THE RELATIONSHIP BETWEEN SCHOOLS TO WATCH© DESIGNATION AND ACADEMIC ACHIEVEMENT: A STUDY OF COLORADO, NEW YORK, OHIO, AND VIRGINIA

by

KRISTINA NICOLE FALBE

A Dissertation Submitted to the Faculty in the Curriculum and Instruction Program of Tift College of Education at Mercer University in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

Atlanta, GA

2014
DEDICATION

This research is dedicated to the John H. Lounsbury School of Education in honor of Dr. John Lounsbury and the young adolescents to whom he devoted his career.
ACKNOWLEDGEMENTS

I would have never been able to complete this dissertation without the help and guidance of my excellent committee, the wonderful women of Cohort 3, and the support of my husband, friends, and DMS family.

I cannot say enough about my committee members, Karen Swanson, Geri Collins, and Justus Randolph. As a team they were dynamic, all bringing to the table their own strengths, and each pushing me to go just a little further and consider a little different angle. Karen, thank you for not only helping me to complete this dissertation, getting me through all of my topic ideas, and helping me to settle on the one I am truly passionate out, but also for working with me to start my professional career. Dr. Collins, you were, and continue to be, my constant cheerleader, getting me through each checkpoint and letting me know that I could do it. I so appreciate you opening up your classroom, so that I try my hand at teaching adults and discover that I truly enjoy it. Dr. Randolph, you have been constantly patient with me while I ask questions over and over, even when I figure out that I can answer them on my own. I am so grateful for your help and for you helping me to find the best way to answer the questions that I had about education.

To the ladies of Cohort 3 (Go Armadillos!), I am honored to have been able to learn and grow with you all. Each of you has not only supported me, but you also have taught me something or forced me to look at something in different. I admire each
of you and the passion you all have. I know that my future children are in good hands with you all as the future leaders in education.

To my husband, Brandon, who has only ever known me as a student, I look forward to our next season. Thank you for always supporting my goal no matter the time, cost, or stress it caused.

To my friends and family, this has been a long journey. For everyone who has said a kind word, prayed for me, understood when I had to miss an event or leave early, overlooked the mess of my house, etc. so that I could study or write, I appreciate you. Part of navigating higher education is balancing your academic responsibilities with your sanity. You all helped keep me healthy and moving forward.

To the staff at Davis Middle School (otherwise known as BFITW), I could not have found a better place to work, with a better set of teachers and supporters. Special thanks to the administration, Mike Mauriello, Randy Goerner, John Martin, and Hillary Meeler, who have not only supported me, but have also given me countless opportunities to lead, share, and serve. You all have written numerous recommendations so that I could continue my education and progress my career. I love being in a place where I am constantly being pushed forward, and not held back.

I cannot end my acknowledgements without recognizing the foundation for my passion for young adolescents, which stems from my education at Georgia College and State University. To Leigh, Karynne, and Nancy, what I said when I graduated from your program, I still recognize it to be true now: you all taught me what it meant to be a
PROFESSIONAL educator. I view this study as just one more way that I serve as an architect of change.

A special thanks also to Jason Jaros and Dr. Elizabeth Kearney, for all of their work editing and formatting my paper.
# TABLE OF CONTENTS

| DEDICATION | iii |
| ACKNOWLEDGEMENTS | iv |
| LIST OF TABLES | x |
| LIST OF FIGURES | xi |
| ABSTRACT | xii |

## CHAPTER

1. INTRODUCTION TO THE STUDY ............................................................ 1
   - Conceptual Framework .......................................................... 3
   - Theoretical Framework .......................................................... 3
     - Middle School Philosophy ............................................... 4
     - Association of Middle Level Education .............................. 5
     - Schools to Watch© Program ............................................... 7
   - Problem Statement .............................................................. 7
     - No Child Left Behind ....................................................... 8
   - Purpose of the Study .......................................................... 9
   - Need for the Study .............................................................. 9
   - Research Questions and Hypotheses .................................. 10
   - Methods ............................................................................ 12
   - Limitations ....................................................................... 13
   - Assumptions ..................................................................... 14
   - Summary ......................................................................... 14

2. REVIEW OF THE LITERATURE ......................................................... 16
   - Search Strategy and Inclusion Criteria ............................... 16
   - Research Questions ............................................................. 17
   - The Adolescent .................................................................. 18
     - Physical Growth ............................................................. 19
     - Intellectual Characteristics ............................................ 20
TABLE OF CONTENTS (Continued)

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other Characteristics</td>
<td>20</td>
</tr>
<tr>
<td>History of the Middle School</td>
<td>21</td>
</tr>
<tr>
<td>The Middle School Concept</td>
<td>23</td>
</tr>
<tr>
<td>Beyond Grade Configuration</td>
<td>24</td>
</tr>
<tr>
<td>Schools to Watch©</td>
<td>24</td>
</tr>
<tr>
<td>Characteristics of Middle School Philosophy</td>
<td>26</td>
</tr>
<tr>
<td>Academic Excellence</td>
<td>26</td>
</tr>
<tr>
<td>Developmentally Responsive</td>
<td>27</td>
</tr>
<tr>
<td>Organizational Practices</td>
<td>34</td>
</tr>
<tr>
<td>School Climate/Social Equity</td>
<td>38</td>
</tr>
<tr>
<td>Middle School Misconceptions and Criticisms</td>
<td>41</td>
</tr>
<tr>
<td>Middle Schools in the Age of Accountability</td>
<td>44</td>
</tr>
<tr>
<td>High-Stakes Testing and the Middle School</td>
<td>44</td>
</tr>
<tr>
<td>No Child Left Behind</td>
<td>48</td>
</tr>
<tr>
<td>Consequences of No Child Left Behind</td>
<td>49</td>
</tr>
<tr>
<td>The Common Core Initiative</td>
<td>50</td>
</tr>
<tr>
<td>Academic Achievement for Adolescence</td>
<td>52</td>
</tr>
<tr>
<td>Middle Level Philosophy and Academic Achievement</td>
<td>55</td>
</tr>
<tr>
<td>High-Stakes Testing</td>
<td>57</td>
</tr>
<tr>
<td>Subgroup Trends in Academic Achievement</td>
<td>61</td>
</tr>
<tr>
<td>Similar Research Studies</td>
<td>64</td>
</tr>
<tr>
<td>Summary</td>
<td>66</td>
</tr>
</tbody>
</table>

3. METHODOLOGY ................................................................................................................. 69

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Questions and Hypotheses</td>
<td>70</td>
</tr>
<tr>
<td>Population and Sample</td>
<td>73</td>
</tr>
<tr>
<td>Research Design</td>
<td>76</td>
</tr>
<tr>
<td>Measures and Covariates</td>
<td>77</td>
</tr>
<tr>
<td>Predictor Variables</td>
<td>77</td>
</tr>
<tr>
<td>Outcome Variables</td>
<td>77</td>
</tr>
<tr>
<td>Instrumentation</td>
<td>77</td>
</tr>
<tr>
<td>Schools to Watch© Rubric</td>
<td>79</td>
</tr>
<tr>
<td>Data Collection Procedures</td>
<td>79</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>80</td>
</tr>
<tr>
<td>Propensity Score Matching</td>
<td>80</td>
</tr>
<tr>
<td>Variable Matching</td>
<td>81</td>
</tr>
<tr>
<td>Ethical Safeguards</td>
<td>82</td>
</tr>
<tr>
<td>Role of the Researcher</td>
<td>82</td>
</tr>
<tr>
<td>Summary</td>
<td>83</td>
</tr>
</tbody>
</table>

4. RESULTS OF THE STUDY ................................................................................................. 84

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propensity Score Matching</td>
<td>86</td>
</tr>
</tbody>
</table>

viii
TABLE OF CONTENTS (Continued)

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results of Propensity Score Matching</td>
<td>86</td>
</tr>
<tr>
<td>Research Question 1</td>
<td>90</td>
</tr>
<tr>
<td>Linear Mixed Modeling</td>
<td>96</td>
</tr>
<tr>
<td>Language Arts/Reading Results</td>
<td>99</td>
</tr>
<tr>
<td>Math Results</td>
<td>100</td>
</tr>
<tr>
<td>Summary</td>
<td>101</td>
</tr>
<tr>
<td>5. DISCUSSION</td>
<td>103</td>
</tr>
<tr>
<td>Summary of the Study</td>
<td>103</td>
</tr>
<tr>
<td>Interpretation of the Results</td>
<td>105</td>
</tr>
<tr>
<td>Interpretation of the Findings</td>
<td>106</td>
</tr>
<tr>
<td>Addressing the Limitations</td>
<td>107</td>
</tr>
<tr>
<td>Schools to Watch©</td>
<td>107</td>
</tr>
<tr>
<td>Matching</td>
<td>107</td>
</tr>
<tr>
<td>New York Assessment System</td>
<td>108</td>
</tr>
<tr>
<td>State Expectations</td>
<td>109</td>
</tr>
<tr>
<td>Comparison to Other Studies</td>
<td>110</td>
</tr>
<tr>
<td>Implications</td>
<td>112</td>
</tr>
<tr>
<td>Future Research</td>
<td>113</td>
</tr>
<tr>
<td>Conclusion</td>
<td>113</td>
</tr>
<tr>
<td>Final Thoughts</td>
<td>114</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>115</td>
</tr>
<tr>
<td>APPENDICES</td>
<td></td>
</tr>
<tr>
<td>A  IRB APPROVAL</td>
<td>126</td>
</tr>
<tr>
<td>B  SCHOOLS TO WATCH© RUBRIC</td>
<td>128</td>
</tr>
<tr>
<td>C  CODE FOR PSM</td>
<td>138</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Practices for Promoting Academic Excellence</td>
<td>26</td>
</tr>
<tr>
<td>2 Organizational Tenets</td>
<td>35</td>
</tr>
<tr>
<td>3 School Climate Tenets</td>
<td>39</td>
</tr>
<tr>
<td>4 Nationwide Schools to Watch© Middle Schools</td>
<td>74</td>
</tr>
<tr>
<td>5 Standardized Test Reliability and Validity Statistics by State</td>
<td>78</td>
</tr>
<tr>
<td>6 Colorado Demographic Information: Before and After Matching</td>
<td>86</td>
</tr>
<tr>
<td>7 New York Demographic Information: Before and After Matching</td>
<td>87</td>
</tr>
<tr>
<td>8 Ohio Demographic Information: Before and After Matching</td>
<td>88</td>
</tr>
<tr>
<td>9 Virginia Demographic Information: Before and After Matching</td>
<td>89</td>
</tr>
<tr>
<td>10 Unstandardized Estimates of Fixed Effects Covariance Parameters for Language Arts/Reading Scores</td>
<td>100</td>
</tr>
<tr>
<td>11 Unstandardized Estimates of Fixed Effects Covariance Parameters for Math</td>
<td>101</td>
</tr>
<tr>
<td>12 Cut Scores for Colorado, New York, Ohio, and Virginia State Achievement Tests</td>
<td>110</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Conceptual Framework of the Study</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>This We Believe: Keys to Educating Young Adolescents</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>Histograms of Colorado Propensity Scores: Before and After Matching</td>
<td>87</td>
</tr>
<tr>
<td>4</td>
<td>Histograms of New York Propensity Scores: Before and After Matching</td>
<td>88</td>
</tr>
<tr>
<td>5</td>
<td>Histograms of Ohio Propensity Scores: Before and After Matching</td>
<td>89</td>
</tr>
<tr>
<td>6</td>
<td>Histograms of Virginia Propensity Scores: Before and After Matching</td>
<td>90</td>
</tr>
<tr>
<td>7</td>
<td>Histogram for Percentage of Passing in Language Arts/Reading in Non Schools to Watch Schools</td>
<td>92</td>
</tr>
<tr>
<td>8</td>
<td>Histogram for Percentage of Passing in Language Arts/Reading in Schools to Watch© Schools</td>
<td>93</td>
</tr>
<tr>
<td>9</td>
<td>Histogram for Percentage of Passing in Math in Non Schools to Watch Schools</td>
<td>93</td>
</tr>
<tr>
<td>10</td>
<td>Histogram for Percentage of Passing in Math in Schools to Watch© Schools</td>
<td>94</td>
</tr>
<tr>
<td>11</td>
<td>ELA/Reading Results for Non Schools to Watch vs. Schools to Watch©</td>
<td>95</td>
</tr>
<tr>
<td>12</td>
<td>Math Results for Non Schools to Watch vs. Schools to Watch©</td>
<td>95</td>
</tr>
<tr>
<td>13</td>
<td>Language Arts/Reading Scatterplot</td>
<td>97</td>
</tr>
<tr>
<td>14</td>
<td>Language Arts/Reading Histogram</td>
<td>98</td>
</tr>
<tr>
<td>15</td>
<td>Math Scatterplot</td>
<td>98</td>
</tr>
<tr>
<td>16</td>
<td>Math Histogram</td>
<td>99</td>
</tr>
</tbody>
</table>
The purpose of this quantitative study was to examine existing data to determine if there was quantitative evidence that the middle school philosophy, when fully implemented, provides an increase in academic achievement. The primary focus of this study was to compare the academic success of schools designated as Schools to Watch© (meaning they have been reviewed, and are following middle school tenets as defined by the National Forum of Middle Grades Reform) to schools without such designation, but with similar demographic parts.

For this study, academic achievement was measured using results from the 2013 state-based criterion-referenced test in the areas of mathematics and reading/Language Arts for Colorado, New York, Ohio, and Virginia. Schools for the study were selected from each of the four participating states using propensity score matching on the criteria of percentage of minority students, percentage of students in poverty, and the total population of the school. A MANCOVA analysis was conducted to look at the multivariate difference in math and language arts/reading scores. A linear mixed model
analysis was conducted to examine the univariate difference in scores for both math and language arts/reading when taking into accounts the clustering of schools within the states.

Results of the study showed small, but not statistically significant, gains in academic achievement in both areas (math, and reading/language arts) when looking at the multivariate analysis. The univariate analysis suggested statistically significant gains in language arts/reading when taking into account the clustering of schools within states. The univariate analysis for math, showed small, but not statistically significant gains when taking into account the clustering of schools within schools. Recommendations for future research included research to help establish a causal relationship between the implementation of the middle level model and academic achievement as well as duplication of study methodology to include more schools and states.
CHAPTER 1
INTRODUCTION TO THE STUDY

Despite continued reform to the education system, the United States continues to find itself in the midst of a "dropout crisis" (p. 5), according to the annual report from the America's Promise Alliance (Balfanz, Bridgeland, Bruce, & Fox, 2013). The 2010 graduation rate was 78.2%, the highest it has ever been, yet this indicates that 22% of students do not complete a high school education (Balfanz et al., 2013). A research study from ACT Inc. suggested a direct link between high school readiness and academic achievement in middle school. Results showed that 80% of middle schoolers lack the skills and knowledge to be successful in high school (ACT, 2008).

In a 2011 address to the Association for Middle Level Education, U.S. Secretary of Education Arne Duncan acknowledged that the adolescent years are critical for development, both physically and academically. Recognizing that the "middle grades can either put students on a path to college and careers—or steer them to dropping out and the unemployment line" (para. 16), he states that middle schools are the "last, best opportunity for educators to reach all students" (Duncan, 2011). Supporting Arne Duncan's assertion, former first lady of the United States, Laura Bush, points out that many adolescents may not physically drop out until high school, but they make up their mind to drop out in middle school (Stengle, 2011). Duncan (2011) also argues that early intervention is the best prevention for reducing high school dropout.
With little empirical data to show a connection between increased academic achievement and the implementation of the middle school model, fewer middle schools implement the middle school model. Some turn to other models such as K-8 and Junior High Schools to educate young adolescents (Chaker, 2005; Elovitz, 2007). To address the increased emphasis on standardized assessment to show proficiency in teaching, school administrators seek proven strategies to increase students’ academic success, particularly as it is reflected in standardized test scores. The middle school model derived from a philosophy that promotes consideration of young adolescents’ social and emotional development; instead, policymakers focus on academic achievement as evidence of academic success.

Research exists to suggest that middle school should support the whole student, not just his or her cognitive development and academic achievement; however, it is a direct link I am looking for. In response to this, the Association of Middle Level Education (AMLE) publicized This We Believe: Keys to Educating Young Adolescents (AMLE, 2013) presenting 16 characteristics to use as a basis for middle school design. These components include: curriculum that is relevant, challenging, integrative, and exploratory; multiple learning and teaching approaches that respond to their diversity; assessment and evaluation programs that promote quality learning; organizational structures that support meaningful relationships and learning; school-wide efforts and policies that foster health, wellness, and safety; and multifaceted guidance and support services.
Conceptual Framework

The conceptual framework of this study is based upon the research conducted by AMLE (2013), formerly the National Middle School Association, and the National Forum to Accelerate Middle-Grades Reform (2010, 2013a). According to these entities, student academic achievement increases when teaching and organizational practices respond to the developmental needs of adolescents; challenge adolescents academically; promote a safe culture and climate, and utilize best practices for young adolescent students. Figure 1 depicts the conceptual framework for this research.

![Conceptual Framework of the Study](image)

**Figure 1.** Conceptual Framework of the Study

Theoretical Framework

According to constructivist learning theory, learning is tied to the development of the student and the practices of the classroom and school. Piaget’s theory settles on the idea that cognitive development goes through fixed stages (Schunk, Meece, & Pintrich, 2012). In his work, Piaget identified four stages: the sensorimotor stage (Birth through age 2); the preoperational stage (age 2 through age 7); the concrete operational stage (age 7 to 11); and the formal operational stage (age 11 through adult). Children in a middle
school fall between the ages of 10 and 15; therefore, middle grades teachers and administrators work with students transitioning from the concrete operational stage into the formal operational stage. Relevant to this study’s conceptual framework is the acknowledgement that age and development are key factors in the type of instruction that students should receive, indicating that appropriate education must be developmentally responsive in order to lead to academic success.

Vygotsky’s theory of social constructivism (1978) is relevant to the concept of developmental responsiveness and its link to the tenets of academically challenging, organizational practices, and school culture. Vygotsky’s theory rests on the idea that “the social environment [is] critical for learning and through that, social interactions transformed learning experiences” (Schunk et al., 2012, p. 243). According to Vygotsky (1978), challenging learners through the zone of proximal development (ZPD) also promotes social interaction. The zone of proximal development can be defined as what one can do independently through problem solving, and what one can do with the assistance or guidance of a more able peer (Vygotsky, 1978). Through ZPD development, students work with more able peers or adults. In this sense, learning is social, and classrooms and schools are designed facilitate challenge and social interaction, key tenets in the middle school model advocated by AMLE (2013) and the National Forum to Accelerate Middle-Grades Reform (2010, 2013a)

Middle School Philosophy

In order to meet the critical developmental needs of young adolescents, the application of the middle school concept has been used as a reform targeted for
adolescent success. The middle school concept is articulated through the 16 characteristics described by AMLE (2013) and the characteristics defined by the National Forum to Accelerate Middle-Grades Reform in their *Schools Criteria Rating Sheet* (National Forum to Accelerate Middle-Grades Reform, 2010). For purposes of this study, the term *middle school* includes schools that serve grades six through eight in any configuration and use the term middle school and/or junior high school to describe themselves.

In-depth discussion of the particular components of AMLE’s tenets and the characteristics set forth by the National Forum to Accelerate Middle-Grades Reform occurs in the literature review of this study. Several attributes overlap: developmental responsiveness, challenging academics, school climate and culture, and school organization practices. For this study, these overlapping attributes serve as the basis for defining a quality middle school. It is the intent of this study to identify a link between simultaneous use of these four tenets and improved academic achievement.

Association of Middle Level Education

AMLE, formerly known as the National Middle School Association, organized the characteristics of the middle school for the practitioner in the publication *This We Believe: Keys to Educating Young Adolescents* (2013). The association recognizes 16 characteristics as key components to middle school. These characteristics are situated inside of the four attributes that they feel distinguish a true middle school. *Figure 2* displays the conceptual framework as laid out in *This We Believe: Keys to Educating Young Adolescents* (AMLE, 2013).
Figure 2. This We Believe: Keys to Educating Young Adolescents. Used with permission from the Association for Middle Level Education (www.amle.org).
Schools to Watch© Program

The middle school tenets have to be fully implemented to realize the full potential, and therefore the academic success, of the philosophy (Beane & Lipka, 2006). The Schools to Watch© program, established by the National Forum to Accelerate Middle Grades, works to recognize schools that are fully implementing the middle school philosophy. On its website, the National Forum to Accelerate Middle-Grades Reform describes themselves as “an alliance of over 60 educators, researchers, national associations, and officers of professional organizations and foundations committed to promoting the academic performance and healthy development of young adolescents” (National Forum to Accelerate Middle-Grades Reform, 2013a, para. 1). The Forum established the Schools to Watch© program (STW) in 1999 to recognize middle schools that successfully implement all of the middle school tenets. The vision of STW© is to distinguish schools that are academically excellent, responsive to the developmental needs and interests of young adolescents, and socially equitable. The program uses a rubric to determine performance in each of these areas. The rubric is available in Appendix B. Currently the National Forum to Accelerate Middle-Grades Reform recognizes 352 schools as Schools to Watch©, spanning 18 states. For purposes of this study, I focus on Colorado, New York, Ohio, and Virginia.

Problem Statement

The middle school philosophy is more than just a name change for it includes the adoption of a philosophy to meet the educational and developmental needs of young
adolescent learners. While many schools have converted to “middle schools” by name, still (years after reform) many of them do not fully implement the philosophy and the characteristics linked to the implementation of the middle school model. The implementation of federal regulations, such as those created with No Child Left Behind, adds pressure to middle schools to meet academic demands. As Meyer (2011) points out though, little quantitative evidence is available to support the middle school as a model for academic growth.

No Child Left Behind

The regulations of No Child Left Behind (NCLB, 2001) increased pressure for schools to prove that they are performing at or above the minimum state competency levels. Intended to increase accountability and the competency of students of all demographic groups, NCLB generated several unintended consequences (Schul, 2011). Schul (2011) acknowledges four unintended consequences that plague the education system at all grade levels: overemphasis of standardized testing, de-professionalization of teachers, disregard for teacher education, and marginalization of social studies education. The 2001 legislation requires all students meet or exceed proficiency levels on state administered academic achievement tests. As accountability increases, reform efforts are typically linked to how to raise test scores. At the middle school level, some schools implement study skills classes in lieu of advisory (meeting one-on-one with students) and include additional academic classes in place of exploratories such as art, industrial arts, and music. Therefore the problem to be studied is if does the middle school model, when fully implemented, have an impact on academic achievement?
Purpose of the Study

The purpose of this quantitative study was to examine existing state data to determine if there was evidence that the middle school philosophy is not only important in building the social and emotional development of young adolescents, but also that it is a worthwhile model to enhance academic achievement. This study primarily focused on comparing the academic performance of schools identified as Schools to Watch© (meaning they have met the characteristics to be deemed a middle school) to schools without such designation. Schools were compared on similar demographic variables: percentage of disadvantaged students, percentage of minority students, and school population.

Need for the Study

There is little experimental data to support the assertion that the implementation of middle school programs affects academic achievement. Mertens (2006) recognizes that middle level education is facing a research crisis. He acknowledges that few reliable studies have been conducted to validate a positive relationship with the components of the middle school to academic achievement, such as teaming, advisory, shared vision, exploratory curriculum, and others (AMLE, 2010b). Findings from this study can influence policy makers and administrators to use middle school components to create a balance between meeting the social and emotional needs of students and improving academic achievement. Furthermore, it provides evidence that supports administrators’ use of middle level practices in an environment of accountability.
Research Questions and Hypotheses

There are three questions and three hypotheses for this study:

R1 Is there an overall multivariate statistically significant difference in mathematics and language arts/reading achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools (middle schools not identified as Schools to Watch©) when controlling for percentage of disadvantaged students, percentage of minority students, and school population?

H1 There will be an overall multivariate statistically significant difference in mathematics and reading achievement of adolescent students in Schools to Watch middle schools and adolescent students in matched traditional middle schools (middle schools not identified as Schools to Watch©) when controlling for percentage of disadvantaged students, percentage of minority students, and school population.

H0 There will not be an overall multivariate statistically significant difference in mathematics and reading achievement of adolescent students in Schools to Watch middle schools and adolescent students in matched traditional middle schools (middle schools not identified as Schools to Watch©) when controlling for percentage of disadvantaged students, percentage of minority students, and school population.

R2 Is there an overall univariate statistically significant difference in the language arts/reading academic achievement of adolescent students in Schools to Watch©
middle schools and adolescent students in matched traditional middle schools when controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states?

H<sub>1</sub> There will be an overall univariate statistically significant difference in the language arts/reading academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools with controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states.

H<sub>0</sub> There will not be an overall univariate statistically significant difference in the language arts/reading academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools with controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states.

R<sub>3</sub> Is there an overall univariate statistically significant difference in the reading academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools when controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states?
H₂ There will be an overall univariate statistically significant difference in the math academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools with controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states.

H₀ There will not be an overall univariate statistically significant difference in the math academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools with controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states.

Methods

For this study, academic achievement was measured using the state-based criterion-referenced test in the areas of mathematics and reading, which is based on each particular state's standards. Tests include the Virginia Standards of Learning Assessment (SOL), the Ohio Achievement Assessment (OAA), the New York Common Core ELA and Mathematics test, and the Colorado Student Assessment Program. To ensure similarity of standards among states and increase generalizability of results, this study included only schools that used the Common Core initiative.
Limitations

This study is limited by the use of the Schools to Watch© program to define middle schools. While it is assumed that all of the schools are implementing the tenets, I am not able to determine to what degree the schools are following the tenets. In the same way it is not fair to assume that schools that do not have a Schools to Watch© designation are not implementing the tenets of the middle school as defined by Schools to Watch.

The inability to find an exact match also limits my study. While propensity score matching allowed for a close match in terms of percent of minority students, percent of disadvantaged students, and total population, there are many other factors which could impact academic achievement. By design, correlational studies do not provide strong support for causation. Field (2009) states, "Causality between two variables cannot be assumed because there may be other measured or unmeasured variables affecting the results" (p. 173). Even though propensity score matching controls for some variables, there are still other variables that are not controlled.

The multi-state approach also presents a limit, as not all states use the same standardized test. This study limits the states used to only states that are currently using the Common Core initiative to influence their state standards, but each state writes their own test. The use of the test as a measurement of academic achievement is also a limitation, as a standardized test cannot measure actual learning. Finally, my personal bias toward the success of the middle school philosophy may have an effect on data interpretation and presentation.
Assumptions

In order to draw conclusions from the data, many assumptions were made. First, I assumed accurate identification of schools designated as Schools to Watch© and carefully monitoring of Schools to Watch© criteria. It is also assumed that these schools are continuously implementing the characteristics of the middle school philosophy, as they were when they were awarded status. Re-evaluation is done every four years, but practices are likely to change in that timeframe.

Assumptions must also be made about the validity of the criterion-based test. Criticisms exist regarding the simplicity of the test and/or its representativeness when assessing required standards. For this study, it was assumed that the state criterion-referenced tests were an accurate representation of what students are expected to know based on state standards.

Summary

Middle school implementation, still in its development phase, has experienced growth in reforming adolescent learning to meet the developmental needs of young adolescent students. While this success has been examined, more research needs to be done to link the particular components of the middle school to academic achievement. Without empirical research, the middle school is in danger of losing its developmental focus.

In Chapter 2, I discuss the timeline of the middle school model, the characteristics of young adolescents, and the middle level philosophy as defined by the AMLE and the National Forum to Accelerate Middle-Grades Reform. Also addressed are the ways in
which academic achievement is defined by the No Child Left Behind Act. Throughout the chapter, I dissect various studies that have been conducted linking middle level components, or the middle level philosophy as a whole to increased academic achievement, or to social and emotional growth. It is important to note that a majority of the studies discussed are outdated because few studies linking academic achievement and middle level philosophy have been conducted recently.
CHAPTER 2
REVIEW OF THE LITERATURE

This literature review examines research conducted on the middle school philosophy since its inception in the 1950s. This chapter has several major sections: the adolescent, history of the middle school, an overview of the middle school concept and characteristics, a review of middle schools in the age of accountability, the literature on academic achievement for adolescence, and a review of similar research studies. The pressure to increase academic achievement has the potential to change the way that middle school is defined, particularly if academic achievement cannot be linked to current components of the middle school concept. No national study of the connection between academic achievement and the implementation of middle level philosophy exists. This study attempts to illuminate the connection.

Search Strategy and Inclusion Criteria

First, I conducted an electronic search of the academic databases Academic Search Elite, Research Library (at ProQuest), and Wilson OmniFile: Full Text Mega Edition. The electronic search used the search terms middle grades philosophy; middle grades concept; middle school and high-stakes testing; middle school and academic achievement; young adolescents; and Schools to Watch©. Inclusion of articles and books depended upon meeting the following criteria:

- The article or book was written in English.
• The article or book dealt with the middle school philosophy or one of the identified components of the middle school philosophy.
• The article or book dealt with young adolescents.
• The article or book focused on assessment or legislation relating to middle schools.
• The article or book focused on academic achievement in the areas of math and reading.

The title of each entry was read to determine if it might meet the criteria for inclusion. If it appeared to fit, the abstract was read to determine if it met the inclusion criteria. References from chosen articles and books identified other articles or experts in the field to consider for my search. This process was repeated until a sufficient amount of evidence was gathered to support my understanding of the research questions.

Seminal works published by the Association of Middle Level Education established the historical and philosophical foundation used in this study for middle level education. The Association of Middle Level Education considers the texts *This We Believe: Keys to Educating Young Adolescents* (AMLE, 2010b) and *Research to Support This We Believe: Keys to Educating Young Adolescents* (AMLE, 2010a) to be its position papers on the middle school concept. Therefore, they were crucial to this research.

**Research Questions**

The research questions for this study are listed below:

R₁ Is there an overall multivariate statistically significant difference in mathematics and language arts/reading achievement of adolescent students in Schools to
Watch© middle schools and adolescent students in matched traditional middle schools (middle schools not identified as Schools to Watch©) when controlling for percentage of disadvantaged students, percentage of minority students, and school population?

R$_2$ Is there an overall univariate statistically significant difference in the language arts/reading academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools when controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states?

R$_3$ Is there an overall univariate statistically significant difference in the reading academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools when controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states?

The Adolescent

Young adolescence is recognized as the human development phase that takes place between ages 10 and 15. In this age range, a myriad of changes occurs, particularly physical and emotional. These types of changes play a large role in student self-perception, a contributing factor in the academic performance of young adolescents (Parker, 2010). Lounsbury (2000) describes the adolescent years as critical years where
students' philosophies, self-concepts, value systems, and beliefs are solidified. As students secure these beliefs, they begin to define themselves as others see them (O'Connor, 2009, p. 23). As adolescents begin to define themselves as others see them, they take less academic risks, fearing the perceptions of their peers (San Antonio, 2006). This unique developmental stage and fragility of adolescents specifically led middle school advocates to design and promote a school philosophy that is unique to the defining characteristics of a middle school student.

Physical Growth

Caskey and Anfara (2007) explain that aside from the growth experienced in the first two years of life, the physical growth experienced in young adolescence is greater than any other time in human development. They further explain that included in this physical growth are the refinement of motor skills, both gross and fine, and biological maturity. The physical changes that occur during young adolescence are varying and often do not happen at the same time for every child. In fact, adolescent females often experience physical maturity one to two years before their male counterparts (Caskey & Anfara, 2007; Salyers & McKee, 2007).

Some specific examples of physical characteristics include awkwardness and coordination issues due to nonsymmetrical growth of the muscles and the bones (Caskey & Anfara, 2007; Salyers & McKee, 2007). Adolescents also begin the process of sexual maturation, which leads to the development of sexual characteristics (Caskey & Anfara, 2007). According to Kellough and Kellough (2008), changes in metabolisms require adolescents to need more nutrition and increased time for physical activity.
Intellectual Characteristics

While intellectual changes are not as noticeable as physical changes, Caskey and Anfara (2007) make the case that these changes are just as intense as the changes adolescents are undergoing physically. Intellectual changes in adolescence are marked by an increased ability to reason, a wider perspective of relevant social issues, an increased ability to participate in metacognition, a curious nature, an increased ability to find relationships between things, a tendency to argue to increase understanding, and motivation to connect learning to immediate goals and interests (Caskey & Anfara, 2007; Salyers & McKee, 2007). These characteristics have implications for teaching in the middle school classroom. According to Caskey and Anfara (2007), the intellectual characteristics of young adolescents are best utilized when instruction is active and not passive. This type of instruction would include real life experiences, connections to the future, and personal goals.

Other Characteristics

Adolescents also undergo changes in moral/ethical, emotional/psychological, and social development. On the moral grounds, young adolescents are idealistic and fair (Kellough & Kellough, 2008). At this age, they are transitioning from their self-centered perspectives and recognizing others (Caskey & Anfara, 2007). For this reason, education for young adolescents should include components of social justice and ask students to make considerations and connections to others. Emotionally, young adolescents are trying to gain their independence from their families and adults, whom they still depend on, and gain acceptance from their peer group. Emotional characteristics resulting from
the transition through adolescence into adulthood include moodiness, erratic behavior, tendencies to exaggerate, a sense of optimism, rebellion with adults, and a seeking of self-esteem and self-concept. These emotional changes lead to changes in social behavior as well. Young adolescents struggle with their desires to please adults and their peer groups, so they often test the limits of satisfactory behavior. Socially, young adolescents are also starting to seek acceptance from the opposite sex and conform to group norms.

History of the Middle School

As early as the 1890s, it became evident that young adolescence is a unique developmental stage that requires specialized educational practices. The 1899 report by the National Education Association states, “The age of adolescence demands new methods and wiser direction” (NEA, 1899, p. 31). In the same report, a model was proposed to reduce the stress of the transition to high school by creating a bridge from the one-teacher classroom of elementary school to the multi-teacher, departmentalized, and multi-classroom model common in high schools. In 1913, nearly 13 years later, this report led the Committee on the Economy of Time and the Commission on the Reorganization of Secondary Education to recommend the introduction of the junior high concept on a national level (Juvonen, Le, Kaganoff, Augustine, & Constant, 2004). The main reorganization came in different grade configurations, changing from a 1st-8th elementary school to a 9th-12th or a 7th-12th secondary school option (Cuban & Usdan, 2003). Cuban and Usdan (2003) acknowledge that the reorganization was intended to “end waste in schooling children, rescue teenage boys and girls from dropping out in the eighth grade, and provide prevocational choices to uncertain youth” (p. 231).
As the name suggests, junior high schools became scaled-down versions of the high schools and did not specialize in meeting the needs of young adolescent learners; rather the junior high schools worked to meet the needs of the society. At the height of the popularity of the junior high school, immigration and urbanization were on the rise. The junior high school attempted to *Americanize* (meaning to assimilate the various cultures) society through the curriculum and programs offered. Junior high schools also had wellness facilities and showers to help with the sanitary concerns posed by the neighborhoods and societies where the schools were located (Juvonen et al., 2004). Juvonen et al. (2004) suggest that junior high schools did not support adolescent development due to the sharp transition; the lack of connection between the curriculum and the daily lives of the young adolescents; teacher-driven instructional strategies; and rigorous academic standards that left many students behind.

In their book *The Modern Junior High School*, Gruhn and Douglas (1956) lay out a vision for a separate junior high school that adhered to a basic philosophy for educating adolescents. Their philosophy was specific that the junior high school was a school for early adolescents, and should have an emphasis on growth and development of students at this age. They also felt that the junior high school should “recognize each child as an individual”, provide “general rather than specialized instruction”, provide “preparation in basic skills and knowledge”, and function as community (p. 30). By the 1960’s, increased political tensions, along with academic failure, at the junior high school, created a climate for a middle school grass roots movement. Alexander (1968), building on work from Gruhn and Douglas (1956), developed a vision for a school that would do
more than mimic the high school; instead it would meet the developmental needs of students and create a community of teachers and learners (Weiss & Kipnes, 2006).

As the movement began to see success, the National Middle School Association (NMSA, 2003) was founded in 1973. This organization, now named the Association for Middle Level Education (AMLE), has been instrumental in centralizing the views of the middle level movement. By 1980, the president of NMSA, John Swaim, understood the need for clarifying the design of the middle school and began the task of developing a position paper. This paper, *This We Believe*, now in its third publication, is a foundational piece of the middle school movement. Another pivotal publication, *Turning Points: Preparing American Youth for the 21st Century*, published by the Carnegie Council (1989) has been instrumental in not only providing support of the middle school movement, but also in describing developmentally appropriate practices to be used in the middle school. Beane and Lipka (2006) acknowledge that these two documents have provided a blueprint for the restructuring of education for young adolescents.

The Middle School Concept

Research from the Association for Middle Level Education and institutions such as the National Forum for Middle Grades Reform recognizes that middle schools are different in that they are designed to operate on a set of principles or ideals that are unique to adolescent learners. While these organizations do not espouse the exact same set of characteristics, a comparable overlap in expectations is present in the work of the National Forum to Accelerate Middle-Grades Reform and the Association for Middle Level Education. *Figure 1* depicts a comparison of the overarching tenets. These beliefs
focus upon adolescent developmental needs, academic challenges, school culture and climate, and organizational practices. Additional comparisons on the more specific features of each are discussed throughout this section.

Beyond Grade Configuration

The idea of a middle school lies beyond just the grade configuration or name of a building. Beane and Lipka (2006) point out that none of the research on the middle school concept is specific in regards to the grade configuration. The ideas of the middle school concept are relevant to good practices across all grades levels, but are specifically designed to meet the developmental needs of young adolescents (Beane & Lipka, 2006). Therefore grade configurations can include K-8, 6-8, 7-9, or any other configuration that includes students in their young adolescent years.

When defining the purpose of middle school, Jackson and Davis (2000) state plainly that middle schools should further the development of adolescents. They expand that to add:

It is to enable every student to think creatively, to identify and solve meaningful problems, to communicate and work well with others, and to develop the base of factual knowledge and skills that is the essential foundation for these 'higher order' capacities. (pp. 10-11)

Schools to Watch©

The Schools to Watch© program, founded by the National Forum to Accelerate Middle-Grades Reform, designated their first four schools in 1999 after inviting 28 schools nationally to apply to the program. Since 1999, the National Forum to
Accelerate Middle-Grades Reform has expanded their program into a state program, which launched in 2002. Currently 18 states are involved in the state program, and 352 schools have been designated Schools to Watch©.

Developed from a vision of the National Forum to Accelerate Middle-Grades Reform, the Schools to Watch© program seeks to identify and define school excellence at the middle level (Lipsitz & West, 2006). Lipsitz and West (2006) explain that the National Forum to Accelerate Middle-Grades Reform holds a central belief that “young adolescents are capable of learning and achieving at high levels”; they are “dedicated to improving schools for middle-grades students across the country” (p. 57). The National Forum to Accelerate Middle-Grades Reform (2013b) perceives three principles as essential to successful schools in the middle grades:

a) They are academically excellent—these schools challenge all students to use their minds well.

b) They are developmentally responsive—these schools are sensitive to the unique developmental challenges of early adolescence.

c) They are socially equitable—these schools are democratic and fair, providing every student with high-quality teachers, resources, and supports.

In order to define these priorities, the forum created a checklist-type rubric that identifies specific actions taken by successful schools for each of the priorities (Appendix B). The rubric is organized in four domains: Academic Excellence, Developmental Responsiveness, Social Equity, and Organizational Support and Processes. These four domains help to organize the characteristics of the middle school philosophy.
Characteristics of Middle School Philosophy

Academic Excellence

The National Forum to Accelerate Middle-Grades Reform and the Association of Middle Level Education define academic excellence by the presence of particular best practices identified as correlating to academic achievement in the middle school. Table 1 presents these practices.

Table 1

Practices for Promoting Academic Excellence

<table>
<thead>
<tr>
<th>National Forum to Accelerate Middle-Grades Reform (2010)</th>
<th>Association of Middle Level Education (2010b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) All students are expected to meet high academic standards.</td>
<td>a) Educators value young adolescents and are prepared to teach them.</td>
</tr>
<tr>
<td>b) Curriculum, instruction, assessment, and appropriate academic interventions are aligned with high standards.</td>
<td>b) Students and teachers are engaged in active, purposeful learning.</td>
</tr>
<tr>
<td>c) The curriculum emphasizes deep understanding of important concepts and the development of essential skills.</td>
<td>c) Curriculum is challenging, exploratory, integrative, and relevant.</td>
</tr>
<tr>
<td>d) Instructional strategies include a variety of challenging and engaging activities that are clearly related to the grade level standards, concepts, and skills being taught.</td>
<td>d) Educators use multiple teaching approaches. e) Varied and ongoing assessments advance learning as well as measure it.</td>
</tr>
<tr>
<td>e) Teachers use a variety of methods to assess and monitor the progress of student learning (e.g., tests, quizzes, assignments, exhibitions, projects, performance tasks, portfolios, student conferences).</td>
<td>f) The faculty and master schedule provide students time to meet rigorous academic standards.</td>
</tr>
<tr>
<td>f) The faculty and master schedule provide students time to meet rigorous academic standards.</td>
<td>g) Teachers know what each student has learned and still needs to learn.</td>
</tr>
<tr>
<td>g) Teachers know what each student has learned and still needs to learn.</td>
<td>h) The adults in the school are provided time and frequent opportunities to enhance student achievement by working with colleagues to deepen their knowledge and to improve their standards-based practice.</td>
</tr>
</tbody>
</table>
Developmentally Responsive

Middle school practices are set apart because of the attention to adolescent development. Adolescents do not just grow physically; they also experience growth in the intellectual, emotional, and social realms. Pitton (2001) asserts, “If teachers do not recognize the impact of these developmental differences, then they will not be able to respond accordingly” (p. 24). In order to develop a curriculum that acknowledges all of these areas, it is imperative to understand what characteristics define this age range.

While the Association of Middle Level Education lists developmental responsiveness as one of its essential attributes, the association does not ascribe particular descriptions to how this looks in a middle school. However, the National Forum to Accelerate Middle-Grades Reform does list particular components of developmental responsiveness in their rubric for Schools to Watch© schools (National Forum to Accelerate Middle-Grades Reform, 2010, pp. 4-5):

1. The staff creates a personalized environment that supports each student’s intellectual, ethical, social, and physical development.
2. The school provides access to comprehensive services to foster healthy physical, social, emotional, and intellectual development.
3. All teachers foster curiosity, creativity, and the development of social skills in a structured and supportive environment.
4. The curriculum is both socially significant and relevant to the personal and career interest of young adolescents.
• Teachers use an interdisciplinary approach to reinforce important concepts, skills and address real-world problems.

• Students are provided multiple opportunities to explore a rich variety of topics and interests in order to develop their identity, learn about their strengths, discover and demonstrate their own competence, and plan for the future.

• Students have opportunities for voice-posing questions, reflecting on experiences, and participating in decisions and leadership activities.

• The school staff members develop alliance with families to enhance and support the well-being of the children.

• Staff members provide all students with opportunities to develop citizenship skills, to use the community as a classroom, and to engage the community in providing resources and support.

• The school provides age-appropriate, co-curricular activities to foster social skills and character and to develop interests beyond the classroom environment.

Cognitive development. The work of Jean Piaget, often linked to constructivism, recognizes that cognitive development occurs in stages, and that these stages require different levels of support when learning. Muth and Alverman (1992) list Piaget’s four stages as sensorimotor thought, preoperational thought, concrete operational thought, and formal operational thought. Learners move through these stages because of biological development and experience. According to Piaget’s theory, the proposed curricula and
practices for the middle level need to provide students opportunities to transition from concrete thought to formal operational thought.

Lev Vygotsky (1978) also proposed a theory of cognitive development relevant to the study of adolescence. Often referred to as the social development theory, Vygotsky’s work emphasizes the importance of social experiences as a learner engages in creating meaning in his or her world. When analyzing Vygotsky’s work in relation to cognition, it seems that Vygotskian theory points specifically to three neurological processes important in cognition: memory, attention, and concept formation.

Vygotsky’s work is founded on the understanding that chronological age does not always directly correlate with development (Vygotsky, 1934/1998). While Vygotsky’s theories do not recognize formal, inflexible stages, Gredler (2009) recognizes four stages in Vygotsky’s work: two pre-mastery stages; a stage requiring external regulation of one’s thinking; and a stage requiring internal regulation of one’s thinking (p. 5). These stages developed from multiple experiments on memory, attention, and concept formation intended to measure self-awareness and concept formation (Gredler, 2009). It is in the internal regulation stage that metacognition, typically defined as thinking about thinking, becomes evident. When considering metacognition, Vygotsky’s theory speaks to ideas about self-regulation and other developmental processes of cognitive behavior. In this stage, learners begin to form new ideas and concepts on their own, based on links and associations to other concepts (Gredler, 2009, p. 7). While interpretations of Vygotsky’s work typically put this internalization stage in adolescence, it is important to note that Vygotskian theory suggests that no one principle can account for development (Tudge &
Winterhoff, 1993). Tudge and Winterhoff (1993) establish that biological maturation is not necessarily the only factor in cognitive development, but development has a close relationship to history, culture, and associations with others and the environment.

Vygotsky (1978) considered these associations when he proposed his theory of the zone of proximal development, often referred to as ZPD. Vygotsky’s work is grounded in the belief that cognitive development is aided when more competent peers or adults work with students within their zone of proximal development (Tudge & Winterhoff, 1993, p. 67). Defined by Tudge and Winterhoff (1993), the zone of proximal development is “the variation of what a person can do on their own versus what he can do with a more abled person” (p. 6). The zone is created based on the activity and the social situation surrounding the activity, and therefore is not a finite zone that can always be identified (Tudge & Winterhoff, 1993). Citing Vygotsky, Tudge and Winterhoff (1993), assert that “only when working in the zone of proximal development, and alongside other people, are certain cognitive processes ignited and challenged” (p. 67). While the theory now includes the help of more abled peers, Vygotsky’s work appears to focus on working with adults to achieve the correct zone of proximal development.

Curriculum and instruction. In the high-pressured education atmosphere, teachers and administrators are constantly trying to find the balance between engaging, developmentally appropriate curriculum and coverage of the state standards in time for the mandated state test. The Association for Middle Level Education recognizes that covering the content and learning the content is not synonymous; furthermore, having credible academic standards neither implies nor demands a uniform, prescribed
curriculum. Therefore, AMLE (2010b) recommends that middle grades teachers and curriculum developers work diligently to provide appropriate educational experiences for young adolescents.

Differentiation. Tomlinson (2001) recognizes the dilemma that classroom teachers today face when she explains that teachers “are admonished to attend to student differences, but they must ensure that every student becomes competent in the same subject matter and can demonstrate the competencies on an assessment that is differentiated neither in form nor in time constraints” (p. 8). This dilemma leads to a curriculum taught one way to all students with no concern for students’ unique learning styles or developmental needs. AMLE (2010b) recognizes that successful education practices for the middle grades involve students personally by applying multiple learning and teaching approaches in the classroom.

The understanding that students in the middle grades need unique instruction leads to multiple learning and teaching approaches in the classroom. The National Forum to Accelerate Middle-Grades Reform (2010) recognizes that some of these approaches include varieties of direct instruction, cooperative learning, project-based learning, simulations, hands-on learning, and technology integration. AMLE (2010b) suggests the use of experiments, demonstrations, surveys and opinion polls, simulations, inquiry-based and group projects, community-based services, and independent studies (p. 23).

The intent of instructional strategies such as these is to keep the student engaged, while adhering to young adolescents developmental needs. Tomlinson (2001) asserts, “Engagement is a nonnegotiable of teaching and learning” (p. 52). Klem and Connell
(2004) acknowledge that student engagement leads to academic achievement in the forms of higher grades and test scores; in contrast, low engagement is accompanied with misbehavior, non-attendance, and school drop outs.

Assessment. Tenets set forth by the Association of Middle Level Education support assessment that, like curricula, is developmentally appropriate. Examples of developmentally appropriate assessments include portfolios, along with product and performance assessments (Schurr, Thomason, & Thompson, 1996). These types of assessments work because they are authentic and require integration of recall in a more meaningful way.

Formative assessment. Formative assessment is assessment used throughout instruction to provide feedback on student learning, which speaks to and informs future instruction. Examples include oral feedback, observations, portfolios, rubrics, and interviews (Airasian, 2001; Shepard, Hammerness, Darling-Hammond, & Rust, 2005). Clark (2010) offers three uses of formative assessment: a) informs teaching practice, b) make instructional decisions based on this information; and c) scaffolding students to improve their work

All three uses seem to make the case that formative assessment is used by classroom teachers to make instructional decisions about what to teach and how, in an effort to promote the academic and personal growth of students (Clark, 2010). Consistent throughout the literature is the fact that formative assessment is not formative assessment unless used to change instruction to meet the needs of students. For this purpose,
formative assessment is typically used in the classroom for classroom instructional purposes.

*Portfolio assessment.* Popham (2010) describes portfolio assessment as a compilation of student work samples collected during a designated timeframe. Portfolios can take numerous forms and have numerous uses. Whatever the intention, portfolios are useful because they can show a student progress over time, and teachers can use this progress to determine how to improve instruction for the student. In addition, portfolios allow students to view and reflect their own products (Airasian, 2001). Teachers can increase the power of portfolios by requiring periodic reflection of progress by students over time (Airasian, 2001).

*Performance assessment.* Airasian (2001) defines performance assessments as “assessments in which pupils carry out an activity or produce a product in order to demonstrate their learning” (p. 228). Performance assessments allow teachers to assess procedural knowledge, as opposed to factual knowledge. McLaughlin, McLaughlin, and Pringle (2013) list examples of performance assessment as oral presentations, debates, exhibits, written products, construction of models, and solutions to problems. They explain that in order to create appropriate performance assessments, teachers need to consider the focus and context of the task, along with the directions and a planned rubric.

*Standardized assessment.* The implementation of No Child Left Behind resulted in increased emphasis on standardized testing, which typically takes the form of a one-size-assesses-all model. The pressure associated with the test is creating what some term as high-stakes assessment. High-stakes assessments are measurements tied to significant
rewards or consequences and linked to important decisions (O’Connor, 2009). These standardized assessments can be normative or criterion. The use of the results of normative assessments is to compare test takers to one another; whereas the use of results of criterion testing is to compare test takers’ mastery level of predetermined standards (Airasian, 2001; Linn & Gronlund, 2000; O’Conner, 2009).

The implementation of No Child Left Behind has increased not only the frequency and impact of standardized testing, but also the assessment practices of the classroom. Turner (2009) points out that the pressures associated with high-stake assessments cause teachers to rely on increased direct instruction of test-taking strategies, such as rote memorization, and multiple choice selection strategies as a way to boost test scores. Use of these strategies in place of varied, differentiated assessment and teaching fails to provide a developmentally responsive curriculum.

Organizational Practices

Schools that implement the components of the middle school philosophy share a set of organizational best practices that defines how the school is structured. Table 2 provides a list the characteristics associated with school structure cited by AMLE and Schools to Watch© (STW). These characteristics overlap with ideas of vision statement, strong leaders, professional development, and common best practices.
Table 2

Organizational Tenets

<table>
<thead>
<tr>
<th>National Forum to Accelerate Middle Grades Reform (2010)</th>
<th>Association of Middle Level Education (2010b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) A shared vision of what a high-performing school is and does drives every facet of school change.</td>
<td>a) A shared vision is developed by all stakeholders and guides every decision.</td>
</tr>
<tr>
<td>b) The principal has the responsibility and authority to hold the school-improvement enterprise together, including day-to-day know-how, coordination, strategic planning, and communication.</td>
<td>b) Leaders are committed to and knowledgeable about this age group, educational research, and best practices</td>
</tr>
<tr>
<td>c) The school is a community of practice in which learning, experimentation, and the opportunity for reflection are the norm.</td>
<td>c) Leaders demonstrate courage and collaboration.</td>
</tr>
<tr>
<td>d) The school and district devote resources to content-rich professional learning, which is connected to reaching and sustaining the school vision and increasing student achievement.</td>
<td>d) Ongoing professional development reflects best educational practices.</td>
</tr>
<tr>
<td>e) The school is not an island unto itself; it is a part of a larger educational system, i.e., districts, networks, and community partnerships.</td>
<td>e) Organizational structures foster purposeful learning and meaningful relationships.</td>
</tr>
<tr>
<td>f) The school staff holds itself accountable for student success.</td>
<td></td>
</tr>
<tr>
<td>g) District and school staff possess and cultivate the collective will to persevere, believing it is their business to produce increased achievement and enhanced development of all students.</td>
<td></td>
</tr>
<tr>
<td>h) The school staff and district staff partner with colleges and universities.</td>
<td></td>
</tr>
<tr>
<td>i) The school includes families and community members in setting and supporting the school's trajectory toward high performance.</td>
<td></td>
</tr>
</tbody>
</table>

Vision statements. Schools across the nation use vision statements as a way to make public their goals and aspirations for their school. While these vision statements are aspirations, some vision statements are created and used to direct every decision, and some are simply never incorporated into planning. According to AMLE (2010b), the vision statements of quality middle schools are created with input from multiple stakeholders, including administration, staff, teachers, students, parents, and community
members. These vision statements are not static, but are referred to when making curriculum decisions.

The vision statements described in middle school publications such as *This We Believe* and *Turning Points 2000* are not visions of one person, but are shared visions created by the entire school staff and even community members (AMLE, 2010b; Hopfenberg, 1993; Jackson & Davis, 2000). DuFour (2004) cautions against lofty statements about creating lifelong learners or successful citizens, as these goals are not easily measured. Instead, vision statements need to be clear and concise with measurable and specific goals with obvious outcomes that support the school mission statement (DuFour, 2004; Jackson & Davis, 2000).

Strong leaders. Leadership can be a significant factor in student learning (Leithwood, Louis, Anderson, & Wahlstrom, 2004). Academically successful schools have capable leaders (Jackson & Davis, 2000). Reynolds and O'Dwyer (2008) point out that, since the passing of the No Child Left Behind Act (2001), school leadership is under immense pressure to enact reforms that lead to measurable and visible results. In order to accomplish this at the middle school level, the AMLE (2010b) argues that the leadership has to know what works for young adolescents. In defining *courageous and collaborative leadership*, AMLE (2010b) recognizes leaders that are willing to acknowledge and work towards things that they believe. This belief helps them to “challenge and change practices that do not serve students’ best interests and confront those issues or situations that are out of alignment with the school vision” (p. 29). While the principal is certainly not the only leader in the school, research supports the fact that
the principal is a key figure, and the leadership of the principal can have direct effects on student achievement (Jackson & Davis, 2000; Kruse & Louis, 1997).

Professional development. While quality pre-service education serves as a critical and much needed source of change for the middle grades, quality professional development is an equal source of change. (Jackson & Davis, 2000) Professional development is the key to keeping teachers up-to-date on best practices. Yet still, Jackson and Davis (2000) recognize that “professional development is a problem in virtually every middle grades school in the country” (p. 110). The authors argue that the problem with the professional development is that it is often ineffective due to its lack of relevance for teachers, delivery by nonprofessionals, or rushed delivery (Jackson & Davis, 2000).

Structures. Middle school advocates understand that success, academic and otherwise, is linked to successful relationships between students and teachers (Sterrett, 2012). The middle school uses several organizational structures to foster effective relationships. The components used in middle schools to facilitate relationship building include advisory, teaming, and interdisciplinary teaching.

One way that middle schools build meaningful relationships is through advisory programs. In a middle school, advisory starts with designating a specified time for teacher guidance (Erb, 2006). The intent of advisory is for students to make a personal connection with an adult in the school (Kellough & Kellough, 1999), which can influence students’ personal and academic endeavors (Jackson & Davis, 2000). Studies investigating the effect of advisory programs discovered a link to decreased dropout
rates, improved school climate, and increased student self-esteem (George & Oldaker, 1985; Ziegler & Mulhall, 1994).

Teaming is also an important component of the implementation of the middle level movement. In teaming, one group of teachers (typically 2-4) shares a group of students and creates their own small learning community. Proponents of teaming acknowledge three advantages of this model: the opportunity to discuss the students' needs, plan interdisciplinary curriculum, and facilitate the development of relationships between students, as well as between students and teachers (Mertens, Flowers, Anfara, & Caskey, 2010). One key part of teaming is that it allows teachers to have a common planning time. This is essential for allowing teachers the opportunity to discuss their students, make individualized plans, and create interdisciplinary units. Common planning time links directly to the common positive effects of teaming, such as increased academic scores, parental involvement, and improved self-esteem (Mertens et al., 2010).

School Climate/Social Equity

The National Forum to Accelerate Middle-Grades Reform (2013b) believes “High-performing schools with middle grades are socially equitable, democratic, and fair. They provide every student with high-quality teachers, resources, learning opportunities, and supports. They keep positive options open for all students” (para. 1) Social equity is also recognized by AMLE (2013) as an essential attribute for middle grades: “Advocating for and ensuring every student's right to learn and providing appropriately challenging and relevant learning opportunities for every student” (p. 2) are indispensable components of the teaching middle grades. Table 3 lists the school climate tenets of
Schools to Watch© (National Forum to Accelerate Middle-Grades Reform, 2010) and AMLE (2010b).

Table 3

*School Climate Tenets*

<table>
<thead>
<tr>
<th>National Forum to Accelerate Middle-Grades Reform (2010)</th>
<th>Association of Middle Level Education (2010b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) To the fullest extent possible, all students, including English learners, students with disabilities, gifted and honors students, participate in heterogeneous classes with high academic and behavioral expectations.</td>
<td>a) A shared vision developed by all stakeholders guides every decision.</td>
</tr>
<tr>
<td>b) Students are provided the opportunity to use many and varied approaches to achieve and demonstrate competence and mastery of standards.</td>
<td>b) Leaders are committed to and knowledgeable about this age group, educational research, and best practices.</td>
</tr>
<tr>
<td>c) Teachers continually adapt curriculum, instruction, assessment, and scheduling to meet their students' diverse and changing needs.</td>
<td>c) Leaders demonstrate courage and collaboration.</td>
</tr>
<tr>
<td>d) All students have equal access to valued knowledge in all school classes and activities.</td>
<td>d) Ongoing professional development reflects best educational practices.</td>
</tr>
<tr>
<td>e) Students have ongoing opportunities to learn about and appreciate their own and others cultures.</td>
<td></td>
</tr>
<tr>
<td>f) The school community knows every student well.</td>
<td></td>
</tr>
<tr>
<td>g) To the fullest extent possible, the faculty welcomes and encourages the active participation of all its families and makes sure that all its families are an integral part of the school.</td>
<td></td>
</tr>
<tr>
<td>h) The school's reward system is designed to value diversity, civility, service, and democratic citizenship.</td>
<td></td>
</tr>
<tr>
<td>i) To the fullest extent possible, staff members understand and support the family backgrounds and values of their students.</td>
<td></td>
</tr>
<tr>
<td>j) The school rules are clear, fair, and consistently applied.</td>
<td></td>
</tr>
</tbody>
</table>
Successful middle schools have a unique school culture and a sense of community, a term used to describe the environment of the middle school. The AMLE (2010b) recognizes that middle schools should be a "joyful community that promotes in depth learning and enhances students' physical and emotional well-being" (p. 33). A community of that type is only possible through the building of relationships between students and students and adults. The National Forum to Accelerate Middle-Grades Reform (2010) describes developmentally responsive middle schools as those that "support each student's intellectual, ethical, social, and physical development" (p. 4). They use indicators such as teaming, mutually respectful relationships, mentor programs, and advisor/advisee programs to assess the personalized environment.

Adult advocacy builds student and adult relationships. While advisory programs described previously are certainly one method of promoting advocacy for young adolescents, AMLE (2010b) acknowledges that "advocacy is not a singular event or a period in the schedule; it is an attitude of caring that translates into actions big and small" (p. 35). Middle schools demonstrate advocacy by creating a community where adults know and recognize all students. These relationships are used to meet the learning needs of students (AMLE, 2013).

Positive relationships between schools, parents, and community members also build effective middle school communities. Typically, middle school is the time when parents give students more independence at school and communicate less with teachers. Many times this independence can be abrupt, leaving the students without the appropriate
support. The Schools to Watch© (National Forum to Accelerate Middle-Grades Reform, 2010) rubric acknowledges that:

- families and community members are informed about the school’s goals for student success and students’ responsibility for meeting these goals; and
- representatives of all stakeholders are engaged in ongoing and reflective conversation and decision-making about governance to promote school improvement.

Middle School Misconceptions and Criticisms

The middle school movement is not without some criticism. Skeptics of the middle school movement challenge the effectiveness of the implementation and the use of the middle school model due to its attention to adolescent development and its non-traditional organization systems. Lounsbury (2000) acknowledges that “our perceived lack of emphasis on traditional college preparatory curriculum makes us suspect to other misguided but very influential folk” (p. 193).

The main criticism of the implementation of the middle school philosophy is that the model, while promoting self-esteem and social development, does not promote academic achievement. Citing evidence from the National Center for Education Statistics, Yecke (2006) also argues that evidence links high school problems to middle school failure:

Although many U.S. middle schools are flourishing with strong and rigorous programs, the middle school concept—the notion that middle schools should be
havens of socialization and not academics of knowledge—has wrought havoc on the intellectual development of the middle school student. (Yecke, 2006, p. 20)

Yecke (2006) is not alone in her belief that middle schools are not meeting the academic demands of society. Mitchell and Willis (1995) also argue that the middle school does nothing to increase the intellectual progress of students as they transition from middle school to high school.

John Lounsbury (2000), a longtime advocate of middle level education and a founder of the middle school, argues:

The purported academic failure of the middle school, it should be noted, is due to the fact that the tenets of the middle school have *not* been sufficiently implemented—not that these tenets *have been* implemented. To blame the rarely implemented middle school concept for unsatisfactory test scores is unfair. (p. 193)

Lounsbury’s (2000) assertion are that the middle school components, when implemented correctly, do in fact lead to higher test scores because they are sensitive to the unique needs of the adolescents. Beane and Lipka (2006) supports Lounsbury’s comments by noting that the middle school concept has not been well implemented; so while there are obvious gaps in middle school education, they are not a result of implementation, but rather a result of non-implementation. The issue with this assumption is that is has not been thoroughly examined, especially in the midst of the current academic accountability movement.
While plenty of research does exist to support the fact that the middle school has a negative effect on academic achievement, Holas and Huston (2012) argue that this outcome is a result of bad research design and ask the question, "Are middle schools ill-suited for early adolescents, or can school characteristics account for differences in student functioning?" (p. 333). In their study, Holas and Huston (2012) investigated the academic achievement of students in middle schools, compared to those that stay in elementary settings. Before beginning their research, Holas and Huston (2012) found that methodologies limited the existing literature, which suggests that middle schools have lower academic achievement. The limitations they noted include lack of equal comparison group, the timing of transitions in seventh grade instead of fifth or sixth, limited information about the classroom quality, limited information about the characteristics of the school, and lack of studies controlling for demographic differences.

In shaping their own study, Holas and Huston (2012) used a cohort of students chosen from birth to participate in a wide scale longitudinal study. In this case, the researchers examined the students' academic achievement, measured by performance on the Woodcock-Johnson Psycho-Educational Battery-Revised in third, fifth, and sixth grade. The study also included classroom observations and surveys of students, teachers, and parents. Holas and Huston (2012) found that the difference in academic achievement was comparable between students in middle school settings and students in elementary settings. They also discovered that there is less quality of instruction in middle schools than in elementary schools, and that discrepancies in scores might be due to the quality of instruction than the middle school structure. Holas and Huston (2012) also pointed out
that middle schools are typically bigger than the elementary schools, and represent a mo
gre heterogeneous group.

Middle Schools in the Age of Accountability

High-Stakes Testing and the Middle School

High-stakes testing is at an all-time high, particularly in the middle grades where ado
elscents are among the most frequently tested groups of students (Turner, 2009). French (2003) acknowledges that this drive for high-stakes testing arises from the No Child Left Behind (2001) legislation. French (2003) states,

> The advent of the No Child Left Behind federal legislation has plunged us into an unprecedented era of high-stakes testing, with the presumption that testing and more testing, coupled with the threat of not being promoted from grade to grade and of not graduating from high school, will be the engine that drives improvement in instruction and student achievement. (p. 1)

Numerous studies have pointed to the realization that high-stakes testing does influence the implementation of middle school components. The McEwin, Dickinson, and Jenkins (2004) study asked teachers to identify changes in instructional practices since the implementation of high-stakes testing. In this study, teachers recognized an increase in student remediation, less time in exploratory classes, and decreased time devoted to advisory. Faulkner and Cook (2006) explain that the implementation of high-stakes testing has led some middle schools, in an effort to build skills useful for test taking such as memorizing facts, to give up instructional strategies that are student-centered.
Unlike the developmentally responsive practices recommended for the middle grades, standardized assessment does little to account for disabilities, development, or language deficits (Lewis, 2004). Often presented in one paper-pencil format, these tests do not represent the creative teaching and assessment methods seen in a developmentally appropriate middle grades classroom. As assessment continues to be presented in this rigid, high-pressure format, standardization of curriculum and instruction is likely to follow.

Musoleno and White (2010) considered the impact on best practices in the middle school since the implementation of high-stakes testing as introduced by No Child Left Behind (2001). In Musoleno and White’s (2010) qualitative study, school staff was asked to consider the use of practices for:

- grouping for instruction;
- developmentally appropriate practices (cooperative learning vs. lecture; interactive vs. rote; discovery/inquiry based learning vs. teacher directed instruction);
- interdisciplinary teaming and planning; and
- inclusion of an activity period for interest-based clubs and advising (p. 4).

Respondents felt strategies that cater to the development of young adolescents were used more before the implementation of No Child Left Behind (Musoleno & White, 2010). Teachers in the study reported dilemmas and frustrations related to high-stakes testing. For example, teachers desired to teach material in a way that would be engaging,
challenging, and demanding, but found that the amount of time required for test preparation left little time for creative instruction.

These dilemmas were echoed in a report by Turner (2009), who recognized that middle school teachers' deliberate preparation of students for high-stakes testing may result in failure to provide developmentally responsive instruction. Many of these teachers struggle with their beliefs about developmentally appropriate curriculum and instruction. However, Turner does offer hope by suggesting that the two (teaching for high-stakes testing and providing developmentally appropriate instruction) are not mutually exclusive, and it is possible to teach ethically, responsibly, and effectively in the middle grades.

In a descriptive study considering the perceptions of middle level teachers regarding the impact of high-stakes testing on classroom instructional strategies, Faulkner and Cook (2006) concluded that the state test dictated the curriculum, despite the beliefs of the teacher on the importance of strategies that utilized student-centered methods. The study was part of a larger, more comprehensive study on the implementation of middle school practices in North Kentucky's schools. Targeting 804 staff/teachers in 17 schools, Faulkner and Cook (2006) asked teachers to complete a Likert format survey to evaluate their school implementation of key middle school practices and support of practices across the region. Teachers reported increased use of "teacher-focused instructional methods" (Faulkner & Cook, 2006, p. 1). The researchers reported, "There was an overwhelming sense that the state assessment dictated their practices and, in a sense,
forced them to use ineffective, teacher-focused instructional strategies” (Faulkner & Cook, 2006, p. 7).

Faulkner and Cook (2006) acknowledge that while the intentions of the tests are good, there are many negative consequences. According to Faulkner and Cook (2006), the positives of testing are that it:

- helps schools set performance goals;
- provides a focus for the curriculum;
- reveals academic progress to the public; and
- potentially provides additional funding support through federal programs (p. 2).

Faulkner and Cook (2006, p. 2) recognize the negatives of testing as:

- administrator and teacher cheating;
- student cheating;
- exclusion of low performing students from testing;
- misrepresentation of student dropouts;
- teaching to the test;
- narrowing the curriculum;
- conflicting accountability ratings;
- questions about the meaning of proficiency;
- declining teacher morale; and
- score reporting errors (p. 2).
As stated by Faulkner and Cook (2006), these negative impacts "far outweigh the positives" (p. 2).

**No Child Left Behind**

In response to a growing imbalance of academic achievement amongst affluent and disadvantaged students and large gaps in achievement between minority and majority ethnicities, President George Bush proposed the No Child Left Behind Act (NCLB) in 2001 (Brown, 2002). The act passed in January of 2002, and increased accountability became a reality in the public school system. George (2002) describes the No Child Left Behind Act as "one of the widest sweeping education bills in history" (p. 5).

With No Child Left Behind came widespread changes to education at all grade levels as schools made efforts to meet new demands to show adequate yearly progress (AYP). The NCLB law requires all students, third through eighth grade, to participate in a standards-based test once a year (Crone, 2004). While students typically take tests in all subjects, initial accountability targeted math and reading scores. Since then, the focus has widened to emphasize language arts and science scores.

Accountability involves the collection and analysis of students' test scores to determine if a school meets the criteria for Adequate Yearly Progress (AYP) (Harriman, 2005). The results of AYP are publically posted and used as a means to compare schools. According to Harriman (2005), the mission of accountability is to bring to light any inequalities in education, particularly in the area of socio-economic status. Public attention increases pressure, but now the question raised by Harriman (2005) that needs
answering is “high-stakes annual standardized testing and public reporting on yearly progress of subgroups effectively addressing the mission?” (p. 65).

Consequences of No Child Left Behind

While change spanned the grade levels, the middle school was particularly hard hit, as NCLB pays little attention to developmental responsiveness that is at the very core of the middle school philosophy (Arth et al., 2004). Opponents of the standards-based movement worry that standards do little to take into account student developmental levels, background knowledge, abilities, and culture (Dudley, Jackson, & Stevens, 2006).

Also of concern is the shift in instructional practices since the passing of the NCLB law. In a qualitative study examining such shifts, Harriman (2005) discovered that teachers now question their teaching methods and often turn to the ones that will get results the fastest. Bracey (2009) notes that “schools under the gun to raise test scores increasingly rely on strategies that get immediate, but short-lived results” (p. 34). Hursh (2008) points out that less content is being covered in favor of covering only what will be on the test. Finn (2008) warns against this, stating, “What gets tested is what gets taught” (p. 250). The problem with this mentality is that not all things can be tested on a paper and pencil test, going back to the question of whether the methodology of NCLB is helping to reach the goal of NCLB.

Powell, Higgins, Aram, and Freed, (2009) assert that the implementation of NCLB has created changes in the classroom. These include an increase in instructional time for reading and a decrease in instructional time for social science and science. Powell et al. (2009) also note a decrease in non-curricular activities like recess and
naptime. Other changes noted included creation of new jobs such as reading coaches, increased use of basal readers, increased language arts professional development, more student standardized testing in the form of benchmarks, and a change in perception among decision makers. Surprisingly, principals noted that "meeting AYP and raising test scores" as the reasoning behind decision-making (Powell et al., 2009, p. 24). This response surpassed "children's success in future schooling" and "children's development as effective citizens" (Powell et al., 2009, p. 24).

Powell et al. (2009) makes a persuasive argument against NCLB through the results of their study. One statement, made by a participant of an interview and repeated several times throughout the article, seems to reflect the sentiments of the researchers on the issue. The participant stated, "True learning is being replaced by a focus on passing the tests, and so students are learning less" (Powell et al., 2009, p. 26). This study provides evidence of the noticeable mark on the implementation of middle school practices caused by the No Child Left Behind Act.

The Common Core Initiative

The Common Core Initiative is changing the perception of curriculum in content area classrooms across America. Adopted by 46 states, the Common Core Standards are a set of standards that provide standard benchmarks of the knowledge and skills that all students in the country should have in language arts and math in order to be college and career ready upon high school graduation (Center for Public Education, 2013). The National Governors Association and the Chief State School Officers released the standards in 2010. While the standards are not federally mandated, they are endorsed at
the federal level. The intention of the Common Core is to increase the quality of education and allow for greater consistency from state to state (Center for Public Education, 2013).

What differentiates the Common Core from the current state standards is the purposeful alignment to college and career skills. *Gifted Child Today* (News Briefs: Common Core Standards, 2010) describes the standards as:

- aligned with college and work expectations;
- clear, understandable and consistent;
- rigorous content and application of knowledge through higher order skills;
- built upon strengths and lessons of current state standards;
- informed by other top performing countries so that students are prepared to succeed in a global economy and society; and
- evidence and research based (p. 9).

The Center for Public Education (2013) is firm in its stance that the Common Core is not a curriculum. They explain that the intention of Common Core is not prescription of curriculum and teaching; instead, its design allows for the state and local districts to create curriculum. This contradicts the critics of Common Core who assert that the implementation of the Common Core will lead to more prescribed methods of teaching.

The standards emphasize cross-curricular study and in-depth/complex reading (Hill, 2011). This call to increase text complexity is in response to numerous studies that suggest the amount of students who still read at a remedial level by high school graduation is shocking (Kern, 2012; Sanacore & Palumbo, 2009). What is notable about
the Common Core standards is the need for all teachers to teach literacy and mathematics/reasoning in their discipline. This is an important shift, one necessitated by the demand for complex informational text in the college and work force. Hill (2011) acknowledges that data from the 2005 ACT test clearly indicate that the difference between “those students who are ready for college versus those who were not was the inability to comprehend complex text” (p. 42).

**Academic Achievement for Adolescence**

In a response to criticism about the academic achievement for adolescence in the middle grades, James Beane, a noted middle grades advocate, stated clearly that “study after study shows when [the middle school philosophy is] implemented well over a period of time, students achieve significant increases in academic achievement” (2006). Despite such statements, very few definitive studies support this statement. When looking at the overall achievement of middle school students, research conclusions are unclear regarding the existence of a positive increase in academics. If anything, many research studies point to a negative trend in the middle grade years.

While the connection between the middle grades and academic success is still unclear, there is evidence that a crucial link exists between academic success in middle school and high school performance. When speaking about the need to enact educational reform, Secretary of Education Arne Duncan (2011) recognized three warning flags for future high school dropout: low scores in language arts and math in middle school, misconduct in and out of school, and lack of attendance.
A study conducted by West and Schwerdt (2012) investigating the effectiveness of the "middle school experiment" (as coined by the researchers) gives evidence that middle school is the time when achievement in literacy and math drops dramatically (2012). West and Schwerdt (2012) argue that students have a significant academic drop upon entering middle school, and the decline continues throughout middle school. In some cases, the researchers suggest that the decline evens out in high school around ninth grade. In other cases, particularly cases of students who went to a middle school as opposed to a K-8 school, the decline does not even out as much for students who attended other configurations besides middle school.

To support their argument, West and Schwerdt (2012) examined the math and reading scores, ranging from August 2000 to May 2009, of Florida third through tenth grade students. The researchers compared achievement of students in middle schools to achievement of students in other grade configurations, such as K-8 school. When looking at student achievement before middle grades and achievement during the first year of middle grades (rather the first year was in sixth grade or seventh grade), the data showed an achievement drop of 0.12 standard deviations in math and a drop of 0.009 standard deviations in reading. Results of this quantitative study suggest a link in improved test scores, but not a positive link in terms of students in a middle school setting. If anything, evidence supports larger drops for students in a middle school as opposed to a K-8 school.

Other large-scale studies in states like California and New York echoed these results. Williams et al. (2010) conducted one of the largest studies on middle grades
practices and student outcomes. In 2008-2009, the team carried out a wide-scale qualitative study of middle schools in the state of California. The research focused on 303 middle schools and utilized three surveys to make inquiries regarding 10 areas of successful middle level practice. The study used self-reports by superintendents, principals, and teachers to measure the level of middle level implementation and ELA and mathematics standards-based test scores from the 2009 administration. The schools in the study were of various compositions including K-8, 6-8, and 7-8. Using multiple regression to control for differences in demographics, the researchers found that some middle schools achieved higher than others, and that implementation of middle school practices possibly explained the differences. In order to take into account all of the different responses and demographic differences, the study included 840 regression models. With such a large number of models, the authors chose to report the substantial conclusions, based on a set of criteria they determined. One of the criteria was that the reported domains and subdomains were all significant at the .05 level of statistical significance. Williams et al. (2010) also reported that that the most consistent predictor of improved academic achievement was “an intense school-wide focus on improving academic achievement” (p. 34). They authors further concluded that other policies correlate with higher academic achievement, particularly practices regarding building school culture and educational professionals.

In summarizing their research, Williams et al. (2010) asserted that the connection between academic achievement and the model of the school was not strong enough to draw clear-cut assumptions. The researchers could not point to the exact middle school
component that increases achievement, but did note that a strong focus on improving academics was a commonality among the high-achieving schools (Williams et al., 2010).

Thinking beyond success in high school, ACT (2008) asserts that success in middle school is also a predictor of college and career readiness, and they contend that intervention in middle school can be useful in producing more college and career ready graduates. Results from the ACT (2008) study indicate that “the level of academic achievement that students attain by eighth grade has a larger impact on their college and career readiness by the time they graduate from high school than anything that happens academically in high school” (p. 2). The ACT report analyzed results from multiple predictive models that compared college and career readiness, based on performance on the ACT, with popular predictor factors such as background characteristics, coursework, grade point average, and testing behavior and compared them with predictions on achievement on eighth grade EXPLORE tests. The models clearly predicted that eighth grade achievement was the most influential factor on future achievement on the ACT. In language arts, the study shows an increase of two-and-a-half times; in Science, an increase of three-and-a-half times is reported.

Middle Level Philosophy and Academic Achievement

Despite concerns that the middle school model as a whole does not meet the academic demands of the young adolescent, studies reveal a positive relationship between academic achievement and components of the middle school concept. For example teaming, one of the most widely recognized components of the middle school concept, has been studied multiple times. Arhar’s (1997) examination of early literature on the
effect of interdisciplinary teaming on teachers and students revealed the most positive effects of teaming were achievement and engagement. However, studies conducted by Arhar in 1992 and 1994 report no significant difference in the area of academic achievement. These early studies did support the role of teaming in increasing engagement and strengthening relationships between students and teachers.

Using data from the Michigan Educational Assessment Program (MEAP), Flowers, Mertens, and Mulhall (1999) concluded that schools that use teaming have higher student achievement scores on state standardized tests. A longitudinal study of MEAP data revealed that not only was academic achievement higher in schools that used teaming versus schools that did not use teaming, the more years that schools had implemented a structure of teaming, the higher achievement of the schools. The researchers determined higher achievement by looking at the percent of students that met or exceeded expectations on the MEAP in the selected middle schools, and comparing the rate of passing with the self-reported levels and years of teaming. When looking at the average percent of meeting on schools that utilize teaming and those that do not the authors reported a 6 -7 percent increase in schools that use teaming in reading, and a 2 – 4 percent increase in mathematics, over the two year period used in this study. The authors did not report the p-values, or effect sizes.

Wilcox and Angelis (2012) discuss another highly correlated component of the middle grades, which is a shared vision. The study, designed to recognize practices that impact student achievement while also considering defining characteristics of a school, was conducted with a mixed methods approach, using regression analysis to choose
schools and a qualitative multiple case study method to make connections between the academic achievement of students and the demographics of the school. Based on the results of the qualitative case studies, Wilcox and Angelis (2012) conclude the element that seemed to be most associated with high academic achievement was the culture of the school. Based on the interview data, high performing middle schools had a pervasive vision for high academic achievement. This pervasive vision spreads through administration, the teachers, and even the district.

High-Stakes Testing

The introduction of the No Child Left Behind Act (2001) required more high-stakes testing in schools across the nation. With this increase in testing, more data was available for researchers to examine the middle grades on a broad scale in an effort to determine if the schools implementing the middle grades philosophy were having the success that advocates in the field proclaimed. Several important, although outdated, studies exist. The studies of *Focus on the Wonder Years: Challenges Facing the American Middle School* and *America's Middle Schools in the New Century: Status and Progress* are two of the most notable large scales studies. While these studies are broad and examine the middle grades on a full-scale perspective, they each have a small focus on academic achievement in middle grades. Each of these studies makes mention that the attention to academic achievement is linked to the No Child Left Behind Act. This section looks at each of these foundational studies, but focuses primarily on the portions concerning academic achievement, as these elements best inform this study.
The RAND Study: The Wonder Years. Published by the RAND Corporation, the literature review conducted by Juvonen et al. (2004) included 20 years of middle grades research and national and international data. Attempting to “paint a comprehensive picture of teaching and learning” in the middle grades, Juvonen et al. (2004) use relevant research and data on the multiple components of the middle grades. Whereas Juvonen et al. (2004) investigated multiple aspects of the middle school (i.e. its history, its theoretical background, the young adolescent, middle school teachers, middle school leadership, and middle school reform), this review looks specifically at the findings on the academic achievement of middle grades students.

Recognizing the often-criticized academic performance of middle grades students, Juvonen et al. (2004) made an effort to consider the accuracy of such allegations by making international comparisons of academic performance on the TIMSS, and TIMMS-R. After examining TIMMS fourth and eighth grade data, Juvonen et al. (2004) determined that American fourth grade students achieved equally with their international peers in math and science, but a decisive achievement gap appears by eighth grade. This data is troubling as it suggests that something happens in the middle grades to create this gap. Juvonen et al. (2004) added that students who experience a poor beginning in middle school will continue to fall behind.

While an international view offers a global perspective, Juvonen et al. (2004) also used the National Assessment of Educational Progress (NAEP) to look at national trends since the introduction of the middle grades, which revealed more favorable academic trends for the middle grades. Juvonen et al. (2004) noted that in math and reading “scores
have generally been increasing since the mid-1980s, and students in 1999 scored significantly better than did students in 1978” (p. 34). Even with reading scores staying stagnant between 1980 and 1999, “13-year-olds in 1999 nevertheless performed significantly better than did 13-year-olds in 1971” (Juvonen et al., 2004, p. 31).

Recognizing the influence of No Child Left Behind (2001), Juvonen et al. (2004) also looked at subgroup trends and data. In the middle grades, the researchers noted that poor performance on standardized tests was not limited to students in low-achieving demographic groups, for half of the students regarded as high economic status failed to meet the prescribed standards.

In summarizing their work, Juvonen et al. (2004) concluded that academic achievement in the middle level is “positive but uneven” (p. 113). The researchers acknowledged a clear need for improvement in academics, but also discussed the lens of achievement as narrow and suggested that the middle grades philosophy has more to offer than academic improvements. They described the work on academic achievement as weak in the sense that little evidence is available and many of the studies generalize evidence. Juvonen et al. (2004) suggested further research in the area of academic achievement in the middle grades.

America’s Middle Schools in the New Century: Status and Progress. McEwin, Dickinson, and Jenkins acknowledge three comprehensive studies conducted prior to their own 2004 study. The first study took place in 1968, followed by one in 1988, and another in 1996. While the 2001 study followed similar methodologies utilized in the
earlier studies, the McEwin et al. study was unique in that it occurred after the implementation of the No Child Left Behind act.

The 2004 McEwin et al. study is a quantitative study collecting surveys from 746 middle schools across the nation in the spring and fall of 2001. The survey questions were not only analyzed, but they were also compared to the similar questions in other studies, particularly the 1968 study (Alexander, 1968), the 1988 study (Alexander & McEwin, 1988), and the 1996 study (McEwin, Dickinson, & Jenkins, 1996). As a result of NCLB and the accompanying increase in high-stakes testing, McEwin et al. included additional questions to gauge the perspectives of teachers and administrators on how the implementation of NCLB changed instruction in the middle grades.

Throughout the study, the McEwin et al. (2004) discuss middle school components. They include the following as components that are unique in a middle school: advisory programs, curriculum, electives, instructional delivery, instructional grouping, remediation, scheduling, school climate, student learning, teacher planning time, and teaming. A majority of survey participants responded that the introduction of high-stakes testing shaped the implementation of the middle grades components at their schools either positively or negatively. In regards to positive influence, the researchers noted a majority of participants reported positive impact in the areas of curriculum and remediation, student learning and instructional delivery, and instructional grouping practices. One-third or more also reported positive impact on teaming (48%), school climate (45%), scheduling (42%), and electives (33%)” (p. 38). The same survey reported
that 41% of participants thought that high-stakes testing had a negative impact on school climate (41%), and electives (31%).

Overall, participants revealed spending increased time on remediation and instruction because of high-stakes testing (McEwin et al., 2004). With this, some participants also noted the loss of time for electives, advisory, and teacher planning. Interpretation of these results could support the fact that components considered foundational to the middle grades are slowly disappearing as high-stakes testing becomes more prevalent.

Subgroup Trends in Academic Achievement

The No Child Left Behind Act (NCLB) of 2001 is a piece of legislation that recognizes a growing academic achievement gap among students in different ethnic groups and at different socio-economic levels. Part of the purpose statement of NCLB (2001) Section.1001. 3 specifically reads, “Closing the achievement gap between high- and low-performing children, especially the achievement gaps between minority and nonminority students, and between disadvantaged children and their more advantaged peers”. As a proposed solution to close the gap, NCLB requires high levels of accountability by requiring schools and districts to track pertinent data and make comparisons about the achievement of all students. Tracked data include achievement according to socioeconomic status, ethnicity, special education, and English as a Second Language learning.

Poverty. Beane and Lipka (2006) acknowledge that poverty is “the single greatest correlate and predictor of school success” (p. 1). Studies on middle school achievement
support the claim that the students lagging behind academically are largely from areas with high-poverty (Balfanz, 2009; Balfanz & Byrnes, 2006; Wenglinsky, 2004). Balfanz and Byrnes (2006) warn that if more is not done in middle school to address the growing gaps between advantaged and disadvantaged schools, than middle schools will become “dropout factories rather than stepping-stones to a strong education” (p. 144).

A 2009 John Hopkins University study conducted by the Everyone Graduates Center and Talent Development acknowledges that the middle grades are the “launching pad” for not only high school education, but also for career and college readiness (Balfanz, 2009). The study recognized that middle grades achievement gaps were much wider in schools with higher poverty populations. This finding prompted an investigation of variables that enabled a close in the achievement gap, and the authors found that attendance, behavior, and effort all play a role on the academic achievement of high poverty students in middle school. The conclusion drawn from the study was that students in high poverty areas need a good education in middle school to increase their success in high school and beyond.

Sirin (2005) conducted a meta-analysis of socioeconomic status (SES) and academic achievement using research articles from 1990-2000. In this analysis, Sirin noted a mean correlation of .29 (SD = .19) between academic achievement and SES. Based on these results, he concluded that there was a medium effect size on poverty and academic achievement. From the data available in the meta-analysis, Sirin (2005) established the existence of correlation between academic achievement and SES for the
student and the school level. He determined that 60% of the variance in standardized tests can be accounted for by the poverty variable.

Achievement gap. Analysis of National Assessment of Educational Progress (NAEP) scores in both reading and mathematics show a statistically significant academic achievement gap between White students and Black students, as well as between White students and Hispanic students (Hemphill, Vanneman, & Rahman, 2010; Vanneman, Hamilton, Anderson, & Rahman, 2009). When looking at the gap between Black and White students, results from the 2007 administration suggests White students score an average of 26 points higher (on a 0-500 scale) than their Black peers in all subject areas. The 2009 data show a 21 to 26 point gap between Hispanic and White students. This analysis is pivotal as it recognizes that the gap is narrowing; however, it also recognizes that a gap exists and has existed since far before the creation of NAEP.

Olszewski-Kubilius, Lee, and Ngor (2004) designate the academic achievement gap between White students and minority students as the most considerable issue in education in the United States. Williams (2011) also recognizes this problem, pointing out that a larger gap appears in math and science achievement between minority and non-minority students. Williams rejects the notion that poverty is solely to blame for the gap and calls for schools to use research-based interventions with minority students. His analysis of the available research identified that some school-based changes were successful in closing the gap (Williams, 2011). These included engaging, standards-based instructional strategies; heterogeneous grouping; a clear vision; focus on high achievement; and small class sizes—factors that have all been successful in decreasing
the achievement gap between minority students and white students. While Williams' (2011) investigation included elementary, middle, and high school levels, the successful interventions he identified are all associated with components of the middle school.

School size. The population of a school is a debatable topic when it is being correlated with academic achievement. Some researchers, such as Weiss, Carolan, and Baker-Smith (2010), make the argument that school size correlates with school climate and interpersonal relationships, which play a huge role in education, particularly for adolescents. In the 2010 quantitative study, the authors concluded that academic achievement could not be linked directly to school size, but student engagement was significantly impacted by the amount of students in the school. Increased engagement is a substantial finding, as research studies support the role of engagement in academic achievement.

Similar Research Studies

While research exists on academic achievement in middle school and on the individual best practices in the middle school, few research studies attempt to link model middle schools, like Schools to Watch©, to increased academic achievement. Cook, Faulkner, and Kinne (2009) wrote one report that details the study of the Kentucky Schools to Watch© program and the link between these schools and academic achievement. The researchers offered a comparison of academic achievement of middle grades students in Schools to Watch© schools on state standardized test compared to students in middle schools not designated Schools to Watch© in the state of Kentucky. Using This We Believe as the conceptual framework, the authors attempted to identify the
level of implementation of the components of the middle level philosophy described by AMLE and the level of academic achievement in the Schools to Watch© school and Non Schools To Watch schools. One purpose of their study was to look at the relationship between academic achievement and Schools to Watch© designation in order to suggest that STW© schools and the STW© rubric could act as a model for all middle schools, particularly struggling middle schools. In their review of the literature, Cook et al. (2009) recognized that very little study on the topic of Schools to Watch© schools exists. They point to studies that investigated academic achievement and the middle grades, but no other studies that specifically examined Schools to Watch© schools.

Cook et al. (2009) defined the middle school as any school that has a seventh grade. Using a stratified random sampling of the 344 schools in Kentucky with a seventh grade, the researchers randomly chose 40 schools and purposefully selected all 10 of the Kentucky Schools to Watch© schools. Following the identification of schools, the researchers administered surveys asking the participants about the implementation of certain middle school practices in the middle school where they worked. Separately, the authors used data available on the Kentucky Department of Education website to examine student achievement in math, reading, and writing.

In order to make a comparison of academic achievement in STW© schools and Non STW schools, Cook et al. (2009) used the overall rank, both numeric and percentile, of the Schools to Watch© schools compared to the state results. Overall, the Schools to Watch© middle schools scored 77.7% better in mathematics, 75.7% better in writing, and 78.1% better in reading. The authors reported the mean and standard deviation for the
numeric rank (1-277) of the Schools to Watch schools for each subject area. In reading, $M = 47.9$, $SD = 36.1$, in math $M = 49.5$, $SD = 41.6$, and in writing $M = 53.9$, $SD = 38.7$.

The researchers did not report an effect size, or a level of significance. Cook et al. (2009) concluded that the assessment data made the case that in Kentucky, students in the Schools to Watch© schools outperformed students in the Non Schools to Watch schools. Despite these findings, Cook et al. (2009) did not feel that their results created a causal link between being a Schools to Watch© school and improved academic achievement. Recognizing a need for continued research in this area, they recommended that forthcoming studies investigate schools in many states and conduct greater in-depth analysis of achievement data to allow for connections that are more robust.

Summary

The roots of the middle school reach back to the 1950s. With the direction of middle school advocates such as William Alexander and John Lounsbery, the middle school has experienced success in reforming adolescent learning and establishing a research base and research groups to advocate for the young adolescent. The Association of Middle Level Education, the National Forum to Accelerate Middle-Grades Reform, the Carnegie Institute, and the Turning Points Foundation are among the key advocacy groups in the middle grades field. The intention of these groups, and of other middle grades reformers, is to develop middle school curricula and best practices that are developmentally appropriate according to the unique physical, emotional, cognitive, and social needs of the learners.
Although proponents of middle grades reform often highlight the success of middle school to meet the emotional and social needs of the adolescent, critics of the movement often cite this as the downfall of middle school. Critics claim that the middle school concept is not conducive to intellectual development and academic achievement. They support their claims with studies linking components of the middle grades with failing schools or schools unable to make the academic improvements seen in other schooling models.

Despite the growing body of research linking the academic achievement and the middle level elements, statistical data is still lacking in this area, for many of the studies are qualitative or mixed method in nature. More research needs to be done to determine the link between the particular components of the middle school and academic achievement. Without such critical research, the middle school is in danger of reverting to old structures seen in the days of the junior high school in an attempt to keep up with the academic pressure created by No Child Left Behind and the publicizing of Adequate Yearly Progress. The Common Core initiative attempts to create a climate more conducive to standardization and prescribed learning as the nation begins to align standards and increase competition across the nation. This type of climate is the antithesis of the tenets recommended by supporters of middle grades education.

Erb (2005) recognizes that “the middle school concept remains unequalled as the most potent factor for improving the performance of young adolescents” (p. 10). The purpose of this study was to provide a link between academic achievement and the full implementation of the middle level model. The literature review highlighted previous
research in this vein, as well as a history of the middle school and legislation that impacts the middle grades.

In Chapter 3, I discuss in detail the quantitative approach used to answer the research question: Is there a statistically significant difference in the academic achievement of middle school students in a Schools to Watch© middle school and similar traditional middle schools? In Chapter 3, I also explain propensity score matching and the criteria used to make comparisons across schools in this study.
CHAPTER 3

METHODOLOGY

Literature exists on the components of the middle level philosophy and its connection to adolescent social and emotional development. Less research is available on the role that this model plays in academic achievement. Since legislation continues to place a large emphasis on academic achievement, schools consistently seek proven practices that increase student academic achievement.

This research adds to the current body of literature on the middle school philosophy by investigating the link between academic achievement and the implementation of the middle level philosophy. Without the support of academic achievement, school personnel who have been successful in putting into place the tenets of middle school feel pressured to make changes even if it means that the changes no longer reflect middle school characteristics. Studies like this quantitative study are necessary in order to provide that empirical link.

This study sought to answer the question: Is there a statistically significant overall difference in the academic achievement of middle school students in a Schools to Watch© middle school and matched traditional middle schools? It also considered the role that the state plays in the academic achievement of middle school students in Schools to Watch schools and Non Schools to Watch schools. Propensity score matching was used to match Schools to Watch© middle schools to middle schools in the same state.
that are not identified Schools to Watch©, but have similar concentrations of minority students, disadvantaged students, and school size, variables suggested by research as influential in student academic achievement (Balfanz, 2009; Hemphill et al., 2010; Sirin, 2005; Vanneman et al., 2009; Weiss, Carolan, & Baker-Smith, 2010).

As explored in Chapter 2, researchers have been successful in describing the advantages of the middle level movement (AMLE, 2010b). Research is available to support each tenet of the movement and its appropriateness of the model for the adolescent learner. What is still missing is a comprehensive piece looking at all of the tenets as one and linking them specifically to academic achievement and not adolescent development.

This chapter provides details on the methodological design used to speak to the research question and the research hypothesis presented in Chapter 1. It also gives a description of the research design with the measures and covariates. The chapter ends with a discussion of the data collection procedures, the method used for data analysis, ethical safeguards, and the role of the researcher.

Research Questions and Hypotheses

There are three questions and three hypotheses for this study:

R₁ Is there an overall multivariate statistically significant difference in mathematics and language arts/reading achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools (middle schools not identified as Schools to Watch©) when controlling
for percentage of disadvantaged students, percentage of minority students, and school population?

H₁ There will be an overall multivariate statistically significant difference in mathematics and reading achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools (middle schools not identified as Schools to Watch©) when controlling for percentage of disadvantaged students, percentage of minority students, and school population.

H₀ There will not be an overall multivariate statistically significant difference in mathematics and reading achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools (middle schools not identified as Schools to Watch©) when controlling for percentage of disadvantaged students, percentage of minority students, and school population.

R₂ Is there an overall univariate statistically significant difference in the language arts/reading academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools when controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states?

H₁ There will be an overall univariate statistically significant difference in the language arts/reading academic achievement of adolescent students in Schools to
Watch© middle schools and adolescent students in matched traditional middle schools with controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states.

**H₀** There will not be an overall univariate statistically significant difference in the language arts/reading academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools with controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states.

**R₃** Is there an overall univariate statistically significant difference in the reading academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools when controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states?

**H₂** There will be an overall univariate statistically significant difference in the math academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools with controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states.
\( H_0 \) There will not be an overall univariate statistically significant difference in the math academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools with controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states.

Population and Sample

The Mercer University Institutional Review Board (IRB) granted their permission to conduct this study (see Appendix A). While a study of all 352 schools would be ideal, the timeframe and resources of this study did not allow for such a wide scale examination since, nationally, there are 352 Schools to Watch© schools in 17 states. Table 4 lists the 17 participating states, along with the number of schools recognized as Schools to Watch© in that state, the region they belong in, the status of use of Common Core, and the year of beginning participation in the Schools to Watch© program. This information served as the basis for deciding which schools to include in the study.
Table 4

*Nationwide Schools to Watch© Middle Schools*

<table>
<thead>
<tr>
<th>State</th>
<th># of STWO Schools</th>
<th>Year Joined STWO</th>
<th>Region</th>
<th>Common Core</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illinois</td>
<td>24</td>
<td>2003</td>
<td>Midwest</td>
<td>Adopted</td>
</tr>
<tr>
<td>Ohio</td>
<td>20</td>
<td>2004</td>
<td>Midwest</td>
<td>Adopted</td>
</tr>
<tr>
<td>Michigan</td>
<td>12</td>
<td>2005</td>
<td>Midwest</td>
<td>Adopted</td>
</tr>
<tr>
<td>Indiana</td>
<td>11</td>
<td>2004</td>
<td>Northeast</td>
<td>Adopted</td>
</tr>
<tr>
<td>New York</td>
<td>25</td>
<td>2004</td>
<td>Northeast</td>
<td>Adopted</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>21</td>
<td>2006</td>
<td>Northeast</td>
<td>Adopted</td>
</tr>
<tr>
<td>New Jersey</td>
<td>8</td>
<td>2007</td>
<td>Northeast</td>
<td>Adopted</td>
</tr>
<tr>
<td>Georgia</td>
<td>20</td>
<td>2002</td>
<td>South</td>
<td>Adopted</td>
</tr>
<tr>
<td>North Carolina</td>
<td>33</td>
<td>2002</td>
<td>South</td>
<td>Adopted</td>
</tr>
<tr>
<td>Kentucky</td>
<td>13</td>
<td>2003</td>
<td>South</td>
<td>Adopted</td>
</tr>
<tr>
<td>Arkansas</td>
<td>10</td>
<td>2005</td>
<td>South</td>
<td>Adopted</td>
</tr>
<tr>
<td>Florida</td>
<td>0</td>
<td>2007</td>
<td>South</td>
<td>Adopted</td>
</tr>
<tr>
<td>Texas</td>
<td>17</td>
<td>2010</td>
<td>South</td>
<td>Not Adopted</td>
</tr>
<tr>
<td>Virginia</td>
<td>28</td>
<td>2003</td>
<td>South</td>
<td>Not Adopted</td>
</tr>
<tr>
<td>California</td>
<td>36</td>
<td>2002</td>
<td>West</td>
<td>Adopted</td>
</tr>
<tr>
<td>Colorado</td>
<td>16</td>
<td>2003</td>
<td>West</td>
<td>Adopted</td>
</tr>
<tr>
<td>Utah</td>
<td>10</td>
<td>2006</td>
<td>West</td>
<td>Adopted</td>
</tr>
<tr>
<td>Oregon</td>
<td>0</td>
<td>2007</td>
<td>West</td>
<td>Adopted</td>
</tr>
</tbody>
</table>
Using GPower software, a power analysis determined that 42 schools were needed to yield a medium effect size, and 602 schools were necessary for a small effect size. An effect size of .25 (.10 = small, .25 = medium, .40 = large), an alpha of .05, and an actual power of .80 served as the basis for the calculation of the number of schools needed. This study included 80 Schools to Watch© schools and 80 matched schools, 160 schools total. This allowed for a medium effect size.

States excluded from this study included Florida, Oregon, Texas, and Virginia. At the time of this study, no Florida and Oregon schools participated in the STW© program, and Texas and Virginia did not use the Common Core Initiative to drive their instruction. Use of Common Core was a criterion for the study as the Common Core curriculum seeks to standardize curriculum from state to state. Using states that adopted Common Core facilitated generalization of data, as the intention of the Common Core is to standardize curriculum across the nation. However, it is important to note, that as of the 2013 administration, not all of the participating states administered an assessment based solely on the Common Core standards. Of the remaining states, the mean number of Schools to Watch© in each state was determined to be 18. The state in each region closest to this average was used in the study. These states include Ohio, Pennsylvania, Georgia, and Colorado.

Access to demographic and testing data was necessary in order to run the needed tests. Information from 2013 was not available from Pennsylvania or Georgia at the time of this study. For this reason, New York served as the Northeast state, and Virginia served as the South state.
Research Design

A quantitative correlational research design was used to determine if Schools to Watch© schools have increased academic achievement on criterion standards-based tests, over schools not labeled as Schools to Watch©. The choice of a quantitative design also fulfilled the purpose of this study, which was to add empirical research on the current body of literature on the effectiveness of the middle school model. Field (2009) describes correlational research as "research that observes what naturally happens, without manipulating some aspect of the environment" (p. 12). For this study, I used publically available data concerning state academic achievement tests, so no manipulation of a variable occurred. To investigate the influence of the middle school model (used in STW© schools) on academic achievement, I retrieved data from state websites and the National Forum to Accelerate Middle-Grades Reform (2013).

In order to ensure comparisons of similarly matched schools, the criteria of percentage of minority students, percentage of disadvantaged students, and school population established propensity score matching. The percentage of minority students included the percentage of all students in the school not labeled White. The percentage of disadvantaged students derived from the number of students receiving free and reduced lunch. Total school population equaled the number of students served at the school. After matching, I used MANCOVA to determine the existence of a correlation between academic achievement and a Schools to Watch© label. Use of linear mixed modeling (LMM) determined the existence of a correlation between academic achievement and a Schools to Watch© label, when talking into account clustering of schools within states.
Measures and Covariates

Predictor Variables

The variables of the first part of the research design consisted of selected school demographic characteristics: disadvantaged students, minority students, and school size. For the second and third part of the research design (the MANCOVA and LMM), Schools to Watch© status (participating or not participating) served as the primary predictor. Appointment of a school as Schools to Watch© depends on meeting of expectations on the Schools to Watch© rubric compiled by the National Forum to Accelerate Middle-Grades Reform (2010).

Outcome Variables

Passing percentages on the eighth grade reading and math state criterion-referenced test served as the outcome variable for this study. While each state reports levels of proficiency at their own discretion, percentages are reported for students who were proficient and not proficient. In some cases, information is also available for students who exceeded proficiency or met other levels of proficiency. No Child Left Behind calls for 100% of students to meet or exceed expectations on the state test. Percentages of students who pass, do not pass, and exceed are available publicly on state websites and the Schools report card. These percentages serve as part of the criteria for Annual Yearly Progress (NCLB, 2001).

Instrumentation

No Child Left Behind requires each state to administer yearly assessments in grades three through eight in the subject areas of math, reading, language arts, math, and
science. States are not required to give the same test; therefore, each state develops their test. While each test is different, NCLB (2001) requires creation of the test by content and assessment experts to ensure reliability and validity. This study used data from the Virginia Standards of Learning (SOL), the Ohio Achievement Assessment (OAA), New York State Testing Program, and the Colorado Student Assessment Program. Each state used a scaled score. In order to compare the states, this study used the percentage of students who passed each test. I obtained data from each state's board of education online website. Table 5 displays the reliability and validity of the states' standardized tests. It should be noted Virginia reported the Cronbach's alpha for each of their three forms, and did not provide a composite. I averaged these to find the alpha for the overall data set. Virginia did not provide a Standard Error of Measurement for the overall data set.

Table 5

*Standardized Test Reliability and Validity Statistics by State*

<table>
<thead>
<tr>
<th>State</th>
<th>Content Area</th>
<th>Cronbach's Alpha</th>
<th>Standard Error of Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virginia</td>
<td>Mathematics</td>
<td>.88</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Reading</td>
<td>.89</td>
<td>--</td>
</tr>
<tr>
<td>Ohio</td>
<td>Mathematics</td>
<td>.91</td>
<td>9.71</td>
</tr>
<tr>
<td></td>
<td>Reading</td>
<td>.87</td>
<td>10.61</td>
</tr>
<tr>
<td>New York</td>
<td>Mathematics</td>
<td>.85</td>
<td>2.51</td>
</tr>
<tr>
<td></td>
<td>English/language arts</td>
<td>.85</td>
<td>2.26</td>
</tr>
<tr>
<td>Colorado</td>
<td>Mathematics</td>
<td>.94</td>
<td>16.35</td>
</tr>
<tr>
<td></td>
<td>Reading</td>
<td>.93</td>
<td>17.41</td>
</tr>
</tbody>
</table>
The Schools to Watch© Self Study and Rating Rubric (Appendix B) is the tool used by each state organization to certify schools. The National Forum to Accelerate Middle-Grades Reform reports that schools should strive to reach the level 4 domain on the rubric consistently (National Forum to Accelerate Middle-Grades Reform, 2010). A search on the National Forum to Accelerate Middle-Grades Reform website and EBSCO revealed no information about the construction, validity, or reliability of the rubric. There is also no information about the fidelity of the scoring from state to state.

Data Collection Procedures

For this study, preexisting data derived from the online websites of the state boards of education of Virginia, Ohio, New York, and Colorado. Demographic data included percentage of students on free and reduced lunch, ethnicity information, and school size information. The National Forum to Accelerate Middle-Grades Reform website provided information designating which schools are Schools to Watch© schools. Additionally, test scores for the middle school standardized achievement test given in each state were collected. Each state administered a different version of the test based on their state’s standards, typically between April and May. For purposes of this study, data derived from the eighth grade scores on math and reading for the 2013 administration of the state test.

Demographic data (including socio-economics, ethnicity, and school size) were obtained from state websites. In all cases, the data, available in an Excel file, included demographic information for all of the schools in the state. In order to sift out the schools
that did not meet the criteria of my study, I filtered the data set to include schools containing either the word "middle" or the word "junior". This was an attempt to narrow the data set to all of the middle schools and junior high schools in the state. Data from Excel were transferred to Matchit for propensity score matching. Using nearest-neighbor matching, I chose the experimental group from the propensity scores.

After choosing the schools for this study, I accessed academic data available on each state’s website in an Excel format. Search of the data set involved identifying the participating schools to get the needed information. I transferred records from Excel Data files into SPSS for statistical analysis. First analysis of data used a MANCOVA test will, followed by a Linear Mixed Models (LMM) test.

Data Analysis

I began the data analysis by first collecting the data and running the propensity score test in R. The results from the propensity score gave me a data set of 160 schools. I then used SPSS to run a MANCOVA test to analyze the first research question and a linear mixed models test to answer questions two and three.

Propensity Score Matching

For the purposes of this study, propensity scores represented the probability of each middle school in the state being a Schools to Watch© school, while taking into account the factors of percentage of minority students, percentage of disadvantaged students, and school size. The intention of matching the schools was to ensure that any differences in academic achievement were directly related to their status as a Schools to Watch© school, and not a result of other demographic characteristics. Ho (2007) claims
that "Matchit implements a wide range of sophisticated matching methods, making it possible to greatly reduce the dependence of causal inferences in hard-to-justify, but commonly make, statistical modeling assumptions" (para. 2).

Performing propensity score matching requires several steps (Ho, 2007). The initial stages require gathering the data to run the test. In the case of this study, I visited each state website and downloaded information on the demographics of the school. I sifted my data to include only schools with the name "middle" and "junior" in their title. I then deleted unnecessary information, keeping only the percentage of minority students, the percentage of economically disadvantaged students, and the total population of the school. In some cases, these percentages had to be determined by adding and subtracting the information available in the spreadsheet. For purposes of this study, minority students were defined as students who were not White, and economically disadvantaged was defined as the students who qualified for free and reduced lunch.

When the data sets were ready, I prepared to match the data using the R program. In R, I used the code, available in Appendix C. The code allows for a one-to-one exact match (also known as nearest-neighbor matching), what Ho (2007) describes as the best type of match.

Variable Matching

For the purposes of this study, propensity scores represented the probability of being a Schools to Watch© school given a set of covariates. In this case, the covariates include disadvantaged students, minority students, and school size. Disadvantaged students were identified by the information provided on students who received free and
reduced lunch. Percentage of minority students derived from ethnicity information. In this study, the minority percentage represented all non-white students. The number of students attending the school served as the sole basis for defining school size. The decision to use this information, available on state websites, was based on previous research that suggested that these factors play a role in academic achievement.

Data were analyzed using the Statistical Program for Social Science Program (SPSS). Using SPSS, a MANCOVA test was run. Data needed to meet several assumptions in order to run the test. According to Field (2009), the assumptions for MANCOVA include independence, random sampling, multivariate normality, and homogeneity of covariance matrices. This assumes that observations are independent, data are sampled randomly, each outcome has multivariate normality within groups, variances for each group are roughly equal, and the correlation between outcome variables is roughly the same for each group.

Ethical Safeguards

An IRB approval, located in Appendix A, was obtained from Mercer University before compiling data from this study. All the information collected in this study is public data available through each state’s website. No humans were at risk in this study at any time.

Role of the Researcher

The researcher was a middle school teacher and a middle school advocate. She attended a teacher preparation program that emphasized the middle level philosophy, which played a large role in her understanding of teaching and learning. At the time of
this study, the researcher sat on the board of directors for the Georgia Middle School Association (GMSA). She was also a member of the Association for Middle Level Educators.

Summary

This study used a quantitative correlational research design to determine if using middle school best practices is correlated with academic achievement, by taking into account the academic achievement of young adolescent students in Schools to Watch© schools and Non Schools to Watch schools. A medium effect size of 80 Schools to Watch© schools and 80 matched schools were selected. Information about school demographics, along with scores in Reading and Math from standardized testing on the Georgia Criterion Referenced Competency Test, the Ohio Achievement Assessment, the Colorado Student Assessment Program, and the Pennsylvania System of School Assessment were used to determine correlations. Data analysis involved the process of propensity score matching, along with MANCOVA and LMM.
CHAPTER 4
RESULTS OF THE STUDY

The purpose of this research study was to use existing state data to provide quantitative evidence that the middle school philosophy is not only important in building the social and emotional development of young adolescents, but also that it is a worthwhile model to enhance academic achievement. As discussed in detail in Chapter 3, I collected data on the middle schools in Ohio, Colorado, New York, and Virginia. The results of this study are based on results of the 2013 criterion-based assessment tests given in each of the states included in this study, as well as the demographic information of each school published on the individual state board of education website. This chapter explains the findings of the study and the results of the research questions:

R₁ Will there be an overall multivariate statistically significant difference in mathematics and reading achievement in adolescent students in Schools to Watch© Middle Schools, and adolescent students in matched traditional middle schools (middle schools not identified as Schools to Watch©) when controlling for percentage of disadvantage students, percentage of minority student, and school population?

R₂ Will there be an overall univariate statistically significant difference in the reading academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools
when controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states?

R₃ Will there be an overall univariate statistically significant difference in the mathematics academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools when controlling for percentage of disadvantaged students, percentage of minority students, and school population, and when taking into account clustering of schools within states?

Schools in Colorado, Ohio, New York, and Virginia, identified as Schools to Watch© schools, comprised the control group. Propensity score matching was conducted using SPSS in order to find a sample that matched the control group in each state. This chapter first discusses the results of the propensity score matching. Second, the assumptions and findings from the MANCOVA are presented, which address research question one regarding the multivariate difference in mathematics and reading achievement of students in Schools to Watch© schools and matched traditional middle schools. Next, the assumptions and results of linear mixed modeling (LMM) are revealed in order to answer research questions three and four, regarding the overall multivariate difference in academic achievement when taking into account clustering of schools within states. The chapter concludes with a summary of the results of the study.
Propensity Score Matching

Results of Propensity Score Matching

Propensity score matching allowed matching of the samples on the demographic data of school population, percentage of disadvantaged students, and percentage of minority students. Tables 6-9 show the demographics for each state before and after matching. The mean difference represented in the last column helps to demonstrate the power of propensity score matching in making accurate comparisons. Austin (2011) explains that propensity score matching is used for balancing the covariates. The mean difference column shows that after matching, balancing of the comparisons increased than when taking the mean before matching. The histograms (Figures 3-6) also illustrate the balance of the model after matching for Colorado, New York, Ohio, and Virginia.

Table 6

*Colorado Demographic Information: Before and After Matching*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean for Schools to Watch©</th>
<th>Mean for Non Schools to Watch</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Population</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before matching</td>
<td>662.00</td>
<td>453.75</td>
<td>208.25</td>
</tr>
<tr>
<td>After matching</td>
<td>662.00</td>
<td>647.14</td>
<td>14.86</td>
</tr>
<tr>
<td>% of Minority</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before matching</td>
<td>38.00</td>
<td>38.00</td>
<td>0.00</td>
</tr>
<tr>
<td>After matching</td>
<td>38.00</td>
<td>42.00</td>
<td>4.00</td>
</tr>
<tr>
<td>% of Disadvantaged</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before matching</td>
<td>40.00</td>
<td>47.00</td>
<td>7.00</td>
</tr>
<tr>
<td>After matching</td>
<td>40.00</td>
<td>39.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Figure 3. Histograms of Colorado Propensity Scores: Before and After Matching

Table 7

New York Demographic Information: Before and After Matching

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean for Schools to Watch©</th>
<th>Mean for Non Schools to Watch</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Population</td>
<td>832.64</td>
<td>568.90</td>
<td>263.74</td>
</tr>
<tr>
<td>Before matching</td>
<td>832.64</td>
<td>830.64</td>
<td>2.00</td>
</tr>
<tr>
<td>After matching</td>
<td>17.00</td>
<td>28.00</td>
<td>11.00</td>
</tr>
<tr>
<td>% of Minority</td>
<td>17.00</td>
<td>18.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Before matching</td>
<td>18.00</td>
<td>41.00</td>
<td>22.00</td>
</tr>
<tr>
<td>After matching</td>
<td>18.00</td>
<td>18.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Figure 4. Histograms of New York Propensity Scores: Before and After Matching

Table 8

Ohio Demographic Information: Before and After Matching

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean for Schools to Watch</th>
<th>Mean for Non Schools to Watch</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Population</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before matching</td>
<td>608.64</td>
<td>516.85</td>
<td>91.79</td>
</tr>
<tr>
<td>After matching</td>
<td>608.64</td>
<td>566.14</td>
<td>42.50</td>
</tr>
<tr>
<td>% of Minority</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before matching</td>
<td>14.00</td>
<td>38.00</td>
<td>4.00</td>
</tr>
<tr>
<td>After matching</td>
<td>14.00</td>
<td>42.00</td>
<td>6.00</td>
</tr>
<tr>
<td>% of Disadvantaged</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before matching</td>
<td>23.00</td>
<td>43.00</td>
<td>21.00</td>
</tr>
<tr>
<td>After matching</td>
<td>23.00</td>
<td>23.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Figure 5. Histograms of Ohio Propensity Scores: Before and After Matching

Table 9

Virginia Demographic Information: Before and After Matching

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean for Schools to Watch©</th>
<th>Mean for Non Schools to Watch</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Population</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before matching</td>
<td>954.81</td>
<td>740.82</td>
<td>213.99</td>
</tr>
<tr>
<td>After matching</td>
<td>954.81</td>
<td>973.30</td>
<td>339.38</td>
</tr>
<tr>
<td>% of Minority</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before matching</td>
<td>42.00</td>
<td>43.00</td>
<td>1.00</td>
</tr>
<tr>
<td>After matching</td>
<td>42.00</td>
<td>44.00</td>
<td>2.00</td>
</tr>
<tr>
<td>% of Disadvantaged</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before matching</td>
<td>31.00</td>
<td>43.00</td>
<td>12.00</td>
</tr>
<tr>
<td>After matching</td>
<td>31.00</td>
<td>34.00</td>
<td>2.00</td>
</tr>
</tbody>
</table>
In order to answer the first research question, the data from Colorado, New York, Ohio, and Virginia were aggregated into one data set. As a whole the data set has a school population of $M = 802.60$, $SD = 350.95$, minority population of $M = 25.25\%$, $SD = 0.22$, and a population of disadvantaged students $M = 31.06\%$, $SD = 0.26$.

Research Question 1

Once a data set was established, a multivariate analysis of covariance (MANCOVA) was used to address the first research question: Will there be an overall multivariate difference in mathematics and reading achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle school (middle schools not identified as Schools to Watch©) when controlling for percentage of disadvantage students, percentage of minority students, and school
population? The outcome variable in this test was the percentage of students who met proficiency or who exceeded on the criterion-referenced test. The predictor variable was being certified as a Schools to Watch© school or not.

Assumptions. Field (2009) recognizes four assumptions for MANCOVA: independence, random sampling, multivariate normality, and homogeneity of covariance matrices. The assumption of independence is met as it is assumed that no one school’s scores have any effect on another school’s scores. Results of Levene’s test showed that for the language arts/reading variable and the math variable there was homogeneity of variances; language arts/reading, $F(1, 158) = 1.61, p = .21$; math variable, $F(1, 158) = 1.73, p = .190$. Box’s Test results again provide evidence for heterogeneity of covariance; Box’s $M = 1.93, F(3, 4493520) = 0.63, p = .594$.

The assumptions of multivariate normality and random sampling were not initially met. The assumptions of multivariate normality were not met when considering just the inspection of histograms and the results of the Kolmogorov-Smirnov test. Examination of the histograms showed nonnormality in all groups. The Kolmogorov-Smirnov test, with Lilliefors significance, revealed nonnormality for both groups on the language arts/reading variable; Non Schools to Watch group, $D(80) = .107, p = .024$; Schools to Watch© group, $D(80) = .104, p = .032$. The results also showed nonnormality on the math variable; Non Schools to Watch, $D(80) = .144, p = .000$; Schools to Watch© variable $D(80) = .111, p = .016$. (The $p$ values are in the lower bound of the true significance for the Kolmogorov-Smirnov test.) While the histogram and the Kolmogorov-Smirnov tests failed to show normality, Field (2009) still argues that the
assumption of normality can be assumed when the sample size is greater than 30. The sample size is 160 in this data set. The assumption of random sampling was also not met, as these schools were not picked at random, and rather were matched on the criteria of total population, percentage of disadvantaged students, and percentage of minority students. Again, Field (2009) allows for a violation of assumptions if the sample sizes are equal, which in this case they are: n = 80 (Schools to Watch©) and n=80 (Non Schools to Watch). Figures 7-10 display the histograms for percentage of achievement on Language Arts/Reading and math assessments in Non Schools to Watch Schools and Schools to Watch© Schools.

*Figure 7. Histogram for Percentage of Passing in Language Arts/Reading in Non Schools to Watch Schools*

![Histogram](image-url)
Figure 8. Histogram for Percentage of Passing in Language Arts/Reading in Schools to Watch© Schools

![Histogram for ELA/Reading](image)

**Note.** ** are expressed in decimal form.

Figure 9. Histogram for Percentage of Passing in Math in Non Schools to Watch Schools

![Histogram for Math](image)

**Note.** ** are expressed in decimal form.
Mancova results. MANCOVA analysis, using Pillai’s trace, indicated that there was not a statistically significant difference between the Non Schools to watch and the Schools to Watch© groups: $V = 0.10$, $F(2, 157) = .79$, $p = .457$, $\eta^2 = .01$. The Schools to Watch© schools had a higher percentage of passing in language arts/reading ($M = .72$, 95% CIs [.69, .77]) than the Non Schools to Watch school ($M = .69$, 95% CIs [.64, .74]). Also, there was not a statistically significant univariate difference in language arts/reading: $F(1, 158) = 1.05$, $p = .040$, $\eta^2 = .01$. In reference to Math, the Schools to Watch© schools was once again higher ($M = .67$, 95% CIs [.62, .72]) than the Non Schools to Watch schools ($M = .65$, 95% CIs [.59, .69]), and the univariate difference was not statistically significant: $F(1, 158) = .43$, $p = .515$, $\eta^2 = .003$. Figures 11 and 12 display the MANCOVA results for Schools to Watch© and Non Schools to Watch for the content areas of language arts/reading and math respectively. The y-axis represents the percentage of students who met or exceeded standards on the assessment.
Figure 11. ELA/Reading Results for Non Schools to Watch vs. Schools to Watch

Figure 12. Math Results for Non Schools to Watch vs. Schools to Watch
Linear Mixed Modeling

The linear mixed models (LMM) test is used in educational research in cases of nested models, where data sets are nested in larger data sets. This type of data can be referred to as hierarchical data. Field (2009) explains that in statistical analysis one must account for the effects of such nesting. In this case, the data for Schools to Watch© and Non Schools to Watch is nested in the larger data set, the state. The linear mixed models test takes into account the impact of the state on academic achievement in Schools to Watch© schools and allows comparisons between the states while considering the role of the state effect.

Assumptions. Field (2009) recognizes that the assumptions for linear mixed modeling are the same as they are for regression with the addition of two others. The following is a list of essential assumptions for linear mixed modeling:

- there are not influential data points;
- homoscedasticity should be present;
- the residual error should be distributed normally;
- multicollinearity should not be present;
- intercepts are distributed normally around the average intercept; and
- the slopes are distributed normally around the average slope (Field, 2009, p. 739).

As there were no indicators otherwise, I assumed the criteria of independence of observation and normally distributed random intercepts and slopes. The data were inspected to reveal that both the outcome and predictor were continuous. Neither the
outcome nor predictors had zero variance. A visual examination of the residuals plots and histograms of standardized residuals was completed to check assumptions regarding homoscedasticity and normal distribution (available in Figures 13-16). The examination indicated multiple striations, and some unequal variance, which called into question the assumption regarding homoscedasticity. The histogram of standardized residuals appeared to be normally distributed upon visual inspection; a visual examination of the data set did not show any residuals above 3.00. A regression test was run in SPSS to test multicollinearity. Results suggest that multicollinearity was not a concern, Total Population, Tolerance = .98, VIF = 1.02; Percent of Minority Students, Tolerance = .71, VIF = 1.4; Percent of Disadvantaged Students, Tolerance = .70, VIF = 1.4.

*Figure 13. Language Arts/Reading Scatterplot*
Figure 14. Language Arts/Reading Histogram

Figure 15. Math Scatterplot
Language Arts/Reading Results

The multilevel linear analysis for language arts/reading can be viewed in Table 10. The table reports the percentage of students meeting or exceeding expectations on the assessment. A statistically significant difference was noted when including the state variable as a random intercept, in the case of controlling for state effect between those schools that are Schools to Watch©, and those who are not Schools to Watch. The state effect accounted for 83% of the variance. A statistically significant difference was found in the restricted log likelihoods between models: $\chi^2(1) = 224.63, p < .001$. 
Table 10

Unstandardized Estimates of Fixed Effects Covariance Parameters for Language

Arts/Reading Scores

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Without State Effect</th>
<th>With State Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fixed effects</td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>82.2%*** (.05)</td>
<td>88.2%** (.10)</td>
</tr>
<tr>
<td>Non Schools to Watch</td>
<td>-2.9% (0.03)</td>
<td>-2.8%* (.01)</td>
</tr>
<tr>
<td>Schools to Watch©</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td>4.0%*** (.00)</td>
<td>1.0%*** (.00)</td>
</tr>
<tr>
<td>State (intercept)</td>
<td>--</td>
<td>4.0% (.03)</td>
</tr>
</tbody>
</table>

-2*restricted log likelihood -48.97 -274.60

Note. N=160. Standard errors are in parentheses. Restricted maximum likelihood estimation was used. * p < .05, **p < .01, ***p < .001.

a Reference group

Math Results

The multilevel linear analysis for math can be viewed in Table 11. The table reports values in terms of percentage of students meeting or exceeding the assessment. A statistically significant difference was not noted when including the state variable as a random intercept, in the case of controlling for the state effect on academic achievement in Math between those who are Schools to Watch© school and those who are not. The state effect accounted for 81.00% of the variance. A statistically significant difference was found in the restricted log likelihoods between models: χ²(1) = 212.05, p < .001.
Table 11

*Unstandardized Estimates of Fixed Effects Covariance Parameters for Math Scores*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Without State Effect</th>
<th>With State Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>72.00%*** (0.05)</td>
<td>81.00%* (0.12)</td>
</tr>
<tr>
<td>Non Schools to Watch©</td>
<td>-2.00% (0.04)</td>
<td>-2.00% (.02)</td>
</tr>
<tr>
<td><strong>Random parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td>5.00%*** (0.01)</td>
<td>1.00%*** (0.00)</td>
</tr>
<tr>
<td>State (intercept)</td>
<td>--</td>
<td>5.00% (0.04)</td>
</tr>
<tr>
<td>-2*restricted log likelihood</td>
<td>11.43</td>
<td>-200.62</td>
</tr>
</tbody>
</table>

*Note. N= 160. Standard errors are in parentheses. Restricted maximum likelihood estimation was used. * p < .05, **p < .01, ***p < .001.
© Reference group

Summary

This study had multiple stages. The first stage was collecting and compiling demographic data from state databases, which included school population, minority student percentage, and disadvantaged student percentage. The next stage was using the gathered information to match the schools using the propensity score matching method. Propensity score matching allowed me to have a data set that included an equal number of Schools to Watch© schools and Non Schools to Watch schools in each state that shared similar characteristics, thus reducing the variability that can be introduced by demographic factors. The Schools to Watch© schools were then compared to the Non Schools to Watch schools using eighth grade achievement scores in math and language
arts/reading. Results of the MANCOVA analysis indicated that there was no statistically significant difference in the language arts/reading scores or on the math scores between the Schools to Watch© schools and the Non Schools to Watch schools. Results of the linear mixed models analysis indicated that the state where a school is located plays a significant role in the variability of performance on the achievement test.
CHAPTER 5
DISCUSSION

This chapter begins with a summary of the study, a restatement of its purpose, and a review of the research questions. Following this is a discussion of the interpretations of the results and findings. Also addressed are the limitations and comparisons to other studies. This chapter ends with a discussion of implications for this research and the need to conduct future research on this topic.

Summary of the Study

The purpose of this study was to use existing data to provide quantitative evidence that the middle level philosophy, when fully implemented, has a positive impact on academic achievement. This study evolved out of an observation that some middle schools, in the quest of meeting nationally mandated academic performance measures, have considered abandoning practices that middle school researchers and advocates describe as developmentally responsive. The addition of this quantitative research may provide could aid decision makers with evidence to support developmentally appropriate practices like those supported by the National Forum to Accelerate Middle-Grades Reform and the Association of Middle Level Education.

This study was a national approach looking at four states (one each to represent the North, South, East, and West) in the continental United States. Based on the National Forum to Accelerate Middle-grades reform rubric (Appendix B), Schools to Watch© schools are deemed to be model middle schools based on a thorough review process. This study’s sample was comprised of all of the Schools to Watch© schools in Colorado, New
York, Ohio, Virginia, as well as middle schools in the state with similar demographics in regards to total population, race/ethnicity, and economically disadvantaged students. These schools were found using propensity score matching on these demographic variables in an effort to reduce their impact on the results, since research has shown that these variables play a noticeable role on academic achievement. After matching the data on these points, SPSS was used to run a multivariate analysis of covariance (MANCOVA) test and a linear mixed models test. These statistical tests answered the research questions as follows:

\[R_1\] There is not an overall multivariate statistically significant difference in mathematics and reading achievement in adolescent students in Schools to Watch© Middle Schools, and adolescent students in matched traditional middle schools (middle schools not identified as Schools to Watch©) when controlling for percentage of disadvantaged students, percentage of minority students, and school population.

\[R_2\] There is an overall univariate statistically significant difference in the language arts/reading academic achievement of adolescent students in Schools to Watch© middle schools and adolescent students in matched traditional middle schools when controlling for percentage of disadvantaged students, percentage of minority students, and school population, when taking into account clustering of schools within states.

\[R_3\] There is not an overall univariate statistically significant difference in the mathematics academic achievement of adolescent students in Schools to Watch©
middle schools and adolescent students in matched traditional middle schools when controlling for percentage of disadvantaged students, percentage of minority students, and school population, when taking into account clustering of schools within states.

Interpretation of the Results

The middle school concept is set apart for its attention to adolescent development and practices that support that development. Some middle level advocates claim developmentally appropriate practices lead to increased academic achievement (AMLE, 2010b; Beane, 2006), but few studies support these claims with quantitative evidence. Overall, findings from this study did not suggest a statistically significant difference in academic achievement in language arts/reading and math for Schools to Watch© schools and Non Schools to Watch schools. Results showed an average 2%-3% increase in mean score in the case of Schools to Watch© schools, however this increase was not high enough to be considered to be statistically significant. When taking into account the state effect, a statistically significant difference was found in language arts/reading. However, there was not a statistically significant difference in math when including the state effect. Results from the linear mixed model test suggest that the nesting of schools within states accounts for approximately 83% of the variance in the percentage of students passing the language arts/reading assessment and 81% of the variance in the percentage of students passing the math assessment.
Interpretation of the Findings

A major component of the middle level philosophy is that curriculum, instruction, and assessment should be designed around the development of young adolescents. The assumption is that instruction designed with the young adolescents in mind will help them to make greater academic gains. While some researchers support this notion (Beane, 2006; Lounsbury, 2000; AMLE, 2010a; AMLE 2010b), few studies investigate academic achievement and the implementation of the middle school components as a whole. Several studies examine components of the middle level philosophy, such as the use of teaming, common planning, and shared vision (Arhar, 2001; Flowers, Mertens, & Mulhall, 1999; Merten et al., 2010; Williams et al.)

Overall, the findings of this study did not definitively support the expectation that students in middle schools implementing all of the components of the middle school philosophy attain higher academic achievement. The findings show that Schools to Watch© schools had a 2-3% higher pass rate than schools Non Schools to Watch. This difference is not considered statistically significant except in the case of language arts/reading when taking into account the state effect. While 2%-3% is not considered statistically significant it can be argued that a difference of that amount still impacts a meaningful amount of students. Considering the total population of all of the middle schools and junior high schools in Colorado, New York, Ohio, and Virginia is 1,074,895, a percentage of 2.5% would be approximately 26,872 additional students meeting or exceeding standards on the state achievement test. Another interesting finding of this study is that the state effect played an 81-83% role in the variance.
Addressing the Limitations

Several factors might have affected the results of this study. Possible limiting factors include:

- the use of the Schools to Watch© program to define model middle schools;
- trying to find exact matches for each model middle school;
- problems associated with the new assessment system of the state of New York; and
- different minimal passing expectations in each state.

Schools to Watch©

The Schools to Watch© program is a national program run by the National Forum to Accelerate Middle-Grades Reform. While the program stems from a national program, each state has a director and governing body that appoint the schools. While each state uses the same rubric developed by National Forum to Accelerate Middle-Grades Reform, it is possible that interpretations differ among states. The rubric components are research-based, but have not yet been validated, nor has it been vetted for construct validity.

Matching

Matching was also a possible limitation in my study. Shadish, Cook, and Campbell (2002) write, “Random assignment balances treatments on both observed and unobserved covariates on expectation; but propensity score adjustments balance treatments only on observed covariates, leaving hidden bias due to unobserved covariates” (p. 164). While propensity score matching rendered the best possible match
from schools in the state, other extraneous school factors not considered might have also influenced academic achievement.

Without visiting every school, it was impossible to determine whether each middle school fully implemented the middle school concept. While the criteria of being a Schools to Watch© school can give some support that the middle school concept was, in fact, being implemented in these schools, it is not fair to say that the same criteria as the Schools to Watch© schools was not being followed in the schools labeled Non Schools to Watch. Not having Schools to Watch© status is not an indication of not implementing middle school tenets. It is possible that there is a lack of knowledge or unavailability of the information needed to apply for the STW© program. The application process can also be time and labor intensive. Not all schools have the resources to allocate for this recognition.

New York Assessment System

When running statistical tests, the state of New York showed some different results than the other states. First, New York did not have a separate reading test like the other states. The domains of the New York test were equivalent to the reading domains of Ohio, Colorado, and Virginia, but the New York test just reports a score for "English". Next, when gathering the data for New York, it became clear that in 2013 New York was the only state in this study that switched their assessment to align more closely to the PARCC assessment. The PARCC is defined by Castlehano (2013) as a group of states that has the goal of creating assessments that align with the Common Core standards. The same author explains that the goal of PARCC is to create assessments for grades K-12,
which are linked to college and career readiness, align to the Common Core state
standards, use more technology in assessment, provides better accountability, and offer
support to teachers and educators.

The New York Times reported that scores plummeted in the 2013 administration
with just 31% of third through eighth grade students passing in English and 31% of third
through eighth grade students passing in math (Hernandez & Gebeloff, 2013). This is a
24% drop in English, and a 34% drop in math from the 2012 test. New York officials
identified the cause of the drop as the alignment of the test with the Common Core
standards. Hernandez and Gebeloff (2013) explain that New York is one of the first states
to align the test with the new standards, and that states that have made the shift, like
Kentucky, have also seen low scores. Since speculation surrounds the low scores,
Hernandez and Gebeloff (2013) suggest that students were not given enough time to
prepare for the assessment, which was more analytical than previous assessments.
State Expectations

Another possible limitation to consider is the cut scores for proficient and
exceeding scores of each state. In this study, I used the percentage of students who scored
proficient or above on the state test. Since each state sets their own cut score (the
percentage of correctly answered questions to be considered proficient), comparing
schools on the national stage can be difficult. Table 12 displays the differing cut scores of
each state.
Table 12

*Cut Scores for Colorado, New York, Ohio, and Virginia State Achievement Tests*

<table>
<thead>
<tr>
<th></th>
<th>Colorado</th>
<th>New York</th>
<th>Ohio</th>
<th>Virginia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale for reading</td>
<td>330-990</td>
<td>100-417</td>
<td>252-507</td>
<td>0-600</td>
</tr>
<tr>
<td>Passing for reading</td>
<td>642-990</td>
<td>316-417</td>
<td>400-507</td>
<td>400-600</td>
</tr>
<tr>
<td>Mathematics</td>
<td>312-890</td>
<td>119-403</td>
<td>252-507</td>
<td>0-600</td>
</tr>
<tr>
<td>Passing for math</td>
<td>577-890</td>
<td>322-403</td>
<td>400-507</td>
<td>400-600</td>
</tr>
</tbody>
</table>

Comparison to Other Studies

A 2014 search of all of the databases available at Mercer University on the EBSCO host yielded just five peer-reviewed articles with a topic of “Schools to Watch©”. Of those five studies, only the study conducted by Cook, Faulkner, and Kinne (2009) examined the academic achievement in middle schools designated Schools to Watch©. The Cook et al. (2009) study is the closest comparison to this study, and it differs in its methodology and results.

The Cook et al. (2009) study examined Schools to Watch© Schools in just one state, Kentucky. Initially, the researchers utilized a quantitative survey designed by the researchers to identify the level of implementation of the middle school tenets. Next, they investigated the academic achievement of Schools to Watch© Schools by looking at the rank of each school in the state in overall achievement, math, and reading. The Cook et al. (2009) study used a stratified random sampling method that included 40 Non Schools to Watch schools and 10 Schools to Watch© schools. While random sampling is often
the accepted methodology for experimental research, the use of propensity score
matching in this present study reduces the variance associated with random sampling.
Shadish, Cook, and Campbell (2002) state, “When either matching or stratifying can be
used, matching is preferred” (p. 304).

Cook, Faulkner, and Kinne’s (2009) study found that Kentucky Schools to
Watch© schools had higher academic achievement than the Non Schools to Watch
schools, but no causal relationship was established between being a Schools to Watch©
school and higher academic achievement. Both the Cook et al. (2009) study and my study
show an increase in academic achievement. Since the Cook et al (2009) study did not
conduct a correlational test; it is difficult to make a comparison. It can be speculated that
the difference can be a result of the increased sample number of this study and a sampling
method that included the same number of Schools to Watch© schools and Non Schools
to Watch schools matched on specific demographic criteria.

In other research on academic achievement and the middle level concept, such as
the Rand Study (Juvonen et al., 2004), and America’s Middle Schools in the New
Century (McEwin, Dickinson, & Jenkins, 2004), the researchers utilized solid
methodological processes; however, these studies are outdated. In addition, these studies
failed to make connections between increased academic achievement and the middle
level concept. Both of these studies, along with my study are all conducted by
researchers who believe in the middle level concept, and the idea that the curriculum
designed with the young adolescent in mind has benefits that extend beyond academic
achievement. However, none of the studies provide a clear link to the middle school
concept and improved academic achievement. Much like this study none of the authors (Cook et al. 2009, Juvonen et al., 2004; McEwin et al., 2001) concluded that academics were not negatively impacted by schools that used the middle school concept.

Implications

This study raised several questions. One thing that was evident from the early stages of the research was that the middle level concept supported by the Association of Middle Level Education, the National Level for Middle Grades Reform, and well-known educational researchers is much more of a philosophy than a theory. The tenets described by middle level advocates are designed with good intentions, but again there has been little to support that all of the components are necessary to create a successful middle school, or to reach the young adolescent. This issue needs more empirical research study.

The fact that this research failed to yield positive results prompts questions regarding the merit of continued use of the middle school model. Research does support that many of the tenets, such as teaming, common planning time, and shared vision, do support the emotional and social well-being of adolescent students. Studies also suggest full implementation of the middle level concept increases student engagement in school and reduces student behavior problems. The results of this study did not show any negative impacts (and in all cases showed a positive mean increase) on academic achievement. That being the case, if it does not negatively impact academics and supports the developmental growth of students in the social, emotional, and cognitive domains, why not use this model?
Future Research

Up to this point no research is available to establish a causal relationship between the implementation of the middle school concept and academic achievement. There is a need for future research investigating the impact of middle school practices on academic achievement. This study should be expanded to include all of the Schools to Watch© schools in the United States. However, a study of this nature should also include some type of investigation into the implementation of the middle level components, similar to the Cook, Faulkner, and Kinne (2009) study. The present study revealed a great deal of variability between states, which may indicate that each of the states' Schools to Watch© programs use the STW© rubric differently. The inclusion of observations of each school, a survey, or the ratings from the Schools to Watch© rubric could offer a better picture of the level of implementation of the middle school concept in each school. It is also suggested that research is conducted to look at the validity and reliability of the Schools To Watch© rubric as a measure of middle school implementation. This research would increase the power of future studies.

Conclusion

This study was designed to investigate the impact of the middle level concept on academic achievement. Overall, the results of this study showed small but statistically insignificant gains in academic achievement in language arts/reading or math between schools labeled Schools to Watch© schools that fully implement the concept and Non Schools to Watch schools. Several limitations for this study were explored including the use of the Schools toWatch© program to define model middle schools, the difficulty in
finding exact matches for each middle school, the differences in the New York Assessment system, and different mastery expectations in each state. It was determined that the state where the schools were located played a significant role in the academic achievement, more so than if they were Schools to Watch© schools or not. Future research should examine academic achievement on a broader, national scale.

Final Thoughts

Like the young adolescents it is designed to teach, the research on the middle level concept is still not fully developed. Advocacy groups like the Association of Middle Level Education, and the National Forum for Middle Grades have proposed specific characteristics that define what it means to fully implement the middle level concept. While these characteristics are linked to adolescent development, little research exists to link them to academic achievement. In this sense the middle school is still in its growing phases.

Despite the fact that this study did not find a definitive impact on academic achievement, I still feel that the best way to reach young adolescent learners is to teach to their developmental needs. The literature is clear that implementation of particular components does improve the social, emotional and behavioral well-being of young adolescents. The study did not show a negative impact on academic achievement and suggested some increases in academics. If we can educate happier children, in a safer environment, the benefits go beyond scores on a test.
REFERENCES


Association for Middle Level Education (AMLE). (2010a). Research supporting this we believe: Keys to educating young adolescents. Westerville, OH: Author.

Association of Middle Level Education (AMLE). (2010b). This we believe: Keys to educating young adolescents. Westerville, OH: Author.

Association for Middle Level Education (AMLE). (2013). This we believe: Keys to educating young adolescents. Retrieved from http://www.amle.org/AboutAMLE/ThisWeBelieve/tabid/1273/Default.aspx


DuFour, R. (2004). What is a" professional learning community"? Educational Leadership, 61(8), 6-11.


APPENDIX A

IRB APPROVAL
01-Aug-2013

Kristina N. Falbe
Tift College of Education - Atlanta
Macon, GA 31207-0001

RE: The Adolescence of Middle Level Implementation and Academic Achievement (H1307209)

Dear Falbe:

Your application entitled: The Adolescence of Middle Level Implementation and Academic Achievement (H1307209) was reviewed by this Institutional Review Board for Human Subjects Research in accordance with Federal Regulations 21 CFR 56.110(b) and 45 CFR 46.110(b) (for expedited review) and was approved under Category 5 per 45 CFR 60364.

Your application was approved for one year of study on 01-Aug-2013. The protocol expires 01-Aug-2014. If the study continues beyond one year, it must be re-evaluated by the IRB Committee.

Item(s) Approved:
New Application

Please complete the survey for the IRB and the Office of Research Compliance. To access the survey, click on the following link: http://https://www.surveymonkey.com/s/KCTT2R

Respectfully,

Ava Chambless-Richardson, M.E.D., CIP, CRM
Member
Institutional Review Board
Mercer University IRB & Office of Research Compliance
Phone (478) 301-4101
Fax (478) 301-2329
ORC_Mercer@Mercer.edu
Instructions:

The Schools To Watch Program is a copyright protected program of the National Forum to Accelerate Middle Grades Reform. Criteria established by the National Forum is used as the basis for the Forum's Schools To Watch Program, which focuses on school improvement efforts characterized by a continuous trajectory toward success. The following Self Study and Rating Rubric may be used freely by any middle grades or secondary school to study and rate its practices. No adaptations to this self-study and rubric may be used without written permission from the Schools To Watch Committee and Board of the National Forum to Accelerate Middle Grades Reform. The rubric will take 20-25 minutes to complete.

This is a continuing-improvement tool as well as a mandatory self-rating for schools interested in applying for a state Schools To Watch designation. Schools applying for a STW state designation should be consistently averaging scores between 3.4 and 4.0 in all sections and components.

The rubric is divided into four domains: Academic Excellence, Developmental Responsiveness, Social Equity, and Organizational Structures and Processes. Under each domain there is general criteria followed by concrete, expected examples of excellence. Self-rate each general and detailed component. Your self-ratings should reflect your perceptions for your whole school, not for your specific classroom, grade level, or subject. The ultimate goal is to be consistently excellent and rate a well-evidenced score point 4 in every component (general and detail) of every section. Even when that ultimate goal is reached, a true high performing middle school will continue to seek ways to improve as new challenges arise.

A 4 in any general and detail component means the practice is highly and completely implemented, systemic, coherent in every classroom, by every teacher, across the school.
A 3 in any general and detail component means that there is a high degree or maturing quality of implementation that is systemic, but it may not be coherent or of the highest quality in every classroom and by every teacher, but certainly by most, 75 percent or better.
A 2 in any general and detail component means that there is a mixed, fair, immature quality of implementation. A 2 also means that practices may include many teachers but not the majority. The program may be too new to have realized accountable results or to be evaluated as effective.
A 1 in any general and detail component means that the practice may just have gotten started, (very immature), or is only practiced by a handful of practitioners.
<table>
<thead>
<tr>
<th>Subject</th>
<th>Grade 1</th>
<th>Grade 2</th>
<th>Grade 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math</td>
<td>85</td>
<td>90</td>
<td>92</td>
</tr>
<tr>
<td>Science</td>
<td>78</td>
<td>85</td>
<td>90</td>
</tr>
<tr>
<td>English</td>
<td>88</td>
<td>92</td>
<td>95</td>
</tr>
</tbody>
</table>

**Math**
- Improved understanding of basic concepts and the development of essential skills.
- The student is approaching the master's level of understanding.

**Science**
- The student is performing at the master's level of understanding.
- The student is approaching the master's level of understanding.

**English**
- Consistent performance at or above grade level standards.
- The student is approaching the master's level of understanding.
- The student is performing at the master's level of understanding.

**General Comments**
- The student demonstrates a strong work ethic and is consistently developing new skills.
- The student shows a strong understanding of the subject matter.
- The student is consistently performing at a high level.

**Detailed Evidence of Growth**

**Additional Information**

The performance of the student in the subject areas is consistent with the expectations for their grade level.

**School**

**District**

**National Rankings**

The school is ranked among the top 10% nationally.

**Academic Excellence (continued)**

1. The student demonstrates a strong work ethic and is consistently developing new skills.
2. The student shows a strong understanding of the subject matter.
3. The student is consistently performing at a high level.

**Academic Excellence (continued)**

1. The student demonstrates a strong work ethic and is consistently developing new skills.
2. The student shows a strong understanding of the subject matter.
3. The student is consistently performing at a high level.

**Academic Excellence (continued)**

1. The student demonstrates a strong work ethic and is consistently developing new skills.
2. The student shows a strong understanding of the subject matter.
3. The student is consistently performing at a high level.
<table>
<thead>
<tr>
<th>General Criteria</th>
<th>Detailed Evidence of Criteria</th>
<th>Self-Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developmental Responsiveness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-performing schools with middle grades are sensitive to the unique developmental challenges of early adolescence.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>School</th>
<th>District</th>
<th>General Criteria</th>
<th>Detailed Evidence of Criteria</th>
<th>Self-Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>Z</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>Z</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>Z</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>Z</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>Z</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>Z</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>Z</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>Z</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>Z</td>
</tr>
</tbody>
</table>

* The above data is provided for planning and program support to ensure student development by.

1. Teachers in the school are provided time and program opportunities to ensure student development by.

2. Teachers are provided the support they need to meet performance standards.

3. Students are provided the support they need to meet performance standards.

4. Teachers and students provide evidence to show more rigorous academic standards.

5. Teachers and students provide evidence to show more rigorous academic standards.

6. Teachers and students provide evidence to show more rigorous academic standards.

7. Teachers and students provide evidence to show more rigorous academic standards.
## National Forum to Accelerate Middle Grades Reform
### Schools To Watch
#### Self-Study and Rating Rubric

<table>
<thead>
<tr>
<th>District:</th>
<th>School:</th>
<th>General Criteria</th>
<th>Detailed Evidence of Criteria</th>
<th>Self-Rating</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The staff creates a personalized environment that supports each student's intellectual, emotional, social, and physical development.</td>
<td>1. The staff creates a personalized environment that supports each student’s intellectual, ethical, social, and physical development.</td>
<td>4 3 2 1</td>
<td></td>
<td>Average</td>
<td></td>
</tr>
<tr>
<td>• Adults and students are grouped into smaller communities (e.g., teams, houses, academies) for enhanced teaching and learning.</td>
<td>• Adults and students are grouped into smaller communities (e.g., teams, houses, academies) for enhanced teaching and learning.</td>
<td>4 3 2 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• These small learning communities are characterized by stable, close, and mutually respectful relationships.</td>
<td>• These small learning communities are characterized by stable, close, and mutually respectful relationships.</td>
<td>4 3 2 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Every student has a mentor, advisor, advocate, or other adult he/she trusts and stays in relationship with throughout the middle school experience.</td>
<td>• Every student has a mentor, advisor, advocate, or other adult he/she trusts and stays in relationship with throughout the middle school experience.</td>
<td>4 3 2 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. The school provides access to comprehensive services to foster healthy physical, social, emotional, and intellectual development.</td>
<td>2. The school provides access to comprehensive services to foster healthy physical, social, emotional, and intellectual development.</td>
<td>4 3 2 1</td>
<td></td>
<td>Average</td>
<td></td>
</tr>
<tr>
<td>• Teachers are trained to recognize and handle student problems.</td>
<td>• Teachers are trained to recognize and handle student problems.</td>
<td>4 3 2 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Students with difficulties, and their families, can get help.</td>
<td>• Students with difficulties, and their families, can get help.</td>
<td>4 3 2 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• The school houses a wide range of support (e.g., nurses, counselors, resource teachers) to help students and families.</td>
<td>• The school houses a wide range of support (e.g., nurses, counselors, resource teachers) to help students and families.</td>
<td>4 3 2 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• School staff members offer parent education activities involving families.</td>
<td>• School staff members offer parent education activities involving families.</td>
<td>4 3 2 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. All teachers foster curiosity, creativity and the development of social skills in a structured and supportive environment.</td>
<td>3. All teachers foster curiosity, creativity and the development of social skills in a structured and supportive environment.</td>
<td>4 3 2 1</td>
<td></td>
<td>Average</td>
<td></td>
</tr>
<tr>
<td>• Teachers enhance standards-based learning by using a wide variety of instructional strategies.</td>
<td>• Teachers enhance standards-based learning by using a wide variety of instructional strategies.</td>
<td>4 3 2 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Teachers incorporate well-developed procedures and routines for effective classroom management.</td>
<td>• Teachers incorporate well-developed procedures and routines for effective classroom management.</td>
<td>4 3 2 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Teachers facilitate learning by deliberately teaching study and organizational skills.</td>
<td>• Teachers facilitate learning by deliberately teaching study and organizational skills.</td>
<td>4 3 2 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Teachers integrate creative activities in the lessons (e.g., current technologies, visual and performing arts, etc.).</td>
<td>• Teachers integrate creative activities in the lessons (e.g., current technologies, visual and performing arts, etc.).</td>
<td>4 3 2 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. The curriculum is both socially significant and relevant to the personal and career interests of young adolescents.</td>
<td>4. The curriculum is both socially significant and relevant to the personal and career interests of young adolescents.</td>
<td>4 3 2 1</td>
<td></td>
<td>Average</td>
<td></td>
</tr>
<tr>
<td>• Students talk about daily issues in their own lives, their community and their world.</td>
<td>• Students talk about daily issues in their own lives, their community and their world.</td>
<td>4 3 2 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Students take action, make informed choices, work collaboratively, and learn to resolve conflicts.</td>
<td>• Students take action, make informed choices, work collaboratively, and learn to resolve conflicts.</td>
<td>4 3 2 1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Developmental Responsiveness (continued)

<table>
<thead>
<tr>
<th>General Criteria</th>
<th>Detailed Evidence of Criteria</th>
<th>Self-Rating</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Teachers use an interdisciplinary approach to reinforce important concepts, skills, and address real-world problems.</td>
<td>5. Teachers use an interdisciplinary approach to reinforce important concepts, skills, and address real-world problems.</td>
<td>4 3 2 1</td>
<td>Average</td>
</tr>
<tr>
<td>• Teachers collaborate to create integrated activities that cross content boundaries.</td>
<td>• Teachers collaborate to create integrated activities that cross content boundaries.</td>
<td>4 3 2 1</td>
<td></td>
</tr>
</tbody>
</table>

4 = High quality, complete, mature, and coherent implementation – NEARLY PERFECT, LITTLE ROOM FOR IMPROVEMENT
3 = Good quality, missing but not fully implemented by all – GOOD QUALITY BUT STILL ROOM FOR REFINEMENT and IMPROVEMENT
2 = Fair quality, minimal implementation, minimal practice, sporadic by some – SIGNIFICANT IMPROVEMENT NEEDED
1 = Poor quality, low level of implementation, new program, by a few – CONSIDERABLE STRATEGIC PLANNING, CONSENSUS BUILDING AND IMPROVEMENT NEEDED
National Forum to Accelerate Middle Grades Reform  
Schools To Watch℠  
Self-Study and Rating Rubric℠

<table>
<thead>
<tr>
<th>District: General Criteria</th>
<th>Detailed Evidence of Criteria</th>
<th>Self-Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>* Students can work on the same project in several different classes.</td>
<td>4 3 2 1 Average</td>
</tr>
<tr>
<td>8. Students are provided multiple opportunities to explore a rich variety of topics and interests in order to develop their identity, learn about their strengths, discover and demonstrate their own competence, and plan for their future.</td>
<td>4 3 2 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* Teachers and counselors push students to challenge themselves and set high academic and career goals for their future.</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td></td>
<td>* Students are provided with multiple opportunities to explore topics of interest to them.</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td></td>
<td>* Students demonstrate competency in a variety of subjects as well as areas of interest.</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>7 Students have opportunities for voice - posing questions, reflecting on experiences, and participating in decisions and leadership activities.</td>
<td>4 3 2 1 Average</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* All students have a real say, or have legitimate representation, in what happens at school.</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td></td>
<td>* School staff members have an &quot;open-door&quot; policy to encourage student involvement and connection.</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td></td>
<td>* Students take an active role in school-family conferences.</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>8. The school staff members develop alliances with families to enhance and support the well-being of the children.</td>
<td>4 3 2 1 Average</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* Parents are more than just volunteers or fund-raisers; they are meaningfully involved in all aspects of the school.</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td></td>
<td>* Parents are informed, included, and involved as partners and decision-makers in their children's education.</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>9. Staff members provide all students with opportunities to develop citizenship skills, to use the community as a classroom, and to engage the community in providing resources and support.</td>
<td>4 3 2 1 Average</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* Students take on projects to improve their school, community, state, nation, and world.</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td></td>
<td>* Members of the community engage in meaningful learning opportunities with the school.</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>10. The school provides age-appropriate, co-curricular activities to foster social skills and character, and to develop interests beyond the classroom environment.</td>
<td>4 3 2 1 Average</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* Student co-curricular activities cover a wide range of interests—team sports, clubs, exploratory opportunities, service opportunities, and a rich program in the visual and performing arts.</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td></td>
<td>* Co-curricular programs are infused with activities that help students develop relevant life skills.</td>
<td>4 3 2 1</td>
</tr>
</tbody>
</table>

Social Equity  
High performing schools with middle grades are socially equitable, democratic, and fair. They provide every student with high-quality teachers, resources, learning opportunities, and supports. They keep positive options open for all students.

4 = High quality, complete, mature, and coherent implementation – NEARLY PERFECT, LITTLE ROOM FOR IMPROVEMENT
3 = Good quality, missing but not fully implemented by all – GOOD QUALITY BUT STILL ROOM FOR REFINEMENT AND IMPROVEMENT
2 = Fair quality, select implementation, innovative practice, sporadic by some – SIGNIFICANT IMPROVEMENT NEEDED
1 = Poor quality, low level of implementation, new program, by a few – CONSIDERABLE STRATEGIC PLANNING, CONSENSUS BUILDING AND IMPROVEMENT NEEDED
National Forum to Accelerate Middle Grades Reform
Schools To Watch ©
Self-Study and Rating Rubric ©

<table>
<thead>
<tr>
<th>District:</th>
<th>School:</th>
</tr>
</thead>
</table>

### General Criteria

<table>
<thead>
<tr>
<th>Detailed Evidence of Criteria</th>
<th>Self-Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. To the fullest extent possible, all students, including English learners, students with disabilities, gifted and honors students, participate in heterogeneous classes with high academic and behavioral expectations.</strong></td>
<td></td>
</tr>
<tr>
<td>• Faculty and administrators are committed to helping each student produce proficient work.</td>
<td>4 3 2 1 Average</td>
</tr>
<tr>
<td>• Evidence of this commitment includes tutoring, mentoring, enrichment assignments, differentiated instruction, special adaptations, supplemental classes and other supports.</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>• Accelerated, short-term interventions for students with similar needs are fluid.</td>
<td>4 3 2 1</td>
</tr>
</tbody>
</table>

| **2. Students are provided the opportunity to use many and varied approaches to achieve and demonstrate competence and mastery of standards.** |
| • Teachers differentiate instruction in order to give each student equal opportunity to comprehend the standards-based curriculum. | 4 3 2 1 |
| • Teachers provide a variety of learning experiences so all students have opportunities to master a challenging curriculum. | 4 3 2 1 |
| • Teachers provide learning activities that represent varying learning styles so all students have opportunities to master standards. | 4 3 2 1 |

| **3. Teachers continually adapt curriculum, instruction, assessment, and scheduling to meet their students' diverse and changing needs.** |
| • The faculty is always seeking ways to improve programs, curriculum, and assessment to better meet student needs. | 4 3 2 1 |
| • Teachers assess mastery continuously and modify their instruction to meet current needs. | 4 3 2 1 |
| • The master schedule is developed in a way that provides flexibility for teachers to meet specific instructional needs on a daily basis. | 4 3 2 1 |

| **4. All students have equal access to valued knowledge in all school classes and activities.** |
| • To the fullest extent possible, students use technology to do research and analyze data, read more than textbooks, and understand how to solve complex problems. | 4 3 2 1 |
| • To the fullest extent possible, students with disabilities are in regular classrooms. | 4 3 2 1 |
| • Students have access to interest-based classes, activities, or opportunities. | 4 3 2 1 |

| **5. Students have ongoing opportunities to learn about and appreciate their own and others' cultures.** |
| • The school values knowledge from the diverse cultures represented in the school, community, and our nation. | 4 3 2 1 |
| • Materials in the media center represent all of the cultures of the students. | 4 3 2 1 |
| • Families often come and share their traditions and beliefs. | 4 3 2 1 |
| • Multiple viewpoints are encouraged. | 4 3 2 1 |

### Social Equity (continued)

<table>
<thead>
<tr>
<th>Detailed Evidence of Criteria</th>
<th>Self-Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>6. The school community knows every student well.</strong></td>
<td>4 3 2 1 Average</td>
</tr>
</tbody>
</table>

---

* 4 = High quality, complete, mature, and coherent implementation – NEARLY PERFECT, LITTLE ROOM FOR IMPROVEMENT
* 3 = Good quality, evolving but not fully implemented by all – GOOD QUALITY BUT STILL ROOM FOR REFINEMENT and IMPROVEMENT
* 2 = Fair quality, mixed implementation, immature practice, sporadic by some – SIGNIFICANT IMPROVEMENT NEEDED
* 1 = Poor quality, low level of implementation, new program, by a few – CONSIDERABLE STRATEGIC PLANNING, CONSENSUS BUILDING AND IMPROVEMENT NEEDED
### National Forum to Accelerate Middle Grades Reform

**Schools To Watch**

**Self-Study and Rating Rubric**

#### District: Schools To Watch*

<table>
<thead>
<tr>
<th>General Criteria</th>
<th>Detailed Evidence of Criteria</th>
<th>Self-Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Each student is appreciated and respected.</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>Staff members do not use negative labels or discuss students in negative ways.</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>Every student has an adult advocate and supporter in the school.</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>To the fullest extent possible, the faculty welcomes and encourages the active participation of all its families and makes sure that all its families are an integral part of the school.</td>
<td>4 3 2 1</td>
<td>Average</td>
</tr>
<tr>
<td>Transportation, meals, childcare, and translation support are provided so all families of diverse cultures and languages can attend school events.</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>Efforts are made to eliminate barriers (e.g., transportation, childcare, translation) to attend school events.</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>Multiple forms of communication are used with families and communication is two way.</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>Families have a voice in the decision-making process of the school.</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>Opportunities are provided for families to engage in supporting student learning (e.g., parenting classes, literacy programs, accessing information about student progress, making meaningful connections to the curriculum).</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>The school's reward system is designed to value diversity, civility, service, and democratic citizenship.</td>
<td>4 3 2 1</td>
<td>Average</td>
</tr>
<tr>
<td>The faculty recognizes the contributions of all its students.</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>Awards are not limited to sports and academic honors.</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>To the fullest extent possible, staff members understand and support the family backgrounds and values of their students.</td>
<td>4 3 2 1</td>
<td>Average</td>
</tr>
<tr>
<td>The school recruits a culturally and linguistically diverse staff.</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>The staff members are a good match to the school's community.</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>The school rules are clear, fair, and consistently applied.</td>
<td>4 3 2 1</td>
<td>Average</td>
</tr>
<tr>
<td>Students and parents are informed of school rules and know exactly what will and does happen if students break the rules.</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>Staff members routinely analyze and act upon referral and suspension data to make sure that no one group of students is unfairly singled out by classroom and school staff.</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
<tr>
<td>The school's disciplinary referrals and suspension rate are low as a result of proactive interventions that keep students engaged, resilient, healthy, safe, and respectful of one another.</td>
<td>4 3 2 1</td>
<td>4 3 2 1</td>
</tr>
</tbody>
</table>

#### Organizational Structures and Processes

High-performing schools with middle grades are learning organizations that establish norms, structures, and organizational arrangements to support and sustain their trajectory toward excellence.

---

4 = High quality, complete, robust, and coherent implementation – NEARLY PERFECT, LITTLE ROOM FOR IMPROVEMENT
3 = Good quality, missing but not fully implemented by all – GOOD QUALITY BUT STILL ROOM FOR REFINEMENT and IMPROVEMENT
2 = Fair quality, mixed implementation, immature practice, sporadic by some – SIGNIFICANT IMPROVEMENT NEEDED
1 = Poor quality, low level of implementation, new program, by a few – CONSIDERABLE STRATEGIC PLANNING, CONSENSUS BUILDING AND IMPROVEMENT NEEDED
### National Forum to Accelerate Middle Grades Reform

#### Schools To Watch

Self-Study and Rating Rubric

<table>
<thead>
<tr>
<th>District</th>
<th>General Criteria</th>
<th>Detailed Evidence of Criteria</th>
<th>School</th>
<th>Self-Rating</th>
</tr>
</thead>
</table>
| 1.       | A shared vision of what a high-performing school is and does drives every facet of school change. | - The shared vision drives constant improvement.  
- Shared, distributed, and sustained leadership propels the school forward and preserves its institutional memory and purpose.  
- Everyone knows what the plan is and the vision is posted and evidenced by actions. | 4 3 2 1 | Average |
| 2.       | The principal has the responsibility and authority to hold the school-improvement enterprise together, including day-to-day know-how, coordination, strategic planning, and communication. | - Lines of leadership for the school's improvement efforts are clear.  
- The school leadership team has the responsibility to make things happen.  
- The principal makes sure that assignments for staff are clear and explicit and are completed in a timely manner. | 4 3 2 1 | Average |
| 3.       | The school is a community of practice in which learning, experimentation, and the opportunity for reflection are the norm. | - School leadership fosters and supports interdependent collaboration.  
- Expectations of continuous improvement permeate the school culture.  
- Learning is ongoing for everyone. | 4 3 2 1 | Average |
| 4.       | The school and district devote resources to content-rich professional learning, which is connected to reaching and sustaining the school vision and increasing student achievement. | - Professional learning is intensive, of high quality, ongoing, and relevant to middle-grades education.  
- Teachers get professional support to improve instructional practice (i.e., classroom visitations, peer coaching, demonstration lessons, etc.).  
- Opportunities for learning increase knowledge and skills, challenge outmoded beliefs and practices, and provide support in the classroom. | 4 3 2 1 | Average |
| 5.       | The school is not an island unto itself; it is a part of a larger educational system (i.e., districts, networks and community partnerships). | - Deliberate vertical articulation and transition programs exist between feeder elementary schools and destination high schools.  
- The district supports through funding and time its schools' participation in best practice networks, associations, learning communities, and professional learning focused on middle grades improvement and achievement.  
- The school and district work collaboratively to bring coherence to curriculum, instruction, assessment, intervention, data collection, analysis, and accountability for student achievement. | 4 3 2 1 | Average |

---

**Organizational Structures and Processes (continued)**

- **4**: High quality, complete, mature, and coherent implementation — NEARLY PERFECT, LITTLE ROOM FOR IMPROVEMENT
- **3**: Good quality, maturing but not fully implemented by all — GOOD QUALITY BUT STILL ROOM FOR REFINEMENT AND IMPROVEMENT
- **2**: Fair quality, mixed implementation, immature practice, sporadic by some — SIGNIFICANT IMPROVEMENT NEEDED
- **1**: Poor quality, low level of implementation, new program, by a far — CONSIDERABLE STRATEGIC PLANNING, CONSSENSUS BUILDING AND IMPROVEMENT NEEDED
## National Forum to Accelerate Middle Grades Reform
### Schools To Watch
#### Self-Study and Rating Rubric

<table>
<thead>
<tr>
<th>District:</th>
<th>School:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>General Criteria</th>
<th>Detailed Evidence of Criteria</th>
<th>Self-Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. The school staff holds itself accountable for student success.</td>
<td>- The school collects, analyzes, and uses data as a basis for making decisions. 4 3 2 1</td>
<td>Average</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 3 2 1</td>
</tr>
<tr>
<td></td>
<td>- School-generated evaluation data is used to identify areas for more extensive and intensive improvement. 4 3 2 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- The school staff intentionally and explicitly reconsiders its vision and practices when data call them into question. 4 3 2 1</td>
<td></td>
</tr>
<tr>
<td>7. District and school staff possess and cultivate the collective will to persevere, believing it is their business to produce increased achievement and enhanced development of all students.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Barriers are viewed as challenges, not problems. 4 3 2 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- District and school staff assess and evaluate current programs regularly and adapt them as needed to maximize the level of student mastery. 4 3 2 1</td>
<td></td>
</tr>
<tr>
<td>8. The school staff and district staff partner with colleges and universities.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- A mentoring program for new teachers is in place. 4 3 2 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- The principal contacts colleges and universities when hiring new teachers. 4 3 2 1</td>
<td></td>
</tr>
<tr>
<td>9. The school includes families and community members in setting and supporting the school's trajectory toward high performance.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Families and community members are informed about the school's goals for student success and students' responsibility for meeting those goals. 4 3 2 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Representatives of all stakeholders are engaged in ongoing and reflective conversation and decision making about governance to promote school improvement. 4 3 2 1</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C

CODE FOR PSM
PSM using nearest neighbor method and 2 controls matched to every treatment. Justus Randolph

Notes: Install and load the MATCHIT package before attempting. Install and upload optmatch if using optimal matching.
Use the forward slash instead of the backslash when assigning a file name.
You can’t duplicate the output file names or you will get an error.
Delete the ratio command if you want one to one matching. Increase the number if you want more controls per treatment.
You might have to replace the word’s quotation mark with R’s to not get an error.

Using the nearest neighbor method

```r
mydata <- read.csv("C:/r/newyork.csv")
attach(mydata)
print(mydata)
m.out <- matchit(stw ~ tot + min + dis, data = mydata, method = "nearest")
summary(m.out)
plot(m.out)
plot(m.out, type = "jitter")
plot(m.out, type = "hist")
m.data1 <- match.data(m.out)
write.csv(m.data1, file = "C:/r/newyork_nearest1.csv")
```

Using the Optimal method:

```r
mydata <- read.csv("C:/r/newyork.csv")
attach(mydata)
print(mydata)
m.out <- matchit(stw ~ tot + min + dis, data = mydata, method = "optimal", ratio = 1)
summary(m.out)
plot(m.out)
plot(m.out, type = "jitter")
plot(m.out, type = "hist")
m.data1 <- match.data(m.out)
write.csv(m.data1, file = "C:/r/newyork1.csv")
```

See http://gking.harvard.edu/matchit for more information.