.6%
Did not attend school due to fear of safety; past 30 days	6.2%	6.6%	-0.4%
Threatened with weapon at school; past 12 months	8.8%	9.3%	-0.5%
Personal property stolen or damaged at school; past 12 mos	33.3%	34.6%	-1.3%
Engaged in physical fight at school; past 12 months	20.9%	20.8%	0.1%
Smoked cigarettes on school property; past 30 days	3.4%	2.8%	0.6%
Used chewing tobacco on school property; past 30 days	3.0%	2.1%	0.9%
Drank alcohol on school property, past 30 days	2.9%	2.9%	0.0%
Used marijuana on school property, past 30 days	2.4%	2.4%	0.0%
Been offered/sold/given illegal drug at school; past 12 mos	13.8%	14.0%	-0.2%
<b>At-Risk Behaviors Not at School</b>			
Felt sad or hopeless and stopped regular activities; past 12 m	23.6%	23.7%	-0.1%
Actually attempted suicide in past 12 months	12.2%	12.9%	-0.7%
Tried cigarette smoking in lifetime	34.9%	33.0%	1.9%
Tried alcohol at least once in lifetime	55.8%	56.7%	-0.9%
Used marijuana at least once in lifetime	15.1%	15.6%	-0.5%
Used inhalants at least once in lifetime	15.0%	16.6%	-1.6%

\*% Diff (Percent Difference) is the difference between the State of Montana and Sample.

Table 4.17 represents a comparison between the sample (students from schools selected for this research) and the population (all students who took the YRBS in 2005). Note the small percent differences between the sample and the population.



Table 4.18 *Comparison Table: All Intermediate and K-8 Schools in Sample*

<b>Descriptor</b>	<b>ALL IS</b>	<b>ALL K8</b>	<b>%DIFF</b>
Number of Students Surveyed	1,147	824	
Number of Schools	11	13	
Average Total Enrollment	590	392	
Free and Reduced Lunch %	31.0%	40.0%	
<b>At-Risk Behaviors at School</b>			
Weapon on school property; past 30 days	4.2%	5.6%	-1.4%
Did not attend school due to fear of safety; past 30 days	6.6%	6.5%	0.1%
Threatened with weapon at school; past 12 months	10.7%	7.3%	3.4%
Personal property stolen or damaged at school; past 12 mos	34.0%	35.4%	-1.4%
Engaged in physical fight at school; past 12 months	18.4%	24.2%	-5.8%*
Smoked cigarettes on school property; past 30 days	3.1%	2.4%	0.7%
Used chewing tobacco on school property; past 30 days	1.9%	2.4%	-0.5%
Drank alcohol on school property, past 30 days	2.7%	3.2%	-0.5%
Used marijuana on school property, past 30 days	2.5%	2.7%	-0.2%
Been offered/sold/given illegal drug at school; past 12 mos	15.7%	11.7%	4.0%
<b>At-Risk Behaviors Not at School</b>			
Felt sad or hopeless and stopped regular activities; past 12 m	24.8%	22.1%	2.7%
Actually attempted suicide in past 12 months	13.8%	11.7%	2.1%
Tried cigarette smoking in lifetime	31.1%	35.6%	-4.5%*
Tried alcohol at least once in lifetime	53.8%	60.8%	-7.0%*
Used marijuana at least once in lifetime	16.4%	14.4%	2.0%
Used inhalants at least once in lifetime	15.8%	17.7%	-1.9%

\*Important percent difference - defined as a difference of at least 5%.

Table 4.18 shows the at-risk comparison between students from the sample attending IS and those attending K8. A negative %DIFF denotes a larger percentage for K8.

Table 4.19 *Comparison Table: All High SES and Low SES Schools in Sample*

<b>Descriptor</b>	<b>HGH SES</b>	<b>LOW SES</b>	<b>%DIFF</b>
Number of Students Surveyed	1,284	687	
Number of Schools	13	11	
Average Total Enrollment	516	444	
Free and Reduced Lunch %	27.0%	46.0%	
<b>At-Risk Behaviors at School</b>			
Weapon on school property; past 30 days	4.3%	5.7%	-1.4%
Did not attend school due to fear of safety; past 30 days	6.0%	7.7%	-1.7%
Threatened with weapon at school; past 12 months	7.9%	11.9%	-4.0%
Personal property stolen or damaged at school; past 12 mos	33.5%	36.7%	-3.2%
Engaged in physical fight at school; past 12 months	20.4%	21.6%	-1.2%
Smoked cigarettes on school property; past 30 days	2.5%	3.3%	-0.8%
Used chewing tobacco on school property; past 30 days	2.3%	1.8%	0.5%
Drank alcohol on school property, past 30 days	2.8%	3.1%	-0.3%
Used marijuana on school property, past 30 days	2.3%	3.1%	-0.8%
Been offered/sold/given illegal drug at school; past 12 mos	14.5%	13.1%	1.4%
<b>At-Risk Behaviors Not at School</b>			
Felt sad or hopeless and stopped regular activities; past 12 m	23.5%	24.0%	-0.5%
Actually attempted suicide in past 12 months	13.1%	12.4%	0.7%
Tried cigarette smoking in lifetime	30.8%	37.2%	-6.4%*
Tried alcohol at least once in lifetime	55.3%	59.4%	-4.1%
Used marijuana at least once in lifetime	14.3%	17.9%	-3.6%
Used inhalants at least once in lifetime	16.4%	17.1%	-0.7%

\*Important percent difference - defined as a difference of at least 5%.

Table 4.19 shows the comparison between students from the sample attending HSES and those attending LSES schools. A negative %DIFF denotes a larger % for LSES.

Table 4.20 *Comparison Table: All Large and All Small Schools in Sample*

<b>Descriptor</b>	<b>LRG SCH</b>	<b>SML SCH</b>	<b>%DIFF</b>
Number of Students Surveyed	1,281	690	
Number of Schools	14	10	
Average Total Enrollment	657	239	
Free and Reduced Lunch %	33.0%	39.0%	
<b>At-Risk Behaviors at School</b>			
Weapon on school property; past 30 days	5.3%	3.8%	1.5%
Did not attend school due to fear of safety; past 30 days	6.5%	6.8%	-0.3%
Threatened with weapon at school; past 12 months	9.4%	9.0%	0.4%
Personal property stolen or damaged at school; past 12 mos	34.4%	35.0%	-0.6%
Engaged in physical fight at school; past 12 months	19.6%	23.1%	-3.5%
Smoked cigarettes on school property; past 30 days	2.4%	3.5%	-1.1%
Used chewing tobacco on school property; past 30 days	1.8%	2.6%	-0.8%
Drank alcohol on school property, past 30 days	2.8%	3.2%	-0.4%
Used marijuana on school property, past 30 days	2.3%	3.0%	-0.7%
Been offered/sold/given illegal drug at school; past 12 mos	14.4%	13.3%	1.1%
<b>At-Risk Behaviors Not at School</b>			
Felt sad or hopeless and stopped regular activities; past 12 m	22.2%	26.5%	-4.3%*
Actually attempted suicide in past 12 months	12.8%	13.2%	-0.4%
Tried cigarette smoking in lifetime	32.8%	33.3%	-0.5%
Tried alcohol at least once in lifetime	55.8%	58.4%	-2.6%
Used marijuana at least once in lifetime	16.6%	13.7%	2.9%
Used inhalants at least once in lifetime	15.7%	18.3%	-2.6%

\*Important percent difference - defined as a difference of at least 5%.

Table 4.20 shows the comparison between students attending large schools with those attending small. A negative %DIFF denotes a higher percentage for small schools group.

Table 4.21 *Comparison Table: High SES Intermediate and High SES K-8 Schools*

<b>Descriptor</b>	<b>HSES IS</b>	<b>HSES K8</b>	<b>%DIFF</b>
Number of Students Surveyed	888	396	
Number of Schools	8	5	
Average Total Enrollment	572	425	
Free and Reduced Lunch %	25.0%	29.0%	
<b>At-Risk Behaviors at School</b>			
Weapon on school property; past 30 days	4.2%	4.5%	-0.3%
Did not attend school due to fear of safety; past 30 days	6.1%	5.8%	0.3%
Threatened with weapon at school; past 12 months	9.5%	4.1%	5.4%*
Personal property stolen or damaged at school; past 12 mos	33.3%	34.0%	-0.7%
Engaged in physical fight at school; past 12 months	17.5%	26.9%	-9.4%*
Smoked cigarettes on school property; past 30 days	2.7%	2.0%	0.7%
Used chewing tobacco on school property; past 30 days	1.9%	3.1%	-1.2%
Drank alcohol on school property, past 30 days	2.5%	3.6%	-1.1%
Used marijuana on school property, past 30 days	2.0%	2.8%	-0.8%
Been offered/sold/given illegal drug at school; past 12 mos	14.9%	13.8%	1.1%
<b>At-Risk Behaviors Not at School</b>			
Felt sad or hopeless and stopped regular activities; past 12 m	23.6%	23.2%	0.4%
Actually attempted suicide in past 12 months	14.3%	10.6%	3.7%
Tried cigarette smoking in lifetime	27.9%	37.2%	-9.3%*
Tried alcohol at least once in lifetime	52.8%	60.8%	-8.0%*
Used marijuana at least once in lifetime	14.3%	14.3%	0.0%
Used inhalants at least once in lifetime	14.4%	20.9%	-6.5%*

\*Important percent difference - defined as a difference of at least 5%.

Table 4.21 shows the comparison between students attending HSES intermediate schools with those from HSES K8. A negative %DIFF denotes a higher percentage for the K8.

Table 4.22 *Comparison Table: Low SES Intermediate and Low SES K-8 Schools*

<b>Descriptor</b>	<b>LSES IS</b>	<b>LSES K8</b>	<b>%DIFF</b>
Number of Students Surveyed	259	428	
Number of Schools	3	8	
Average Total Enrollment	637	371	
Free and Reduced Lunch %	47.0%	46.0%	
<b>At-Risk Behaviors at School</b>			
Weapon on school property; past 30 days	4.2%	6.5%	-2.3%
Did not attend school due to fear of safety; past 30 days	8.5%	7.2%	1.3%
Threatened with weapon at school; past 12 months	14.6%	10.3%	4.3%
Personal property stolen or damaged at school; past 12 mos	36.7%	36.7%	0.0%
Engaged in physical fight at school; past 12 months	21.3%	21.7%	-0.4%
Smoked cigarettes on school property; past 30 days	4.3%	2.8%	1.5%
Used chewing tobacco on school property; past 30 days	1.6%	1.9%	-0.3%
Drank alcohol on school property, past 30 days	3.5%	2.8%	0.7%
Used marijuana on school property, past 30 days	3.9%	2.6%	1.3%
Been offered/sold/given illegal drug at school; past 12 mos	18.5%	9.9%	8.6%*
<b>At-Risk Behaviors Not at School</b>			
Felt sad or hopeless and stopped regular activities; past 12 m	28.9%	21.0%	7.9%*
Actually attempted suicide in past 12 months	12.1%	12.6%	-0.5%
Tried cigarette smoking in lifetime	42.2%	34.2%	8.0%*
Tried alcohol at least once in lifetime	57.0%	60.9%	-3.9%
Used marijuana at least once in lifetime	23.6%	14.5%	9.1%*
Used inhalants at least once in lifetime	21.0%	14.7%	6.3%*

\*Important percent difference - defined as a difference of at least 5%.

Table 4.22 shows the comparison between students attending LSES intermediate schools with those from LSES K8. A negative %DIFF denotes a higher percentage for the K8.

Table 4.23 *Comparison Table: Large Intermediate and Large K-8 Schools*

<b>Descriptor</b>	<b>LRG IS</b>	<b>LRG K8</b>	<b>%DIFF</b>
Number of Students Surveyed	817	464	
Number of Schools	9	5	
Average Total Enrollment	634	699	
Free and Reduced Lunch %	31.0%	36.0%	
<b>At-Risk Behaviors at School</b>			
Weapon on school property; past 30 days	4.9%	6.0%	-1.1%
Did not attend school due to fear of safety; past 30 days	6.1%	7.1%	-1.0%
Threatened with weapon at school; past 12 months	10.6%	7.4%	3.2%
Personal property stolen or damaged at school; past 12 mos	33.5%	36.1%	-2.6%
Engaged in physical fight at school; past 12 months	16.6%	24.9%	-8.3%*
Smoked cigarettes on school property; past 30 days	2.6%	2.2%	0.4%
Used chewing tobacco on school property; past 30 days	2.0%	1.5%	0.5%
Drank alcohol on school property, past 30 days	2.8%	2.6%	0.2%
Used marijuana on school property, past 30 days	2.2%	2.4%	-0.2%
Been offered/sold/given illegal drug at school; past 12 mos	15.0%	13.5%	1.5%
<b>At-Risk Behaviors Not at School</b>			
Felt sad or hopeless and stopped regular activities; past 12 m	23.5%	19.9%	3.6%
Actually attempted suicide in past 12 months	13.0%	12.4%	0.6%
Tried cigarette smoking in lifetime	30.3%	37.3%	-7.0%*
Tried alcohol at least once in lifetime	51.3%	63.9%	-12.6%*
Used marijuana at least once in lifetime	17.0%	15.9%	1.1%
Used inhalants at least once in lifetime	16.0%	15.2%	0.8%

\*Important percent difference - defined as a difference of at least 5%.

Table 4.23 shows the comparison between students attending large intermediate schools with those from large K8. A negative %DIFF denotes a higher percentage for the K8.

Table 4.24 *Comparison Table: Small Intermediate and Small K-8 Schools*

<b>Descriptor</b>	<b>SML IS</b>	<b>SML K8</b>	<b>%DIFF</b>
Number of Students Surveyed	330	360	
Number of Schools	2	8	
Average Total Enrollment	392	200	
Free and Reduced Lunch %	31.0%	42.0%	
<b>At-Risk Behaviors at School</b>			
Weapon on school property; past 30 days	2.4%	5.0%	-2.6%
Did not attend school due to fear of safety; past 30 days	7.9%	5.8%	2.1%
Threatened with weapon at school; past 12 months	10.9%	7.2%	3.7%
Personal property stolen or damaged at school; past 12 mos	35.5%	34.5%	1.0%
Engaged in physical fight at school; past 12 months	22.8%	23.3%	-0.5%
Smoked cigarettes on school property; past 30 days	4.2%	2.8%	1.4%
Used chewing tobacco on school property; past 30 days	1.5%	3.6%	-2.1%
Drank alcohol on school property, past 30 days	2.4%	3.9%	-1.5%
Used marijuana on school property, past 30 days	3.0%	3.0%	0.0%
Been offered/sold/given illegal drug at school; past 12 mos	17.5%	9.5%	8.0%*
<b>At-Risk Behaviors Not at School</b>			
Felt sad or hopeless and stopped regular activities; past 12 m	28.1%	24.9%	3.2%
Actually attempted suicide in past 12 months	15.7%	10.8%	4.9%*
Tried cigarette smoking in lifetime	33.1%	33.5%	-0.4%
Tried alcohol at least once in lifetime	60.0%	56.9%	3.1%
Used marijuana at least once in lifetime	14.9%	12.5%	2.4%
Used inhalants at least once in lifetime	15.5%	20.8%	-5.3%*

\*Important percent difference - defined as a difference of at least 5%.

Table 4.24 shows the comparison between students attending small intermediate schools with those from small K8. A negative %DIFF denotes a higher percentage for the K8.

## *Hypothesis Analysis for Incidence of At-Risk Behavior*

### *Introduction*

Student at-risk behavior was defined as the occurrence of certain at-risk behaviors on the YRBS. In the State of Montana, the YRBS was given to 9,178 seventh and eighth grade students in the spring of 2005. This research examined a sample of 1,971 students attending Montana K-8 and intermediate schools in 2005. Demographic for the sample is listed in Table 4.15. Although there were more than 80 questions on the YRBS, this research focused on 16 questions from the survey (See Table 4.17 for a list of questions). These questions were divided into at-risk behaviors occurring at school and at-risk behaviors not occurring at school.

The results from the schools responding to this survey were divided into the following categories: (a) K-8 and intermediate schools, (b) small and large schools, (c) high SES and low SES schools. In addition, the results were divided into the following sub-categories: (a) large K-8 and intermediate schools, (b) small K-8 and intermediate schools, (c) low SES K-8 and intermediate schools, (d) high SES K-8 and intermediate schools.

The following hypotheses were presented in the methodology section of this research. Each hypothesis compares at least two categories or sub-categories. For each hypothesis, an important difference was defined *a priori* as a difference of 5% for any of the YRBS questions.



*Hypothesis #1*

There will be an important difference in the incidence of student at-risk behavior between schools with an intermediate school configuration and those schools with a K-8 configuration.

Table 4.18 compares the YRBS results for all the intermediate schools and all K-8 schools in the sample. From the sample of 1,971 students, 1,147 were attending intermediate schools and 824 were attending K-8 schools. Eleven of the sample schools were intermediate schools and 13 of the sample schools were K-8 schools.

In comparing the two groups (K-8 and intermediate schools) two YRBS results had greater than a 5% difference. First, 5.8% more K-8 than intermediate school students reported that they had engaged in a physical fight at school in the past 12 months. Second, 7.0% more K-8 than intermediate school students reported that they had tried alcohol at least once in their lifetime. In addition there were two YRBS questions which came close to reaching the *a priori* level of a 5% difference. There were 4% more intermediate school than K-8 students who reported that they had been offered/sold/or given an illegal drug at school in the past 12 months. In addition, 4.5% more K-8 than intermediate school students reported that they had tried cigarette smoking at least once in their lifetime.

From these results, the appropriate conclusion would be to accept hypothesis #1. For at least two of the YRBS results, there was an important difference in the incidence of student at-risk behavior between students attending schools with an intermediate school configuration, and those attending schools with a K-8 configuration.

*Hypothesis #2*

There will be an important difference in the incidence of student at-risk behavior between large schools with an intermediate school configuration and large schools with a K-8 configuration.

Table 4.23 compares the YRBS results for all large intermediate schools and all large K-8 schools in the sample. From the sample of 1,971 students, 817 were attending large intermediate schools and 464 were attending large K-8 schools. Nine of the sample schools were large intermediate schools and five of the sample schools were large K-8 schools.

In comparing the two groups (large K-8 and intermediate schools) three YRBS results had greater than a 5% difference. The first was an at-risk behavior reported at school. The other two results were at-risk behavior reported as non-school related behaviors. For those students attending large schools, 8.3% more K-8 than intermediate school students reported that they had engaged in a physical fight while at school in the past twelve months. For those attending large schools, 7% more K-8 than intermediate school students reported that they had tried cigarette smoking at least once in their lifetime. For those attending large schools, 12.6% more K-8 than intermediate school students reported they had tried alcohol at least once in their lifetime.

From these results, the appropriate conclusion would be to accept hypothesis #2. For at least three of the YRBS comparisons, there was an important difference in the incidence of student at-risk behavior between those attending large schools with an intermediate school configuration, and those attending large schools with a K-8 configuration.

*Hypothesis #3*

There will be an important difference in the incidence of student at-risk behavior between small schools with an intermediate school configuration and small schools with a K-8 configuration.

Table 4.24 compares the YRBS results for all small intermediate schools and all small K-8 schools in the sample. From the sample of 1,971 students, 330 were attending small intermediate schools and 360 were attending small K-8 schools. Two of the sample schools were small intermediate schools and eight of the sample schools were small K-8 schools.

In comparing the two groups (small K-8 and intermediate schools) there were three YRBS results that had greater than a 5% difference. The first was an at-risk behavior reported at school. The other two results were at-risk behavior reported as not necessarily school related. For those students attending small schools, 8% more intermediate school than K-8 students reported that they had been offered/sold/or given an illegal drug at school in the past 12 months. For those students attending small schools, 4.9% more intermediate school than K-8 students reported they had actually attempted suicide in the past 12 months. For those attending small schools, 5.3% more K-8 than intermediate school students reported they had used inhalants at least once in their lifetime.

From these results, the appropriate conclusion would be to accept hypothesis #3. For at least three of the YRBS results, there was an important difference in the incidence of student at-risk behavior between those attending small schools with an intermediate school configuration, and those attending small schools with a K-8 configuration.

*Hypothesis #4*

There will be an important difference in the incidence of student at-risk behavior between high SES schools with an intermediate school configuration and high SES schools with a K-8 configuration. (high SES schools are defined as those with <40% F/RL)

Table 4.21 compares the YRBS results for all high SES intermediate schools and all high SES K-8 schools in the sample. From the sample of 1,971 students, 888 were attending high SES intermediate schools and 396 were attending high SES K-8 schools. Eight of the sample schools were high SES intermediate schools and five of the sample schools were high SES K-8 schools.

In comparing the two groups (high SES K-8 and high SES intermediate schools) five YRBS comparisons had greater than a 5% difference. Two were reported as at school behaviors. The other three results were non-school related at-risk behaviors. For those students attending high SES schools, 5.4% more intermediate school than K-8 students reported that they had been threatened with a weapon at school in the past 12 months. For those attending high SES schools, 9.4% more K-8 than intermediate school students reported that they had engaged in a physical fight while at school in the past 12 months. For those attending high SES schools, 9.3% more K-8 than intermediate school students reported that they had tried cigarettes at least once in their lifetime. There were 8.0% more K-8 students who had reported that they had tried alcohol at least once in their lifetime. For those attending high SES schools, there were also 6.5% more K-8 than intermediate school students who reported they had used inhalants at least once in their lifetime.

From these results, the appropriate conclusion would be to accept hypothesis #4. For at least five of the YRBS comparisons, there was an important difference in the incidence of student at-risk behavior between those attending high SES schools with an intermediate school configuration, and those attending high SES schools with a K-8 configuration.

#### *Hypothesis #5*

There will be an important difference in the incidence of student at-risk behavior between students attending low SES schools with an intermediate school configuration and those attending low SES schools with a K-8 configuration (Low SES >40% F/RL).

Table 4.22 compares the YRBS results for all high SES intermediate schools and all high SES K-8 schools in the sample. From the sample of 1,971 students, 259 were attending low SES intermediate schools and 428 were attending low SES K-8 schools. Three of the sample schools were low SES intermediate schools and eight of the sample schools were low SES K-8 schools.

In comparing the two groups (low SES K-8 and intermediate schools) five YRBS results had greater than a 5% difference. One comparison was reported as an at school at-risk behavior. The other four comparisons were reported as non-school related at-risk behaviors. For those students attending low SES schools, 8.6% more intermediate school than K-8 students reported that they had been offered/sold/or given an illegal drug at school in the past 12 months. For those attending low SES schools, 7.9% more intermediate school than K-8 students reported that they had felt sad or hopeless and stopped regular activities in the past 12 months. For those attending low SES schools, 8.0% more intermediate school than K-8 students reported that they had tried cigarettes at

least once in their lifetime. There was 9.1% more intermediate school than K-8 students who had reported that they had used marijuana at least once in their lifetime. For those attending low SES schools, 6.3% more intermediate school than K-8 students reported they had used inhalants at least once in their lifetime.

From these results, the appropriate conclusion would be to accept hypothesis #5. For at least five of the YRBS results, there was an important difference in the incidence of student at-risk behavior between those attending low SES schools with an intermediate school configuration, and those attending low SES schools with a K-8 configuration.

The null hypothesis analysis of CRT scores showed very few differences among schools of various configurations. The differences that did exist were more directly tied to SES rather than configuration. Similarly, the analysis of incidence of at-risk behaviors as measured by the YRBS found very few differences based solely on configuration. From these two results, the conclusion might be to assume that there was no relationship between school configuration and student success. The final qualitative analysis may have helped to explain and support this conclusion.

### *Results of One-on-One Comparison Interviews*

#### *Introduction*

The purpose of the qualitative phase of this study was to examine the potential similarities and differences in school characteristics, which may exist between intermediate schools and K-8 schools. The qualitative data was gathered through one-on-one interviews with educators from both types of configurations. The data from the interviews was used to explore specific school characteristics related to configuration and

school organization. This qualitative information may help explain the relationship between grade configuration and student success.

Educators were selected from eight schools in one Montana community. The purpose for selecting only one community was to limit the effects of environmental differences among the schools. For each school, one teacher and one administrator were interviewed. The interviews took place in the school setting. All interviews were conducted during the school day. The timeline for the interviews spanned from May to September of 2008. Each subject signed a consent form to participate in the interviews (See Appendix C). In addition, each subject was asked the same questions (See Appendix D). Beyond the questions related to configuration, each subject was asked two questions particular to their experience. First, each subject was asked to clarify their current position, previous work experience, and years of experience in the field of public education. Second, each subject was asked to comment about their familiarity with configurations other than the school where they currently worked.

### *Participants*

Interview subjects were chosen from eight schools within the county of Missoula (Montana). Schools represented a variety of configurations, sizes, and socio-economic levels. School demographic information is presented in Table 4.25.

Table 4.25 *Demographic Information: Schools Participating in Interviews*

School Name	Configuration	Size	Free/Reduced Lunch Percentage
K-8 School A	K-8	378	55%
K-8 School B	K-8	187	43%
K-8 School C	K-8	1,200	31%
K-8 School D	K-8	595	41%
K-8 School E	K-8	378	29%
Intermediate School A	Grades 6-8	546	52%
Intermediate School B	Grades 6-8	507	27%
Intermediate School C	Grades 6-8	599	32%
Average		549	39%

There were three types of educators interviewed for this research. From each school, one teacher and one administrator were selected for an interview. In addition, five educators who had previous experience in both grade configurations were also selected for an interview. These educators were all currently working as administrators in other settings (K-5) or they were currently serving as superintendents. In total there were twenty-two interviews conducted. Ten of these interviews were completed with K-8 teachers and administrators. Seven of these interviews were completed with intermediate school teachers and administrators. Five of these interviews were completed with educators who had work experience in both settings.



*Teacher demographics.* The teachers were selected by the school administrator. The teachers and administrators were asked the same questions, however the interviews occurred separately. Teachers were reminded that their answers would not be shared with the administration at their school. There were only two requirements for teachers to participate in this research. First, the teachers were required to be assigned as intermediate-grade teacher (grade 6, 7, or 8). Second, the selected teachers were required to have at least five years of teaching experience. From the nine teachers selected, two were sixth grade teachers, two were seventh grade teachers, and two were eighth grade teachers. The remaining three teachers were assigned to teach both seventh and eighth grade. With regard to curriculum, the subjects' teaching assignments varied. Each subject was assigned to teach at least one of the following content areas: reading, mathematics, social studies, communication arts, and health enhancement. Table 4.26 presents basic demographic information for all the subjects interviewed.

Table 4.26 *Demographic Data: Subjects Participating in Comparison Interviews*

Group	Number of Subjects	Years of Experience	School Size	Free or Reduced Lunch Percentage
K-8 Teachers	5	Average 16 (9 – 30)*	Average 540 (187 – 1,200)	Average 39% (29 – 55%)
K-8 Administrators	5	Average 21 (9 – 32)	Same as K-8 Teachers	Same as K-8 Teachers
Intermediate School Teachers	4	Average 12 (6 – 17)	Average 551 (507 – 600)	Average 37% (32 – 52%)
Intermediate School Administrators	3	Average 28 (25 – 32)	Same as IS Teachers	Same as IS Teachers
Educators with experience in both settings	5	Average 22 (15 – 33)	Average 560 (292 – 1,200)	Average 34% (29 – 51%)
Total number of Interviews	22			

\* Numbers in parenthesis represent range of data.

### *Interview Analysis*

The data from the comparison interviews was analyzed using a multi-step process outlined by Creswell (1998). Although the complete process is outlined in the methodology section of this research, the process can be summarized in the following four steps: (a) organizing the data, (b) clarifying the data by reading and memoing, (c) classifying and interpreting the data, and (d) representing the data in a narrative and visual form. After classifying and interpreting the data, responses from those educators who had worked in both settings was used to verify the responses given by the educators who had experience with just one configuration.

*Categories and sub-categories.* The purpose of completing interviews with educators working in the field was to explore and compare the characteristics of the intermediate schools with those of the K-8 schools. Through the process of comparison, the following major categories emerged: (a) organizational structure, (b) barriers to student success, (c) student characteristics, and (d) advantages and disadvantages for each configuration.

### *Organizational Structure*

Each educator who participated in an interview was asked to describe the organization or structure of their intermediate grades (6,7,8). The structures were described by one of the following; a team structure or a departmentalized structure. Schools who used the teaming concept generally defined their teams as a small group of teachers (2-3) who had a small group of students in common. Schools who defined their upper grades as departmentalized generally had subject area specialists who taught a specific subject to seventh and eighth grade students. Overall the intermediate school educators report their schools to be structured around a team concept. Most had at least two teams at the sixth and seventh grade level. One school had multiple teams at the eighth grade level, while the other two intermediate schools report just one team at the eighth grade level. In the K-8 schools, educators generally reported the organization of their upper grades to be defined by a department structure. Most all the K-8 schools in the sample had content area specialists teaching one subject at the seventh and eighth grade level. Table 4.27 presents the organizational structure of each school participating in the comparison interviews.

Table 4.27 *Organization of Schools Participating in Comparison Interviews*

School Name	Structure of Lower Grades (K-5)	Structure of Upper Grades (6-8)
K-8 School A	Self Contained Elementary Model	Content Area Specialists (6,7,8) Teach one subject to all three grades
K-8 School B	Self Contained Elementary Model	Content Area Specialists (6,7,8) Multiple subjects to all three grades
K-8 School C	K-2, 3-5 Separate buildings Self contained classrooms	Grade 6 – Multiple Teams, Generalists Grade 7/8 – Specialists teaching one subject to both grade levels
K-8 School D	Self Contained Elementary Model	Grade 6 – 2 Teacher Team Grade 7/8 – Specialists teaching one subject to both grade levels.
K-8 School E	Self Contained Elementary Model	Grade 6 – 2 Teacher Team Grade 7/8 – Specialists teaching one subject to both grade levels.
Intermediate School A	No Lower Grades	Grade 6 – 2 Teams, Generalists Grade 7 – 2 Teams, Some Specialists Grade 8 – 1 Team, Specialists
Intermediate School B	No Lower Grades	Grade 6 – 2 Teams, Generalists Grade 7 – 2 Teams, Generalists Grade 8 – 1 Team, Specialists
Intermediate School C	No Lower Grades	Grade 6 – 3 Teams, Generalists Grade 7 – 2 Teams, Generalists Grade 8 – 2 Teams, Generalists

*Specialists and generalists.* The structure of the schools in the sample is defined by the characteristics of the intermediate-grade teachers. In the K-8 schools, the majority of upper grade level teachers are considered content area specialists. These teachers focus on one subject area and teach that subject to multiple grade levels. Most often

these specialists have received college training in one subject area and generally hold a secondary teaching certification. By contrast, in the intermediate schools the majority of teachers are considered content generalists. These teachers focus on multiple subjects and teach those subjects to one grade level or one team of students. The generalists have typically received their training in multiple content areas and hold elementary teaching certification.

Those interviewed reported advantages and disadvantages related to the use of specialists and generalists. Most K-8 educators reported value in the use of specialists. By contrast, intermediate school educators reported the value of generalists. Educators who had experience in both configurations verified the advantages and disadvantages to the use of specialists in both the K-8 and intermediate school setting.

According to K-8 principal #1:

Sixth grade was self-contained when I first got here. We switched my second year because I wanted our students taught by a certified specialist. ...they attend my math specialist, my English specialist, my science specialist, and my social studies specialist. Five years ago when I first came out to (school name), not one of my junior high teachers at that time were subject endorsed.

K-8 teacher #1 reported the following:

I am a firm believer in secondary certification and it is kind of on its way out unfortunately, in some schools. I believe that I have tons of knowledge on that subject and if you have an elementary degree and you are suppose to teach math, science, social studies, you are not going to have nearly the depth of knowledge that teachers who are trained in the subject do. I think that benefits the kids.

When commenting on the use of specialists, intermediate school teacher #1 stated:

You see too many kids in one day. We have one specialist on our team who can only teach one subject and that really dictates how our teams then get organized because she can only teach H/PE. So that locks you into some things, where as people could teach more things and you could get your team numbers down.

Intermediate school principal #1 reported:

I believe that the team structure does (promote close teacher-student relationships) because our teachers teach more than one subject, so the majority of our teachers have the same student more than once a day, which allows them ...to get to know the kids sooner. If you are in a traditional junior high where you see 150 kids, once a day, it takes longer to get to know them.

An educator (#1) with experience in both settings provided the following verification:

[In the K-8] there was, I think a greater emphasis on subject area specialty. One of the things that I've seen, as our district has transitioned to the middle school, was that there seems to be a greater emphasis on hiring people with an elementary endorsement who maybe don't have the subject area specialty. I do think that in some respects it is beneficial to have people with that greater expertise particularly with seventh and eighth grade. I think when we have generalists teaching, for example some of the more advanced sciences and math, the quality of the instructional program is not great.

Educator #1 also believed there was some value in using generalists.

That being said I think there is a different approach to teaching from an elementary (certified) person that is more of a hands-on, more tolerant and more

accepting. I do think there is a greater emphasis in bringing the difficult topics together and not teaching in isolation. I think this is beneficial.

### *Barriers to Student Success*

Those interviewed were asked to identify specific challenges or barriers to success among intermediate-grade students. From their responses the following themes were common among both types of configurations. Students in the intermediate grades possess the following characteristics which can create challenges in school: (a) a need for independence, (b) a lack of self-esteem or self-concept, (c) social pressures related to conformity and acceptance.

*Independence.* Teachers and administrators who were interviewed reported that intermediate grade students typically struggle with the concept of independence. They are consistently seeking independence by testing their boundaries and questioning adults. This behavior can create problems for those students who do not possess appropriate decision making skills to cope with the responsibility of independence.

K-8 teacher #1 stated the following:

(They) are trying to figure out what independence means. They don't want their teachers or parents babysitting them at this point. Yet they don't necessarily have the skills to take care of themselves. ...at seventh grade we really baby them. We are calling home all the time. Then at eighth grade we are trying to let them go a little more. Their parents are trying to let them have a little more independence and so are we. Some can handle it and some can't.

Intermediate school teacher #2 agreed that seeking independence was an issue.

You know that is probably one of the biggest challenges. Finding out who they are and they are trying to make their own decisions. Getting away from their parents a little bit and wanting that freedom. Pressure from social, not just the drug and alcohol choices, but also it is the do I do my homework or go out and play? Do I go out for this sport or that sport? Some have it down and some come to me and ask for help.

An educator (#2) who had experience in both settings verified the complex issue of independence in the following statement.

I think their biggest issue is experimentation with life. Experimentation with academics and pushing myself harder or not pushing myself at all. Where some students just excel as hard as they can and some suddenly decide I'll check out. That is a challenge, how do we allow kids to take risks that allow them to experience something but not cause a life changing experience. They need opportunity to explore without getting hurt.

*Self-esteem or self-concept.* Teachers and administrators agreed that a lack of self-esteem or a poor concept of oneself was an issue that created challenges in educating intermediate-grade students. Those interviewed believe that the majority of intermediate grade students do not have a positive concept of themselves. This lack of self-esteem can affect the students' attitude towards success.

K-8 teacher #2 stated:

I think self-esteem is a huge issue. Trying to maneuver their way to what they have been taught as being right and then kind of experimenting with is it right or



not and they put all the things they know and what their parents believe to be true and right whether it be political or religious...and then they begin to play around with [reality] and sometimes it doesn't come out the way it should.

Intermediate school teacher #3 reported:

So many middle school kids are basically unclear, uncertain, and dissatisfied with who they are. They come into a whole group of kids like that and when they are all like that, no one feels safe. I am consistently surprised by seeming competent teenagers who still tell me how isolated they felt, whether it's around a physical thing, they're taller or shorter or fatter. Whether it is about a social economics factor, an intellectual factor, or something at home. All of those things start really bothering them at once and it starts bothering all of them.

Intermediate school principal #1 stated the following:

I think the biggest issue is self-perception. It goes from being accepted at home to being accepted as part of a peer group. They are moving from the focus of family to the focus of friends. Developmentally it is right where they are supposed to be. So I think that is the number one issue for kids and it goes along with the development of self-concept and all that is intertwined into social pressures.

*Conformity and acceptance.* Teachers and administrators interviewed believe that the pressure to conform and be accepted by the peer group creates challenges for educating the intermediate-grade student. As already stated, students are at a developmental stage where they are struggling with self-concept and independence. These issues manifest into a desire to be accepted by their peers. This peer pressure can lead to making poor decisions and failure.

Intermediate school principal #1 stated:

Acceptance is a major issue. They believe it is more important to be accepted by peer group rather than family. This creates more susceptibility to peer pressure and creates more social issues. They need to be accepted, they need to conform and not be different. That can affect a very bright student who does not want to be perceived as being bright because that would be different.

Intermediate school teacher #4 reported:

I think by the time they reach eighth grade, they have definitely formed their cliques. There is a lot more problems with the socializing of all the kids because they've had time to know who is who. There are problems with isolation for those who don't fit in. Sometimes you see loners or just a couple of pairs who stick out. There is diversity in physical size and intellectual development. These differences show and some kids are cruel and point out differences.

According to K-8 teacher #3:

There are big social issues. They are experiencing lots of changes in themselves. They are growing, their voices are changing, and they have stress related to their physical appearance. I just think it is a real awkward stage for them. They constantly seem to be trying to adjust and fit in. I think that has to do with the media and all the stuff out there today that we did not have to deal with.

### *Characteristics of Successful Students*

Those educators interviewed were asked to identify characteristics of successful students. The purpose was to determine if student success had any relationship to grade configuration. By comparing their answers, it was obvious that all successful student

possessed specific characteristics, regardless of the organization or structure of the school. From their responses the following themes were common among both types of configurations. Successful students in the intermediate grades possess the following characteristics: (a) personal support, (b) learner qualities, and (c) personal competence. Two of these characteristics, personal support and competence were reported as important findings in a previous dissertation that compared grade configurations (Freeman, 2005). Another characteristic which was only reported by the K-8 educators was the issue of transient students. A family that moved around a lot was reported as a characteristic which made it difficult for the student to experience success.

*Personal support.* Consistently, both intermediate school and K-8 educators reported the level of personal support a student received from outside of school was a key indicator for success in school. Typically the personal support was expressed as support from parents and other family members. Also, successful students had a supportive home-life outside of school. Those students, who struggled with chaos outside of school, typically struggled with school expectations. These findings were consistent across schools, regardless of configuration or organizational structure.

Intermediate school principal #2 defined support from home as follows:

The kids that succeed here are the kids who succeed in any school. It comes down to the support that they get from home and what skills they bring to school. It is the parent who is not afraid to call and ask. The parent who makes sure that if they are not seeing homework, they call and ask why. It is the parent who has a study time set up or asks where is the homework? It is totally related to the support of parents.

According to K-8 principal #1, personal support is as follows:

I don't see it as any different than any other school configuration. Students that do best are the ones who come from a family background that values education and supports education. [A family] that is able to give both academic and social/emotional support at home. The ones who do best are the ones who are prepared to meet any of those given challenges.

Intermediate school teacher #2 described personal support as follows:

The student who does best is the one who has the self confidence and the support from home in all ways. The support from home in terms of basic emotional, mental, and physical needs being met. Those are things we can't provide.

K-8 teacher #4 also agrees that personal support starts at home. He stated:

I would say that here in our school district we definitely have a pretty good share of kids that come from those tough home environments. I think when certain needs aren't met at home, it obviously makes our jobs as educators that much more challenging.

K-8 teacher #5 describes a characteristic that may be unique to the K-8 setting.

The [student] that would succeed the best is the one who comes with generational support or a family that has lived here and attended this school. When grandparents can come into the school and say "I remember when I was here," it builds support for the student and the school. There is a connection that I feel is critical to getting education accomplished because if they have questions they feel they can come in and ask. It helps students feel more secure. The family connection is huge.

An educator (#1) who has experience in both configurations verified that home support was important for student success in both settings. He stated:

[Successful kids] have the appropriate support at home and have a good academic foundation. The kids that do not have support at home and struggle academically will have difficulty in both settings. The kids that have appropriate support and social skills will flourish in either setting.

*Learner qualities.* Educators agree that successful students possess specific learner qualities that enable the student to be successful in the academic setting, regardless of the grade organization or the teaching structure. Both K-8 and intermediate school educators agreed that some of the following learner qualities are essential to academic success. Successful students are organized, can manage their time, can manage multiple expectations, take responsibility for their own learning, and can adjust to different learning environments.

In describing a successful student, intermediate school teacher #1 reported:

[They are] attentive. I mean they look at you and they listen to instructions, consistently day in and day out. They are focused. They listen well. More organized. They are able to separate themselves from distractions or if they are distracted they are able to pull themselves back.

A K-8 teacher (#1) described the successful student as “programmed.” She stated:

The word I keep using is they are programmed. They have had eight years of sitting in the desk quietly doing their homework raising their hands. They are totally programmed. I think this school creates programmed children. I think whatever is happening here we are very good at creating a good student.

The K-8 principal (#2) from this same school agreed with the teacher's assessment.

The academic rigor is higher here than other schools I've been in. A [successful] student would be in that sort of perfectionist, type-A personality area. They may spend way too much time on academic pursuit to the detriment of other things because there is so much work to be done. Those who strive on structure and academic rigor would be successful here.

*Personal competence.* Educators from both K-8 and intermediate schools agreed that successful students had a high degree of personal competence defined by a positive self image and appropriate attitude towards academics. Several educators described this student as "well-adjusted" with appropriate social skills and level of maturity. The educators who participated in this research believed that a lack of self-esteem could lead to issues with acceptance and conformity. The pressure to conform to standards set by the peer group could cause students to lose focus on academic expectations.

Intermediate school teacher #3 reported the following:

Students who lack self confidence do not feel safe and are unable to focus on learning. They are not comfortable in their own skin and therefore they are not safe with their peers because they know somebody is going to find something that is wrong with them. Something they already knew about themselves or something that they did not know. They are afraid to take any risks.

K-8 teacher #2 described how a lack of self-esteem could impact academic performance.

A specific example, is there was a girl that came in part way through last year and she described her experience as kids looked at her and made fun of her hair and

her clothes, and she looked just like everyone else. She was an athlete. That attack on her self-esteem carried over to this school year and caused some issues. An educator (#2) who had experience in both settings described social isolation issues that can develop from a lack of self-esteem.

In either setting, you have some students that are just alone. They do not have anybody they are connected to. I guess you could find that in either setting because you just have some students who can't find their way or their social skills are not good enough to help them establish relationships. Students develop a lack of self-esteem from experiencing a poor track record in building relationships.

*Transient student.* Most of the K-8 educators agreed that students from transient families had a particularly low level of success because they were expected to conform into a long established peer group. This characteristic may be unique to the K-8 setting. K-8 teacher #5 reported the following:

The student who does not do well is one that comes from our transient population. I think it is very difficult to come into a school that has been very close knit and you are the new kid. A new seventh grade student has to establish his reputation and our students don't really care where he has come from. I think it is hard for the new student because our kids have known each other for so long.

K-8 principal #2 agreed that the K-8 setting can be difficult for new students. He stated:

There is a function of that because we are smaller and the kids that have been here for awhile already know each other, there is not a lot of change in process when a new person arrives. There is not melting pot as students enter the upper grades. The students already know each other very well. So as a new person coming in

you are the only person experiencing that newness. It depends on their personality. If they are flexible and adaptable, they will fit in.

K-8 principal #3 reported the following concern related to transient students:

We have a transient rate of 28 to 33%, so if you get kids that have been in five or six different schools and have only been there for a year or two, you see that often those kids don't get along as well. I don't think it is because our kids aren't nice or accepting. I think we get kids who transfer in and we know they may only be here for two or three months and they may be packing some external baggage. Those kids struggle because our other kids have been together for 8 years and there is a real bond.

#### *Advantages of Both Configurations.*

Each educator who participated in an interview was asked to provide examples which clarify the advantages of their individual school configuration. The advantages described by the K-8 educators were divided into the following four categories: relationships, mentoring opportunities, communication, and looping. By contrast, as reported by the intermediate school educators, the advantages of the intermediate school configuration could be explained by one of the following categories: organizational structure, developmentally responsive program, and teaming.

#### *K-8 Advantages*

The advantages of the K-8 structure could be summarized generally by the power of the relationships within the K-8 school. Seven of the ten K-8 educators believe that the major strength of the K-8 school was the potential to develop long-term relationships. As reported by the K-8 educators, these long term relationships exist among students,



staff, and families. Educators reported that strong relationships exist for several reasons. First, the small size of the school promotes a sense of community where every student is known by everyone else. Second, students attend school for nine years with the same group of students. Third, students can maintain connections with previous teachers. Fourth, siblings can attend school together, which helps create a family connection to the school.

A K-8 principal (#3) described the strength of the relationships as follows:

I guess at the K-5 level it is pretty much like any other place. I believe just the smallness with about 230 students provides an environment for great relationships. We have a situation where every kid knows every other kid. I have always thought that our strengths are related to our small school set up, with one sixth, one seventh, and one eighth grade class. You talk about teachers getting to know kids. It really happens here.

A K-8 teacher (#4) describes the benefit of teacher and student relationships in the K-8.

When an older kid walks down the halls they can see a previous teacher. It's not just one set of eyes. It's the whole staff looking out for the growth of the child. We always hear stories how kids were five years ago and how they are now and how they (the previous teacher) have had a chance to see them change.

Educators from the K-8 schools noted three other advantages directly tied to the strength of relationships in the K-8 setting. The other advantages were: older students mentoring younger students, increased opportunities for communication, and a looping organizational structure for students in the upper grade levels.

*Mentoring.* The majority (8 out of 10) of the K-8 educators reported an advantage related to opportunities for older students to mentor younger students. Those interviewed generally report that these mentoring opportunities benefit not only the younger students but also the older students. Younger students are exposed to a positive role model. For the older students, mentoring provides an opportunity for success and encourages responsibility and a sense of ownership in the school as a community.

K-8 teacher #4 reported the following:

Eighth graders are always looking down, but it (mentoring) is an opportunity for them to teach and model. And I have seen some older students who are absolutely fantastic with younger kids. Even for those who have a tough time in the classroom, when they get one on one with some younger students it's just amazing to see them shine.

Another K-8 teacher (#5) described the benefits of mentoring as follows:

I have seen an older child who was struggling academically and emotionally. Sometimes they were allowed to tutor a younger student or be a classroom helper for a lower grade teacher. I could just really see the differences in the older student, giving them responsibility and many of the younger children loved them unconditionally which may have been the first time in their life they've experienced that. These are at-risk kids who have benefited.

*Communication.* Most all of the K-8 educators reported an advantage related to opportunities for communication in the K-8 setting. Eight out of ten of the educators reported that the K-8 setting helped improve communication between the school and families. Five out of ten reported that the K-8 setting helped to foster communication

among the teaching staff. Three of the K-8 educators reported that the K-8 setting provides consistency for parents with regard to rules and procedures. This consistency helps to ease parent concerns and increase the sense of community at the school.

A K-8 principal (#4) summarizes the benefits of communication among the staff.

We regularly communicate with staff at the lower grade levels concerning families and students. We have our school psychologist, school nurse, school counselor, and special education director meet often so we have communication and follow through with students and families. This communication helps ease transition between the lower and upper grades.

Another K-8 principal (#2) reported the following:

I tell parents the benefits of our school. Basically what they can expect is some level of consistency that they have their kids in the same building through their entire elementary years. They see that as comforting. They get to be here from kindergarten through eighth grade which means they not only know the teachers well but also the teachers know them well. Teachers working in the middle school can lean on information given to them by the K-5 teachers. And because they are all in the same building, that flow of information and basic collaboration does lend itself to a better serving of the students both academic and behavior needs.

*Looping structure at the upper grade levels.* Ninety percent of the K-8 educators report that the organizational structure of their upper grade levels helps to improve student – teacher relationships. For all of the K-8 schools in this research, students receive specific subject area instruction from the same teacher for both seventh and

eighth grade. For example, the math teacher at K-8 School E teaches math to both seventh and eighth grade students. This process of experiencing the same teacher for two or more years is described as looping. At least at one school, students receive instruction from the same teacher for sixth, seventh, and eighth grade. The K-8 educators believed that the process of looping helped build strong relationships between students and teachers. In addition, teachers believed that looping helped improve communication between students, teachers, and families. As students transitioned between seventh and eighth grade, there was no break-in period. Students seem to know the expectations from teachers and seemed to understand specific classroom procedures.

A K-8 principal (#5) stated the following:

I think the student – teacher relationships are pretty good. I think it is dependent on the amount of time the teacher has to spend getting to know each student. In the seventh and eighth grade most teachers have those students for two years. It (the looping structure) lends itself to the teachers being able to get to know their kids a lot better than if they would only see the student once per day for one year.

K-8 teacher #1 commented on the benefits of looping.

We all teach the seventh and eighth grade. So we have them for two years. The difference in my relationship with a seventh grader compared to an eighth grader is night and day because they have had me for a full year. It is the first time in their educational career where they come back from summer break and they know all their teachers. Nothing is new and nothing to be scared of and they feel comfortable with their teachers. I think our relationship with eighth graders is very good because we have already had a year to try and work things out.

### *Intermediate School Advantages*

Intermediate school educators who were interviewed agree that the strength of the grade configuration could be summarized in the following categories: organizational structure, creating programs which are developmentally responsive to the needs of the young adolescent student and teaming. Educators reported that the team structure in the intermediate school was effective for building positive relationships between students, teachers, and families. In addition, educators believed that having a separate school for sixth, seventh, and eighth graders helped to focus programs to meet the developmental needs of this level of student.

*The power of teaming.* The intermediate school educators who participated in this research believed that the team organization in the intermediate school was an effective method of delivering instruction and creating a positive learning environment for all. They reported several advantages related to the team structure. First, having a small group of students work with a team of two or three teachers helped foster communication. Teachers could discuss specific learning behaviors of each student and tailor instruction to meet student needs. Second, the teaming process helped to build rapport and trust between students and teachers in a short period of time. Third, the collaboration process between teachers provided support for at-risk students who may be lacking support from home. Fourth, the team structure allowed for interdisciplinary instruction and flexible use of time. These last two factors helped students to build connections among curriculum and helped to improve achievement. The following quotes from intermediate school educators help to clarify the advantages of teaming.

Intermediate school educator #1 stated:

I think the team approach is certainly far better than a traditional junior high approach. It's not departments. We get to meet as a team of teachers and we know what one another is doing. We plan units together. We get to do all that planning together so we can try to make the connection between the disciplines. These connections are certainly helpful.

Intermediate school teacher #2 stated:

The rapport is probably the greatest advantage to our structure. Teaming helps us get to know the students better, it helps build a safe environment, and it students feel like they can trust you. Building that mutual respect is important. I think it is a big key to the success of students in our school. Teaming helps create better communication with the kids and parents. The teaming (structure) allows flexibility in our teaching. We can change the schedule when we need to spend more time on a particular project or subject.

*Developmentally responsive programs.* Intermediate school educators who were interviewed believed that the grade configuration and school structure allowed for the creation of programs which were developmentally responsive. For example, when the team was small (2 or 3 teachers), then students worked with a small group of teachers each day. Developmentally, students are not ready to transition from one teacher in fifth grade to seven teachers in the sixth grade. As intermediate school students progressed through the system, they experienced more teachers and larger teams. By the time they reached high school, they were prepared to meet the challenge of working with seven or eight different teachers each day.

In commenting on the structure issue, intermediate school principal #2 stated:

One of the things that we really try to do is in sixth (grade) it is a little more self-contained because they are coming out of elementary and there isn't quite as much movement for them between teachers. Seventh grade we have two teams of three teachers and kids move within that team and still go to an exploratory. When we hit eighth grade, it is configured to have kind of a high school look to prepare for the next step.

Every intermediate school that participated in this research had developed a program of exploratory courses. As reported by the intermediate school educators, these courses were designed to provide instruction which was appropriate to the ability level and interest of the student. The intermediate school educators believed that their grade configuration allowed the school to focus programs specifically designed for the intermediate-grade student. The following quote helps clarify this issue.

Intermediate school principal #1 said:

I think the advantages of having six, seven, and eight in the same building allows us to develop a program that is specifically designed to meet the needs of those kids. Everything from an exploratory program to having a PE program that is specifically designed to meet the needs of that age group. I think just the overall program and again the teaming is a huge part of that development.

#### *Disadvantages of Both Configurations*

In the interview process, the K-8 educators were asked to identify disadvantages and challenges related to their grade configuration or organizational structure. Their responses could be summarized in the following two general categories: (a) the K-8

grade configuration is not developmentally responsive to the needs of the young adolescent and (b) scheduling problems inherent to the K-8 structure create a situation which can be challenging for students and staff.

*K-8 not developmentally responsive.* The K-8 educators that were interviewed cite some examples which demonstrate the lack of developmentally responsive programming in the K-8 structure. The K-8 educators reported some problems with the age difference between the upper and lower grade students. Specifically, there are some problems with keeping the younger students separate from the older students at all times of day. Also, many of the school-wide rules are designed for the safety of the younger students. In some instances these rules are not developmentally appropriate for older students.

With regard to differences in age level, K-8 teacher #5 stated the following:

There is always the issue of puberty and I was just telling our superintendent that if we have to build another school building, build it for the junior high students.

Older students can be mean to the younger students. There is only one time in the day that they are together and that is before school on the playground and that can create issues. Sometimes the differences in age level can cause problems.

K-8 teacher #1 commented on the length of time spent in one school:

I think for some kids, nine years in one place is too long. Transition between grades consists of students walking one door down each year until they reach the end. I think some students are ready to move on. I think that perhaps they feel they are treated like they are littler kids because they are attending school with the little kids. But they have never been in another middle school so they have



nothing to base that on. I think they believe that if they are with the little kids then we (the school) must be treating them like little kids. We try to keep rules different for the middle-grades, but they are always testing their limits.

The K-8 educators also report that social issues can develop for those students who earn a negative reputation in the elementary grades. Typically the K-8 student carries this negative label for their entire nine years of enrollment.

A K-8 principal (#2) gave an example to illustrate this issue:

The student-teacher relationship depends on the reputation of the student. As a student moves up to the next level, all the teacher knows is here comes student X, who has all the behavior issues. In the middle school structure, when you switched from an elementary to a middle school between fifth and sixth grade, you got a clean slate. Here (at the K-8) you don't get that clean slate. As you move between grades and if you have that big X on your head because you have been labeled, you really don't escape that.

Another issue, the size of most K-8 schools, limits the opportunities for some students. Students have fewer opportunities to make new friends, request a different teacher, and take exploratory courses.

A K-8 principal (#1) provided the following example which relates to the lack of opportunity to request a different teacher.

The disadvantage is that there is not an option. There is only one math teacher, so if there is a teacher where a parent says I did not have a good experience with that teacher, there is no option for them to go across the hall. There is less flexibility to fix a problem with a personality conflict.

A K-8 teacher (#4) described the problem related to the limited number of students:

The first disadvantage that comes to mind is a lot of these kids grew up together. They spend their first nine years of schooling together and for some that is too much time with the same group. For example, we have some classes where there is a small group of females. There is always the chance that you have a problem with someone else in your class and there is a limited opportunity to make other friends.

*Scheduling issue in the K-8 structure.* Those K-8 educators who were interviewed report some issues related to their school schedule. Teachers and administrators report that the seventh and eighth grade looping model creates an atmosphere where teachers are subject focused and not student focused. This situation has been described as a departmentalized structure, where teachers are divided into departments based on subject area. Some believe that the departmentalized structure lacks opportunity for teachers to plan together and integrate curriculum. Curriculum integration is seen as an important component of the intermediate school model, where educators believe that this process helps students see connections between subject areas.

K-8 teacher # 2 stated:

It is much more difficult to integrate curriculum. I mean I can teach social studies within my language arts but we use it as a vehicle certainly. But the opportunity to integrate or to create cross curricular units is difficult and the scheduling also then becomes a problem as well.

Although the majority of K-8 educators cited the process of looping seventh and eighth graders as a way to build strong student-teacher relationships in the K-8 school, some also felt that the looping structure created some problems. As a subject area specialist, an upper grade teacher in a K-8 school would teach one subject to both seventh and eighth grade students. In the course of a day, this teacher may see over 100 students. Some believed that this was too many students. By contrast, an intermediate school teacher who is part of a small team may only see 60 to 80 students in one day.

K-8 principal #5 stated the following:

When you have that many students whether it be 120 or 140, in any given day, I am not convinced that we are doing the best job of teaching those students. We are meeting as many individual needs as possible, but I think it (seeing too many students) is a disadvantage. Also, for each teacher to truly evaluate or assess that many students, it is not possible. We may say that we are focused on individual student needs, but that statement is misleading.

Finally, many of the K-8 educators (six out of ten) cited a scheduling issue related to a lack of resources. Many K-8 schools share teachers across many grade levels. For example, one school reported that they share an H/PE teacher, a computer teacher, and a music teacher between the upper and lower grades. These subject area specialists drive the schedule for the entire building. If you are required to provide physical education for multiple grade levels, that teacher needs to be scheduled at specific times during the day. This fact leads to a lack of flexibility in the overall building schedule. For example, if a group of upper grade teachers wanted to flex their schedule to spend more time on a particular project or activity, they would be constrained by the building schedule.

A K-8 teacher #5 stated the following: “Well the disadvantage would be the scheduling. The junior high tends to dictate all scheduling. For instance, PE and Music pullouts. Junior high is first and then everybody fits in around that.”

Another K-8 teacher (#1) reported the following concern related to the schedule:

One of the things that we have struggled with in becoming more of a true middle school is that we have to share some of the resources. We share a PE teacher and computer lab. We are much more limited in being able to do more middle school types of things because our hands are tied. At another school where they don't have to share resources, if they wanted to stop ringing the bell for the day and totally change their schedule and spend three hours on a particular activity, they have those options, where it is much tougher here.

#### *Intermediate School Disadvantages*

Seven intermediate school educators were interviewed for this research.

Although the list of potential disadvantages of the intermediate school configuration was diverse, there was some agreement on two general issues. First, intermediate school educators believed that the strength of the student-teacher relationship was directly tied to the team organization. As the team size increased (more students and more teachers) then the strength of the student-teacher relationship declined. Second, intermediate school educators believed that there were problems related to the schedule and how teacher teams were organized. Some believed the lack of teacher subject area specialists was a problem, while others reported concerns with the use of subject area specialists. In addition, despite the fact that teacher teams were allowed some autonomy with the time schedule, some still believed that the schedule lacked flexibility and they had to make

compromises with other teachers. The following are some quotes from educators who saw disadvantages in the organization and structure of the intermediate school.

Intermediate school teacher #1 (who taught in a team of five teachers) stated the following:

I prefer the two teacher teams. I like the set up they have because they have fewer transitions. We have five, six, or seven transitions in a day and that causes problems. Time is lost not only to the transition, but also the time needed for the kids to settle down for that next class period. There is also a problem building relationships when we have that many students in one day. I work as hard as I can on knowing as many kids as I can. But in reality I have close to 80 to 90 kids in a day that that is a lot of kids to get to know.

Intermediate school teacher #4 (who taught on a team of three teachers) saw a problem with the lack of specialists. She stated:

I think we need more specialists (teachers). If [there] was a middle school endorsement or certification it would be better for students. A lot of times teachers have to teach in areas that they are not comfortable with and you know they don't have the passion or expertise and so everyone kind of suffers a little bit from that. When you run small teams of teachers, you are often forced to teach two or three different content areas.

Intermediate school teacher #3 (who taught on a team of three teachers) cited an issue related to the lack of flexibility of the schedule.

Scheduling remains a huge problem here. So if I wanted to teach a really in depth subject in social studies that involved a lot of writing and then maybe have an art

project attached to it, I would have to ask other teachers to give up curriculum for that day to do that. I can't say, you know what guys we are doing Egypt for two weeks and then we are going back to science a week from now and it will be just fine. If we could do that it would be nice.

Finally, two intermediate school educators saw a potential disadvantage related to the lack of opportunities for intermediate school students to mentor younger students.

Intermediate school principal #3 stated:

One disadvantage would be when you have just sixth, seventh, and eighth grade students there are less opportunities on site for older kids to mentor younger kids. I don't think that it is a critical issue, just one of those things you don't have available.

Intermediate school teacher #2 (who attended a K-8 school as a child) stated the following:

I think the biggest disadvantage is not having younger kids close by. With my previous experience it was such an amazing school because the older kids got to take that role of a leader and teach to the younger kids. It really put a boost in their self esteem and self confidence. The little kids just loved it.

#### *Summary of Advantages and Disadvantages for Each Configuration*

Most K-8 educators agreed that the advantages of the structure of their school could be summarized in the general theme of positive relationships. The K-8 structure allowed for mentoring opportunities, increased communication between staff, students, and families, and a looping structure for seventh and eighth grade students. All of these advantages created an environment where it was easy to build relationships. By contrast,

K-8 educators agreed that the primary disadvantage of the structure could be summarized by the general theme of lack of opportunity. This disadvantage was defined by a lack of opportunities for enrichment, a lack of opportunities for new friendships, and a lack of flexibility in the daily schedule.

The intermediate school educators agreed that the advantages of their configuration could be summarized by the following theme: the power of team structure. They believed that the team structure allowed for the creation of developmentally responsive programs. Depending on the size of the team, educators believe that the team structure increased communication opportunities among teachers, students, and parents. In comparison, most intermediate school educators saw potential disadvantages in the organizational structure. If team sizes were too big (more than three teachers) then there were problems with communication, building relationships, and flexibility of time. Intermediate school educators also saw disadvantages in the lack of content area specialists and the lack of opportunities for older students to mentor younger students.

#### *Analysis of Responses from Educators with Experience in Both Settings*

In the process of interviewing educators, there were five who had experience in both intermediate school and K-8 settings. Some had been teachers in both settings. Others had been administrators in both settings. All of these educators were no longer working in an intermediate school or K-8 setting. Four out of five had received administrative appointments in K-5 schools. The fifth was currently serving as a K-12 superintendent. These educators were asked to comment on the advantages and disadvantages of the K-8 and the intermediate school configuration. Their responses

verify the information that was collected from the educators currently working in a K-8 or intermediate school setting.

The educators with experience in both settings agreed that the strength of the K-8 school was the opportunities for students, staff, and parents to build positive relationships. They believed that the K-8 setting provided a greater sense of community. They saw benefits of older students being able to mentor younger students. In addition, having students attend the same school for nine years fostered stronger student-teacher relationships than existed in the intermediate school setting. Beyond these advantages, one educator who had worked in both settings commented on the advantage of having content area specialists in the K-8 setting.

By contrast, these educators agreed that there were potential disadvantages inherent to the K-8 structure. Due to the size of the K-8 schools, there were fewer opportunities for students. They reported that K-8 schools could not provide as many course opportunities as intermediate schools. They also said that some students lacked opportunities to make new friends. They believed there was a difference in student – student relationships in the K-8 setting. There were more students who were socially isolated and were not able to break out of the reputation created at a younger age.

When comparing the K-8 setting to the intermediate school, the educators who had experience in both settings saw some advantages related to an increase of opportunities for students attending the intermediate school. Generally, the intermediate schools were larger and students had more opportunities for enrichment. They could participate in courses designed to explore possible areas of interests. They also had opportunities to take advanced courses to prepare for high school, such as advanced



eighth grade math. There were also more opportunities for co-curricular and extra-curricular activities.

The educators with experience in both settings reported some potential disadvantages in the intermediate school configuration. They believed that the power of the team structure was dependent on the size of the team. If the teams were large there was potential for problems with communication and transition between grades. Large teams did not help to foster student-teacher relationships. In addition, as reported by one educator, “the pace of middle school is harried.” Educators reported, at the intermediate school there is too much movement in one day. This fact can create chaos and a feeling of disconnect for students. However, these problems can be mitigated by reducing the size of the team. Table 4.28 summarizes the advantages and disadvantages of each structure as well as the verification by the educators who have worked in both settings.

### *Summary*

The analysis of the CRT results showed that there were many consistent findings of the comparison of scores among various configuration groups. However, not many of these consistent comparisons reached the *a priori* level of an important difference. For those that did have an important mean difference, many of the differences could be attributed to differences in SES level. The analysis of YRBS results found some important differences in incidence of at-risk behaviors between the two types of configurations. However, at-risk behavior was discovered in both configurations, so it would be difficult to conclude one was better than the other. Finally, the one-on-one comparison interviews showed similar results. Analysis from the interview responses showed advantages and disadvantages of both settings. So the quantitative findings were

supported by the qualitative findings. Even though Chapter Two outlined several research studies which concluded one configuration was better than the other, those findings were not supported by this research. The implications of these results will be discussed further in Chapter Five.

Table 4.28 *Summary of Advantages and Disadvantages*

	<b>Theme</b>	<b>Categories</b>	<b>Educators with experience from both settings</b>
<b>K-8 Advantages</b>	Relationships	<ul style="list-style-type: none"> <li>• Mentoring</li> <li>• Communication</li> <li>• Community</li> <li>• Looping</li> </ul>	<ul style="list-style-type: none"> <li>√ Sense of community</li> <li>√ Older students mentoring younger</li> <li>√ Content area specialists</li> </ul>
<b>K-8 Disadvantages</b>	Lack of Opportunity	<ul style="list-style-type: none"> <li>• Limited enrichment</li> <li>• Fewer students</li> <li>• Fewer resources</li> <li>• Schedule not flexible</li> </ul>	<ul style="list-style-type: none"> <li>√ Less opportunities</li> <li>√ Social isolation</li> <li>√ Reputation hard to break</li> </ul>
<b>IS Advantages</b>	Teaming	<ul style="list-style-type: none"> <li>• Programs are age appropriate</li> <li>• Teacher collaboration</li> <li>• Communication</li> <li>• Relationships</li> </ul>	<ul style="list-style-type: none"> <li>√ More relationship opportunities</li> <li>√ Exploratory courses</li> <li>√ More extra-curricular opportunities</li> </ul>
<b>IS Disadvantages</b>	Structure	<ul style="list-style-type: none"> <li>• Too many kids</li> <li>• Lack of flexibility in schedule</li> <li>• Lack of content specialists</li> </ul>	<ul style="list-style-type: none"> <li>√ Relationships - size of team</li> <li>√ Pace is too harried</li> <li>√ Chaos and disconnect</li> </ul>

## CHAPTER FIVE

### Conclusion

### *Introduction*

Within the last 10 years, the middle school grade configuration has been criticized for focusing too much attention on social development while failing to help students succeed academically (Yecke, 2006). Approximately seven years ago school district officials around the country began restructuring their middle schools in preference of the K-8 configuration. Some districts have claimed that the K-8 learning environment provides a better atmosphere than the large, impersonal middle schools (Beane & Lipka, 2006). Some middle school experts believe that the middle school concept has proven to work when all the components are fully implemented. They also contend that some of the districts who have made the decision to return to K-8 schools have done so without analyzing the status of the implementation of all recommended middle school practices (Swaim, 2004).

This mixed-methods research was completed in two stages. In the first stage, quantitative research questions or hypotheses sought to address the potential relationship between school grade configurations, student performance, and student at-risk behavior for intermediate level students in the State of Montana. In the second phase, qualitative interviews were conducted to compare characteristics of schools serving intermediate level students with various grade configurations.

To explore the potential advantages and disadvantages between K-8 and intermediate schools, this research examined three types of indicators. First, student performance on the state-wide CRT (MontCAS) was analyzed for all eighth grade

students who took the test in 2005. Second, results from the YRBS for a sample of seventh and eighth graders who took the test in 2005 were analyzed. Third, responses from questions asked during one-on-one interviews with educators from both types of configurations were analyzed. From this qualitative analysis, common themes and categories were evident in the interview responses.

The next step of the analysis process was to examine the results and describe conclusions as they relate to each type of datum. In Chapter Four, results were separated into the following three sections: (a) the null hypotheses for student performance on the MontCAS test, (b) the hypotheses for incidence of at-risk behavior as reported on the YRBS, (c) the summary of themes and categories that emerged from the comparison interviews. As a final step, the conclusions from the qualitative and quantitative results were integrated. In this step the qualitative research helped to explain and explore the quantitative results.

#### *Summary of Null Hypotheses for CRT Scores*

Quantitative analysis of this study, examined five null hypotheses related the potential relationship between the independent variable of grade configuration and dependent variable of student performance on the MontCAS test. The ANOVA analysis testing controlled for the potential effects of the interacting variables of school size and SES. For the purpose of this research, an experimentally important difference was defined to be a difference in the mean score of at least three points on the MontCAS test. In addition, experimental consistency was set at the .05 level. In order to reject the null hypothesis, both *a priori* conditions must be met.

Given the five null hypotheses, considering student performance in both math and reading, there were 46 possible comparisons among the groups based on configuration, size, and SES. Although almost every comparison reached the *a priori* condition of a consistent mean difference, only a small minority of the comparisons reached the *a priori* requirement of an important mean difference. In general, those comparisons which were based solely on configuration or school size, there was not enough difference in the means to reject the null hypothesis (see Table 5.1, null hypothesis #1, #2, and #3). For these comparisons, the p-value was generally very low ( $p < .0001$ ). With a very low p-value and a small mean difference, it would be appropriate to conclude that these comparisons were consistently unimportant. In other words, comparing schools solely on the basis of configuration and/or school size consistently showed no important difference among the means.

Those comparisons which incorporated SES level with configuration generally showed an important difference among the means (see Table 5.1, null hypotheses #4 and #5). For these comparisons, the p-value was very low and the difference in the means was at least three points. Based on configuration and SES, it would be appropriate to conclude that the comparison of the means among the schools was important and consistent. Generally this important and consistent mean difference was seen in reading and math scores. These results reinforce the powerful influence of SES level on student performance.

Table 5.1 presents a summary of all null hypothesis conclusions. Overall, there was little difference among student test scores based on solely on configuration. Although the ANOVA analysis showed experimental consistency in several areas, there

were very few comparisons which showed experimental importance. There was only one set of comparisons which showed an experimentally important and consistent difference potentially related to configuration. In comparing high SES school groups, the mean score for reading and math was higher for intermediate school group and elemiddle group than for the K-8 group.

Table 5.1 *Summary of Null Hypothesis Comparisons and Conclusions*

Null Hypothesis	MontCAS Test	Experimental Consistency	Experimental Importance	Conclusion
#1 No difference among three configuration groups: K8, EM, IS	Reading	Yes ( $p < .01$ )	No	Fail to reject
	Math	Yes ( $p < .05$ )	No	Fail to reject
#2 No difference between the two school size groups: small (<400) and large (>400) schools.	Reading	Yes ( $p < .01$ )	No	Fail to reject
	Math	Yes ( $p < .01$ )	No	Fail to reject
#3a No difference among the three configuration groups for small schools: SMLK8, SMLEM, SMLIS	Reading	Yes ( $p < .01$ )	No	Fail to reject
	Math	Yes ( $p < .01$ )	No	Fail to reject
#3b No difference among the four groups defined by configuration and size: SMLEM, SMLIS, LRGEM, LRGIS	Reading	Yes ( $p < .01$ )	Yes LRGIS = 42 SMLIS = 39	Reject for LRGIS vs. SMLIS
	Math	Yes ( $p < .01$ )	No	Fail to reject
#4 No difference between the two school groups defined by SES: HSES and LSES	Reading	Yes ( $p < .01$ )	Yes HSES = 41 LSES = 38	Reject for HSES vs. LSES
	Math	Yes ( $p < .01$ )	Yes HSES = 31 LSES = 28	Reject for HSES vs. LSES

Table 5.1 (continued) *Summary of Null Hypothesis Comparisons and Conclusions*

Null Hypothesis	MontCAS Test	Experimental Consistency	Experimental Importance	Conclusion
#5a No difference among the three high SES configuration groups (HSES K8, EM, and IS)	Reading	Yes (p<.05)	Yes HSES EM= 42 HSES IS = 42 HSES K8 = 39	Reject for EM vs. K8 IS vs. K8
	Math	Yes (p<.01)	Yes HSES EM= 32 HSES IS = 32 HSES K8 = 28	Reject for EM vs. K8 IS vs. K8
#5b No difference among the three low SES configuration groups (LSES K8, EM, and IS)	Reading	Yes (p<.05)	No	Fail to reject
	Math	Yes (p<.01)	No	Fail to reject
#5c No difference between the two K-8 groups defined by SES. (HSES and LSES K8)	Reading	No	No	Fail to reject
	Math	No	No	Fail to reject
#5d No difference between the two elemiddle groups defined by SES (HSES and LSES EM)	Reading	Yes (p<.01)	Yes HSES EM= 42 LSES EM= 39	Reject for HSES EM vs. LSES EM
	Math	Yes (p<.01)	Yes HSES EM= 32 LSES EM= 28	Reject for HSES EM vs. LSES EM
#5e No difference between the two intermediate school groups defined by SES (HSES and LSES IS)	Reading	Yes (p<.01)	Yes HSES IS = 42 LSES IS = 39	Reject for HSES IS vs. LSES IS
	Math	Yes (p<.01)	Yes HSES IS = 32 LSES IS = 27	Reject for HSES IS vs. LSES IS

This table shows the comparisons among subgroups created by the five null hypotheses set forth in the methodology of this research. Experimental consistency was set at the .05 level and a difference of at least three points was required to conclude an

important mean difference. The next section summarizes the results of the analysis of the incidence of the at-risk behavior as reported on the YRBS.

*Summary of Hypotheses Related to the YRBS.*

Student at-risk behavior was defined as the occurrence of certain behaviors on the YRBS. In the State of Montana, the YRBS was given to 9,178 seventh and eighth grade students in the spring of 2005. This research examined a sample of 1,971 students attending Montana K-8 and intermediate schools in 2005. Demographic for the sample is listed in Table 4.15. Although there were more than 80 questions on the YRBS, this research focused on 16 questions from the survey (See Table 4.17 for a list of questions). These questions were divided into at-risk behaviors occurring at school and at-risk behaviors not occurring at school.

The YRBS comparisons show that there were at-risk behaviors prevalent in both configurations. Fighting in school was reported as an issue more often for K-8 students than intermediate school students. This issue showed up as an important difference for schools defined as large K-8 schools as well as schools defined as high SES K-8 schools. The use of alcohol was more reported more often by the K-8 students in the sample. Drugs and depression issues seemed to be more prevalent among students attending intermediate schools in the sample. Two subcategories (low SES and small intermediate schools) were more likely to report being offered/sold/or given illegal drugs at school. These two intermediate school subcategories were also more likely to report feelings of depression than their counterparts in the K-8 schools. The use of cigarettes and inhalants seemed to be an at-risk behavior reported equally by both intermediate school and K-8



students in the sample. Table 5.2 presents a summary of important differences for the YRBS.

Table 5.2 *A Summary of Important Differences for the YRBS Results*

		<u>K-8 Schools</u> At-Risk Behaviors that were reported at least 5% higher than the intermediate school.	<u>Intermediate Schools</u> At-Risk Behaviors that were reported at least 5% higher than the K-8 school.
Size	Large Schools	<ul style="list-style-type: none"> <li>➤ Engaged in fight at school (8.3% higher than IS)</li> <li>➤ Tried cigarettes and alcohol (7% &amp; 12.6%)</li> </ul>	<ul style="list-style-type: none"> <li>➤ For large intermediate schools, there were no behaviors that were notably* higher than large K-8 schools.</li> </ul>
	Small Schools	<ul style="list-style-type: none"> <li>➤ Used inhalants in lifetime (5.3%)</li> </ul>	<ul style="list-style-type: none"> <li>➤ Been offered drugs at school (8% higher than K-8)</li> <li>➤ Attempted suicide in past 12 mos. (4.9%)</li> </ul>
SES	High SES	<ul style="list-style-type: none"> <li>➤ Engaged in fight at school (9.4% higher than IS)</li> <li>➤ Tried cigarettes, alcohol and inhalants in lifetime (9.3%, 8.0%, &amp; 6.5%)</li> </ul>	<ul style="list-style-type: none"> <li>➤ Been threatened with a weapon at school (5.4% higher than the K-8)</li> </ul>
	Low SES	<ul style="list-style-type: none"> <li>➤ For low SES K-8 schools there were no behaviors that were notably* higher than low SES intermediate schools.</li> </ul>	<ul style="list-style-type: none"> <li>➤ Been offered drugs at school (8.6%)</li> <li>➤ Felt sad or hopeless (7.9%)</li> <li>➤ Used marijuana in lifetime (9.1%)</li> <li>➤ Used cigarettes and inhalants in lifetime (8.0% and 6.3%)</li> </ul>

\*Notably higher was defined as a difference of greater than 5% between the groups.

The first two sections of this chapter presented summary results of the quantitative analysis. In comparing student performance on the MontCAS for schools of various configurations, there appeared to be a small mean difference among the scores. For those comparisons which had an important mean difference, the difference could have been explained by school differences in SES. The comparisons of YRBS results showed that there were at-risk behaviors prevalent in both types of configurations. The qualitative analysis of the one-on-one comparison interviews seems to support the quantitative findings.

#### *Summary of Comparison Interviews*

Comparison interviews were conducted with educational professionals who were currently working in the K-8 or intermediate school. In order to verify the responses from the initial interviews, further interviews were conducted with professionals who had experience with both types of configurations. The purpose of these interviews was to explore the following central question: how was the grade configuration of a school related to student success as perceived by educators working in both configurations? This question was supported by the following sub-questions: (a) how did the characteristics of the school grade configuration help to address students' developmental and academic needs? (b) how were student-teacher and student-student relationships defined in each type of grade configuration? (c) what were the advantages associated with each type of grade configuration? (d) what were the disadvantages associated with each type of grade configuration?

### *Student Success*

Through the interview process, the characteristics of successful students were defined and clarified. The purpose was to determine if characteristics of successful students varied between grade configurations. Characteristics of successful students were not dependent on grade configuration based on the interview responses. Educators from both settings (K-8 and intermediate school) agreed that successful students possessed the following characteristics: personal support, learner qualities, and personal competence. Educators from both settings agreed that student success was at least somewhat dependent on the support that they received outside of school. Successful students were also organized, able to manage their time, able to manage expectations, willing to take responsibility, and able to adjust to a variety of learning environments. Finally, successful students had a positive self-image which allowed them to take academic risks in the classroom and learn from their mistakes. With regard to characteristics of successful students, there was no difference between grade configurations.

### *Developmental and Academic Needs*

When asked how their grade configuration addressed developmental and academic needs of the individual students, educators from both settings believed that their grade configuration was responsive to the needs of the young adolescent. In both settings the capability to address student needs was defined by the organizational structure. Even though this structure was different for each setting, educators from both settings believed that their school was organized to address student needs. The intermediate school educators believed that their team structure, where each student was well known by a small group of adults, helped to respond to individual needs. The K-8

educators believed that their looping structure, where students stayed with the same teacher for two years, helped to address developmental and academic needs. Even though the organizational structure was different for each grade configuration, both K-8 and intermediate school educators were able to clarify how their structure was responsive to the needs of the young adolescent.

### *Relationships*

When asked to describe a potential benefit of their configuration, educators from both settings believed that the strength of the student-teacher and student-student relationships was a major advantage. Intermediate school educators believed that their team structure helped to foster strong relationships between students and teachers. The educators from the K-8 schools believed that their relationships were strong because students and teachers had been together for an extended length of time (9 years). In addition, K-8 educators believed that their student-teacher relationships were strengthened by the looping structure that existed in the seventh and eighth grade. Both school settings believed their organizational structure created a sense of community.

### *Advantages for Each Configuration*

Educators from both settings believed that the advantages were directly tied to their organizational structure. Those from the K-8 schools believed that their school structure fostered positive relationships. Specifically, the opportunities for mentorship, communication between teachers, strong connection to the community, and the process of looping seventh and eighth graders helped to build strong relationships in the K-8 school. Those from intermediate schools believed that team structure was the primary advantage of their configuration. The team structure helped to create developmentally

responsive programs, foster teacher collaboration, and provide more opportunities for communication among students, teachers, and parents.

#### *Disadvantages for Each Configuration*

Educators from both configurations described disadvantages which were not necessarily a product of the configuration, but more directly tied to issues of school size. The K-8 educators believed that their small schools provided fewer opportunities for students as compared to opportunities believed to be available in the intermediate schools. For example, K-8 educators stated that their students had fewer opportunities to make new friends and their schools had fewer resources to help at-risk students. Conversely, intermediate school educators believed that the large size of their schools created some disadvantages. Even though teachers and students were divided into teams, several educators believed that the teams were too large to really develop strong relationships with all students. In addition, intermediate school educators believed that their schedule was inflexible. For example, the structure of the team was sometimes defined by the certification of individual teachers on the team.

#### *Integration of Quantitative and Qualitative Conclusions*

A final step of the analysis process was to integrate the quantitative and qualitative results. The quantitative inquiry process was the primary research method and the qualitative was secondary. In the integration phase, connections were made between the qualitative and quantitative conclusions.

Although the quantitative data showed some differences in the results among the various grade configurations, the majority of hypothesis comparisons showed more similarities than differences. In examining the null hypothesis testing of the MontCAS

scores in reading and math, there were 46 comparisons which could have shown an important and consistent mean difference (See Table 5.1). Among these 46 comparisons, there were only 11 comparisons which reached the *a priori* level for an important and consistent difference in the means. From these 11, only four could be directly related to differences in school configuration. The other five were related to differences in size and SES.

Much like the MontCAS comparisons, the results of the YRBS comparisons showed some important differences. However, the majority of comparisons based on the YRBS results showed that the two grade configurations were more similar than different. There were five hypotheses comparisons developed to determine if there was an important difference between the percentage of students reporting participation in at-risk behaviors at the K-8 school and the percentage reporting at the intermediate schools. Each hypothesis was checked for 16 different at-risk behaviors. In total there were 80 (5 hypotheses and 16 at-risk behaviors) different comparisons made between K-8 and intermediate schools. From these 80 possible comparisons, only 18 reached the level of an important difference between the two grade configurations. At-risk behaviors were separated into school related and non-school related behaviors. Of the 18 comparisons which reached the level of an important difference, only six comparisons were identified as school related behaviors.

Based on the quantitative results, only a small minority of the comparisons between the K-8 and intermediate schools reached the level of an important difference. In addition, for the majority of comparisons which did reach the level of importance, the differences could be explained by factors not related to grade configuration. For example

some of the differences in the MontCAS scores were more directly related to school size and SES level rather than configuration. Most of the differences in the YRBS results were among the non-school related at-risk behaviors. From these results the logical conclusion would be to assume that the two school configurations were more similar than different. The results of the qualitative interview data support this assumption.

In the qualitative phase of this research, educators from both K-8 and intermediate schools were asked to identify strengths and weaknesses of their particular grade configuration. Analysis of the interview results produced similar results for both grade configurations. Educators from both settings describe the power of their school organizational structure in developing strong relationships. As evidence, the K-8 educators cite their looping process, while the intermediate school educators cite their team structure. Educators from both settings agree on the characteristics of a successful student. Each configuration has specific programs and processes to foster student success. When asked to clarify the disadvantages of their school configuration, responses from K-8 and intermediate school educators differed slightly. However, the list of potential disadvantages was more directly tied to the issue of school size rather than configuration.

#### *Verification of the Qualitative Analysis*

It is important to note the agreement between the qualitative and quantitative results demonstrates dual verification of both findings. In other words, the qualitative data was verified by the quantitative data. In addition, the qualitative results support the findings of the quantitative analysis. As mentioned previously, the quantitative results showed few differences among the various configurations as related to student

performance on the CRT. Although the YRBS data had shown some important differences in at-risk behavior between the grade configuration groups, it was interesting to note that the results found at-risk behavior in both types of configurations. The qualitative interviews showed some advantages and disadvantages for both types of configurations. The similarities in these results demonstrate verification of the qualitative data through support from the quantitative results.

### *Recommendations*

This research examined the potential relationship between the configuration of a school and student success for intermediate-grade students in the State of Montana. Quantitative data examined potential differences among configuration groups based on reading and math test scores for all eighth grade students in the State. In addition, the occurrence of at-risk behaviors for a sample of seventh and eighth grade students from across the State was examined for potential differences based on school configuration. Finally, one-on-one comparison interviews were conducted with a sample of K-8 and intermediate school educators from one community in the State to determine potential differences between the configurations. From the various research strategies a large volume of comparisons were made. The majority of these comparisons showed no important difference between the configurations. The comparisons which did show a difference were related not only related to configuration, but also school size and SES. The following recommendations were developed from those comparisons which showed an important difference.

Considering student performance on the MontCAS (Montana's CRT) there was no important and consistent difference in mean reading and math scores for eighth grade



students attending schools of various grade configurations. However, when configuration groups were stratified by school size and SES level, there were some notable differences in the mean MontCAS scores.

For those eighth grade students attending intermediate schools in the State of Montana, the size of the school may be an important factor in determining student success on the reading MontCAS. Students attending large intermediate schools (> 400 students) received a mean reading score three points higher than those attending small intermediate schools. From this result it may be prudent for school districts with intermediate school configuration to consider the impact of school size. Some may argue that small schools are more effective for student performance; however these results would suggest the opposite. From the review of literature it would seem that large schools may be able to provide more opportunities for students. In addition, this research shows that large intermediate schools may help improve reading scores on the CRT.

Among high SES (< 40% free or reduced lunch) schools in the State of Montana, the configuration may be an important factor in determining student success on the reading and math MontCAS. In 2005, eighth grade students attending high SES elemiddle and intermediate schools outperformed those students attending high SES K-8 schools. This difference in average score was important for both reading and math. This finding was not consistent with much of the research presented in Chapter Two, which found higher scores in K-8 schools. After controlling for SES, eighth grade students attending intermediate schools scored higher than their K-8 counterparts. Among high SES schools, it may be advantageous to incorporate the intermediate configuration.

Much like previous research, the findings of this study show that SES has an influence on student performance on the CRT. Among those eighth grade students attending all school configurations in the State of Montana, those attending high SES schools outperformed those attending low SES schools in both math and reading CRT. From these results it would be important for schools in Montana to focus attention and efforts on strategies that could mitigate the impact of SES. These SES results were consistent for elemiddle and intermediate schools. From the null hypothesis analysis it would appear that the scores for those students attending K-8 schools were not influenced by the effects of SES. However, these results did not meet the *a priori* requirement for a consistent mean difference. In other words, even though the mean difference for HSES and LSES K-8 schools was low, the high p-value would suggest that these results were not consistent.

Among those students attending elemiddle (K-6/7-8) schools in the State of Montana, the SES level of the school may be an important factor in determining student success on the reading and math MontCAS. In general, students attending high SES (<40% F/RL) elemiddle schools scored notably higher (greater than a three point difference) than those students attending low SES (>40% F/RL) elemiddle schools.

Among those students attending intermediate schools (6-8) in the State of Montana, the SES level of the school may be an important factor in determining student success on the reading and math MontCAS. In general, students attending high SES intermediate schools scored notably higher than those students attending low SES intermediate schools.

Considering the occurrence of at-risk behaviors among seventh and eighth grade students attending a sample of schools in the State of Montana, the configuration, size and SES level of the school may be an important predictor of certain types of at-risk behaviors. Those students attending a large K-8 or a high SES K-8 school were more likely to report that they had engaged in a physical fight at school than those attending large or high SES intermediate schools. Those attending small or low SES intermediate schools were more likely to report that they have been offered drugs at school than those attending small or low SES K-8 schools.

The one-on-one comparison interviews showed that there were advantages and disadvantages related to both configuration settings. The educators who were working within the K-8 or intermediate school were able to cite examples specific to their own school configuration. Although these advantages were different for each configuration, it would be difficult to argue that one school's advantage was more important than another school's advantage. In considering specific recommendations for this research, it may be more appropriate to examine the responses from those educators who have experience in both settings.

After working as a K-8 and an intermediate school teacher, one particular educator believed that both settings had advantages related to teacher certification. The K-8 school had more content area specialist teaching in the intermediate-grades, while the intermediate school had more generalists. This particular educator found advantages in both scenarios. The specialists had more training and education in their specific content, which translated to higher level of academic focus. Conversely, the generalists

teaching in the intermediate school had a stronger focus on students and were able to connect with students better than the specialists in the K-8 setting.

A second educator with experience in both settings also saw advantages in both the K-8 and intermediate school. This educator believed that students and teachers were able to build stronger relationships in the K-8 school. This advantage was a direct result of school size and length of time students spent in one school. According to this educator, the strength of the intermediate school was found in the scope of opportunities provided to students. The organization and size of the intermediate school allowed for more course opportunities for students.

When asked to compare the configurations, educators who worked in both settings suggested that the best configuration may have qualities of both K-8 and intermediate school. The most appropriate setting may be the school which could provide the strong relationship atmosphere as seen in the K-8, while also providing the breadth of opportunities most often seen in the intermediate school. Researcher Hough (2005) defined the ideal setting as a school which was able to provide both an elementary setting for younger students, while providing an intermediate school setting for older students. He defined this setting as an *elemiddle* school. The findings of the interview research would suggest that the ideal setting for the young adolescent would contain the relationship advantages of the K-8 setting along with the course opportunities of the intermediate school.

#### *Summary of Recommendations*

Although the differences among test scores and at-risk behaviors were important, these differences represent a small percentage of the total possible differences between

the various configurations. In addition, the differences were generally not related differences in configuration but more so related to differences in SES. Based on these results, in conjunction with the results from the comparison interviews, the appropriate conclusion would be to not recommend one grade configuration over another. The number of similarities among the various configurations outweighed the number of differences. The ideal configuration may be an amalgamation of both the K-8 and intermediate school. In addition, schools would be wise to focus attention on strategies that might mitigate the influence of SES, as this factor was shown to influence the differences in student performance. Finally, school district officials who are considering changes in configuration should make sure to incorporate community desires in their decision. This recommendation was supported by the results of the one-on-one comparison interviews. Summary results showed that either configuration (K-8 or intermediate) can work, when there is support from the community.

#### *Implications of the Results*

Several large school districts around the country have abandoned their middle school configuration in favor of a K-8 school. Although their decision to return to the K-8 model may not have been founded in research, the majority of research which has compared the two configuration models would support the decision. In Chapter Two of this research, the review of literature presented a summary of relevant studies which compared K-8 schools with other configurations, typically 5-8, 6-8 or 7-9. Of the more than 20 comparison research studies presented in Chapter Two, 14 studies found that students attending K-8 schools typically outperformed their counterparts attending a school with an intermediate grades configuration.

The majority of comparison research studies found that intermediate-grade students attending K-8 schools typically performed better on academic measures than those attending schools with other configurations (Abella, 2005; Alspaugh, 1998a; Becker, 1987; Comer, 2006; Cook, 2005; Franklin & Glascock, 1998; Freeman, 2005; Gronna, 1998; Moore, 1984; Offenber, 2001; Poncelet, 2004; Vaccaro, 2000; Wihry, Coladarci, & Meadow, 1992; Yakimowski & Connolly, 2001). Student performance was typically measured by state level testing in the areas of reading and math, while some studies measured other content areas. A variety of areas were represented in the comparison research studies, such as Missouri, Tennessee, and North Carolina.

In addition to the issue of student performance on state level testing, several of the comparison studies found other advantages for students attending K-8 schools. Two comparison studies (Moore, 1984; Simmons & Blyth, 1987) found that students attending K-8 schools had higher levels of self-esteem than their counterparts attending schools with a 7-9 configuration. Three comparison studies (Abella, 2005; Cook, 2005; Franklin & Glascock, 1998) found that students attending middle schools or junior highs were more likely to be suspended from school than those attending K-8 schools. Five of the comparison studies (Abella, 2005; Comer, 2006; Cook, 2005; Freeman, 2005; Moore, 1984) found that students at K-8 schools were more likely to attend school more often than those students at intermediate-grades school (5-8, 6-8, or 7-9). Two of the comparison studies found that the advantages carried over as K-8 students transitioned to high school in the form of lower drop out rates and higher grades for those students who had attended a K-8 before high school (Alspaugh, 1998a; Offenber, 2001).

In comparing schools of various configurations in the State of Montana, this research found very few differences between K-8 and intermediate schools for student performance on a CRT and incidence of at-risk behaviors. Qualitative comparison interviews conducted with educators from both grade configurations (K-8 and intermediate school) found advantages and disadvantages for both types of configurations. Among these advantages and disadvantages there was evidence of high quality instructional practices for both types of configurations. The qualitative data supported the results from the quantitative data in that there were very few differences in the schools based solely on grade configuration. Both quantitative and qualitative data support the perception that the schools of various grade configurations were more similar than dissimilar.

Based on the findings from this research in the State of Montana, there were not enough differences between the K-8, elemiddle, and intermediate schools to conclude that one configuration should be supported over another. In Montana, middle-grade students attending a K-8 school did not experience more success than those attending elemiddle or intermediate schools. This conclusion has been supported by previous research.

In Chapter Two, five of the comparison research studies (Blair, 2007; Cox, 1996; Nobles, 2008; Peterson, 2002; Weiss & Kipnes, 2006) found no important differences among schools serving middle-grade students. After controlling for extraneous variables such as school size and SES, three studies (Cox, 1996; Peterson, 2002; Weiss & Kipnes, 2006) found little difference between schools with K-8 and middle school configuration. Considering other extraneous variables such as gender and ethnicity, one of the

comparison research studies found very few student achievement differences based on configuration (Nobles, 2008). One researcher (Blair, 2007) defined student achievement based on growth on state level test from 2002 to 2005. In this study, Blair discovered student achievement advantages for both types of grade configurations (K-8 and middle school).

### *Summary of Implications*

The drive to do away with the middle school has been fueled by the perception that K-8 schools outperform the middle schools and some of the comparison research supports this perception. However, for many urban school districts the conversion to a K-8 school model has been related to factors other than student performance. School districts faced with declining enrollment and dwindling budgets have been forced to consider school closure. The process of school closure can be contentious, creating community dissatisfaction. The attack on the middle school has been used as an easy way to deflect community attention while accomplishing the goal of closing schools. Lacking appropriate information, school districts are spending money and resources to convert their middle schools to K-8 schools. Results from the present study would encourage school officials to focus attention on other factors, as the focus on appropriate grade configuration appears to be insignificant.

In 1990, researcher Joyce Epstein proposed that the debate over appropriate grade span was misguided. Her contention was that intermediate level educators should focus on instructional practices rather than trying to determine which grade span was better for students. In conjunction with the Northwest Regional Education Laboratory, Paglin and Fager (1997) conducted a meta-analysis study related to intermediate level grade



configuration. Similar to Epstein, Paglin and Fager reached the following conclusion: practices and programs were more important than grade span. In Montana, based on this research, it was apparent that the K-8 configuration did not necessarily translate to better student performance or a better environment for student learning. For schools serving middle-grade students in Montana there appeared to be little connection between configuration and student performance. Policy makers and educators who expected to improve student performance or school environment by simply changing the configuration may be disappointed in their efforts. Perhaps, as Epstein recommends, Montana educators would be wise to focus their efforts on implementing the most effective instructional practices for middle-grade students rather than spending resources on changing school configuration.

#### *Suggestions for Further Research*

This research was limited to examining student success on a CRT and incidence of at-risk behaviors among middle-grade students. Further research should examine other measures of student success among middle-grade students in the State of Montana. Some examples of other measures of success are attendance rates, retention rates, discipline statistics, performance in 9<sup>th</sup> grade, and high school drop out rates. In addition, other indicators should be examined beyond the YRBS. For example, measures of self-esteem, participation in school activities, and level of individual depression.

With regard to the MontCAS (CRT) further research could examine two important factors not considered in this study. First, longitudinal data should be tested. For example, test data from other years should be tested using a similar process as outlined in this study. This analysis may help to verify the results of this study. Another

possible longitudinal test would be to examine test data for a cohort group of students from fourth grade and eighth grade. A comparison should be made to determine if growth from fourth to eighth grade has a potential connection to configuration. Second, test subgroup data should be used to determine if there are important differences. For example, a possible comparison could be to examine scores for special education students attending K-8 schools with those attending intermediate schools.

Finally, further comparison interviews should be conducted with other groups to determine if there are any important differences in the two configurations. The perceptions of students, parents and community members should be examined. There may be other factors to consider from those who do not work within the school setting. In addition, interviews with high school teachers who receive students from both configurations may help to verify the similarities or differences.

## References

- Abella, R. (2005). The effects of small K-8 centers compared to large 6-8 schools on student performance. *Middle School Journal*, 37(1), 29-35.
- Allen, R. (2002). Big schools: the way we are. *Educational Leadership*, 59(5), 36-41.
- Alspaugh, J. W. (1998a). Achievement loss associated with the transition to middle school and high school. *Journal of Educational Research*, 92(1), 20-25.
- Alspaugh, J. W. (1998b). The relationship of school-to-school transitions and school size to high school dropout rates. *High School Journal*, 81(3), 154-160.
- Alspaugh, J. W. (2000). The effect of transition grade to high school, gender, and grade level upon dropout rates. *American Secondary Education*, 29(1), 2-9.
- Anfara, V. A., & Buehler, A. (2005). Grade configuration and the education of young adolescents. *Middle School Journal*, 37(1), 53-59.
- Anfara, V. A., & Schmid, J. B. (2007). School transitions: jeopardy or wheel of fortune? *Middle School Journal*, 39 (1), 60-67.
- Atanda, R. (1999). *Do gatekeeper courses expand education options?* (Publication No. 1999-303). Washington, DC: National Center for Educational Statistics.
- Barth, J. J. (2001). The investigation of the relationship between middle school organizational health, school size, and school achievement in the areas of reading, mathematics, and language. Ed.D. dissertation, West Virginia University, United States – West Virginia. Retrieved September 13, 2007, from ProQuest Dissertations database. (Publication No. AAT 3012818)

- Barton, P. E. (2005). *One-third of a nation: Rising dropout rates and declining opportunities* (Rep.). Princeton, NJ: Educational Testing Service, Policy Information Center.
- Beane, J. A., & Lipka, R. (2006). Guess again: Will changing the grades save middle-level education? *Educational Leadership*, 63(7), 26-30.
- Beane, J. A. (1999a). Middle schools under siege: points of attack. *Middle School Journal*, 30(4), 3-9.
- Beane, J. A. (1999b). Middle schools under siege: responding to the attack. *Middle School Journal*, 30(5), 3-6.
- Becker, H. J. (1987). *Addressing the needs of different groups of early adolescents: Effects of varying school and classroom organizational practices on students from different social backgrounds and abilities* (Rep.). Baltimore, MA: Center for Research on Elementary and Middle Schools. (ERIC Document Reproduction Service No. ED291506)
- Bickel, R., Howley, C., Williams, T., & Glascock, C. (2000). *Will the real "Texas Miracle in Education" please stand up? Grade span configuration, achievement, and expenditure per pupil* (Rep.). Randolph, VT: Rural School and Community Trust. (ERIC Document Reproduction Service No. ED447995)
- Blair, N.C. (2007). A comparative study of the effects of grade configuration on middle school and K-8 school value added scores. Ed.D. dissertation, Tennessee State University, United States – Tennessee. Retrieved October 24, 2008, from ProQuest Digital Dissertations database. (Publication No. AAT 3260215).

- Bradley, A. (1998, April 15). Muddle in the middle. *Education Week*, 17(31), 38-42.
- Bushaw, W.J. (2007). From the mouths of middle-schoolers: important changes for high school and college. *Phi Delta Kappan*, 89(3), 189-193.
- Calhoun, F. S. (1983). *Organization of the Middle Grades: A summary of research* (Rep.). Arlington, VA: Educational Research Service, Inc. (ERIC Document Reproduction Service No. ED227581)
- Calsyn, C., Gonzales, P., & Frase, M. (1999). *Highlights from TIMMS: Overview and key findings across grade levels* (Rep. No. NCES 1999-081). Washington, D.C.: Office of Educational Research and Improvement, U.S. Department of Education.
- Coladarci, T., & Hancock, J. (2002). *The (limited) evidence regarding effects of grade span configurations on academic achievement: What rural educators should know* (Rep.). Charleston, WV: ERIC Clearinghouse on Rural Education and Small Schools. (ERIC Document Reproduction Service No. ED467714)
- Comer, H. B. (2006). The effects of grade span configurations on fifth-grade students. Ed.D. dissertation, Union University, United States – Tennessee. Retrieved August 4, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 3232346)
- Cook, H. G. (2005). *What's best in the middle? Student engagement, achievement, attainment, and growth differences between K-8 and middle school grade configurations at Milwaukee Public Schools* (Rep. No. 0501). Milwaukee, WI: Milwaukee Public Schools, Division of Assessment & Accountability.

- Cooney, S. (1998). *Education's weak link: student performance in the middle grades* (Rep.). Atlanta, GA: Southern Regional Education Board. Retrieved November 20, 2007, from <http://www.sreb.org>
- Cotton, K. (1996). *School size, school climate, and student performance*. American Association of School Administrators. Arlington, VA: Retrieved June 30, 2004, from [http://www.aasa.org/issues\\_and\\_insights/cotton.htm](http://www.aasa.org/issues_and_insights/cotton.htm)
- Cox, B. E. (1996). School grade span and configurations which include the eighth grade. Ph.D. dissertation, The University of Southern Mississippi, United States – Mississippi. Retrieved August 4, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 9718174)
- Creswell, J.W. (1998). *Qualitative inquiry and research design: choosing among the five traditions*. Thousand Oaks, CA: Sage Publications.
- Creswell, J. W. (2003). *Research design: qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Cromwell, S. (1998, April 13). *Where does sixth grade belong?* Retrieved November 21, 2006, from <http://www.educationworld.com/admin060.html>
- Cuban, L. (1992). What happens to reforms that last? The case of the junior high school. *American Educational Research Journal*, 29(2), 227-251.
- Deyoung, A. J., Howley, C., & Theobald, P. (1995). The cultural contradictions of middle schooling for rural community survival. *Journal of Research in Rural Education*, 11(1), 24-35.

- Eccles, J. S., Wigfield, A., Midgley, C., Reuman, D., Mac Iver, D., & Feldlaufer, H. (1993). Negative effects of traditional middle schools on students' motivation. *The Elementary School Journal*, 93(5), 553-574.
- Epstein, J. L. (1990). What matters most in the middle grades - grade span or practices? *Phi Delta Kappan*, 71, 438-444.
- Felner, R. D., Jackson, A. W., Kasak, D., Mulhall, P., Brand, S., & Flowers, N. (1997). The impact of school reform for the middle years. *Phi Delta Kappan*, 78, 528-532, 541-550.
- Franklin, B. J., & Glascock, C. H. (1998). The relationship between grade configuration and student performance in rural schools. *Journal of Research in Rural Education*, 14 (3), 149-153.
- Freeman, C. T. (2005). A comparison of K-8 schools, junior high schools, and middle schools with respect to student academic achievement and attendance in North Carolina. Ed.D. dissertation, Western Carolina University, United States – North Carolina. Retrieved August 4, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 3194231).
- George, P. S. (1988). Education 2000: Which way the middle school? *The Clearing House*, 62 (1), 14-17.
- George, P. S. (2005). K-8 of Not? Reconfiguring the middle grades. *Middle School Journal*, 37(1), 6-13.
- Gewertz, C. (2004, May 19). City districts embracing K-8 schools: Higher student scores, better discipline cited. *Education Week*, 23 (37), 1,20.

- Gronna, S.S. (1998). Effects of grade of school transition and school characteristics on eighth grade achievement: a multilevel analysis. Ph.D. dissertation, University of Hawai'i, United States – Hawaii. Retrieved October 24, 2008, from ProQuest Digital Dissertations database. (Publication No. AAT 9913936).
- Guilfoyle, C. (2006). NCLB: Is there life beyond testing? *Educational Leadership*, 64(3), 8-13.
- Hall, G. S. (1904). *Adolescence, its psychology, and its relations to physiology, anthropology, sociology, sex, crime, religion, and education* (I,II ed.). New York: D. Appleton Company.
- Heller, R., Calderon, S., Medrich, E., Bottoms, G., Cooney, S., & Feagin, C. (2003). *Academic achievement in the middle grades: What does research tell us?* (Rep.). Atlanta, GA: Southern Regional Education Board. Retrieved November 15, 2007, from <http://www.sreb.org>
- Herman, B. E. (2004). The revival of the K-8 schools. *Phi Delta Kappan Fastbacks*, 519, 7-37.
- Hough, D. L. (1995). The elemiddle school: a model for middle grades reform. *Principal*, 74(3), 6-9.
- Hough, D. L. (2004). *Grade span does make a difference* (Issue Brief). Springfield, MO: Institute for School Improvement Southwest Missouri State University.
- Hough, D. L. (2005). The rise of the 'Elemiddle' school. *The School Administrator*. American Association of School Administrators. Arlington, VA: Retrieved July 7, 2006, from <http://www.aasa.org/publications/saarticledetail.cfm>



- Howell, D. C. (2002). *Statistical methods for psychology* (Rev. ed.). Pacific Grove, CA: Duxbury/Thomson Learning.
- Howley, C. B. (1996). *Sizing up schooling: a West Virginia analysis and critique* (Rep.). Morgantown, WV: West Virginia University. (ERIC Document Reproduction Service No. ED396880)
- Howley, C. B. (1999). *The Matthew project: state report for Montana* (Rep.). Randolph, VT: Rural Challenge Policy Program. (ERIC Document Reproduction Service No. ED433173)
- Jackson, A., Davis, G. A., Abeel, M., & Bordonaro, A. (2000). *Turning points 2000 educating adolescents in the 21st century*. New York: Teachers College Press.
- Jenkins, D. M., & McEwin, C. K. (1992). Which school for the fifth grader? Programs and practices in three grade organizations. *Middle School Journal*, 23(4), 8-13.
- Jennings, J. & Rentner, D.S. (2006). Ten big effects of the no child left behind act on public schools. *Phi Delta Kappan*, 88(2), 110-113.
- Juvonen, J., Le, V., Kaganoff, T., Augustine, C., & Constant, L. (2004). *Focus on the Wonder Years: Challenges facing the American middle school*. Santa Monica, CA: Rand Corporation.
- Lee, V. E., & Smith, J. B. (1993). Effects of school restructuring on the achievement and engagement of middle-grades students. *Sociology of Education*, 66, 164-187.
- Look, K. F. (2001). The great K-8 debate. *The Philadelphia Public School Notebook*. Philadelphia, PA: The Philadelphia Fund. Retrieved July 8, 2006, from <http://philaedfund.org/notebook/TheGreatK8Debate.htm>

- Mac Iver, D. J., & Epstein, J. L. (1993). Middle grades research: not yet mature, but no longer a child. *The Elementary School Journal*, 93, 519-533.
- McEwin, C. K., Dickinson, T. S., & Jacobson, M. G. (2004). *Programs and Practices in K-8 Schools: Do They Meet the Educational Needs of Young Adolescents*. Westerville, Ohio: National Middle School Association.
- McEwin, C. K., Dickinson, T. S., & Jenkins, D. M. (2003). *America's middle schools in the new century status and progress*. Westerville, Ohio: National Middle School Association.
- Mertens, S. B., & Anfara, V. A. (2006). Student achievement and the middle school concept. *National Middle School Association: Research Summaries*. Retrieved April 14, 2007, from <http://www.nmsa.org/ResearchSummaries/StudentAchievement/tabid/276/Default.aspx>
- Mizell, H. (2005). Grade configurations for educating young adolescents are still crazy after all these years. *Middle School Journal*, 37 (1), 14-23.
- Montana Office of Public Instruction. (2006). *MontCAS, Phase 2, Criterion Referenced Test, Technical Manual*. Helena, MT: Montana OPI. Retrieved November 20, 2007, from, <http://www.opi.mt.gov/PDF/Assessment/CRT/06techmanual.pdf>
- Montana Office of Public Instruction. (2005a). *Montana School Accreditation Manual*. Helena, MT: Montana OPI. Retrieved November 25, 2007, from <http://www.opi.mt.gov/pdf/accred/05accredmanual.pdf>

- Montana Office of Public Instruction. (2005b). *Montana Youth Risk Behavior Survey Report*. Helena, MT: Montana OPI. Retrieved November 19, 2007, from [http://www.opi.mt.gov/PDF/YRBS/2005Report7\\_8.pdf](http://www.opi.mt.gov/PDF/YRBS/2005Report7_8.pdf)
- Moore, D. W. (1984). *Impact of school grade-organization patterns on seventh and eighth grade students in K-8 and junior high schools* (Rep.). Rockport, ME: New England Educational Research Organization. (ERIC Document Reproduction Service No. ED245346)
- National Middle School Association. (1996). Research Summary #5: Young adolescents' developmental needs. *NMSA Research Summaries*. Retrieved September 25, 2004, from <http://www.nmsa.org/research/ressum5.htm>
- National Middle School Association. (2001). Research Summary #1: Grade configuration. *NMSA Research Summaries*. Retrieved September 25, 2004, from <http://www.nmsa.org/research/ressum1.htm>
- Neill, M. W. (1999). An analysis of the implementation of the essential elements in accredited Montana middle schools. Ed.D. dissertation, University of Montana, United States – Montana. Retrieved August 4, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 9939782)
- Nelson, A. (2006). Overcoming the Income Gap. *Information Brief of the Association for Supervision and Curriculum Development*, 47, 1-8.
- Nichols, S.L., Glass, G.V., & Berliner, D.C. (2005). *High-stakes testing and student achievement: problems for the No Child Left Behind Act*. Tempe, AZ: Education Policy Studies Laboratory, Arizona State University. Retrieved March 15, 2008, from <http://www.asu.edu/educ/epsI/EPRU/documents/EPSTL-0509-105-EPRU.pdf>

- Nobles, A.R. (2008). Grade configuration and the achievement of 8<sup>th</sup> grade students on the North Carolina end of grade tests in reading and mathematics. Ed.D. dissertation, East Carolina University, United States – North Carolina. Retrieved October 24, 2008, from ProQuest Digital Dissertations database. (Publication No. AAT 3302345).
- Offenberg, R. M. (2001). The efficacy of Philadelphia's K-to-8 schools compared to middle grades schools. *Middle School Journal*, 32(4), 23-29.
- Paglin, C., & Fager, J. (1997). *Grade configuration: who goes where?* (Rep.). Portland, OR: Northwest Regional Educational Laboratory.
- Pamperien, K. C. (1997). Student academic achievement in middle-level schools. Ed.D. dissertation, University of Missouri – Columbia, United States – Missouri. Retrieved August 4, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 9841357)
- Pardini, P. (2002). Revival of the K-8 school. *School Administrator*, 59(3), 6-12.
- Peterson, D. J. (2002). The impact of school configuration and school size on students' self-concept, participation, and performance. Ed.D. dissertation, East Carolina University, United States – North Carolina. Retrieved September 12, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 3056813)
- Poncelet, P. (2004). Restructuring schools in Cleveland for the social, emotional, and intellectual development of early adolescents. *Journal of Education for Students Placed at Risk*, 9(2), 81-96.
- Pulliam, J. D. (1987). *History of Education in America*. Columbus: Merrill Pub. Co.

- Roberts, J. E. (2002). The relationship of public middle school size, student achievement, and per-pupil expenditures in South Carolina. Ph.D. dissertation, University of South Carolina, United States – South Carolina. Retrieved September 13, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 3076787)
- Sergiovanni, T. J. (1995). Small schools, great expectations. *Educational Leadership*, 53(3), 48-52.
- Seidman, E., Allen, L., Aber, J. L., Mitchell, C., & Feinman, J. (1994). The impact of school transitions in early adolescence on the self-system and perceived social context of poor urban youth. *Child Development*, 65, 507-522.
- Simmons, R. G., & Blyth, D. A. (1987). *Moving into adolescence the impact of pubertal change and school context*. New York: A. de Gruyter.
- Stetson, P. C. (1917). A statistical study of the scholastic records of 404 junior and non-junior high school students. *School Review*, 25, 617-636.
- Swaim, S. (2004, April 21). Strenth in the middle: New grade configurations won't improve the education of young adolescents. *Education Week*, 23(32), 32,35.
- This we believe: successful schools for young adolescents* . (2003). Westerville, Ohio: National Middle School Association.
- Tucker, C. G., & Andrada, G. N. (1997). *Accountability works: analysis of performance by grade span of school* (Rep.). Harford, CT: Connecticut State Department of Education. (ERIC Document Reproduction Service No. ED411278)
- Tucker, M. S., & Coddling, J. B. (1998). *Standards for our schools how to set them, measure them, and reach them*. San Francisco, CA: Jossey-Bass Publishers.

- U.S. Department Of Education. (2001). *Elementary and secondary education act, Title I, Part A*. Washington, D.C.: U.S. Department of Education. Retrieved November 25, 2007, from <http://www.ed.gov/programs/titleiparta/legislation.html>
- U.S. Department of Education. (2002, January 7). Executive summary of No Child Left Behind Act of 2001. *U.S. Department of Education Press Release*. Retrieved August 4, 2007, from <http://www.ed.gov/nclb/overview/intro/execsumm.doc>
- Vaccaro, D.A. (2000) A comparative study of middle school and K-8 school value added gains in the areas of reading, language, math, science, and social studies. Ed.D. dissertation, The University of Tennessee, United States – Tennessee. Retrieved October 24, 2008, from ProQuest Digital Dissertations database. (Publication No. AAT 9973519).
- Vassallo, P. (1990). Muddle in the Middle: Are middle schools the right choice for the wrong reason? *The American School Board Journal*, 177(9), 26-27, 37.
- Wainer, H., & Zwerling, H. L. (2006). Evidence that smaller schools do not improve student achievement. *Phi Delta Kappan*, 88(4), 300-303.
- Wallis, C. (2005, August 8). Is Middle School Bad For Kids? *Time*, 166, 48-51.
- Wallis, C. & Steptoe, S. (2007, June 4). How to fix no child left behind. *Time*, 169, 34-41
- Weiss, C. C., & Kipnes, L. (2006). Reexamining middle school effects: A comparison of middle grades students in middle schools and K-8 schools. *American Journal of Education*, 112(2), 239-272.
- Wihry, D. F., Coladarci, T., & Meadow, C. (1992). Grade span and eighth-grade academic achievement: Evidence from a predominantly rural state. *Journal of Research in Rural Education*, 8(2), 58-70.

- Wren, S. D. (2004). The effects of grade span configuration and school-to-school transition on student achievement. *The Journal of At-Risk Issues*, 10 (1), 7-10.
- Yakimowski, M. E., & Connolly, F. (2001). *An examination of K-5, 6-8 versus K-8 grade configurations* (Rep.). Baltimore, MD: Baltimore City Public School System, Division of Research, Evaluation, and Accountability.
- Yecke, C. P. (2003). *The War Against Excellence: The Rising Tide of Mediocrity in America's Middle Schools*. Westport, CT: Praeger.
- Yecke, C. P. (2006). Mayhem in the Middle: Why we should shift to K-8. *Educational Leadership*, 63 (7), 20-25.

## Appendix A

### Definition of Terms

*Criterion referenced test (CRT).* A criterion referenced test measures an individual student's performance against a predetermined set of standards which are based on the curriculum (Blair, 2007). In Montana, the CRT is also known as the Montana Comprehensive Assessment System (MontCAS).

*Elemiddle school.* A school that includes both primary and middle grades where there is a specific focus on implementing a middle-level program for those students in the higher grade levels (Hough, 1995). For the purpose of this research, an elemiddle school was defined as a school which housed elementary and middle-grade students in the same building or on the same campus. In addition, an elemiddle was defined by a school where the educational program for the middle-grades differed from the traditional self-contained program offered at the elementary.

*Grade configuration.* The grade spans housed in any one-school setting. For example, a middle school could have a 5-8, 6-8, or 7-8 configuration (Calhoun, 1983).

*Intermediate school.* For the purpose of this research, the intermediate school was defined by any school which was organized in one of the following configurations: 5-8, 6-8, 7-8, or 7-9. In addition as defined by the State of Montana Accreditation Manual, the intermediate school must follow either the middle school philosophy or the departmentalized philosophy often seen in junior high settings (Montana Office of Public Instruction [OPI], Administrative Rules of Montana, Section 10.55.902, 2005a).

*K-8 school.* A school is said to have K-8 configuration when it serves students from kindergarten through eighth grade within the same facility or on the same campus



(McEwin et al., 2004). For the purpose of this research a K-8 school was defined as a school which had the characteristics of an elementary school for all grade levels.

*MontCAS.* This acronym was an abbreviation for Montana Comprehensive Assessment System. In the State of Montana, MontCAS was used to identify the NCLB criterion referenced test.

*School size.* The actual number of students housed in one school building. School size can be further be defined by the number of students at a particular grade level. For example, a large school could be defined as any school having at least 100 students at each grade level (Cotton, 1996). For the purpose of this research a large school was identified as one with more than 400 students (Mertens & Anfara, 2006).

*School socioeconomic status (SES).* For the purpose of this research, SES was defined by the percentage of students within a school who participated in the free or reduced lunch program. A school with a high percentage of students on free or reduced lunch was said to have a low SES. A school with a low percentage of students on free or reduced lunch was said to have a high SES (Barth, 2001). For the purpose of this research, a low SES school was defined as a school that had more than 40% of its student population who participated in the free or reduced lunch program. A high SES school was defined as a school that had less than 40% of its student population who participated in the free or reduced lunch program (US Department of Education, 2001, Section 1114, Title I, Part A).

*Student at-risk behavior.* For the purpose of this research, student at-risk behavior was defined as the occurrence of at-risk behaviors as self-reported on the Youth Risk Behavior Survey (YRBS) given to seventh and eighth grade students in the spring of

2005. The YRBS was a survey established by the U.S. Centers for Disease Control and Prevention to help monitor the prevalence of behaviors that not only influence youth health, but also put youth at risk for the most significant health and social problems that can occur during adolescence and adulthood (Montana OPI, 2005b). For the purpose of this research, the indicator of at-risk behavior was reported as a school average based on the percentage of students who indicated participation in the at-risk behavior.

*Student performance.* For the purpose of this research, student performance was measured by individual student scores on the Montana CRT given to eighth grade students in the spring of 2005. The CRT was the main component of the Montana Comprehensive Assessment System (MontCAS). The CRT test measured student performance in the areas of mathematics and reading comprehension. Results from the CRT were used to assess progress towards completing the goals defined by the No Child Left Behind Act (Montana OPI, 2006). Individual student performance was reported as a raw score which was based on the total number of questions answered correctly.

*Student success.* Although there were numerous methods which could have been used to determine student success in school, for the purpose student success will be defined by the following indicators: student performance on the MontCAS (Montana's CRT) and incidence of at-risk behaviors as measured by the YRBS.

Appendix B  
Sample of Schools

	N	Low Pop	High Pop	Average		Low F/RL	High F/RL	Average
Elemiddle	8	378	1203	803		6%	47%	33%
Middle Schools	17	417	1013	643		13%	52%	29%
K-8's	44	8	573	170		0%	66%	30%

\*School to be used in qualitative purposeful sample.

<b>ELEMIDDLE SCHOOLS</b>							
Low_grade	High_grade	School_name	Agency_name	City	Total Pop	F/RL	F/RL%
PK	06	TARGET RANGE SCHOOL	TARGET RANGE ELEM	MISSOULA	378	108	29%
07	08	TARGET RANGE 7-8*	TARGET RANGE ELEM	MISSOULA			
PK	05	LOWER GRADE HELLGATE	HELLGATE ELEM	MISSOULA	1203	368	31%
06	08	HELLGATE MIDDLE SCHOOL*	HELLGATE ELEM	MISSOULA			
PK	05	LOLO ELEMENTARY	LOLO ELEM	LOLO	595	246	41%
06	08	LOLO MIDDLE SCHOOL*	LOLO ELEM	LOLO			
PK	06	FRENCHTOWN SCHOOL	FRENCHTOWN K-12 SCHOOLS	FRENCHTOWN	852	238	28%
07	08	FRENCHTOWN 7-8*	FRENCHTOWN K-12 SCHOOLS	FRENCHTOWN			
PK	04	EAST EVERGREEN SCHOOL	EVERGREEN ELEM	KALISPELL	769	360	47%
05	06	EVERGREEN 5-6 SCHOOL	EVERGREEN ELEM	KALISPELL			
07	08	EVERGREEN 7-8	EVERGREEN ELEM	KALISPELL			
06	08	MONTANA CITY MIDDLE SCHL	MONTANA CITY ELEM	CLANCY	390	22	6%
KG	05	MONTANA CITY SCHOOL	MONTANA CITY ELEM	CLANCY			
06	08	EAST VALLEY MIDDLE SCHOOL	EAST HELENA ELEM	EAST HELENA	1063	401	38%
PK	02	EASTGATE SCHOOL	EAST HELENA ELEM	EAST HELENA			
03	05	RADLEY ELEMENTARY SCHOOL	EAST HELENA ELEM	EAST HELENA			
03	05	LOCKWOOD INTERMEDIATE	LOCKWOOD ELEM	BILLINGS	1173	518	44%
06	08	LOCKWOOD MIDDLE SCHOOL	LOCKWOOD ELEM	BILLINGS			

<b>MIDDLE-GRADES SCHOOLS</b>							
<b>Low_grade</b>	<b>High_grade</b>	<b>School_name</b>	<b>Agency_name</b>	<b>City</b>	<b>Total Pop</b>	<b>F/RL</b>	<b>F/RL%</b>
07	08	EAST MIDDLE SCHOOL	GREAT FALLS ELEM	GREAT FALLS	851	347	41%
07	08	NORTH MIDDLE SCHOOL	GREAT FALLS ELEM	GREAT FALLS	837	223	27%
06	08	MEADOW HILL MIDDLE SCHOOL*	MISSOULA ELEM	MISSOULA	507	136	27%
06	08	PORTER MIDDLE SCHOOL*	MISSOULA ELEM	MISSOULA	546	283	52%
06	08	WASHINGTON MIDDLE SCHOOL*	MISSOULA ELEM	MISSOULA	599	189	32%
07	08	EAST MIDDLE SCHOOL	BUTTE ELEM	BUTTE	824	294	36%
07	07	LINDERMAN-JR HIGH ANNEX	KALISPELL ELEM	KALISPELL	664	175	26%
08	08	KALISPELL JR HIGH SCHOOL 8TH	KALISPELL ELEM	KALISPELL			
06	08	C R ANDERSON MIDDLE SCHL	HELENA ELEM	HELENA	1013	222	22%
06	08	HELENA MIDDLE SCHOOL	HELENA ELEM	HELENA	720	230	32%
07	08	BELGRADE MIDDLE SCHOOL	BELGRADE ELEM	BELGRADE	417	98	24%
06	08	CHIEF JOSEPH MIDDLE SCHOOL	BOZEMAN ELEM	BOZEMAN	557	119	21%
06	08	SACAJAWEA MIDDLE SCHOOL	BOZEMAN ELEM	BOZEMAN	614	79	13%
07	08	CASTLE ROCK 7-8	BILLINGS ELEM	BILLINGS	704	194	28%
07	08	LEWIS & CLARK 7-8	BILLINGS ELEM	BILLINGS	576	193	34%
07	08	RIVERSIDE 7-8	BILLINGS ELEM	BILLINGS	514	247	48%
07	08	WILL JAMES 7-8	BILLINGS ELEM	BILLINGS	573	76	13%
06	08	LAUREL MIDDLE SCHOOL	LAUREL ELEM	LAUREL	418	99	24%

\* School to be used in qualitative purposeful sample.



07	08	AUGUSTA 7-8	AUGUSTA ELEM	AUGUSTA	52	29	56%
PK	06	AUGUSTA ELEMENTARY SCHOOL	AUGUSTA ELEM	AUGUSTA			
07	08	LINCOLN 7-8	LINCOLN K-12 SCHOOLS	LINCOLN	122	70	57%
PK	06	LINCOLN ELEMENTARY SCHOOL	LINCOLN K-12 SCHOOLS	LINCOLN			
07	08	ANDERSON 7-8	ANDERSON ELEM	BOZEMAN	192	11	6%
PK	06	ANDERSON SCHOOL	ANDERSON ELEM	BOZEMAN			
07	08	GALLATIN GATEWAY 7-8	GALLATIN GATEWAY ELEM	GALLATIN GATEWAY	137	49	36%
KG	06	GALLATIN GATEWAY SCHOOL	GALLATIN GATEWAY ELEM	GALLATIN GATEWAY			
07	08	LAMOTTE 7-8	LAMOTTE ELEM	BOZEMAN	58	0	0%
PK	06	LAMOTTE SCHOOL	LAMOTTE ELEM	BOZEMAN			
07	08	MANHATTAN 7-8	MANHATTAN ELEM	MANHATTAN	355	76	21%
PK	06	MANHATTAN SCHOOL	MANHATTAN ELEM	MANHATTAN			
07	08	OPHIR 7-8	OPHIR ELEM	GALLATIN GATEWAY	142	0	0%
PK	06	OPHIR ELEMENTARY SCHOOL	OPHIR ELEM	GALLATIN GATEWAY			
07	08	THREE FORKS 7-8	THREE FORKS ELEM	THREE FORKS	394	114	29%
PK	06	THREE FORKS ELEM SCHOOL	THREE FORKS ELEM	THREE FORKS			
07	08	WILLOW CREEK 7-8	WILLOW CREEK ELEM	WILLOW CREEK	33	18	55%
PK	06	WILLOW CREEK SCHOOL	WILLOW CREEK ELEM	WILLOW CREEK			
07	08	MONFORTON 7-8	MONFORTON ELEM	BOZEMAN	175	38	22%
03	06	MONFORTON SCHOOL	MONFORTON ELEM	BOZEMAN			
PK	02	MONFORTON PRIMARY	MONFORTON ELEM	BOZEMAN			
PK	08	PASS CREEK SCHOOL	PASS CREEK ELEM	BELGRADE	15	0	0%
PK	08	SPRINGHILL SCHOOL	SPRINGHILL ELEM	BELGRADE	13	0	0%
KG	08	COTTONWOOD SCHOOL	COTTONWOOD ELEM	BOZEMAN	16	0	0%
PK	08	MALMBORG SCHOOL	MALMBORG ELEM	BOZEMAN	20	0	0%
07	08	BROADVIEW 7-8	BROADVIEW ELEM	BROADVIEW	130	58	45%
PK	08	BROADVIEW SCHOOL	BROADVIEW ELEM	BROADVIEW			
07	08	CANYON CREEK 7-8	CANYON CREEK ELEM	BILLINGS	210	52	25%
PK	06	CANYON CREEK SCHOOL	CANYON CREEK ELEM	BILLINGS			
07	08	CUSTER 7-8	CUSTER K-12 SCHOOLS	CUSTER	54	19	35%
KG	06	CUSTER SCHOOL	CUSTER K-12 SCHOOLS	CUSTER			
07	08	ELDER GROVE 7-8	ELDER GROVE ELEM	BILLINGS	340	78	23%
KG	06	ELDER GROVE SCHOOL	ELDER GROVE ELEM	BILLINGS			
07	08	ELYSIAN 7-8	ELYSIAN ELEM	BILLINGS	123	61	50%
PK	06	ELYSIAN SCHOOL	ELYSIAN ELEM	BILLINGS			
07	08	HUNTLEY PROJECT 7-8	HUNTLEY PROJECT K-12 SCHOOLS	WORDEN	479	191	40%
PK	06	HUNTLEY PROJECT ELEM K-6	HUNTLEY PROJECT K-12 SCHOOLS	WORDEN			
07	08	SHEPHERD 7-8	SHEPHERD ELEM	SHEPHERD	573	161	28%
PK	06	SHEPHERD ELEMENTARY	SHEPHERD ELEM	SHEPHERD			

\* School to be used in qualitative purposeful sample.

Appendix C  
Consent Form

**SUBJECT INFORMATION AND CONSENT FORM**

**TITLE:** Comparison of characteristics of Missoula middle level schools.

**PROJECT DIRECTOR(S):**

University Supervisor: Dr. John Matt, University of Montana, (406)243-5610

Researcher: Robert Watson, Doctoral Student, U of M, (406)728-0389

**Special instructions to the potential subject:**

\* This consent form may contain words that are new to you. If you read any words that are not clear to you, please ask the person who gave you this form to explain them to you.

**Purpose:**

\* You are being asked to take part in a research study comparing middle-level schools in Missoula, Montana based on grade configuration.

\* You have been chosen because you have experience working in a middle-level school within Missoula or within a surrounding community.

\* The purpose of this research study is to examine the potential similarities and differences in school characteristics, which may exist between middle-grade schools and K-8 schools. The qualitative data will be gathered through interviews with educators from both types of configurations. The interviews will be used to explore specific school characteristics related to configuration, which may help explain the relationship between grade configuration and student success.

**Procedures:**

\* If you agree to take part in this research study you will be asked to participate in a one-on-one interview process. The interview consists of seven open-ended questions. During the interview, your responses to questions will be recorded in a written format as well as an audio format. The interviewer will ask the questions and record your response in written format. In addition, an audio tape recorder will be used to record your response. The audio tape will be used to verify your response.

\* The study will take place at a location which is mutually agreed upon by both the interviewer and the interviewee.

\* The interview session will last for approximately 25-35 minutes.

**Risks/Discomforts:**

\* There are no known risks and/or discomforts associated with participation in this study.

**Benefits:**

\* Although you may not benefit personally from your participation in this study, your participation may help to define and explain the relationship between school configuration and student success in school. The findings from this study may help educators and policy makers when making configuration decisions.

**Confidentiality:**

\* Your records will be kept private and will not be released without your consent except as required by law. Only the researcher and her faculty supervisor will have access to the files. Your identity will be kept confidential. If the results of this study are written in an educational journal or presented at a meeting, your name will not be used. The data will be stored in a locked file cabinet. Your signed consent form will be stored in a cabinet separate from the data. The audiotape will be transcribed without any information that could identify you. The tape will then be erased.

**Compensation for Injury:**

\* Although we do not foresee any risk in taking part in this study, the following liability statement is required in all University of Montana consent forms. *In the event that you are injured as a result of this research you should individually seek appropriate medical treatment. If the injury is caused by the negligence of the University or any of its employees, you may be entitled to reimbursement or compensation pursuant to the Comprehensive State Insurance Plan established by the Department of Administration under the authority of M.C.A., Title2, Chapter 9. In the event of a claim for such injury, further information may be obtained from the University's Claims representative or University Legal Counsel. (Reviewed by University Legal Counsel, July 6, 1993)*

**Voluntary Participation/Withdrawal:**

\* Your decision to take part in this research study is entirely voluntary. You may refuse to take part in or you may withdraw from the study at any time without penalty or loss of benefits to which you are normally entitled.

**Questions:**

\* You may wish to discuss this with others before you agree to take part in this study. If you have any questions about the research now or during the study contact: Robert Watson, 728-0389 or 240-1995.

\* If you have any questions regarding your rights as a research subject, you may contact the Chair of the IRB through The University of Montana Research Office at 243-6670.

**Subject's Statement of Consent:**

\* I have read the above description of this research study. I have been informed of the risks and benefits involved, and all my questions have been answered to my satisfaction. Furthermore, I have been assured that any future questions I may have will also be answered by a member of the research team. I voluntarily agree to take part or to take part in this study. I understand I will receive a copy of this consent form.

---

Printed (Typed) Name of Subject

---

Subject's Signature

Date



## Appendix D Interview Protocol

### Interview Protocol

Project: Comparison of characteristics of Missoula middle level schools.

Time of interview:

Date:

Place:

Interviewer:

Interviewee:

Position of Interviewee:

(Briefly describe the project.)

This study will examine the potential similarities and differences in school characteristics, which may exist between middle-grade schools and K-8 schools. The qualitative data will be gathered through interviews with educators from both types of configurations. The interviews will be used to explore specific school characteristics related to configuration, which may help explain the relationship between grade configuration and student success.

Questions: (Adapted From Freeman, 2005)

Introductory information – Please describe your experience...current position, grade and subject level, previous work experience, and your experience with other grade configurations.

1. Describe how your school is organized. How might you best describe the grade configuration of your school? Does the grade configuration address or respond to students' developmental needs? Does the configuration promote closer student-teacher relationships? Can you describe why or why not?
2. What do you see as the advantages of the grade configuration of this school? Why do you feel these are advantages? Give some examples that you feel illustrate this.

3. What do you see as the disadvantages of the grade configuration of this school? Why do you feel these are disadvantages? Give some examples that you feel illustrate this.
4. How would you describe the student-teacher relationships that exist in the eighth grade? How would you describe the student-student relationships that exist?
5. What do you think are some of the biggest issues for students at the middle-grade level? Give some examples of these issues.
6. What type of student do you feel does best with this type of grade configuration? Describe the characteristics of these students.
7. What type of student do you feel does not do as well with this type of grade configuration? Describe the characteristics of these students.
8. Please describe the advantages or disadvantages related to curriculum articulation and communication at your school.

(Thank individual for participating in the interview. Assure him or her of confidentiality of responses and potential future interviews.)

## Appendix E Youth Risk Behavior Survey

The following questions taken from the 2005 Youth Risk Behavior Survey will be used to describe student at-risk behavior indicators between various grade configurations.

### Item #2

What is your sex?

- a. Female      b. Male

### Item #3

In what grade are you?

- a. 7<sup>th</sup> grade      b. 8<sup>th</sup> grade

### Item #14

During the past 30 days, on how many days did you carry a weapon such as a gun, knife, or a club on school property?

- a. 0 days      b. 1 day      c. 2 or 3 days      d. 4 or 5 days      e. 6 or more

### Item #15

During the past 30 days, how many days did you not go to school because you felt you would be unsafe at school or on your way to or from school?

- a. 0 days      b. 1 day      c. 2 or 3 days      d. 4 or 5 days      e. 6 or more

### Item #16

During the past 12 months, how many times has someone threatened or injured you with a weapon such as a gun, knife, or club on school property?

- a. 0 times      b. 1 time      c. 2 or 3 times      d. 4 or 5 times  
e. 6 or 7 times      f. 8 or 9 times      g. 10 or 11 times      h. 12 or more times

### Item #17

During the past 12 months, how many times has someone stolen or deliberately damaged your property such as your car, clothing, or books on school property?

- a. 0 times      b. 1 time      c. 2 or 3 times      d. 4 or 5 times  
e. 6 or 7 times      f. 8 or 9 times      g. 10 or 11 times      h. 12 or more times

### Item #20

During the past 12 months, how many times were you in a physical fight on school property?

- a. 0 times      b. 1 time      c. 2 or 3 times      d. 4 or 5 times  
e. 6 or 7 times      f. 8 or 9 times      g. 10 or 11 times      h. 12 or more times

### Item #23

During the past 12 months, did you ever feel so sad or hopeless almost every day for two weeks or more in a row that you stopped doing some usual activities?

- a. Yes      b. No

### Item #26

During the past 12 months, how many times did you actually attempt suicide?

- a. 0 days      b. 1 day      c. 2 or 3 days      d. 4 or 5 days      e. 6 or more

## Item #28

Have you ever tried cigarette smoking, even one or two puffs?

- a. Yes            b. No

## Item #33

During the past 30 days, on how many days did you smoke cigarettes on school property?

- a. 0 days            b. 1 or 2 days            c. 3 to 5 days            d. 6 to 9 days  
e. 10 to 19 days    f. 20 to 29 days    g. All 30 days

## Item #37

During the past 30 days, on how many days did you use chewing tobacco or snuff, or dip on school property?

- a. 0 days            b. 1 or 2 days            c. 3 to 5 days            d. 6 to 9 days  
e. 10 to 19 days    f. 20 to 29 days    g. All 30 days

## Item #39

During your life, on how many days did you have at least one drink of alcohol?

- a. 0 days            b. 1 or 2 days            c. 3 to 9 days            d. 10 to 19 days  
e. 20 to 39 days    f. 40 to 99 days    g. 100 or more days

## Item #43

During the past 30 days, on how many days did you have at least one drink of alcohol on school property?

- a. 0 days            b. 1 or 2 days            c. 3 to 5 days            d. 6 to 9 days  
e. 10 to 19 days    f. 20 to 29 days    g. All 30 days

## Item #44

During your life, how many times have you used marijuana?

- a. 0 days            b. 1 or 2 days            c. 3 to 9 days            d. 10 to 19 days  
e. 20 to 39 days    f. 40 to 99 days    g. 100 or more days

## Item #47

During the past 30 days, how many times did you use marijuana on school property?

- a. 0 times            b. 1 or 2 times            c. 3 to 9 times            d. 10 to 19 times  
e. 20 to 39 times            f. 40 or more times

## Item #50

During your life, how many times have you sniffed glue, or breathed the contents of aerosol spray cans, or inhaled any paints or sprays to get high?

- a. 0 times            b. 1 or 2 times            c. 3 to 9 times            d. 10 to 19 times  
e. 20 to 39 times            f. 40 or more times

## Item #56

During the past 12 months, has anyone offered, sold, or given you an illegal drug on school property?

- a. Yes            b. No