Seventh Grade Mathematics Results in a Standards-Based Learning Environment: Examining Teacher Collaboration and Leadership

Demetria C. Moon

Follow this and additional works at: https://red.library.usd.edu/diss-thesis

Part of the Educational Leadership Commons
Seventh Grade Mathematics Results in a Standards-Based Learning Environment: Examining Teacher Collaboration and Leadership

By

Demetria C. Moon

B.S., South Dakota State University – 1997
M.A., Black Hills State University – 2007
Ed.S., University of South Dakota – 2019

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Education

Division of Educational Administration

Educational Administration and Leadership Program
In the Graduate School
The University of South Dakota
August 2022
The members of the Committee appointed to examine the dissertation of Demetria C. Moon find it satisfactory and recommend that it be accepted.

Chairperson, Dr. Erin Lehmann

Dr. David De Jong

Dr. Diana Messick

Dr. Jane Stavem
ABSTRACT

The pressure on schools to improve student performance in middle school mathematics combined with the movement towards standards-based instruction leaves teachers searching for the most effective instructional practice and administration examining their role in improvement efforts. This study examines the effect on student achievement in middle school mathematics as a function of the quality of teacher collaboration interacting with administration presence during collaboration. This study used a quantitative research design in which seventh-grade mathematics teachers from two different middle schools in the same district with similar demographics taught the same standards-based unit of study and administered a common summative assessment at the end of the unit. Additionally, all teachers evaluated the quality of their job-alike collaboration during this unit using the Teacher Collaboration Assessment Rubric (TCAR). The results of this study suggest direct involvement of leadership during collaboration and instruction serves as a strong predictor of student achievement.

Dissertation Advisor  __________________________
Dr. Erin Lehmann
ACKNOWLEDGMENTS

For as long as my younger daughters can remember, Mom has been pursuing a degree and taking classes that have drawn my attention away from family at inopportune times. Although I recognize the sacrifices my husband, Dan, and my four children have made and the support they have shown throughout the journey, I am most grateful they value education and personal accomplishment enough to never question if the destination has been worth the journey. I hope this accomplishment is an example for Cassidy, Nathan, Lilly and Sophie of what grit and perseverance can achieve. Despite my dislike of being at the center of attention, I will walk through the graduation ceremony as a visual reminder of this valuable life lesson.

In addition to the enormous support from my entire family, immediate and extended, I am forever in debt to Dr. Erin Lehmann and Dr. David DeJong for their consistent encouragement to make this dissertation and doctorate degree a reality. They both saw possibility where my imagination ceased and were relentless with support from application to the final revision. The University of South Dakota is blessed to have two professors who are passionate about growing leadership in others and have a vision for continuous improvement in the program. For all the support, time dedicated to providing feedback, and genuine interest in my dissertation process, I would also like to thank the other two members of my committee, Dr. Jane Stavem and Dr. Diana Messick. Sitting on my personal board of directors, all four committee members contribute a special perspective as models of professional excellence.

I recognize each day how fortunate I am to work in a district surrounded by professionals who celebrate the growth and successes of others. Dr. Teresa Boysen granted permission for this research and has cheered me on through every step. My instructional coach team has listened to me tell stories and vent frustrations and celebrated every milestone throughout this process. I
hope they see this process as an example of how to grow their own leadership capacities and know I will support them in their own aspirations. I am so appreciative of every professional peer with their doctorate degree who has shared stories and reminded me time after time that I will finish, and the pursuit is worth every moment of doubt and frustration.

No one told me one of the best parts of this adventure would be the people I would walk beside as cohort colleagues. Chris, Abi, Renee, and Melissa are new friends who I will always share a special bond as we have endured all the low points and every celebration together. Every milestone, from successful statistical scriptwriting to setting dates for proposals and defenses, has shown me first-hand how vital it is to have a strong support system. We are the Covid cohort who have done things unconventionally at times, but we are committed to getting everyone to the finish line.

Finally, to my dear friend and pedagogical compass, Cassandra Erkens, you are the reason this research and journey are possible. You have exposed me to a new understanding of standards-based methods of teaching and learning and lit a fire inside me. The passion I have for using assessment to propel learning forward for all students is a direct reflection of the wisdom and energy you bring to your work across the country. I am so fortunate to learn from you each time our paths cross, but I am beyond blessed to have a personally and professionally supportive friend. I cannot wait to support you in this journey, as well.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doctoral Committee</td>
<td>i</td>
</tr>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>iii</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>v</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xii</td>
</tr>
<tr>
<td><strong>Chapter</strong></td>
<td></td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Background of the Problem</td>
<td>2</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>3</td>
</tr>
<tr>
<td>Contextual Framework</td>
<td>4</td>
</tr>
<tr>
<td>Conceptual Framework</td>
<td>5</td>
</tr>
<tr>
<td>Theoretical Framework</td>
<td>8</td>
</tr>
<tr>
<td>Research Questions</td>
<td>10</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>11</td>
</tr>
<tr>
<td>Assumptions</td>
<td>13</td>
</tr>
<tr>
<td>Limitations</td>
<td>13</td>
</tr>
<tr>
<td>Delimitations</td>
<td>14</td>
</tr>
<tr>
<td>Summary</td>
<td>14</td>
</tr>
<tr>
<td>2. Review of Related Literature</td>
<td>16</td>
</tr>
<tr>
<td>Systematic Review</td>
<td>16</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Background and Purpose</td>
<td>16</td>
</tr>
<tr>
<td>Subjects</td>
<td>17</td>
</tr>
<tr>
<td>Methods</td>
<td>17</td>
</tr>
<tr>
<td>Results</td>
<td>18</td>
</tr>
<tr>
<td>Conclusion</td>
<td>18</td>
</tr>
<tr>
<td>Standards-Based Teaching and Learning</td>
<td>19</td>
</tr>
<tr>
<td>Standards-Based Curriculum</td>
<td>20</td>
</tr>
<tr>
<td>Learning Targets</td>
<td>23</td>
</tr>
<tr>
<td>Success Criteria</td>
<td>24</td>
</tr>
<tr>
<td>Assessment Literacy</td>
<td>25</td>
</tr>
<tr>
<td>Formative Assessments</td>
<td>27</td>
</tr>
<tr>
<td>Summative Assessments</td>
<td>28</td>
</tr>
<tr>
<td>Teacher Collaboration</td>
<td>30</td>
</tr>
<tr>
<td>Teacher Collective Efficacy</td>
<td>30</td>
</tr>
<tr>
<td>Principal Leadership</td>
<td>33</td>
</tr>
<tr>
<td>Professional Learning Communities Model</td>
<td>34</td>
</tr>
<tr>
<td>Gaps in Research</td>
<td>36</td>
</tr>
<tr>
<td>Summary</td>
<td>37</td>
</tr>
<tr>
<td>3. Methodology</td>
<td>39</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>39</td>
</tr>
<tr>
<td>Research Questions</td>
<td>39</td>
</tr>
<tr>
<td>Review of Related Literature</td>
<td>40</td>
</tr>
<tr>
<td>Research Design</td>
<td>42</td>
</tr>
</tbody>
</table>
Assumptions.............................................................................................................59
Analysis....................................................................................................................60
Research Question Four..........................................................................................61
Assumptions..............................................................................................................61
Analysis....................................................................................................................62
Research Question Five...........................................................................................63
Assumptions..............................................................................................................63
Analysis....................................................................................................................65
Summary...................................................................................................................68
References (Ch. 1, 2, 3, & 4)......................................................................................70
5. Manuscript .........................................................................................................80
   Abstract...............................................................................................................80
   Introduction.........................................................................................................81
      Background of the Problem .............................................................................82
      Conceptual Framework.......................................................................................82
      Research Questions............................................................................................85
      Significance of the Study....................................................................................85
   Review of Related Literature................................................................................86
      Standards-Based Teaching and Learning..........................................................86
      Assessment Literacy...........................................................................................89
      Teacher Collaboration.........................................................................................90
   Methodology..........................................................................................................92
   Research Design..................................................................................................92
List of Tables

Tables                          PAGE
1. Schoolwide Demographic Data ...........................................................................43
2. Student Participants Demographic Data ...............................................................44
3. Factors Considered in Interaction Analysis .........................................................46
4. Student Participants Demographic Data ...............................................................51
5. Number of Teachers by Years of Experience Teaching Middle School Mathematics .....51
6. Student Participation by Experience Teaching Middle School Mathematics ..........52
7. Student Participation by Teachers’ Years of Collaborating as a Team .................52
8. Number of Teachers by TCAR Rating .................................................................53
9. Student Participation by Teacher TCAR Rating ....................................................53
10. Student Participation by Administration Presence ..............................................53
11. Differences in Significances in Mean Values between Pairings .........................57
12. Differences in Significances in Mean Values between Pairings ..........................59
13. Differences in Significances in Mean Values between Pairings ..........................61
14. Differences in Significances in Mean Values between Pairings ..........................63
15. Student Participation by Administration Presence ..............................................64
16. Shapiro-Wilk Test for Normality on Pairings .....................................................64
17. Differences and Significances in Mean Values between Pairings .........................66
18. Effect Sizes of Interaction Comparisons ............................................................67
1. Differences in Significances in Mean Values between Pairings ............................99
2. Differences in Significances in Mean Values between Pairings ............................101
3. Student Participation by Administration Presence ..............................................102
4. Shapiro-Wilk Test for Normality on Pairings................................................................. 103

5. Differences, Significances, and Effect Sizes of Mean Values between Pairings .......... 105
## List of Figures

<table>
<thead>
<tr>
<th>Figures</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Student Summative Assessment Scores</td>
<td>54</td>
</tr>
<tr>
<td>2. Homogeneity of Variance in RQ1 Subsets</td>
<td>56</td>
</tr>
<tr>
<td>3. Homogeneity of Variance in RQ2 Subsets</td>
<td>58</td>
</tr>
<tr>
<td>4. Homogeneity of Variance in RQ3 Subsets</td>
<td>60</td>
</tr>
<tr>
<td>5. Homogeneity of Variance in RQ4 Subsets</td>
<td>62</td>
</tr>
<tr>
<td>6. Homogeneity of Variance in RQ5 Subsets</td>
<td>65</td>
</tr>
<tr>
<td>7. Plot of Factorial ANOVA</td>
<td>68</td>
</tr>
<tr>
<td>1. Student Summative Assessment Scores</td>
<td>97</td>
</tr>
<tr>
<td>2. Homogeneity of Variance in RQ1 Subsets</td>
<td>98</td>
</tr>
<tr>
<td>3. Homogeneity of Variance in RQ2 Subsets</td>
<td>100</td>
</tr>
<tr>
<td>4. Homogeneity of Variance in RQ3 Subsets</td>
<td>103</td>
</tr>
<tr>
<td>5. Plot of Factorial ANOVA</td>
<td>106</td>
</tr>
</tbody>
</table>
CHAPTER 1

Introduction

Maximizing student achievement in mathematics to stay competitive on a global scale has been the primary focus for the American educational community for years (Bush, 2019; Gravemeijer et al., 2017; Ropohl et al., 2018; Wheat, 2021). The drive to compete has enacted many governmental initiatives, such as the reauthorization of the Elementary and Secondary Education Act in 2001, No Child Left Behind (NCLB) of 2002, and Every Student Succeeds Act (ESSA) of 2015, calling for education agencies to prove the effectiveness of their instructional practices through state-wide, and in many instances, nation-wide, summative assessments (Dennis, 2017). This call for one-time summative assessments as the leading measure of school quality and student success has shown little improvement and leaves federal and state policymakers at odds with educational professionals across the nation (Hemelt & Jacob, 2017). Reeves (2004) likens annual summative state assessments to autopsies. They provide standards-aligned information after the instructional year has ended, but they do not provide snapshots of learning along the way. Hill (2013) fittingly states “unfortunately, profound credence has been placed on standardized summative assessments instead of promoting the use of ongoing formative assessments to guide student learning and to foster timely instructional adjustments” (p. 2).

Despite the emphasis placed on standardized tests, members of the educational community versed in standards-based teaching and learning practices understand how a clear picture of the expected learning, communicating those learning intentions to students using success criteria, and using assessments to redirect student learning is the ultimate leverage to higher student achievement (Hillman & Stalets, 2021; Moss et al., 2011; Rinkema & Williams,
These practices are elevated when performed in professional learning communities (PLC) by teams of teachers with common teaching assignments (Burns et al., 2018; DuFour et al., 2008). Young and Kin (2010) found efficient use of data to guide instruction rests on the formative assessment practices of teachers, the usefulness of formative data, and the collective content and pedagogical knowledge of teachers—all of which are strengthened through collaboration in PLCs (Burns et al., 2018; Sutula, 2017). Although not enough on its own, school leaders can further elevate the impact of teacher collective efficacy by supporting and guiding the work of these teams (Sutula, 2017; Tichnor-Wagner et al., 2016).

**Background of the Problem**

As middle school teachers in a mid-size K12 public school district in the Midwest have invested their efforts over the past three years into unpacking standards, writing learning targets, scaling learning targets, and backward designing curriculum from common summative assessments directly related to those standards, crossing the bridge from curriculum to instructional practice continues to need improvement. Through this curriculum creation, teachers have deepened their understanding of content standards and coordinating skills necessary to reach these standards. The need for improvement in designing instructional opportunities which lead to proficiency of the standards on summative assessments continues to exist across the district. Many teachers struggle to leverage this new standards-based curriculum with a learning environment embedded in effective teacher collaboration around the results of common formative assessments, and they often question the significance of collaboration.

These middle school teachers have been provided with professional development centering around writing quality common formative assessments at the district and building levels. Additionally, each building is staffed with an instructional coach to assist school leaders
in guiding teachers through the collaboration process around the results of these common formative assessments. While some teacher collaboration groups are finding it easier to write these assessments and share their data, along with samples of student work, improvement in collaboration quality and an increase leadership involvement remain necessary across the district.

With six middle schools participating in the transition to standards-based teaching and learning, the job-alike collaboration groups experience a variety in the level of involvement from their building-level instructional leaders. In some schools the leadership is present in these collaboration group meetings, but the leaders lack the confidence, knowledge, or belief in the impact of teacher efficacy, so they are more of a silent observer than a guiding light. Donohoo et al. (2018) suggests school leaders play key roles on the instructional environment when involved in teacher collaboration. The support and guidance from the instructional leader or building principal is a strong predictor of the value teachers find in the collaboration process and the academic gains which result from this instructional technique (Donohoo, 2018).

Statement of the Problem

As educational systems maintain continued efforts for improvement in the areas of utilizing best practices and achieving academic success, John Hattie (2021) continues to examine which elements of the learning environment are most predictive of high academic achievement. Improving student achievement results in schools continues to be a focal point at the local, state, and federal level throughout the United States. Additionally, the movement toward a standards-based education system, where each content area has defined, grade-level standards for essential learning, continues to shift the instructional practices for teachers (Pak et al., 2020; Schimmer et al., 2018; Schmoker & Marzano, 1999). No longer does a textbook, where instruction starts on page one, suffice to drive the sequence or the topics covered in the classroom. The pressure for
improvement and the emphasis on a standards-based curriculum leave school leaders wondering what instructional practices produce the greatest impact on student achievement. This study was conducted to contribute to existing research by examining the impact on student achievement in middle school mathematics as a function of effective teacher collaboration on formative assessments in a standards-based environment when it interacts with the presence of administration during collaboration. In addition to studying how the collective efficacy of teachers within collaborative groups relates to the academic achievement of all students at the summative or standard level, it is important to investigate the results of subgroups such as males, females, and student with disabilities.

**Contextual Framework**

The school district in this study is in a rural, Midwest state with a land area of 75,811 square miles and a population of 886,667 inhabitants as of April 1, 2020 (United States Census Bureau, 2020). Located in the state’s most urban city, the district is landlocked by seven neighboring school districts with portions of four of those districts located inside of the city limits of the district investigated in this study. Most of the school district’s attendance area is within city limits and rapid economic growth is spurring housing construction in many undeveloped areas.

Twenty-three elementary schools, six middle schools, four comprehensive high schools, one career and technical academy, two specialty schools, one virtual academy, and four off-campus locations make up the physical areas of educational offerings within this school district. The elementary level serves students in grades early childhood through five. There are nineteen home attendance centers where students attend based on their address and six schools where specialty programs are offered in the areas of language immersion, gifted education, arts
education, parent involvement, and behavior modification. Almost every elementary school is led by a principal, an administrative intern or assistant principal, and an instructional coach. The six middle schools serve students based on home attendance centers, are organized into multidisciplinary teams at each of the sixth, seventh, and eighth grades, and host two specialty programs. These schools are led by a principal and an assistant principal with three also self-funding an administrative intern. Each middle school has an instructional coach, but coaches are also responsible for mentoring and coaching in at least two content areas across all middle schools. The four comprehensive high schools and the career and technical academy serve students in grades nine through twelve based on residency within home attendance areas and are led by a principal and assistant principal(s). Instructional coaches at this level are assigned by content area district-wide and are not connected to a specific building. English learners and students with disabilities at all grade levels are served at their home attendance center or provided transportation to a center site nearby.

Scores on state summative mathematics assessments have declined over the last five years in grades six through eight. Additionally, the school district recently studied, adopted, and implemented the Illustrative Mathematics curriculum at the middle school level. With a focus on improving scores on the state assessment in mathematics, the curriculum was chosen for its design around student discourse and investigative approach to meeting the mathematics practices and standards. At the same time the mathematics curriculum study began, the teachers and leadership at the middle school level began an intense 5-year transition to standards-based teaching and learning. Part of this teaching and learning approach includes a focus on collaboration through the Professional Learning Communities Model.

Conceptual Framework
A standards-based learning environment for students exists when assessment, instruction experiences, and the content standards function as an interdependent system (Schimmer et al., 2018). In this system, teachers go beyond introducing standards within their content and they establish a culture of learning where a balanced assessment system continually provides opportunities for students to demonstrate their proficiency at the intended depth of knowledge (DOK) level within each standard (Chattergoon & Marion, 2016; Coladarci, 2002; Marion et al., 2019; Schimmer et al., 2018). Learners are required to provide evidence of learning on standards related to the content of a particular course. When Lopez et al. (2017) claims “competency-based learning is not about learning skills instead of content; it’s about learning critical skills that empower learners to seek out and engage with content more deeply, meaningfully, and productively” (p. 40), he is describing a learning environment for students centered around evidence-based proficiency of standards.

Hillman and Stalets (2021) describe a system where the quality of teachers’ assessment literacy is predictive of their ability to provide this type of learning environment for students. Teachers must use assessment as a tool to provide direction for instruction, clarity for learning, and hope for student success. Furthermore, when teachers engage in these practices as a team, their collective efficacy creates a deeper understanding of the standards, an increased value in assessment, and more equitable learning experiences for all students (Hillman & Stalets, 2021).

The inner workings of the system of standards-based learning hinge around four major practices which align to the four critical questions of a professional learning community (DuFour & DuFour, 2013; Erkens, 2016; Hillman & Stalets, 2021; Rinkema & Williams, 2018; Schimmer et al., 2018). First, teachers must be able to identify what we want students to know and be able to do. Teams must identify and unpack the priority standards for the content or course which
they are teaching. These priority standards are essential to designing quality instruction, which is clear, concise, and focused on the ultimate destination, proficiency at the standard level. Schimmer et al. (2018) points out the mere existence of standards does not guarantee teachers are designing instruction around them. As teachers begin to unpack the standards, define what the standard means for student learning in terms of learning targets or learning progressions, skills necessary for students to be successful on the standard, and DOK level of each standard, they create a clearer picture of what students must be able to know, understand, and do to be proficient (Rinkema & Williams, 2018). The impact of identifying and unpacking non-negotiables as priority standards is strengthened when teachers engage in these activities as teams of content experts (Schimmer et al., 2018). Teacher collaboration strengthens each members’ individual understanding of proficiency as they discuss ideas and challenge each other’s thinking (Burns et al., 2018).

Next, methods are needed to know if students have learned the desired outcomes. There are two practices within the assessment structure, formative assessments and summative assessments, which provide teachers with real-time evidence of student learning (Erkens, 2016). Assessments are formative or summative based on how they are utilized by the teacher and student. Formative assessments can be formal, informal, individual, or common among a teaching team but are always used to monitor in-progress learning and help students and teachers redirect instruction based on the results (Hillman & Stalets, 2021). Schimmer et al. (2018) states “summative assessment completes a balanced approach to classroom assessment and makes teaching and reporting seamless” (p. 125). When created prior to the planning of instruction, Hillman and Stalets (2021) say summative assessments give direction and understanding of what proficiency looks like for a student. The tool used for a summative assessment should match the
DOK level of the standard which means not all standards can be assessed with multiple-choice items (Schimmer et al., 2018). Schimmer et al. (2018) goes on to explain how teachers must be careful not to confuse standards with standardization and allow for students to demonstrate proficiency in a variety of ways including writing, debate, portfolios, demonstrations, or discussions.

The review of student results on these formative and summative assessments by collaborative teams of teachers is the essential transition to answering the final two critical questions of professional learning communities: What will we do if some students have not learned and how will we extend the learning for those who are proficient? Erkens (2016) explains how the power of common formative assessments comes when teachers collaborate over the creation of the assessment, administer the assessment in proximity of the other team members, and collaborate over the examples of student evidence to calibrate understanding of the student errors. When these events occur in teams of job-alike teachers, programs improve and student learning increases (Burns et al., 2018).

While the power behind standards-based learning environments stems from the collective efficacy of teaching teams, the role of the principal in these teams is essential (Goddard et al., 2010; Hite & Donohoo, 2020). Principals must provide leadership and guidance to these teams as they work through the components of the balanced assessment system for the improvement of student learning (Goddard et al., 2010; Hite & Donohoo, 2020).

**Theoretical Framework**

Recognizing the effectiveness of a team requires unpacking the definition of a team, understanding what it means to have team leadership, and acknowledging the elements needed to maximize the leadership capacity. While Yukl (2013) defines leadership as “the process of
facilitating individual and collective efforts to accomplish shared objectives” (p. 23), the leadership theory referred to as team leadership requires further expansion.

A team is a group of people who are working together with a specific goal to accomplish (Northouse, 2021). Northouse (2021) goes on to explain not all groups of people are automatically a team, rather, certain characteristics must be present to qualify as a team. First, a sense of interdependence must be observable. As a team works toward their common goal, the success of individual team members depends on the success of the whole group and vice versa. Teams also display a commitment to the work of the group through communication, collaboration, and coordination of efforts. Finally, the team needs to be accountable as a unit to the organization, or a larger team. As academic, job-alike groups of teachers come together each week for the purpose of increasing student achievement by learning from the strengths of each other, within their school and district, they fully meet the definition of a team.

Highly functioning teams contain an element of leadership, and it is predictive to their success. Leaders within teams can be a formal, identified person or the leadership role can be shared by the members of the team (Northouse, 2021). Even when a formal leader is identified within a team, the other members of the team can still take on leadership roles which is described as distributed leadership. In the framework of a team of teachers, the principal often is identified as the formal leader of the team as the other members are all equal in rank and may be bound to their role through negotiated contract language. However, as is commonly seen with teachers, elements of leadership exist in many teachers, and they often execute these skills in the absence of the principal. Whether formal, shared, or distributed, the capacity of the team is dependent on the strength of the leadership guiding the group (Northouse, 2021).
Leadership is called to task within a team when a characteristic of team excellence is missing (Northouse, 2021). These characteristics of highly effective teams include elevation of goals, a results-driven structure, competent team members, a unified commitment, a collaborative climate, standards of excellence, external support or recognition, and principled leadership (Northouse, 2021, pp. 441-445). With collaborative teams of teachers, if one or more of these elements is missing, it becomes the leader’s job to intervene (Northouse, 2021). The leader evaluates the missing element and decides if intervention must occur internally or externally. If the intervention is internal, the missing element is either task-related or relational in nature and the intervention must match the need. However, if the environment where the team operates is not supporting the team, an external intervention is needed. Quality team leadership will recognize the need to intervene and match the intervention to the classification of the missing element impeding the success of the team.

As education systems become flush with professional learning communities (Brown et al., 2018; Burns et al., 2018) and administrators work to carve out time in the daily or weekly schedule for teams of teachers to meet and work toward a common goal of increased student performance, it is critical these teams understand the components which make them effective. Further, building leadership or administrators must recognize the significance of their role within the team as the formal team leader (Brown et al., 2018). As team leader, the team needs guidance on their academic goals of student improvement, but they also require identification and intervention when missing a key element of team excellence.

**Research Questions**

The following research questions guided this study:
1. To what extent does student achievement on a summative assessment differ as a function of the years of experience teaching mathematics at the middle school level in this district?

2. To what extent does student achievement on a summative assessment differ as a function of the years of experience as a team participating in collaboration over formative assessments throughout a unit of study?

3. To what extent does student achievement on a summative assessment differ as a function of the quality of involvement by the classroom teacher in collaboration over formative assessments throughout the unit of study, as determined by self-evaluation using the Teacher Collaboration Assessment Rubric (TCAR)?

4. To what extent does student achievement on a summative assessment differ as a function of involvement of leadership in the teacher collaboration process throughout the unit of study?

5. To what extent does student achievement on a summative assessment differ as a function of the interaction between the quality of involvement by the classroom teacher in collaboration, as determined by self-evaluation using the Teacher Collaboration Assessment Rubric (TCAR) and the involvement of leadership in the collaboration process throughout the unit of study?

Definitions of Terms

**Criterion-Referenced Assessment:** Assessments are criterion-referenced if they are scored against a fixed set of criteria or required knowledge, understanding, or skill. These assessments offer greater validity and reliability for measuring student proficiency and identifying gaps in learning (Burton, 2006).
**Depth of Knowledge:** These degrees or complexities of knowledge demanded by content standards were created by Norman Webb from the Wisconsin Center for Education Research. There are four levels in Webb’s Depth of Knowledge including basic recall, skill or concept, strategic thinking, and extended thinking. Educators often look to the verbs within content standards to reveal the intended complexity of the standard (Hess et al., 2009).

**Formative Assessment:** An assessment is formative not because it is given to students during a unit of study but because of how it is used by the teacher. Assessments are formative when they provide information allowing a teacher to make an instructional response, when students are given feedback showing them the next steps in their learning, and the feedback is infused into the learning process for students and teachers (Hillman & Stalets, 2021).

**Learning Targets:** When standards are unpacked, a set of clear and specific skills are defined which provide students with the tools necessary to reach proficiency of the standard. These skills are often transferable from one course to another within a content or across different content (Rinkema & Williams, 2018).

**Standards-based curriculum:** A standards-based curriculum is designed using the content standards as a framework for the instructional journey. The path to student proficiency is clearly defined and teachers can make precise and efficient instructional decisions. In this type of curriculum, students’ grades are communication tools solely about their proficiency of the standards and do not contain any negative or positive reflections of student behavior (Schimmer et al., 2018).

**Summative Assessment:** An assessment is considered summative when it confirms the student’s learning and proficiency. These types of assessments should come at a point in the learning journey after students have had an opportunity to practice their skills and develop proficiency.
Often, summative assessments become formative if the teacher chooses to intervene instructionally based on the evidence of learning provided (Hillman & Stalets, 2021).

**Teacher Collaboration:** When a group of teachers meet for the purpose improving student learning and they are focused on student results, the organized meetings are referred to as teacher collaborations or professional learning communities (PLCs) (DuFour & DuFour, 2013). These collaborative meetings happen in a timely manner so discussions around instructional planning, student assessment results, and intervention is current and accelerates learning.

**Assumptions**

The following assumptions were made for the consideration of this study:

1. The summative assessment used was criterion-referenced and aligned to the standards within the unit of study.
2. The summative assessment used was identical from group to group within the comparisons of this study.
3. The results of this study may benefit building leadership by providing them information about the impact of leadership within teacher collaboration and may help them make decisions about the role they will play within those job-alike collaboration groups.

**Limitations**

The following limitations were factors outside of the researcher’s control and could pose possible threats to the validity of the study.

1. All data is gathered from two schools within one school district.
2. The data was collected during the 2021-2022 school year following the heavy impact of the COVID-19 global pandemic. All students in this study were in face-to-face
instruction but some students may have been absent from school for extended periods of time due to infection of COVID-19.

3. Negotiated language within the certified teacher contract states activities within teacher collaboration time are at the discretion of the participating teachers and may not be assigned or led by principals without the invitation of the participating teachers.

4. The researcher is an employee of the school district where the study is being completed.

Delimitations

The following delimitations are factors within the researcher’s control and could pose possible threats to the validity of this study:

1. One teacher team chosen for this study has been a job-alike team for eight years and believes in the power of teacher collaboration. Their comparative results may not generalize the same results of other teams without a similar culture of collaboration.

2. The study focuses on the results of student achievement in seventh grade mathematics and may not be generalizable to other grades or other content areas.

Summary

The pressure on schools to improve the results of student performance on the state summative assessment has left teachers and principals looking for a sense of hope in research-based practices which will lead to greater student achievement. Standards-based teaching and learning practices provide a foundation for student learning making proficiency visible to both student and teachers. The PLC model enhances the collaboration efforts of teachers and allows both teachers and students to reap the rewards of teacher collective efficacy. Both teachers and school leaders need research results justifying the investment of time in collaborative teams to improve student achievement. Furthermore, teachers and school leaders need to know,
understand, and implement the parameters of effective teams resulting in improved student learning.
CHAPTER 2

Review of Related Literature

The purpose of this chapter is to describe the current research on the elements of standards-based teaching and learning, a balanced assessment system, leadership, and teacher collaboration using the professional learning communities (PLC) model as it relates to student achievement. The current research details how utilization of these instructional practices influences student achievement. Chapter 2 provides an extensive review of the literature associated with the development and utilization of a standards-based curriculum, including the components of learning targets and success criteria. Within a standards-based environment, assessment, teacher collaboration, and the role of leadership emerge as components worthy of investigation by the researcher. Through the research, studies are presented which examine the relationship between the elements of standards-based teaching and learning and student achievement. Literature suggests a standards-focused curriculum, a high knowledge and usage of assessment literacy, and a PLC model accompanied with effective leadership can increase student learning and student achievement on summative assessments aligned to content standards (Erkens, 2016; Hillman & Stalets, 2021; Rinkema & Williams, 2018; Schimmer et al., 2018).

Systematic Review

Background and Purpose

Richard DuFour authored his first outline of the PLC model in 1998 about the same time content standards were popping up in every content area (DuFour et al., 2008; Marzano & Haystead, 2008). Over the last 20 years, school leaders have also continued to search for instructional best practices for academic improvement conjuring studies on standards-based curriculum, professional learning communities, and leadership. The purpose of this systematic
literature review was to identify current research on key curriculum and instruction components related to increased student achievement, including standards-based curriculum, assessment literacy, teacher collaboration, and leadership involvement.

Subjects

This review of literature explores research published from 1967 through 2021. Sources were chosen during these years to create the full timeline of standards development in education from their origin to present.

Methods

A systematic review of the literature was conducted during the summer and fall of 2021 to pinpoint the major and minor themes around standards-based instruction and the PLC model. Within standards-based instruction, themes such as standards-based curriculum, learning targets and success criteria, assessment literacy, formative assessment, summative assessment, and backward design quickly emerged. Search terms including standards-based curriculum, teaching, and learning also generated results centered around standards-based grading. Although the differentiation between standards-based curriculum and standards-based grading is addressed in this review of the literature, most findings on grading practices are not included as they do not impact the study.

Narrowing the search for current and relevant research required combining these themes with student achievement. Results began to include instructional elements and a new theme of teacher collective efficacy. When searching for studies involving the relationship between teacher efficacy and student achievement, the impact of leadership at the organizational and team level also appeared.
Results

The review of literature identified 87 sources relating to the study, including 47 peer-review journal articles, 26 books, four published dissertations, three legislative acts, and seven scholarly articles posted on the web. Thirty-three of the references directly address standards-based curriculum, learning targets, success criteria, or instruction involving these elements. Forty references focused on the importance of assessment and backward design as it relates to student achievement. Thirty-seven studies concentrated on the influence of collective teacher efficacy and teacher collaboration. Finally, 27 of the studies discussed the vital role of leadership within teams and how it impacts team efficacy and ultimately, student achievement. Throughout the quest of literature to review, searches failed to produce results showing the relationship between teacher collaboration around formative assessment results and student achievement on summative assessments.

Conclusion

Strong evidence supports the necessity for a standards-based curriculum allowing teachers and students to map their journey toward proficiency. The elements of standards-based learning include learning targets to identify critical skills, success criteria to help students visualize what proficiency looks like, and criterion-referenced assessments aligning to the intended complexity and depth of the standard. These elements all show a positive relationship with increased student achievement when implemented as part of the classroom culture. Adding collaboration among teams of teachers with common job assignments in this standards-based environment layers teacher collective efficacy within the elements predicting higher results for students on standardized assessments aligned to standards. The effectiveness of the team
increases when team leadership, either by a formal leader, distributed leadership, or shared leadership, is part of the PLC model for teacher teams.

**Standards-Based Teaching and Learning**

When standards-based teaching, learning, or curriculum arises in a conversation with educators, most minds go straight to the concept of standards-based grading. While standards-based grading can be an element of the standards-based teaching and learning process, the latter can exist within a traditional grading system (Schimmer et al., 2018). Furthermore, since grading practices are only a mechanism for reducing student proficiency to a single symbol, the type of grading utilized does not automatically imply students are learning in a standards-based environment (Schimmer et al., 2018). Rinkema and Williams (2018) claim the work of Wiggins and McTighe (2008) holds true today when deciding what works in curriculum design and instruction. Identifying desired results, determining assessment evidence, and planning the learning experiences and instruction are the stages of backwards design most effective to elicit desired student learning (Wiggins & McTighe, 2008). A teaching and learning environment where a common understanding of the standards exists, the destination of learning is clear to the students, and supports are in place linked to the needs identified by a balanced assessment system more accurately defines standards-based teaching and learning than the utilized grading practices (Erkens, 2016; Hillman & Stalets, 2021; Rinkema & Williams, 2018; Schimmer et al., 2018). When teachers use a standards-based curriculum and learning happens through a standards-based instructional model, there are significant gains in student achievement on assessments aligned to the rigor of the content standards (McCarthy, 2020).

Standards-based teaching and learning practices are multi-dimensional and require the systemic, interdependent use of three major factors for the implementation to be executed with
fidelity. A few of those who invest time researching the impact of standards-based teaching practices, including Schimmer et al. (2018) and Erkens (personal communication, October 28, 2021), relate this multi-dimensional system to planning a trip. First, the deep understanding or unpacking of the content standards provides the clear picture of the destination or evidence of learning. Proficiency on content standards also requires understanding the skills, or learning targets, needed to reach the desired level of rigor or complexity of the standard (Hillman & Stalets, 2021; Rinkema & Williams, 2018). Making this picture clear to students means their learning must be visible; the standards, the learning targets, exemplars of proficiency, and success criteria must be transparent to the learner throughout the instructional journey (Hattie, 2012; Schimmer et al., 2018). Next, like checkpoints on a trip or when a GPS assistant adjusts the route during a trip, a balanced assessment system needs to be in place to help teachers and students monitor students’ learning progress and assist teachers in providing interventions to get students back on track if they stray off course (Erkens, 2016; Schimmer et al., 2018). Finally, maximizing teacher effectiveness with student interventions happens when common assessments can be used by teacher teams, and the teams can collaborate over the results using the PLC model (DuFour et al., 2008; Hillman & Stalets, 2021). The following sections will expand on the research related to standards-based curriculum, learning targets, and success criteria, and their impact on student achievement. The research will also examine what happens when these components are used interdependently, and the teacher does more than deliver content but rather serves as a facilitator of learning (Rinkema & Williams, 2018).

**Standards-Based Curriculum**

The standards movement in the United States is traced back to the early 1980s when the publication by the National Commission on Excellence in Education (1983) left the nation’s
citizens pleading for reform to K-12 education and improvement in academic performance (Marzano & Haystead, 2008). Marzano and Haystead (2008) list dozens of events between the publication in 1983 and the end of the century which helped shape the current standards guiding instruction and assessment today, including an assembly of governors by the President of the United States, new laws, and publications of standards. While presenting a common framework for teachers to calibrate what they teach in their classroom, Marzano and Haystead (2008) maintain the newly produced standards consistently contain far more content than can be learned in one academic year and the standards are not unidimensional which makes assessing them as they are written almost impossible. Hillman and Stalets (2021) more recently suggest prioritizing a select number of standards from each grade and content as a first step to any work with standards-based planning as the all-encompassing collection of standards is often far too extensive.

A standards-based curriculum uses the adopted standards for a course, content, or grade-level as the framework for its development and clearly articulates the desired results for student learning (Hillman & Stalets, 2021; Rinkema & Williams, 2018). Although a well-designed curriculum framed around standards can have a positive impact on student achievement and a standards-based curriculum is often experienced by higher achieving students, Shoen et al. (2003) claims how the curriculum is implemented by teachers is the most influential element. Some of the implementation strategies used by mathematics teachers most strongly associated with higher student achievement include eliciting and building on student thinking using conceptual issues, keeping instructional tasks at a high cognitive level, and forcing students to make meaning of their thinking, using group work to encourage engagement and entertain alternative problem-solving methods, and utilizing strategies to keep student discussions and
problem-solving at the higher depth of knowledge (DOK) levels (National Council of Teachers of Mathematics, 2014; Shoen et al., 2003). These reform teaching practices were all implemented using a standards-based curriculum.

Standards define the evidence of learning needed by the student at the end of their learning journey (Erkens, 2016; Hillman & Stalets, 2021; Schimmer et al., 2018). To create a curriculum out of standards, they must be unpacked into learning progressions which outline the skills and sequencing of instruction needed to obtain proficiency (Erkens, 2016; Hillman & Stalets, 2021; Rinkema & Williams, 2018; Schimmer et al., 2018). When learning is designed around proficiency of a standard, grades and feedback must be criterion-referenced to have meaning within the system (Rinkema & Williams, 2018; Schimmer et al., 2018). Therefore, telling a student their score was at the top of their class only tells them how they performed compared to their peers but does not help them understand the level of proficiency at which they performed around the standard. For this reason, the more experience educators have practicing instruction around a standards-based curriculum, the less meaningful the traditional grading system becomes in their classroom (Schimmer et al., 2018).

Within a standard, the verbs provide the teacher with clues to the complexity or rigor of the standard and indicate the type of evidence of learning needed by students (Erkens, 2016; Hillman & Stalets, 2021; Rinkema & Williams, 2018; Schimmer et al., 2018). Teachers must fully understand and calibrate their thinking about what the standards are soliciting from students before they can elicit evidence of learning from students (Schimmer et al., 2018). To reference the earlier parallel of planning a trip, if teachers do not have a solid understanding of the standard’s complexity, they are simply traveling without a destination and will not know when they have arrived.
Learning Targets

When unpacking a standard, teachers must work together to identify the necessary progression of learning to reach proficiency on a standard. These learning progressions are laced with skills, or learning targets, needed by the students to achieve at the level of complexity identified by the standard. Rinkema and Williams (2018) and Schimmer et al. (2018) claim the learning targets are the *Ds* from the *know (K)*, *understand (U)*, and *do (D)* (K-U-D) framework Tomlinson and Strickland (2005) created. The framework assists teachers in planning what will be taught, practiced, and assessed throughout the unit of study.

Moss et al. (2011) defines shared learning targets as critical information for students: what to learn, how deeply to learn it, and how to demonstrate their learning. Learning targets can be expressed through words, pictures, or actions depending on the age of the students but always provide clarity and direction as the teacher plans and as the student engages in their learning (Brookhart, 2012). A research study found New York City’s Public School 13 made significant gains in student success while simultaneously narrowing the achievement gap through a formative assessment process which included clarifying learning targets and utilizing success criteria (Martuccio & Bloomberg, 2020).

Still, teachers need to do more than compose a learning target and announce it to students at the beginning of a class period. Teachers need to communicate learning targets in a way students will understand and recognize throughout the teaching and learning process. To help students focus on what they can do and present the targets in student-friendly language, “*I can*” statements are often used (Erkens, 2016; Hillman & Stalets, 2021; Rinkema & Williams, 2018; Schimmer et al., 2018). Ferrara et al. (1995) determined teachers may be lacking the skills necessary to effectively communicate clear learning intentions to their students. Crichton and
McDaid (2016) studied both teachers’ and students’ perceptions of learning intentions, finding both teachers and students recognized the importance of clear learning intentions, but rarely found them discussed in the classroom. Further, when learning targets were presented to students, they commented on their frustration at the way in which teachers attempted to communicate them. Ferrara et al. (1995) noted training may need to be implemented with preservice teacher programs on the value and mechanics of integrating learning intentions into their instruction in meaningful ways.

**Success Criteria**

Rinkema and Williams (2018) differentiate learning scales from rubrics, a more commonly identified evaluation tool in education, by their purpose. While rubrics were primarily used for scoring an assessment after the learning was complete, the learning scale guides instruction and provides feedback to students as they progress through their learning. To maximize the usage of learning targets to advance student learning, targets should be scaled to show an increase in complexity (Erkens, 2016; Rinkema & Williams, 2018; Schimmer et al., 2018). Within these scales, the descriptions of learning should be clear and concise and should “move beyond quantifiable requirements and guidelines of compliance to a true description of the learning” (Schimmer et al., 2018, p. 171). Success criteria should focus on what students can do and should refrain from specifics which limit ways students can demonstrate proficiency (Erkens, 2016; Schimmer et al., 2018).

Paterson (2014) found end-of-unit achievement was higher for students when their learning was outlined with a clear roadmap consisting of learning targets and success criteria. These elements make their learning visible and provide a meaningful model for their expected success. When learning is visible, standards are prioritized with clear learning targets scaled with
concise success criteria, teacher clarity increases, feedback is tangible, and classroom discussions are framed with academic vocabulary; all influences Hattie (2015) describes as highly impactful on student achievement. Additionally, when Hattie (2015) summarized all the influences impacting student learning to the top six key findings, the fourth most impactful is “when teachers explicitly inform the students about what success looks like near the start of a series of learning (.77)” (p. 81). Hattie (2015) explains that any factor with an effect size over 0.6 has a large effect on educational outcomes as opposed to factors with medium effect size around 0.4 and those with a low effect being 0.2 or lower.

Assessment Literacy

Within a standards-based learning environment, assessment serves as a communication tool between teachers, students, and parents (Erkens, 2016). Once the learning goals are prioritized and learning targets are unpacked with clearly defined success criteria, teachers plan instruction with opportunities to collect evidence of student learning using a variety of assessment structures (Marion et al., 2019). Assessments, or evidence of learning, should link to the prioritized standards (Marion et al., 2019), be collected from a variety of sources (Schimmer, 2018), empower students to improve their learning through feedback (Hillman & Stalets, 2021), and guide teachers on what should come next in the learning process (Erkens, 2016). When these criteria are met consistently and continuously throughout the academic cycle, the assessment system is said to be balanced (Chattergoon & Marion, 2016; Marion et al., 2019).

Within a balanced assessment system, a variety of assessments are utilized (Marion et al., 2019). In addition to having a variety of ways in which evidence of learning is elicited from students, variety also comes from the purpose of the assessments (Chattergoon & Marion, 2016; Marion et al., 2019). Most often these purposes can be categorized into formative, interim, and
summative, and they happen at the classroom, district, and state levels (Marion et al., 2019). Assessments at the classroom level include both formative and summative which are used to monitor and adjust instruction while providing progress feedback to parents and students (Marion et al., 2019). At the district level, interim and common summative assessments are utilized to predict performance on state assessments, evaluate curriculum and resources, and inform placement decisions into special programs (Marion et al., 2019). Finally, at the state level, assessments are almost always summative and assist in the evaluation of student learning and school quality (Marion et al., 2019).

Comprehensiveness, coherence, continuity, efficiency, and usefulness are common characteristics found in any balanced assessment system (Chattergoon & Marion, 2016; Marion et al., 2019). A comprehensive assessment system ensures all standards are taught and the intended DOK level is reached when students are proficient (Marion et al., 2019). Balanced assessment systems are coherent when they align the curriculum and instruction for students, making the learning process seamless (Chattergoon & Marion, 2016; Marion et al., 2019; Pellegrino, 2001) and continuous when done routinely over the academic year (Chattergoon & Marion, 2016; Marion et al., 2019). Finally, a balanced assessment system is efficient when the fewest number of assessments are given which yield the highest amount of information in return and when used for the purpose in which it was administered (Chattergoon & Marion, 2016; Marion et al., 2019).

While both formative and summative assessments play a vital role in the balanced assessment system, Erkens et al. (2017) explain how formative assessments carry the greatest burden of guiding instruction in and out of the professional learning community (PLC) model so summative assessments can be a celebration of the learning.
**Formative Assessments**

Rather than by design, an assessment becomes formative based on how it is integrated into the teaching and learning process and how critical the results are to the next steps of the learning process (Erkens et al., 2017; Hillman & Stalets, 2021; Rinkema & Williams, 2018; SCASS, 2018; Schimmer et al., 2018). When used formatively, assessment results allow the teacher to provide adaptive instruction to better meet students’ needs (Wiliam, 2018). Students are provided with clear learning targets and success criteria during instruction and quality formative assessment aligns to these targets to determine the exact area where students need intervention (Wiliam, 2018).

Since the purpose of formative assessment results is to guide the next instructional steps, grades are not an effective form of feedback with these types of assessments (McMillan et al., 2017). Grades signal an end to the learning process for students and often stop the action of learning for students when received on formative assessments (McMillan et al., 2017). Rather, teachers need to focus on the feedback they are providing students and move their learning forward to the next place in the learning progression (Erkens, 2016, Erkens et al., 2017; Hillman & Stalets, 2021; McMillan et al., 2017; Rinkema & Williams, 2018; SCASS, 2018; Schimmer et al., 2018). Feedback on formative assessments can stem from three separate sources including the teacher, the student, and peers (Heritage, 2021). For the teacher, the feedback relates to the learning targets and success criteria (Chappuis & Stiggins, 2017; Erkens, 2016; Heritage, 2021; Hillman & Stalets, 2021). A key element in the feedback process comes from the results of students’ self-assessment. This process forces metacognition strategies which assist in the adjustment of students’ own learning goals (Chappuis & Stiggins, 2017; Erkens, 2016; Heritage, 2021; Hillman & Stalets, 2021). Occasionally, in an established culture of high-quality, student-
to-student relationships, it is also appropriate for peers to provide each other feedback to increase clarity in the progress of their current learning and motivate movement to the next step in the learning progression (Chappuis & Stiggins, 2017; Heritage, 2021; Slavin et al., 2003).

As teachers set out to elicit evidence of learning from students and identify the gap between what students know and what they need to know, formative assessment can take on many different forms (Heritage, 2021; Schimmer et al., 2018). Questioning, discussions, observations, activities, constructions, writing, and student-selected response questions are all examples of techniques which assist teachers in identifying the next steps for each student’s learning (Chappuis & Stiggins, 2017; Heritage, 2021; Schimmer et al., 2018). Despite the assessment technique chosen, students’ thinking should be made visible, teachers should be more informed about the new direction needed in the learning progression (Wiliam, 2018), and the learning process should not be interrupted by the assessment for it to be considered formative (Schimmer et al., 2018).

**Summative Assessments**

The modern combination of formative and summative assessments within the learning process creates a balanced, seamless, parallel, and mutually supportive system (Black & McMillan, 2012; Erkens et al., 2017; Schimmer et al., 2018). While formative assessments inform the learning of specific skills at the learning target level, the purpose of summative assessments shifts to reporting on proficiency of learning at the standard level (Black & McMillan, 2012; Erkens et al., 2017). McIntosh and Milam (2016) found assessment techniques modeling the language of the standard are most effective in determining a student’s level of proficiency on the standard.
Due to the nature of summative assessments usually happening at the end of the unit, although they can occur at other points in time, the reporting mechanism generally shifts from feedback to grades (Erkens et al., 2017). However, as instruction shifts to align with standards-based curriculum, reporting traditional grades on summative assessments often seems clunky or disconnected from the learning culture (Erkens et al., 2017). In most cases, when implemented in a logical manner, standards-based grading stems from the teacher’s desire for alignment between assessment practices and grading practices (Erkens et al., 2017; Hillman & Stalets, 2021; Rinkema & Williams, 2018; Schimmer et al., 2018). Since grades are such a personal reflection of a student’s performance, they are most accurate when they depict what a student knows by the end of a grading term and not the speed at which they become proficient (Erkens et al., 2017; Schimmer et al., 2018). In many cases, for grades to be an accurate reflection of student proficiency, students may require opportunities for reassessment, which Lewis (2020) found females took greater advantage of than males.

The proficiency results determined by summative assessments are used for different purposes at different levels. At the classroom level, a summative assessment and the grade assigned to the level of proficiency demonstrated by the student serves as a communication tool between teacher, student, and parent or guardian (Chattergoon & Marion, 2016; Marion et al., 2019). For buildings and districts, the summative assessment often serves as a tool to evaluate the effectiveness of curriculum, instructional practices, programming, and resources (Chattergoon & Marion, 2016; Marion et al., 2019). Finally, at the state level, the composite results of annual summative assessments are used to evaluate district quality (Chattergoon & Marion, 2016; Marion et al., 2019).
Teacher Collaboration

The results of student performance on formative and summative assessments play an important role in the collaboration efforts of teachers. While teachers of various profiles participate in a variety of collaboration experiences, Ronfeldt et al. (2015) found when teachers collaborate in instructional teams of common content, there is a high association with increased student achievement. Additionally, when those teams focus their conversations on backward lesson design and assessment results, the highest positive correlation to student achievement is found (Ronfeldt et al., 2015). Instructional teams collaborating over common curriculum, instruction, and assessment increases the individual teacher’s confidence in their personal capacity and higher teacher efficacy strengthens the performance of the entire team (Donohoo et al., 2018).

Teacher Collective Efficacy

When teachers believe the combined efforts of their team increase their ability to overcome obstacles, the efforts of the individual team members and the efforts of the group are more impactful (Hattie, 2012; Donohoo et al., 2018). Hite and Donohoo (2020) define a model for teacher collective efficacy containing four key components of teacher behavior for teachers on the team and a fifth key component contributed by leadership. Goal consensus, empowered teachers, cohesive teacher knowledge, and embedded reflective practices are the four components within the teacher team presenting evidence of teacher collective efficacy (Hite & Donohoo, 2020). Supportive leadership completes the model and strengthens the other four components, correlating to higher student achievement (Eells, 2011; Hite & Donohoo, 2020). High teacher collective efficacy results in teacher teams with increased persistence and greater
motivation to reach all students through evidence-based practices (Donohoo & Katz, 2019; Eells, 2011).

Goal consensus exists when teams of teachers collectively agree on the goals of the group and commit to them as individuals (Hite & Donohoo, 2020; Kurz & Knight, 2004). While state education agencies publish content standards, the list of standards is often far too extensive for an academic year (Hillman & Stalets, 2021; Marzano & Haystead, 2008). Teachers need to work together to establish priorities within these lists of standards, unpack the standards to set learning targets and success criteria, and create assessments to focus the work of the group (Hite & Donohoo, 2020). When created together, there is increased teacher buy-in and commitment to the goals (Hite & Donohoo, 2020; Kurz & Knight, 2004).

Voelkel and Chrispeels (2017) found perceived control, or teacher empowerment, regarding decisions being made correlated to increased teacher investment and intention regarding instructional decisions. Teacher self-efficacy, or attribution theory, creates teacher empowerment and motivates teachers to work collectively with their team to create high expectations and pathways for student success (Eells, 2011; Gusky, 2021; Hattie, 2012; Hite & Donohoo, 2020; Ninković & Knežević Florić 2018). Self-efficacy of teachers grows from major sources: mastery experiences, vicarious experiences, verbal and social persuasion, and emotional and physiological states (Gusky, 2021; Hattie, 2012; Mielke, 2021). Of these sources, mastery experiences are the most influential way to increase efficacy beliefs (Gusky, 2021; Hattie, 2012; Mielke, 2021). Teachers often use common formative assessments to backward-design lessons and collectively reinforce the belief by each member of the team that all students can achieve at high standards (Hite & Donohoo, 2020).
For teachers to invest in the collective work of collaboration over common formative assessments, they must trust and believe in the instructional practices and current knowledge of their team members (Hite & Donohoo, 2020; Truijen et al., 2013) and see their actions make a difference (Gusky, 2021). Specifically, when teachers possess shared knowledge, participate in collaboration activities focused on assessment results, and utilize a standards-based curriculum, it is a strong predictor of high student achievement (Akiba & Liang, 2016). When teachers possess cohesive teacher knowledge and experience a high return on investment, school culture tends to be characterized by high expectations and student success (Donohoo, 2018; Eells, 2011; Gusky, 2021; Mielke, 2021; Ninković & Knežević Florić, 2018). Donohoo et al. (2018) find there is a ripple effect regarding teacher collective efficacy and its impact on school culture. As teachers experience success with their collective efforts, team confidence increases and generates further success (Gusky, 2021).

Recognizing success as a direct result of the collective efforts of teachers and increasing the self-efficacy of the group stems from reflective practices embedded into collaboration routines (Donohoo et al., 2018; Gusky, 2021; Preston & Donohoo, 2021). Preston and Donohoo (2021) refer to these types of teams as “mindfully organized” (p. 28), claiming they are committed to resilience as they learn by failure. Highly effective teams are data-focused and look for evidence to support decisions (Hite & Donohoo, 2020).

For all the components of collective teacher efficacy to flourish, the environment must be rich with supportive leadership (Goddard et al., 2017; Hite & Donohoo, 2020; Meilke, 2021; Preston & Donohoo, 2021). Strong supportive leadership is a predictor of high collective teacher efficacy (Cansoy & Parlor, 2018; Goddard et al., 2017). Preston and Donohoo (2021) believe supportive leaders grow collective efficacy in their teacher teams by ensuring teams experience
success on tasks they believe are beyond their capacity, sharing similar experiences of other teams, setting high expectations with positive reinforcement, and creating and maintaining a positive atmosphere. Progress must be monitored by supportive leadership to help avoid pitfalls, like hidden biases or low expectations, which may work against the benefits of collective teacher efficacy (Hite & Donohoo, 2020; Preston & Donohoo, 2021). A culture of niceness will not sustain teacher collective efficacy if a culture of transparency and a focus on effective practices does not accompany the collegial atmosphere (Preston & Donohoo, 2021).

**Principal Leadership**

Although not reserved solely for the building principal, the role of leadership directly impacts the instructional practices of teachers and improved instructional practices directly improve the achievement of students (Goddard et al., 2010). As the building principal serves as the instructional leader, influences the best practices of teachers, and impacts student achievement, equally important is the leader’s ability to create structures and systems which allow for a culture of trust and communication (Goddard et al., 2015; Gray et al., 2016; Meyer et al., 2020; Tulowitzki & Pietsch, 2018; Warwas et al., 2019). Goddard et al. (2010) found when principals are involved in the creation of collaboration time, team structure, and goals combined with their direct involvement in the instructional work of collaborative teams, it serves as strong predictor of student achievement.

Tichnor-Wagner et al. (2016) reports several factors are present in highly effective schools and one of the most reported is high quality principal leadership. While not effective without other key factors such as formal collaboration systems, universal high expectations, opportunities for participation in leadership, and deliberate supports for students, teachers report
school leaders must play an active role in the school culture and guide instructional best practices (Tichnor-Wagner et al., 2016).

When principals work directly with collaborative teams to modify their instructional practices and these practices directly result in improved student performance, the individuals on the team value the experience and the self-efficacy of the team increases (Meyer et al., 2020). Teacher teams experiencing the impact of leadership involvement will continue to welcome the input of leadership, and there will be a culture of trust allowing the leader opportunities for continued influence. The role of principal leadership in collaborative teams is a key component of how leadership impacts student achievement both directly and indirectly (Eells, 2011; Hite & Donohoo, 2020).

**Professional Learning Communities Model**

Collective teacher efficacy and utilization of the model for professional learning communities (PLC) outlined by DuFour (DuFour & DuFour, 2013) serve as strong predictors of student success (Voelkel & Chrispeels, 2017). The three big ideas representing the PLC model and its core principles include: ensuring students learn, building a culture of collaboration, and focusing on results (DuFour et al., 2016). A true PLC model exists at the organizational level and the building blocks or components of the model are the individual collaborative teams within the organization (DuFour et al., 2016). The PLC model is a complex idea requiring clarity before competence of its execution can be reached.

Utilization of the PLC model within a school means teachers share purpose, vision, commitment, and goals to collaborate in recurring cycles of collective inquiry (DuFour et al., 2016). Sharing a common purpose means teams know why they exist. The primary answer, learning for all students, helps give focus to a team’s work (DuFour et al., 2016). A common
vision helps teams identify where they are headed and what they want to become. A clear vision builds trust and influences decisions made by team members leading to improved student achievement (Bullock & Moyer-Packenham, 2019; DuFour et al., 2016). When collective commitments are shared by all team members, behaviors are guided by these principles and common goals provide the targets and timelines necessary to establish the incremental steps of improvement (Donohoo, 2018). Hattie (2009) describes how these PLC elements are key to school improvement and can be maximized when teachers collaborate about best practices.

PLCs are the most effective way to improve teacher impact and student learning (DuFour & Mattos, 2013). However, these PLCs require strong leadership to make sure they are structured and supported in a way which allows teachers to find value in them (DuFour & Mattos, 2013). PLCs are most effective when school culture is authentically centered around student success, staff are organized into meaningful PLC groups, there is a guaranteed and viable curriculum, evidence of student learning is at the heart of decision-making, and there is a coordinated intervention plan for students who have not reached mastery from core instruction (DuFour & Mattos 2013).

As teachers collaborate in a PLC structure, their conversations should focus around the four critical questions popularized by DuFour (DuFour et al., 2008):

1. What do we want all students to know and be able to do?
2. How will we know if they learn it?
3. How will we respond when some students do not learn?
4. How will we extend the learning for students who are already proficient?

The answer to the first question is embedded in the content standards (DuFour et al., 2016; Erkens et al., 2017; Schimmer et al., 2018). Content standards serve as the guidepost for what
successful learning looks like and the depth of knowledge students should demonstrate (Erkens, 2016; Hillman & Stalets, 2021; Rinkema & Williams, 2018; Schimmer et al., 2018).

Collaborative teams should use unpacked standards in all their work around student learning. The second questions can be answered with a high-quality, balanced assessment system rich in formative, interim and summative assessments directly aligned with content standards and learning targets (Erkens, 2016). With these results and high-quality collaboration, teachers can work on the responses to questions three and four as they plan the interventions and extensions for all students (Erkens, 2016; Hillman & Stalets, 2021; Rinkema & Williams, 2018; Schimmer et al., 2018).

**Gaps in Research**

Researchers have established standards-based teaching and learning (Erkens, 2016; Hillman & Stalets, 2021; Rinkema & Williams, 2018; Schimmer et al., 2018), a balanced assessment system (Chattergoon & Marion, 2016; Erkens, 2017; Marion et al., 2019; Ronfeldt et al., 2015), strong teacher collaboration (Akiba & Liang, 2016; Gusky, 2021; Hite & Donohoo, 2020; Truijen et al., 2013), and principal leadership (Goddard et al., 2015; Gray et al., 2016; Meyer et al., 2020; Tulowitzki & Pietsch, 2018; Warwas et al., 2019) as strong predictors of high student achievement. Studies, both qualitative and quantitative, have examined the impact on student achievement of each of these factors in isolation and in various combinations (McCarthy, 2020). However, in all the studies on teacher collaboration as a predictor of student achievement, the evidence supporting the specific focus on formative assessment results inside of teacher collaboration remains limited. Further, there is even less evidence examining the results on summative assessments when teacher collaboration within a unit of study focuses on the formative assessment results. A need exists to study the impact of effective teacher collaboration
on the results of summative assessments aligned to content standards when teachers use the PLC model, are supported by principal leadership, and focus on formative assessment results aligned to proficiency on learning targets.

**Summary**

Current research on the elements of standards-based teaching and learning, a balanced assessment system, the role of building leadership, and teacher collaboration using the professional learning communities (PLC) model all serve as predictors of student achievement (Eells, 2011; Goddard et al., 2010; Hite & Donohoo, 2020; Shoen et al., 2003; Tchnor-Wagner et al., 2016; Voelkel & Chrispeels, 2017; Wiliam, 2018). The current research on each of these factors describes necessary components for their effectiveness to be measurable. Various studies even show how the combined use of specific components serves as a stronger predictor of student achievement than when the components are studied in isolation (Martuccio & Bloomberg, 2020; McCarthy, 2020; Ronfeldt et al., 2015; Tchnor-Wagner et al., 2016). A strong presence of teacher collaboration frequently coincides with a well-designed, standards-based curriculum and teachers view both as critical to student success.

Assessment, teacher efficacy, and the role of leadership are all structures within teacher collaboration or the PLC model which are identified and studied as critical elements for increasing student achievement (DuFour et al., 2016). Literature and studies show when a teacher utilizes standards-based instruction and curriculum, experience the benefits of a balanced assessment system, engages in a PLC model, and is supported with effective leadership, instruction improves and student achievement increases (Erkens, 2016; Hillman & Stalets, 2021; Rinkema & Williams, 2018; Schimmer et al., 2018). This study will address the gap in literature
which evaluates students’ success on summative assessments when formative assessment results are the focal point for teacher collaboration and drive student intervention.
Chapter 3

Methodology

This chapter outlines the methods and procedures which guided this research study. Chapter 3 includes the following sections: a) the purpose of the study, b) research questions, c) review of related literature, d) research design, e) population and participants, f) instrumentation, g) data collection and analysis, h) procedure, i) context and conditions, and j) summary.

Purpose of the Study

The purpose of this study was to examine student results on a summative unit assessment when teacher collaboration during the unit of study focused on formative assessment results with and without leadership support. Common summative assessment results were gathered from two middle schools, with similar demographics, within the same school district in a rural Midwest state. Teachers in these two schools evaluated the quality of their collaboration during this unit of study using the Teacher Collaboration and Assessment Rubric (TCAR) instrument (see Appendix A). One school had their building leadership actively participate in teacher collaborations and the other school had no direct leadership support during this unit of study. The outcome of this study could assist administrators in making informed decisions about prioritizing their own involvement in teacher collaboration and the design and expectations for teachers within the collaboration times.

Research Questions

This study was guided by the following questions:

1. To what extent does student achievement on a summative assessment differ as a function of the years of experience teaching mathematics at the middle school level in this district?
2. To what extent does student achievement on a summative assessment differ as a function of the years of experience as a team participating in collaboration over formative assessments throughout a unit of study?

3. To what extent does student achievement on a summative assessment differ as a function of the quality of involvement by the classroom teacher in collaboration over formative assessments throughout the unit of study, as determined by self-evaluation using the Teacher Collaboration Assessment Rubric (TCAR)?

4. To what extent does student achievement on a summative assessment differ as a function of involvement of leadership in the teacher collaboration process throughout the unit of study?

5. To what extent does student achievement on a summative assessment differ as a function of the interaction between the quality of involvement by the classroom teacher in collaboration, as determined by self-evaluation using the Teacher Collaboration Assessment Rubric (TCAR) and the involvement of leadership in the collaboration process throughout the unit of study?

**Review of Related Literature**

The review of literature related to standards-based teaching, learning and curriculum, assessment literacy, teacher collaboration using the Professional Learning Communities (PLC) model, and the impact of principal leadership supporting these components consisted of a computer search using Google Scholar, I.D. Weeks library through the University of South Dakota website, and a personal collection of books. Literature reviews, research studies and dissertations were located and reviewed using using EBSCOhost databases which included Education Research Complete (ERIC), ProQuest and PsychINFO.
The researcher used *Google Scholar* and *The Publication Manual of the American Psychological Association* (7th edition, 2020) for formatting purposes. Key words used in this search for relevant literature included: standards-based teaching, standards-based learning, standards-based curriculum, learning targets, success criteria, assessment literacy, formative assessment, summative assessment, teacher collective efficacy, teacher collaboration, leadership, and student achievement.

The review of literature suggests standards-based teaching and learning, a balanced assessment system, the role of building leadership, and teacher collaboration using the professional learning communities (PLC) model all serve as predictors of student achievement (Eells, 2011; Goddard et al., 2010; Hite & Donohoo, 2020; Shoen et al., 2003; Tichnor-Wagner et al., 2016; Voelkel & Chrispeels, 2017; Wiliam, 2018). Each of these elements describes necessary components for their effectiveness to be measurable. Various studies show how the combined use of specific components serves as a stronger predictor of student achievement than when the components are studied in isolation. A strong presence of teacher collaboration frequently utilizes a well-designed, standards-based curriculum with teachers relying on both components to strengthen student achievement.

Assessment, teacher efficacy, and the role of leadership are all structures within teacher collaboration or the PLC model that are identified and studied as critical elements for increasing student achievement (DuFour et al., 2016). Literature and studies show when a teacher utilizes standards-based instruction and curriculum, experiences the benefits of a balanced assessment system, engages in a PLC model, and is supported with effective leadership, instruction improves and student achievement increases (Erkens, 2016; Hillman & Stalets, 2021; Rinkema & Williams, 2018; Schimmer et al., 2018). This study will address the gap in literature which
evaluates students’ success on summative assessments when formative assessment results are the focal point for teacher collaboration and drive student intervention.

Research Design

This study was a quantitative analysis using a factorial ANOVA to study the variance in student summative assessment results when considering the factors of quality of collaboration as measured by the TCAR, the use of common formative assessments during collaboration, and the support of building leadership during collaboration. The factorial AVOVA analysis provided an opportunity to examine how the factors interact and how changes in one factor can influence changes in another factor (Warner, 2013).

Population and Participants

This study took place in two middle schools, School A and School B, within the same school district in a rural, Midwest state. Students at both middle schools have their schedules organized by multidisciplinary core content teams containing one teacher each of math, science, social studies, and English language arts.

Teacher Population

At School A, there are three teams each at sixth, seventh, and eighth grade. School A hosts a specialized language immersion program so one team at each grade level are language immersion learners. At School B, there are four teams each at sixth, seventh, and eighth grade. School B hosts an additional thirteenth team for a specialized program for students with disabilities. Core content teachers on these teams engage in collaboration three times each week with their team peers and two times each week with their job-alike peers.
**Student Population**

The study analyzed the results of all seventh-grade students enrolled in the general seventh grade mathematics class at these two middle schools. Table 1 depicts the composition and demographics of the populations of each school used in this study.

**Table 1**

*Schoolwide Student Demographic Data*

<table>
<thead>
<tr>
<th>Demographics</th>
<th>School A</th>
<th>School B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Total Students</td>
<td>903</td>
<td>100</td>
</tr>
<tr>
<td>Male</td>
<td>434</td>
<td>48.1</td>
</tr>
<tr>
<td>Female</td>
<td>469</td>
<td>51.9</td>
</tr>
<tr>
<td>Caucasian</td>
<td>539</td>
<td>59.7</td>
</tr>
<tr>
<td>American Indian, Alaska Native or Native Hawaiian</td>
<td>42</td>
<td>4.7</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>159</td>
<td>17.6</td>
</tr>
<tr>
<td>Black or African American</td>
<td>86</td>
<td>9.5</td>
</tr>
<tr>
<td>Students with Disability</td>
<td>117</td>
<td>13.0</td>
</tr>
<tr>
<td>English Learner</td>
<td>74</td>
<td>8.2</td>
</tr>
</tbody>
</table>

*Note.* Data was retrieved from Infinite Campus Student Information System.

**Teacher Participants**

This study used students’ summative assessment results on a common unit of study from one of six different math teachers in one of the two identified middle schools who met the requirements of collaborating full time during the unit of study. Three seventh grade mathematics teachers from School A, team A, are in their first year of collaborating as a job-alike team and did not have building leadership present at their collaboration meetings. Three seventh grade mathematics teachers from School B, team B, are in their eighth year of collaborating as a job-like team and routinely has leadership present at their collaboration meetings. During this unit of study, a building administrator in School B was present for the collaboration on the results of common formative assessments given during the unit.
**Student Participants**

This study examines the summative assessment results of 424 graders enrolled in seventh-grade mathematics at one of the two identified middle schools. Student participants were both male and female with race and socio-economic demographics similarly distributed from class to class as enrollment in courses was randomly generated by the scheduler within the student information system. All students who attended these middle schools came from housing areas with common average home values and average family incomes. Table 2 depicts the composition and demographics of the student participants of each school used in this study.

**Table 2**

<table>
<thead>
<tr>
<th>Demographics</th>
<th>School A</th>
<th></th>
<th>School B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Students enrolled in 7th Grade Mathematics</td>
<td>183</td>
<td>100</td>
<td>241</td>
<td>100</td>
</tr>
<tr>
<td>Male</td>
<td>85</td>
<td>46.4</td>
<td>114</td>
<td>47.3</td>
</tr>
<tr>
<td>Female</td>
<td>98</td>
<td>53.6</td>
<td>127</td>
<td>52.7</td>
</tr>
<tr>
<td>Gifted</td>
<td>2</td>
<td>1.1</td>
<td>6</td>
<td>2.5</td>
</tr>
<tr>
<td>English Learners</td>
<td>4</td>
<td>2.2</td>
<td>9</td>
<td>3.3</td>
</tr>
<tr>
<td>Students with Disability</td>
<td>16</td>
<td>8.7</td>
<td>26</td>
<td>10.8</td>
</tr>
</tbody>
</table>

*Note.* Data was retrieved from Performance Matters Comparative Results.

**Instrumentation**

Teachers in this study evaluated the quality of their collaboration during this unit of study using the TCAR (see Appendix A). The rubric was replicated into a Google Form, completed online by teachers, and results were obtained through a Google Sheet.

The common summative assessment (see Appendix B) was created by the seventh-grade mathematics assessment team consisting of a seventh-grade mathematics teacher from each school in the district. Students completed this test using Performance Matters, the district’s summative assessment platform. Student-selected responses were scored by the platform and
teachers scored the student-created responses in the platform utilizing a common rubric for each question (see Appendix C). All questions on the summative assessment align to the seventh-grade state mathematics standards covered in this unit of study. All teachers in the district use the same curriculum resources, standards and learning targets as they have been mapped, unpacked, and prioritized at the district level.

**Data Collection and Analysis**

Data was downloaded from Performance Matters Reports, the district’s summative and high stakes testing platform, via a spreadsheet. The spreadsheet was stored in a password-protected Google Drive folder. Using the *R Studio* software, the data was loaded with values in three columns. The first column coded the students’ scores to align with a nominal value indicating leadership presence during teacher collaborations, 0 = leadership not present and 1 = leadership present, the second column coded the students’ scores to a nominal value aligned with the overall teacher rating on the TCAR, 1 = beginning, 2 = emerging, 3 = developing, 4 = advancing, and 5 = proficient, the third column contained the ordinal value of the students’ percent score on the common summative assessment. Data was analyzed for descriptive statistics and assumptions of homogeneity of variance were tested both by statistical tests and graphically.

To analyze the variance in student scores between the ten different groups shown in Table 3, from two leadership factors and five different teacher ratings on the TCAR, a factorial ANOVA was performed to test for variance in mean values when the variables interacted, and a Cohen’s D test was run to compare the effect size on student achievement of one group against another group.
Table 3

Factors Considered in Interaction Analysis

<table>
<thead>
<tr>
<th>Factor B (Nominal value)</th>
<th>Factor A (Nominal value)</th>
<th>Leadership Present (n)</th>
<th>Leadership Not Present (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCAR overall rating = 1</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TCAR overall rating = 2</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TCAR overall rating = 3</td>
<td></td>
<td>73</td>
<td>44</td>
</tr>
<tr>
<td>TCAR overall rating = 4</td>
<td></td>
<td>161</td>
<td>63</td>
</tr>
<tr>
<td>TCAR overall rating = 5</td>
<td></td>
<td>0</td>
<td>65</td>
</tr>
</tbody>
</table>

Note. Data was retrieved from Performance Matters Comparative Results.

Procedure

To determine if leadership involvement interacting with teacher collaboration over formative assessments during a unit of study is a predictor of student achievement, two independent variables and one dependent variable were used. The independent variables in this study were leadership presence or absence at collaborations during this unit of study and the teachers’ ratings on collaboration activities and quality during the unit of instruction. The dependent variable is the student score on the common summative unit assessment.

Two independent teams of teachers from two schools worked through a common unit of study in seventh grade mathematics, used common student and teacher resources, followed a common curriculum guide, and administered a common summative assessment to all students at the end of the unit. Both teams were given no additional expectations for collaboration outside of the ongoing district-level training on the PLC model and using formative assessments to guide instruction and intervention. Accountability for meeting these expectations lies at the discretion of the building principal. The district expectation is for teachers to collaborate as a job-alike team on seventh grade mathematics two times per week. The quantity and quality of the common formative assessments used during the unit are determined by the individual teams and only
common among team members, not between teams. All teachers on both teams administered the same common summative assessment at the end of the unit of study. The unit of study lasted approximately 4-5 weeks but varied slightly from teacher to teacher.

Teachers were given the TCAR via a Google Form after all student results were finalized on the common summative assessment and asked to evaluate their own perceptions of collaboration time during the unit of study.

With the large effect size and sample size collected by the teachers for the ten groups, the post hoc statistical power calculation exceeded 0.999 meaning there is less than 1% chance of failing to detect an effect exists.

**Context of the Conditions**

The participating school district in this study was midway through a 5-year implementation of standards-based teaching and learning. A large component of this implementation relied on teachers applying their new knowledge to their independent practices. Particularly, when instructional leadership could not be present during a lesson, collaboration meeting, or assessment delivery, district leadership trusted teachers were implementing the components deemed as best practices. The results of this study will provide teachers with data showing how the quality of collaboration when interacting with leadership support is a predictor of student achievement. The results on how predictive leadership involvement in teacher collaboration is for higher student achievement will also be useful to principals as they make instructional leadership decisions.

**Summary**

The purpose of this quantitative analysis of student achievement data was to examine how the interaction between leadership involvement in teacher collaboration and teachers’
assessment of the quality and activities of their collaboration serves as a predictor of student achievement in a rural, Midwest state. Teachers used a common curriculum guide, common instructional resources, and administered a common summative assessment aligned to state standards. During the unit of study, teachers participated in collaboration with or without their leadership’s support and involvement. After completing the unit of study, teachers self-reported their evaluation on the quality and activities of their collaboration using the TCAR. The TCAR scores were collected through a Google Form and student assessment scores were downloaded from Performance Matters, the district’s summative assessment platform. While Chapter 3 described the methodology and research design for this study, Chapter 4 will outline the results and findings of the study.
Chapter 4

Results

Chapter 4 provides an overview and results of the study. The chapter is organized into five sections: a) purpose of study, b) research questions, c) demographic data, e) findings, and f) summary.

Purpose of the Study

The purpose of this study was to examine student results on a summative unit assessment when teacher collaboration during the unit of study focused on formative assessment results with and without leadership support. Common summative assessment results were gathered from two middle schools, with similar demographics, within the same school district in a rural Midwest state. Teachers in these two schools evaluated the quality of their team’s collaboration during this unit of study using the Teacher Collaboration and Assessment Rubric (TCAR) instrument (see Appendix A). One school had their building leadership actively participate in teacher collaborations and the other school had no direct leadership support during this unit of study. The outcome of this study could assist administrators in making informed decisions about prioritizing their own involvement in teacher collaboration and the design and expectations for teachers within the collaboration times.

Research Questions

This study was guided by the following questions:

1. To what extent does student achievement on a summative assessment differ as a function of the years of experience teaching mathematics at the middle school level in this district?
2. To what extent does student achievement on a summative assessment differ as a function of the years of experience as a team participating in collaboration over formative assessments throughout a unit of study?

3. To what extent does student achievement on a summative assessment differ as a function of the quality of involvement by the classroom teacher in collaboration over formative assessments throughout the unit of study, as determined by self-evaluation using the Teacher Collaboration Assessment Rubric (TCAR)?

4. To what extent does student achievement on a summative assessment differ as a function of involvement of leadership in the teacher collaboration process throughout the unit of study?

5. To what extent does student achievement on a summative assessment differ as a function of the interaction between the quality of involvement by the classroom teacher in collaboration, as determined by self-evaluation using the Teacher Collaboration Assessment Rubric (TCAR) and the involvement of leadership in the collaboration process throughout the unit of study?

**Demographic Data**

Overall, 424 seventh graders enrolled in seventh-grade mathematics at one of the two identified middle schools participated in a common summative assessment at the end of a common unit of study. Student participants were both male and female with race and socio-economic demographics similarly distributed among class periods as enrollment in classes was randomly generated by the scheduler within the student information system. All students who attended these middle schools resided in housing areas with common average home values and
average family incomes. Table 4 depicts the composition, demographics, and average score on the summative assessment of the student participants of each school used in this study.

Table 4

Student Participants Demographic Data

<table>
<thead>
<tr>
<th>Demographics</th>
<th>School A</th>
<th></th>
<th>School B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>Avg. Score</td>
<td>n</td>
</tr>
<tr>
<td>Total enrolled in 7th Grade Mathematics</td>
<td>183</td>
<td>100%</td>
<td>66.5%</td>
<td>241</td>
</tr>
<tr>
<td>Male</td>
<td>85</td>
<td>46.4%</td>
<td>63.6%</td>
<td>114</td>
</tr>
<tr>
<td>Female</td>
<td>98</td>
<td>53.6%</td>
<td>68.6%</td>
<td>127</td>
</tr>
<tr>
<td>Gifted</td>
<td>2</td>
<td>1.1%</td>
<td>67.2%</td>
<td>6</td>
</tr>
<tr>
<td>English Learners</td>
<td>4</td>
<td>2.2%</td>
<td>62.6%</td>
<td>9</td>
</tr>
<tr>
<td>Students with Disability</td>
<td>16</td>
<td>8.7%</td>
<td>54.3%</td>
<td>26</td>
</tr>
</tbody>
</table>

Note. Data was retrieved from Performance Matters Comparative Results.

Teachers leading the instruction of the student-participants in this study varied in years of experience teaching mathematics at the middle school level. These teachers also vary in years they have collaborated as a team and how they view the quality of their collaboration as measured by the Teacher Collaboration Assessment Rubric (TCAR). Additionally, frequency of leadership presence during collaboration varied from one team to the other during this study.

Table 5 shows the number of teachers in each 5-year band of years of experience teaching middle school mathematics.

Table 5

Number of Teachers by Years of Experience Teaching Middle School Mathematics

<table>
<thead>
<tr>
<th>Years of Experience Teaching Middle School Mathematics</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 5</td>
<td>2</td>
</tr>
<tr>
<td>5 to 9</td>
<td>2</td>
</tr>
<tr>
<td>10 to 14</td>
<td>1</td>
</tr>
<tr>
<td>15 to 19</td>
<td>0</td>
</tr>
<tr>
<td>20 to 24</td>
<td>1</td>
</tr>
</tbody>
</table>

Note. Data was retrieved from Infinite Campus student information system.
Table 6 shows the number of students enrolled with teachers based on the teachers’ years of experience teaching middle school mathematics in bands of five years.

**Table 6**

*Student Participation by Experience Teaching Middle School Mathematics*

<table>
<thead>
<tr>
<th>Years of Experience Teaching Middle School Mathematics</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 5</td>
<td>113</td>
</tr>
<tr>
<td>5 to 9</td>
<td>161</td>
</tr>
<tr>
<td>10 to 14</td>
<td>80</td>
</tr>
<tr>
<td>15 to 19</td>
<td>0</td>
</tr>
<tr>
<td>20 to 24</td>
<td>70</td>
</tr>
</tbody>
</table>

*Note.* Data was retrieved from Infinite Campus student information system.

Table 7 shows the number of students enrolled with teachers based the teachers’ years of collaborating as their current seventh grade mathematics job-alike team.

**Table 7**

*Student Participation by Teachers’ Years of Collaborating as a Team*

<table>
<thead>
<tr>
<th>Years Collaborating as Team</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Experience: Less than 5</td>
<td>183</td>
</tr>
<tr>
<td>High Experience: 5 or more</td>
<td>241</td>
</tr>
</tbody>
</table>

*Note.* Data was retrieved from Performance Matters Comparative Results.

Table 8 shows the number of teachers who evaluated the quality of collaboration using the TCAR by rating value. The TCAR has nine indicators with three levels of success criteria for each indicator. Logic rules were used to convert teacher responses to a 5-point rating scale.
Table 8

*Number of Teachers by TCAR Rating*

<table>
<thead>
<tr>
<th>TCAR Score</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Low</td>
<td>0</td>
</tr>
<tr>
<td>2: Low_Avg</td>
<td>0</td>
</tr>
<tr>
<td>3: Avg</td>
<td>2</td>
</tr>
<tr>
<td>4: Avg_High</td>
<td>3</td>
</tr>
<tr>
<td>5: High</td>
<td>1</td>
</tr>
</tbody>
</table>

*Note.* Data was retrieved from TCAR Google Form.

Table 9 shows the number of students enrolled by teacher TCAR rating.

Table 9

*Student Participation by Teacher TCAR Rating*

<table>
<thead>
<tr>
<th>TCAR Score</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Low</td>
<td>0</td>
</tr>
<tr>
<td>2: Low_Avg</td>
<td>0</td>
</tr>
<tr>
<td>3: Avg</td>
<td>127</td>
</tr>
<tr>
<td>4: Avg_High</td>
<td>232</td>
</tr>
<tr>
<td>5: High</td>
<td>65</td>
</tr>
</tbody>
</table>

*Note.* Data was retrieved from TCAR Google Form and Performance Matters Comparative Results.

Table 10 shows the number of students enrolled with teachers based the presence or absence of administration during teacher collaboration meetings throughout this unit of study.

Table 10

*Student Participation by Administration Presence*

<table>
<thead>
<tr>
<th>Administration Present during Teacher Collaboration</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>241</td>
</tr>
<tr>
<td>No</td>
<td>183</td>
</tr>
</tbody>
</table>

*Note.* Data was retrieved from Performance Matters Comparative Results.
Findings

Overall, 424 seventh grade students in mathematics class participated in the common summative assessment after the completion of Unit 4. To meet the assumption of relatively equal sample sizes for the pairing of independent variables, a random sample of 81 student scores were taken from the population of students whose teacher had a TCAR rating of 4 (Avg_High) and administration present and results were analyzed to address five research questions in this study. The summative assessment scores were ratio values of raw scores ranging from a minimum of 0 to a maximum of 1.0, with an average score of 0.65 (Mdn = 0.69) and a standard deviation of 0.23.

Assumptions

Results of the Shapiro-Wilk test of normality suggest overall student achievement on the summative assessment deviates from normality (W=0.919 and p-value < 0.001). However, a histogram of assessment scores shown in Figure 1 shows a visual of the moderate skew and kurtosis (skewness = -0.82, kurtosis = -0.03) combined with the truncation of data on the right side due to a maximum score of 1.0. Examining these results indicated little skew was present and the scores are approximately normally distributed.

Figure 1

*Student Summative Assessment Scores*
Homogeneity of variance across subgroups of the dependent variable, student achievement on the summative assessment, was analyzed and is described within the context of each research question.

Based on the design of this study, the assumption of independence was met for the overall data set of student scores on the summative assessment, as well as each subset of scores. There was no interaction between participants as each student was only assigned to one teacher and a student score could only correspond to one teacher.

**The extent student achievement differs as a function of the years of experience teaching mathematics at the middle school level.** Research question one investigated the extent which student achievement on a summative assessment differs as a function of the years of experience a student’s teacher has teaching mathematics at the middle school level in this district. The student assessment scores were gathered from the assessment platform and data on experience teaching mathematics at the middle level was gathered from the student information system. Years of teaching experience for student sample were ratio values ranging from a minimum of 1 to a maximum of 23, with an average of 9.38 (Mdn = 7) and a standard deviation of 7.93. Years of teaching experience were banded into incremental groups of five years.

**Assumptions**

Results from the Levene’s test on the equality of variances between groups suggested there is equality in the variances of each subgroup in the sample (p = 0.301). This assumption can be validated by the Bartlett test (p = 0.267), and we fail to reject the null hypothesis. Visual inspection of the variances in Figure 2 confirms the variances.
Figure 2

*Homogeneity of Variance in RQ1 Subsets*

![Homogeneity of Variance in RQ1 Subsets](image_url)

**Analysis**

Results from the analysis of variance (ANOVA) suggested there was enough evidence to reject the null hypothesis that the mean values of student achievement on the summative assessment in seventh grade mathematics as a function of years of experience teaching middle school mathematics in this district were significantly different ($F(3, 335) = 20.65$, $p < 0.001$). Overall, years of experience teaching middle school mathematics in this district accounted for a large amount of variance in math achievement, or 15.6%, with a $\eta^2 = 0.156$.

Given a significant omnibus test (the initial ANOVA), post-hoc comparisons were examined to determine the extent of variability in student summative assessment scores, as a function of the teacher’s years of experience teaching middle school mathematics in this district. Results from the Tukey’s HSD, shown in Table 11, suggested there is a strongly significant difference in three of the comparison groups of years of teaching experience ($p$ adj $< 0.1$).
Table 11

*Differences and Significances in Mean Values between Pairings*

<table>
<thead>
<tr>
<th>Group 1 (Years of Experience)</th>
<th>Group 2 (Years of Experience)</th>
<th>diff</th>
<th>p adj.</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 to 9</td>
<td>Less than 5</td>
<td>-0.031</td>
<td>0.670</td>
</tr>
<tr>
<td>10 to 14</td>
<td>Less than 5</td>
<td>0.039</td>
<td>0.759</td>
</tr>
<tr>
<td>20 or more</td>
<td>Less than 5</td>
<td>-0.229</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>10 to 14</td>
<td>5 to 9</td>
<td>0.070</td>
<td>0.290</td>
</tr>
<tr>
<td>20 or more</td>
<td>5 to 9</td>
<td>-0.198</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>20 or more</td>
<td>10 to 14</td>
<td>-0.268</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

*Note.* Table compares the percent of differences in the mean values of the different pairings and their adjusted p-value for significance using Tukey HSD test.

The extent student achievement differs as a function of the years collaborating as seventh grade mathematics team. Research question two investigated the extent which student achievement on a summative assessment differs as a function of the years of experience the student’s teacher has collaborating as a seventh-grade job-like team. The student assessment scores were gathered from the assessment platform and data on years of collaborating as a job-like team was gathered from the student information system. Years of experience collaborating as a team were ratio values ranging from a minimum of 1 to a maximum of 7, with an average of 3.73 (Mdn = 1) and a standard deviation of 2.99. Years of experience collaborating as a team were banded into two incremental groups.

*Assumptions*

Results from the Levene’s test on the equality of variances between groups suggested there is enough variance in the sample to reject the null hypothesis of homogeneity (p = 0.004). While p < 0.05, the exact p-value indicates there is slight significance, and it would make sense
to fail to reject the null hypothesis of homogeneity based on the unequal sample sizes from the two groups. Figure 3 shows a visual representation of variance by group.

**Figure 3**

*Homogeneity of Variance in RQ2 Subsets*

![Figure 3: Homogeneity of Variance in RQ2 Subsets](image)

**Analysis**

Results from the analysis of variance (ANOVA) suggested there is a statistically significant variance in the means of student achievement on the summative assessment in seventh grade mathematics as a function of years of collaborating as a job-like team ($F(1, 337) = 8.631, p = 0.004$). Overall, years of collaborating as a job-like team accounted for a small effect, or 2.5%, on student achievement in summative mathematics assessments with an $\eta^2 = 0.025$.

Given a significant omnibus test (the initial ANOVA), post-hoc comparisons were examined to determine the extent of variability in student summative assessment scores, as a function of the years of experience collaborating as a team. Results from the Tukey’s HSD, shown in Table 12, suggested there is a strongly significant difference, 7.4%, in the comparison groups of years of experience collaborating as a team ($p \text{ adj} < 0.05$).
Table 12

*Differences and Significances in Mean Values between Pairings*

<table>
<thead>
<tr>
<th>Group 1 (Years of Experience)</th>
<th>Group 2 (Years of Experience)</th>
<th>diff</th>
<th>p adj.</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Experience</td>
<td>Low Experience</td>
<td>0.075</td>
<td>0.004</td>
</tr>
</tbody>
</table>

*Note.* Table compares the percent of differences in the mean values of the different pairings and their adjusted p-value for significance using Tukey HSD test.

**The extent student achievement differs as a function of the evaluated quality of collaboration by the individual teacher.** Research question three investigated the extent which student achievement on a summative assessment differs as a function of individual teacher’s evaluation of the quality of collaboration as determined by the TCAR. The student assessment scores were gathered from the assessment platform and quality of collaboration was determined by the TCAR administered via a Google Form. The results of the TCAR were converted into a rating scale of 1 to 5 using logic rules. Quality of collaboration, as determined by the teacher results of the TCAR, were ratio values ranging from a minimum of 3 to a maximum of 5, with an average of 3.83 (Mdn = 4) and a standard deviation of 0.73. Additionally, the administrator from each school completed the TCAR evaluating the quality of collaboration for their seventh-grade mathematics team. The administrator TCAR rating was 3 at the building with administration present during collaboration and 2 at the building with the administration not present during collaboration.

**Assumptions**

Results from the Levene’s test on the equality of variances between groups provided significant evidence to not reject the null hypothesis and verified all subgroups had equal
variance \((p = 0.077)\). This assumption was validated by the Bartlett test with a \(p\)-value = 0.058. Homogeneity of variance in the subgroups by TCAR score can be seen in Figure 4.

**Figure 4**

*Homogeneity of Variance in RQ3 Subsets*

Analysis

Results from the analysis of variance (ANOVA) suggested there is a statistically significant variance in the means of student achievement on the summative assessment in seventh grade mathematics as a function of teacher self-evaluation of quality of collaboration as determined by the TCAR \((F_{2,336} = 4.628, p = 0.010)\). Overall, teacher self-evaluation of quality of collaboration as determined by the TCAR accounted for a small amount, 2.7%, of variance in the mean values of mathematics achievement with a \(\eta^2 = 0.027\).

Given a significant omnibus test (the initial ANOVA), post-hoc comparisons were examined to determine the extent of variability in student summative assessment scores, as a function of individual teacher’s evaluation of the quality of collaboration as determined by the
TCAR. Results from the Tukey’s HSD, shown in Table 13, suggested there is a significant
difference, 9.4%, in one of the comparison groups of TCAR ratings (p adj < 0.05).

Table 13

*Differences and Significances in Mean Values between Pairings*

<table>
<thead>
<tr>
<th>Group 1 (TCAR rating)</th>
<th>Group 2 (TCAR rating)</th>
<th>diff</th>
<th>p adj.</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>3</td>
<td>-0.064</td>
<td>0.670</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>0.030</td>
<td>0.759</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>0.094</td>
<td>0.019</td>
</tr>
</tbody>
</table>

*Note.* Table compares the percent of differences in the mean values of the different pairings and their adjusted p-value for significance using Tukey HSD test.

The extent student achievement differs as a function of the presence of administration during teacher collaboration. Research question four investigated the extent which student achievement on a summative assessment differs as a function of the presence or absence of administration during the unit of study. The student assessment scores were gathered from the assessment platform and the presence of administration at collaboration during this unit was determined by an email to each principal. Presence of administration during team collaboration were nominal values of 1 = yes, present, and 2 = no, not present.

**Assumptions**

Results from the Levene’s test on the equality of variances between groups suggested there is enough variance in the sample to reject the null hypothesis of homogeneity (p = 0.004). While p < 0.05, the exact p-value indicates there is slight significance, and it would make sense to fail to reject the null hypothesis of homogeneity based on the unequal sample sizes from the two groups. Figure 3 shows a visual representation of variance by group.
Analysis

Results from the analysis of variance (ANOVA) suggested there is not statistically significant variance in the means of student achievement on the summative assessment in seventh grade mathematics as a function of administration present during collaboration ($F_{(1, 337)} = 8.631, p = 0.004$). Overall, years of collaborating as a job-like team accounted for a small amount of variance, 2.5%, in math achievement with a $\eta^2 = 0.025$.

Given a significant omnibus test (the initial ANOVA), post-hoc comparisons were examined to determine the extent of variability in student summative assessment scores, as a function of administration presence during team collaboration. Results from the Tukey’s HSD, shown in Table 14, suggested there is a strongly significant difference, 7.4%, in the comparison groups of administration presence during team collaboration (p adj < 0.05).
Table 14

*Differences and Significances in Mean Values between Pairings*

<table>
<thead>
<tr>
<th>Group 1 (Admin Presence)</th>
<th>Group 2 (Admin Presence)</th>
<th>diff</th>
<th>p adj.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Yes</td>
<td>0.075</td>
<td>0.004</td>
</tr>
</tbody>
</table>

*Note.* Table compares the percent of differences in the mean values of the different pairings and their adjusted p-value for significance using Tukey HSD test.

The extent student achievement differs as a function of the interaction between quality of collaboration and presence of administration during teacher collaboration.

Research question five investigated the extent which student achievement on a summative assessment differs as a function of the interaction between the quality of involvement by the classroom teacher in collaboration, as determined by self-evaluation using the Teacher Collaboration Assessment Rubric (TCAR) and the involvement of leadership in the collaboration process throughout the unit of study. The student assessment scores were gathered from the assessment platform, the quality of collaboration was determined by the TCAR administered via a Google Form, and the presence of administration at collaboration during this unit was determined by an email to each principal.

**Assumptions**

Sample sizes in the different pairings have a similar n-size. These n-sizes can be seen in Table 15.
Table 15

Student Participation by Administration Presence

<table>
<thead>
<tr>
<th>Administration Present:TCAR Rating</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes:Avg</td>
<td>73</td>
</tr>
<tr>
<td>Yes:Avg_High</td>
<td>81</td>
</tr>
<tr>
<td>Yes:High</td>
<td>0</td>
</tr>
<tr>
<td>No:Avg</td>
<td>50</td>
</tr>
<tr>
<td>No:Avg_High</td>
<td>70</td>
</tr>
<tr>
<td>No:High</td>
<td>65</td>
</tr>
</tbody>
</table>

Note. Data was retrieved from Performance Matters Comparative Results.

Most pairings (administration presence X TCAR rating) produced students summative assessment scores which suggest deviation from normality when applying the Shapiro-Wilk Test. All pairings produced p-values smaller than 0.05. However, when considering the maximum score of a 1.0 on the summative assessment, the small sample sizes, and the fact that all Shapiro-Wilks tests had a W-value > 0.856, normality in the pairings exists and is parallel to the normality in the overall scores on the summative assessment. Table 16 shows the W-value and p-value for each pairing.

Table 16

Shapiro-Wilk Test for Normality on Pairings

<table>
<thead>
<tr>
<th>Pairing</th>
<th>W-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Admin.Presence:TCAR Rating)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes:Avg</td>
<td>0.937</td>
<td>0.001</td>
</tr>
<tr>
<td>Yes:Avg_High</td>
<td>0.908</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Yes:High</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>No:Avg</td>
<td>0.856</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No:Avg_High</td>
<td>0.920</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No:High</td>
<td>0.921</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Student scores on the summative assessment do not violate the assumption of homogeneity of variance as Levene’s test produced a p-value = 0.143. Homogeneity of variance exists and can be seen in Figure 6.
By design of the study, there was not a connection between the participants of the five pairings created by grouping administration presence with TCAR rating. No student was enrolled in more than one pairing.

Analysis

Results from the analysis of variance (ANOVA) suggested there is statistically significant variance in the means of student summative assessment scores as a function of administration presence when interacting with evaluation of collaboration according to the TCAR rating \( (F_{4.334} = 15.85, \ p < 0.001) \). Overall, evaluation of collaboration according to the TCAR rating interacting with administration presence accounted for a medium amount of variance in the mean values of student summative assessment scores with an \( \eta^2 = 0.089 \), meaning 8.9% of the variance could be explained.
Given a statistically significant omnibus test (the initial ANOVA), post-hoc comparisons were examined to determine the extent of variability in student summative assessment scores, as a function of all the interactions between Administration presence when interacting with evaluation of collaboration according to the TCAR rating. Results from the Tukey’s HSD suggested there is a significant difference in the means for student summative assessment scores in four of the pairing comparisons. Results of the differences in mean values between all pairings, along with their adjusted p-valued can be seen in Table 17.

**Table 17**

*Differences and Significances in Mean Values between Pairings*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg:No</td>
<td>Avg:Yes</td>
<td>0.052</td>
<td>0.778</td>
</tr>
<tr>
<td>Avg_High:Yes</td>
<td>Avg:Yes</td>
<td>0.074</td>
<td>0.275</td>
</tr>
<tr>
<td>Avg_High:No</td>
<td>Avg:Yes</td>
<td>-0.178</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>High:No</td>
<td>Avg:Yes</td>
<td>0.051</td>
<td>0.736</td>
</tr>
<tr>
<td>Avg_High:Yes</td>
<td>Avg:No</td>
<td>0.022</td>
<td>0.993</td>
</tr>
<tr>
<td>Avg_High:No</td>
<td>Avg:No</td>
<td>-0.229</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>High:No</td>
<td>Avg:No</td>
<td>&lt;0.001</td>
<td>1.000</td>
</tr>
<tr>
<td>Avg_High:No</td>
<td>Avg_High:Yes</td>
<td>-0.252</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>High:No</td>
<td>Avg_High:Yes</td>
<td>-0.023</td>
<td>0.988</td>
</tr>
<tr>
<td>High:No</td>
<td>Avg_High:No</td>
<td>0.229</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

*Note.* Table compares the percent of differences in the mean values of the different pairings and their adjusted p-value for significance using Tukey HSD test.

When comparing the pairings of administration presence and a teacher’s self-evaluation of the quality of collaboration during the unit of study with the Cohen’s D test, the effective size in summative assessment scores can be measured. In Table 18, you can view the effect size of
the ten comparisons. Any Cohen’s D value between 0.20 and 0.49 is considered a small effect, between 0.50 and 0.79 is considered a medium effect size and any value 0.80 and over is considered a large effect size. These values should be considered in their absolute value as a negative value only determines the second group has a larger effect than the first group listed.

### Table 18

*Effect Sizes of Interaction Comparisons*

<table>
<thead>
<tr>
<th>Group 1 (TCAR:Admin Presence)</th>
<th>Group 2 (TCAR:Admin Presence)</th>
<th>Cohen’s D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg_High:Yes</td>
<td>High:No</td>
<td>0.116</td>
</tr>
<tr>
<td>Avg:Yes</td>
<td>High:No</td>
<td>-0.265</td>
</tr>
<tr>
<td>Avg_High:No</td>
<td>High:No</td>
<td>-1.054</td>
</tr>
<tr>
<td>Avg:No</td>
<td>High:No</td>
<td>0.004</td>
</tr>
<tr>
<td>Avg:Yes</td>
<td>Avg_High:Yes</td>
<td>-0.377</td>
</tr>
<tr>
<td>Avg_High:No</td>
<td>Avg_High:Yes</td>
<td>-1.153</td>
</tr>
<tr>
<td>Avg:No</td>
<td>Avg_High:Yes</td>
<td>-0.097</td>
</tr>
<tr>
<td>Avg_High:No</td>
<td>Avg:Yes</td>
<td>-0.833</td>
</tr>
<tr>
<td>Avg:No</td>
<td>Avg:Yes</td>
<td>0.233</td>
</tr>
<tr>
<td>Avg:No</td>
<td>Avg_High:No</td>
<td>0.927</td>
</tr>
</tbody>
</table>

*Note.* Table compares the effect size of two different pairings of the independent variables.

The summary of mean values of student achievement on the summative assessment when pairing administration presence and teacher self-evaluation of the quality of collaboration can be viewed in Figure 7.
Chapter 4 provided results of the analysis performed on the quantitative data in this study. The purpose of this study was to examine student results on a summative unit assessment when teacher collaboration during the unit of study focused on formative assessment results with and without leadership support. The study explored effect of teacher’s years of experience teaching at the middle school level in this district, years of experience each team has collaborating, administration’s presence or absence during the unit of study, teacher’s evaluation of collaboration using the TCAR rubric, and the interaction between administration presence and the teacher’s TCAR rating.

This study was guided by five research questions. Data on the student summative assessment scores was downloaded from the district assessment platform. Each student score was paired with the years of experience teaching at the middle school in this district, the years of experience working at a collaborative teaching team, the administration’s presence during the

---

**Figure 7**

*Plot of Factorial ANOVA*

---

**Summary**

Chapter 4 provided results of the analysis performed on the quantitative data in this study. The purpose of this study was to examine student results on a summative unit assessment when teacher collaboration during the unit of study focused on formative assessment results with and without leadership support. The study explored effect of teacher’s years of experience teaching at the middle school level in this district, years of experience each team has collaborating, administration’s presence or absence during the unit of study, teacher’s evaluation of collaboration using the TCAR rubric, and the interaction between administration presence and the teacher’s TCAR rating.

This study was guided by five research questions. Data on the student summative assessment scores was downloaded from the district assessment platform. Each student score was paired with the years of experience teaching at the middle school in this district, the years of experience working at a collaborative teaching team, the administration’s presence during the
unit of study, and the TCAR rating of each student's teacher. Data indicated some significant relationships between student achievement on the summative assessment and teachers’ years of experience teaching mathematics at the middle school level in this district. When administration presence interacted with a teacher’s TCAR rating, large effect sizes were detected when comparing the different pairing combinations.

Although there were significant results when analyzing student achievement as a function of years collaborating as a team and administration presence, their results were identical causing an unreliable conclusion as to which variable is affecting the change.
References


http://www.ascd.org/publications/educational-leadership/sept12/vol70/num01/Preventing-Feedback-Fizzle.aspx


https://doi.org/10.1080/09585176.2015.1103278


Erkens, C., Schimmer, T., & Vagle, N. D. (2017). Essential assessment: Six tenets for bringing hope, efficacy, and achievement to the classroom. Solution Tree Press, the Solution Tree Assessment Center.


Hill, T. C. (2013). Common formative assessments developed through professional learning communities (PLCs): A case study to analyze the alignment of curriculum, assessment, and instruction in a math PLC at a Title I middle school in the southern United States. (3527150) [Doctoral dissertation, Texas A&M University]. ProQuest Dissertations Publishing.


Lopez, N., Patrick, S., & Sturgis, C. (2017). Quality and Equity by Design: Charting the Course for the Next Phase of Competency-Based Education. CompetencyWorks. iNACOL.

https://digitalcommons.georgiasouthern.edu/nyar_savannah/2020/2020/96/


http://www.ascd.org/publications/educational_leadership/mar99/vol56/num06/Realizing_the_Promise_of_Standards-Based_Education.aspx


https://doi.org/10.14507/epaa.v18n19.2010

Chapter 5

Seventh Grade Mathematics Results in a Standards-Based Learning Environment:
Examining Teacher Collaboration and Leadership

By: Demetria C. Moon

Abstract:

The pressure on schools to improve student performance in mathematics combined with the movement to standards-based instruction leaves teachers searching for the most effective instructional practices and administration examining their role in improvement process. This study examines the effect on student achievement in middle school mathematics when the quality of teacher collaboration interacts with administration involvement during collaboration. This study used a quantitative research design to analyze student results on a common summative assessment in seventh-grade mathematics. Additionally, all teachers evaluated the quality of their job-alike collaboration during this unit using a provided rubric. The results of this study suggest direct involvement of leadership during collaboration and instruction serves as a strong predictor of student achievement.
Introduction

Maximizing student achievement in mathematics to stay competitive on a global scale has been the primary focus for the American educational community for years (Bush, 2019; Gravemeijer et al., 2017; Ropohl et al., 2018; Wheat, 2021). The drive to compete has enacted many governmental initiatives, such as the reauthorization of the Elementary and Secondary Education Act in 2001, No Child Left Behind (NCLB) of 2002, and Every Student Succeeds Act (ESSA) of 2015, calling for education agencies to prove the effectiveness of their instructional practices through state-wide, and in many instances, nation-wide, summative assessments (Dennis, 2017). This call for one-time summative assessments as the leading measure of school quality and student success has shown little improvement and leaves federal and state policymakers at odds with educational professionals across the nation (Hemelt & Jacob, 2017). Reeves (2004) likens state assessments to autopsies providing standards-aligned information after the instructional year has ended, but no snapshots of learning along the way.

Despite the emphasis placed on standardized tests, members of the educational community versed in standards-based teaching and learning practices understand how a clear picture of the expected learning, communicating those learning intentions to students using success criteria, and using assessments to redirect student learning is the ultimate leverage to higher student achievement (Hillman & Stalets, 2021; Moss et al., 2011; Rinkema & Williams, 2018; Schimmer et al., 2018). These practices are elevated when performed in professional learning communities (PLC) by teams of teachers with common teaching assignments (Burns et al., 2018; DuFour et al., 2008). Young and Kin (2010) found efficient use of data to guide instruction rests on the formative assessment practices of teachers, the usefulness of formative data, and the collective content and pedagogical knowledge of teachers—all of which are
strengthened through collaboration in PLCs (Burns et al., 2018; Sutula, 2017). Although not enough on its own, school leaders can further elevate the impact of teacher collective efficacy by supporting and guiding the work of these teams (Sutula, 2017; Tichnor-Wagner et al., 2016).

**Background of the Problem**

As middle school teachers in a mid-size K12 public school district in the Midwest have invested their efforts over the past three years into unpacking standards, writing learning targets, scaling learning targets, and backward designing curriculum from common summative assessments directly related to those standards, crossing the bridge from curriculum to instructional practice continues to need improvement. Many teachers struggle to leverage this new standards-based curriculum with a learning environment embedded in effective teacher collaboration around the results of common formative assessments, and they often question the significance of collaboration.

With six middle schools participating in the transition to standards-based teaching and learning, the job-alike collaboration groups experience a variety in the level of involvement from their building-level instructional leaders. In some schools the leadership is present in these collaboration group meetings, but the leaders lack the confidence, knowledge, or belief in the impact of teacher efficacy, so they are more of a silent observer than a guiding light. Donohoo et al. (2018) suggests school leaders play key roles on the instructional environment when involved in teacher collaboration and their guidance is a strong predictor of the value teachers find in the collaboration process and the academic gains of students (Donohoo, 2018).

**Conceptual Framework**

A standards-based learning environment for students exists when assessment, instruction experiences, and the content standards function as an interdependent system (Schimmer et al.,
In this system, teachers go beyond introducing standards within their content and they establish a culture of learning where a balanced assessment system continually provides opportunities for students to demonstrate their proficiency at the intended depth of knowledge (DOK) level within each standard (Chattergoon & Marion, 2016; Coladarci, 2002; Marion et al., 2019; Schimmer et al., 2018). Learners are required to provide evidence of learning on standards related to the content of a particular course. When Lopez et al. (2017) claims “competency-based learning is not about learning skills instead of content; it’s about learning critical skills that empower learners to seek out and engage with content more deeply, meaningfully, and productively” (p. 40), he is describing a learning environment for students centered around evidence-based proficiency of standards.

Hillman and Stalets (2021) describe a system where the quality of teachers’ assessment literacy is predictive of their ability to provide this type of learning environment for students. Teachers must use assessment as a tool to provide direction for instruction, clarity for learning, and hope for student success. Furthermore, when teachers engage in these practices as a team, their collective efficacy creates a deeper understanding of the standards, an increased value in assessment, and more equitable learning experiences for all students (Hillman & Stalets, 2021).

The inner workings of the system of standards-based learning hinge around four major practices which align to the four critical questions of a professional learning community (DuFour & DuFour, 2013; Erkens, 2016; Hillman & Stalets, 2021; Rinkema & Williams, 2018; Schimmer et al., 2018). First, teachers must be able to identify what we want students to know and be able to do. Teams must identify and unpack the priority standards for the content or course which they are teaching. As teachers begin to unpack the standards, define what the standard means for student learning in terms of learning targets or learning progressions, skills necessary for
students to be successful on the standard, and DOK level of each standard, they create a clearer picture of what students must be able to know, understand, and do to be proficient (Rinkema & Williams, 2018). Teacher collaboration strengthens each members’ individual understanding of proficiency as they discuss ideas and challenge each other’s thinking (Burns et al., 2018).

Next, methods are needed to know if students have learned the desired outcomes. There are two practices within the assessment structure, formative assessments and summative assessments, which provide teachers with real-time evidence of student learning (Erkens, 2016). Formative assessments can be formal, informal, individual, or common among a teaching team but are always used to monitor in-progress learning and help students and teachers redirect instruction based on the results (Hillman & Stalets, 2021). Schimmer et al. (2018) states “summative assessment completes a balanced approach to classroom assessment and makes teaching and reporting seamless” (p. 125). When created prior to the planning of instruction, Hillman and Stalets (2021) say summative assessments give direction and understanding of what proficiency looks like for a student. The tool used for a summative assessment should match the DOK level of the standard which means not all standards can be assessed with multiple-choice items (Schimmer et al., 2018).

The review of student results on these formative and summative assessments by collaborative teams of teachers is the essential transition to answering the final two critical questions of professional learning communities: What will we do if some students have not learned and how will we extend the learning for those who are proficient? Erkens (2016) explains how the power of common formative assessments comes when teachers collaborate over the creation of the assessment, administer the assessment in proximity of the other team members, and collaborate over the examples of student evidence to calibrate understanding of
the student errors. When these events occur in teams of job-alike teachers, programs improve and student learning increases (Burns et al., 2018).

**Research Questions**

The following research questions guided this study:

1. To what extent does student achievement on a summative assessment differ as a function of the years of experience teaching mathematics at the middle school level in this district?

2. To what extent does student achievement on a summative assessment differ as a function of the quality of involvement by the classroom teacher in collaboration over formative assessments throughout the unit of study, as determined by self-evaluation using the Teacher Collaboration Assessment Rubric (TCAR)?

3. To what extent does student achievement on a summative assessment differ as a function of the interaction between the quality of involvement by the classroom teacher in collaboration, as determined by self-evaluation using the Teacher Collaboration Assessment Rubric (TCAR) and the involvement of leadership in the collaboration process throughout the unit of study?

**Significance of the Study**

As educational systems maintain continued efforts for improvement in the areas of utilizing best practices and achieving academic success, John Hattie (2021) continues to examine which elements of the learning environment are most predictive of high academic achievement. Improving student achievement results in schools continues to be a focal point at the local, state, and federal level throughout the United States. Additionally, the movement toward a standard-based education system, where each content area has defined, grade-level standards for essential
learning, continues to shift the instructional practices for teachers (Pak et al., 2020; Schimmer et al., 2018; Schmoker & Marzano, 1999). The pressure for improvement and the emphasis on a standards-based curriculum leave school leaders wondering what instructional practices produce the greatest impact on student achievement. The results of this study contribute to existing research by examining the impact on student achievement in middle school mathematics when teachers engage in effective teacher collaboration in a standards-based environment.

**Review of Related Literature**

**Standards-Based Teaching and Learning**

When standards-based teaching, learning, or curriculum arises in a conversation with educators, most minds go straight to the concept of standards-based grading. While standards-based grading can be an element of the standards-based teaching and learning process, the latter can exist within a traditional grading system (Schimmer et al., 2018). Furthermore, since grading practices are only a mechanism for reducing student proficiency to a single symbol, the type of grading utilized does not automatically imply students are learning in a standards-based environment (Schimmer et al., 2018). A teaching and learning environment where a common understanding of the standards exists, the destination of learning is clear to the students, and supports are in place linked to the needs identified by a balanced assessment system more accurately defines standards-based teaching and learning than the utilized grading practices (Erkens, 2016; Hillman & Stalets, 2021; Rinkema & Williams, 2018; Schimmer et al., 2018).

When teachers use a standards-based curriculum and learning happens through a standards-based instructional model, there are significant gains in student achievement on assessments aligned to the rigor of the content standards (McCarthy, 2020).
Standards-based teaching and learning practices are multi-dimensional and require the systemic, interdependent use of three major factors for the implementation to be executed with fidelity. A few of those who invest time researching the impact of standards-based teaching practices, including Schimmer et al. (2018) and Erkens (personal communication, October 28, 2021), relate this multi-dimensional system to planning a trip. First, the deep understanding or unpacking of the content standards provides the clear picture of the destination or evidence of learning. Proficiency on content standards also requires understanding the skills, or learning targets, needed to reach the desired level of rigor or complexity of the standard (Hillman & Stalets, 2021; Rinkema & Williams, 2018). Next, like checkpoints on a trip or when a GPS assistant adjusts the route during a trip, a balanced assessment system needs to be in place to help teachers and students monitor students’ learning progress and assist teachers in providing interventions to get students back on track if they stray off course (Erkens, 2016; Schimmer et al., 2018). Finally, maximizing teacher effectiveness with student interventions happens when common assessments can be used by teacher teams, and the teams can collaborate over the results using the PLC model (DuFour et al., 2008; Hillman & Stalets, 2021).

**Standards-Based Curriculum**

A standards-based curriculum uses the adopted standards for a course, content, or grade-level as the framework for its development and clearly articulates the desired results for student learning (Hillman & Stalets, 2021; Rinkema & Williams, 2018). Although a well-designed curriculum framed around standards can have a positive impact on student achievement and a standards-based curriculum is often experienced by higher achieving students, Shoen et al. (2003) claims how the curriculum is implemented by teachers is the most influential element. Some of the implementation strategies used by mathematics teachers most strongly associated
with higher student achievement include eliciting and building on student thinking using conceptual issues, keeping instructional tasks at a high cognitive level, and forcing students to make meaning of their thinking, using group work to encourage engagement and entertain alternative problem-solving methods, and utilizing strategies to keep student discussions and problem-solving at the higher depth of knowledge (DOK) levels (National Council of Teachers of Mathematics, 2014; Shoen et al., 2003). These reform teaching practices were all implemented using a standards-based curriculum.

**Learning Targets**

When unpacking a standard, teachers must work together to identify the necessary progression of learning to reach proficiency on a standard. These learning progressions are laced with skills, or learning targets, needed by the students to achieve at the level of complexity identified by the standard. Rinkema and Williams (2018) and Schimmer et al., (2018) claim the learning targets are the *Ds* from the *know* (K), *understand* (U), and *do* (D) (K-U-D) framework Tomlinson and Strickland (2005) created. The framework assists teachers in planning what will be taught, practiced, and assessed throughout the unit of study.

Moss et al. (2011) defines shared learning targets as critical information for students: what to learn, how deeply to learn it, and how to demonstrate their learning. Learning targets always provide clarity and direction as the teacher plans and as the student engages in their learning (Brookhart, 2012). A research study found New York City’s Public School 13 made significant gains in student success while simultaneously narrowing the achievement gap through a formative assessment process which included clarifying learning targets and utilizing success criteria (Martuccio & Bloomberg, 2020).
Success Criteria

Rinkema and Williams (2018) differentiate learning scales from rubrics, a more commonly identified evaluation tool in education, by their purpose. While rubrics were primarily used for scoring an assessment after the learning was complete, the learning scale guides instruction and provides feedback to students as they progress through their learning. To maximize the usage of learning targets to advance student learning, targets should be scaled to show an increase in complexity (Erkens, 2016; Rinkema & Williams, 2018; Schimmer et al., 2018). Within these scales, the descriptions of learning should be clear and concise and should “move beyond quantifiable requirements and guidelines of compliance to a true description of the learning” (Schimmer et al., 2018, p. 171). Success criteria should focus on what students can do and should refrain from specifics which limit ways students can demonstrate proficiency (Erkens, 2016; Schimmer et al., 2018).

Assessment Literacy

Within a balanced assessment system, a variety of assessments are utilized (Marion et al., 2019). In addition to having a variety of ways in which evidence of learning is elicited from students, variety also comes from the purpose of the assessments (Chattergoon & Marion, 2016; Marion et al., 2019). Most often these purposes can be categorized into formative, interim, and summative, and they happen at the classroom, district, and state levels (Marion et al., 2019). Assessments at the classroom level include both formative and summative which are used to monitor and adjust instruction while providing progress feedback to parents and students (Marion et al., 2019). At the district level, interim and common summative assessments are utilized to predict performance on state assessments, evaluate curriculum and resources, and inform placement decisions into special programs (Marion et al., 2019). Finally, at the state level...
level, assessments are almost always summative and assist in the evaluation of student learning and school quality (Marion et al., 2019).

While both formative and summative assessments play a vital role in the balanced assessment system, Erkens et al. (2017) explain how formative assessments carry the greatest burden of guiding instruction in and out of the professional learning community (PLC) model so summative assessments can be a celebration of the learning.

**Teacher Collaboration**

The results of student performance on formative and summative assessments play an important role in the collaboration efforts of teachers. While teachers of various profiles participate in a variety of collaboration experiences, Ronfeldt et al. (2015) found when teachers collaborate in instructional teams of common content, there is a high association with increased student achievement. Additionally, when those teams focus their conversations on backward lesson design and assessment results, the highest positive correlation to student achievement is found (Ronfeldt et al., 2015). Instructional teams collaborating over common curriculum, instruction, and assessment increases the individual teacher’s confidence in their personal capacity and higher teacher efficacy strengthens the performance of the entire team (Donohoo et al., 2018).

**Teacher Collective Efficacy**

When teachers believe the combined efforts of their team increase their ability to overcome obstacles, the efforts of the individual team members and the efforts of the group are more impactful (Hattie, 2012; Donohoo et al., 2018). Hite and Donohoo (2020) define a model for teacher collective efficacy containing four key components of teacher behavior for teachers on the team and a fifth key component contributed by leadership. Goal consensus, empowered
teachers, cohesive teacher knowledge, and embedded reflective practices are the four components within the teacher team presenting evidence of teacher collective efficacy (Hite & Donohoo, 2020). Supportive leadership completes the model and strengthens the other four components, correlating to higher student achievement (Eells, 2011; Hite & Donohoo, 2020). High teacher collective efficacy results in teacher teams with increased persistence and greater motivation to reach all students through evidence-based practices (Donohoo & Katz, 2019; Eells, 2011).

For all the components of collective teacher efficacy to flourish, the environment must be rich with supportive leadership (Goddard et al., 2017; Hite & Donohoo, 2020; Meilke, 2021; Preston & Donohoo, 2021). Preston and Donohoo (2021) believe supportive leaders grow collective efficacy in their teacher teams by ensuring teams experience success on tasks they believe are beyond their capacity, sharing similar experiences of other teams, setting high expectations with positive reinforcement, and creating and maintaining a positive atmosphere. Progress must be monitored by supportive leadership to help avoid pitfalls, like hidden biases or low expectations, which may work against the benefits of collective teacher efficacy (Hite & Donohoo, 2020; Preston & Donohoo, 2021). A culture of niceness will not sustain teacher collective efficacy if a culture of transparency and a focus on effective practices does not accompany the collegial atmosphere (Preston & Donohoo, 2021).

**Principal Leadership**

Although not reserved solely for the building principal, the role of leadership directly impacts the instructional practices of teachers and improved instructional practices directly improve the achievement of students (Goddard et al., 2010). As the building principal serves as the instructional leader, influences the best practices of teachers, and impacts student
achievement, equally important is the leader’s ability to create structures and systems which allow for a culture of trust and communication (Goddard et al., 2015; Gray et al., 2016; Meyer et al., 2020; Tulowitzki & Pietsch, 2018; Warwas et al., 2019). Goddard et al. (2010) found when principals are involved in the creation of collaboration time, team structure, and goals combined with their direct involvement in the instructional work of collaborative teams, it serves as strong predictor of student achievement.

Professional Learning Communities Model

Collective teacher efficacy and utilization of the model for professional learning communities (PLC) outlined by DuFour (DuFour & DuFour, 2013) serve as strong predictors of student success (Voelkel & Chrispeels, 2017). The three big ideas representing the PLC model and its core principles include: ensuring students learn, building a culture of collaboration, and focusing on results (DuFour et al., 2016). A true PLC model exists at the organizational level and the components of the model are the individual collaborative teams within the organization (DuFour et al., 2016). Hattie (2009) describes how these PLC elements are key to school improvement and can be maximized when teachers collaborate about best practices. Additionally, PLCs require strong leadership to make sure they are structured and supported in a way that allows teachers to find value in them (DuFour & Mattos, 2013). The PLC model is a complex idea requiring clarity before competence of its execution can be reached.

Methodology

Research Design

This study was a quantitative analysis using a factorial ANOVA to study the variance in student summative assessment results when considering the interaction between collaboration quality as measured by the TCAR and the support of building leadership during collaboration.
The factorial AVOVA analysis provided an opportunity to examine how the factors interact and how changes in one factor can influence changes in another factor (Warner, 2013).

**Population and Participants**

This study took place in two middle schools, School A and School B, within the same school district in a rural, Midwest state. Students at both middle schools have their schedules organized by multidisciplinary core content teams containing one teacher each of math, science, social studies, and English language arts.

**Teacher Population**

At School A, there are three teams each at sixth, seventh, and eighth grade. School A hosts a specialized language immersion program so one team at each grade level are language immersion learners. At School B, there are four teams each at sixth, seventh, and eighth grade. School B hosts an additional thirteenth team for a specialized program for students with disabilities. Core content teachers on these teams engage in collaboration three times each week with their team peers and two times each week with their job-alike peers.

**Student Population**

The study analyzed the results of all seventh-grade students enrolled in the general seventh grade mathematics class at these two middle schools.

**Teacher Participants**

This study used students’ summative assessment results on a common unit of study from one of six different math teachers in one of the two identified middle schools who met the requirements of collaborating full time during the unit of study. Three seventh grade mathematics teachers from School A are in their first year of collaborating as a job-alike team and did not have building leadership present at their collaboration meetings. Three seventh grade
mathematics teachers from School B are in their eighth year of collaborating as a job-like team and routinely has leadership present at their collaboration meetings.

**Student Participants**

This study examines the summative assessment results of four hundred forty-two seventh graders enrolled in seventh-grade mathematics at the two identified middle schools. Student participants were both male and female with race and socio-economic demographics similarly distributed from class to class as enrollment in courses was randomly generated by the scheduler within the student information system.

**Instrumentation**

Teachers in this study evaluated the quality of their collaboration during this unit of study using the TCAR (see Appendix A). The rubric was replicated into a Google Form, completed online by teachers, and results were obtained through a Google Sheet.

The common summative assessment (see Appendix B) was created by the seventh-grade mathematics assessment team consisting of a seventh-grade mathematics teacher from each school in the district. Students completed this test using Performance Matters, the district’s summative assessment platform. Student-selected responses were scored by the platform and teachers scored the student-created responses in the platform utilizing a common rubric for each question. All questions on the summative assessment align to the seventh-grade state mathematics standards covered in this unit of study. All teachers in the district use the same curriculum resources, standards and learning targets as they have been mapped, unpacked, and prioritized at the district level.
**Data Collection and Analysis**

Data was downloaded from Performance Matters Reports, the district’s summative and high stakes testing platform, via a spreadsheet. The spreadsheet was stored in a password-protected Google Drive folder. Using the *R Studio* software, the data was loaded with values in five columns: a) 0 = leadership not present and 1 = leadership present, b) TCAR ratings of 1 = beginning, 2 = emerging, 3 = developing, 4 = advancing, and 5 = proficient, c) the students’ decimal score on the common summative assessment, d) the student’s teachers years of experience teaching middle school mathematics in this district, and e) the number of years the student’s teacher has collaborated with their current team. Data was analyzed for descriptive statistics and assumptions of homogeneity of variance were tested both by statistical tests and graphically.

To analyze the variance in student scores between the ten different groups resulting when the independent variables of TCAR rating and leadership presence interacted, a factorial ANOVA was performed to test for variance and a Cohen’s D test was run to compare the effect size on student achievement of one group against another group.

**Procedure**

Two independent teams of teachers from two schools worked through a common unit of study in seventh grade mathematics, used common student and teacher resources, followed a common curriculum guide, and administered a common summative assessment to all students at the end of the unit. The district expectation is for teachers to collaborate as a job-alike team two times per week. All teachers on both teams administered the same common summative assessment at the end of the unit of study and completed the TCAR via a Google Form after all student results were finalized.
Context of the Conditions

The participating school district in this study was midway through a 5-year implementation of standards-based teaching and learning. A large component of this implementation relied on teachers applying their new knowledge to their independent practices. Particularly, when instructional leadership could not be present during a lesson, collaboration meeting, or assessment delivery, district leadership trusted teachers were implementing the components deemed as best practices. The results of this study will provide teachers with data showing how the quality of collaboration when interacting with leadership support is a predictor of student achievement.

Findings

Overall, 424 seventh grade students in mathematics class participated in the common summative assessment after the completion of Unit 4. To meet the assumption of relatively equal sample sizes for the pairing of independent variables, a random sample of 81 student scores were taken from the population of students whose teacher had a TCAR rating of 4 (Avg_High) and administration present and results were analyzed to address five research questions in this study. The summative assessment scores were ratio values of raw scores ranging from a minimum of 0 to a maximum of 1.0, with an average score of 0.65 (Mdn = 0.69) and a standard deviation of 0.23.

Assumptions

Results of the Shapiro-Wilk test of normality suggest overall student achievement on the summative assessment deviates from normality (W=0.919 and p-value < 0.001). However, a histogram of assessment scores shown in Figure 1 shows a visual of the moderate skew and kurtosis (skewness = -0.82, kurtosis = -0.03) combined with the truncation of data on the right.
side due to a maximum score of 1.0. Examining these results indicated little skew was present and the scores are approximately normally distributed.

**Figure 1**

*Student Summative Assessment Scores*

![Histogram of student summative assessment scores](image)

Homogeneity of variance across subgroups of the dependent variable, student achievement on the summative assessment, was analyzed and is described within the context of each research question.

Based on the design of this study, the assumption of independence was met for the overall data set of student scores on the summative assessment, as well as each subset of scores. There was no interaction between participants as each student was only assigned to one teacher and a student score could only correspond to one teacher.

**The extent student achievement differs as a function of the years of experience teaching mathematics at the middle school level.** Research question one investigated the extent which student achievement on a summative assessment differs as a function of the years of experience a student’s teacher has teaching mathematics at the middle school level in this district. The student assessment scores were gathered from the assessment platform and data on experience teaching mathematics at the middle level was gathered from the student information system. Years of teaching experience for student sample were ratio values ranging from a
minimum of 1 to a maximum of 23, with an average of 9.38 (Mdn = 7) and a standard deviation of 7.93. Years of teaching experience were banded into incremental groups of five years.

**Assumptions**

Results from the Levene’s test on the equality of variances between groups suggested there is equality in the variances of each subgroup in the sample (p = 0.301). This assumption can be validated by the Bartlett test (p = 0.267), and we fail to reject the null hypothesis. Visual inspection of the variances in Figure 2 confirms the variances.

**Figure 2**

*Homogeneity of Variance in RQ1 Subsets*

![Homogeneity of Variance in RQ1 Subsets](image)

**Analysis**

Results from the analysis of variance (ANOVA) suggested there was enough evidence to reject the null hypothesis that the mean values of student achievement on the summative assessment in seventh grade mathematics as a function of years of experience teaching middle school mathematics in this district were significantly different ($F_{(3, 335)} = 20.65, p < 0.001$). Overall, years of experience teaching middle school mathematics in this district accounted for a large amount of variance in math achievement, or 15.6%, with a $\eta^2 = 0.156$. 

98
Given a significant omnibus test (the initial ANOVA), post-hoc comparisons were examined to determine the extent of variability in student summative assessment scores, as a function of the teacher’s years of experience teaching middle school mathematics in this district. Results from the Tukey’s HSD, shown in Table 1, suggested there is a strongly significant difference in three of the comparison groups of years of teaching experience (p adj < 0.1).

**Table 1**

*Differences and Significances in Mean Values between Pairings*

<table>
<thead>
<tr>
<th>Group 1 (Years of Experience)</th>
<th>Group 2 (Years of Experience)</th>
<th>diff</th>
<th>p adj.</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 to 9</td>
<td>Less than 5</td>
<td>-0.031</td>
<td>0.670</td>
</tr>
<tr>
<td>10 to 14</td>
<td>Less than 5</td>
<td>0.039</td>
<td>0.759</td>
</tr>
<tr>
<td>20 or more</td>
<td>Less than 5</td>
<td>-0.229</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>10 to 14</td>
<td>5 to 9</td>
<td>0.070</td>
<td>0.290</td>
</tr>
<tr>
<td>20 or more</td>
<td>5 to 9</td>
<td>-0.198</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>20 or more</td>
<td>10 to 14</td>
<td>-0.2680</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

*Note.* Table compares the percent of differences in the mean values of the different pairings and their adjusted p-value for significance using Tukey HSD test.

**The extent student achievement differs as a function of the evaluated quality of collaboration by the individual teacher.** Research question two investigated the extent which student achievement on a summative assessment differs as a function of individual teacher’s evaluation of the quality of collaboration as determined by the TCAR. The student assessment scores were gathered from the assessment platform and quality of collaboration was determined by the TCAR administered via a Google Form. The results of the TCAR were converted into a rating scale of 1 to 5 using logic rules. Quality of collaboration, as determined by the teacher results of the TCAR, were ratio values ranging from a minimum of 3 to a maximum of 5, with an
average of 3.83 (Mdn = 4) and a standard deviation of 0.73. Additionally, the administrator from each school completed the TCAR evaluating the quality of collaboration for their seventh-grade mathematics team. The administrator TCAR rating was 3 at the building with administration present during collaboration and 2 at the building with the administration not present during collaboration.

**Assumptions**

Results from the Levene’s test on the equality of variances between groups provided significant evidence to not reject the null hypothesis and verified all subgroups had equal variance (p = 0.077). This assumption was validated by the Bartlett test with a p-value = 0.058. Homogeneity of variance in the subgroups by TCAR score can be seen in Figure 3.

**Figure 3**

*Homogeneity of Variance in RQ2 Subsets*

![Image](https://example.com/image.png)

**Analysis**

Results from the analysis of variance (ANOVA) suggested there is a statistically significant variance in the means of student achievement on the summative assessment in seventh grade mathematics as a function of teacher self-evaluation of quality of collaboration as determined by the TCAR (F(2, 336) = 4.628, p = 0.010). Overall, teacher self-evaluation of quality
of collaboration as determined by the TCAR accounted for a small amount, 2.7%, of variance in the mean values of mathematics achievement with a $\eta^2 = 0.027$.

Given a significant omnibus test (the initial ANOVA), post-hoc comparisons were examined to determine the extent of variability in student summative assessment scores, as a function of individual teacher’s evaluation of the quality of collaboration as determined by the TCAR. Results from the Tukey’s HSD, shown in Table 2, suggested there is a significant difference, 9.4%, in one of the comparison groups of TCAR ratings ($p_{adj} < 0.05$).

**Table 2**

* Differences and Significances in Mean Values between Pairings

<table>
<thead>
<tr>
<th>Group 1 (TCAR rating)</th>
<th>Group 2 (TCAR rating)</th>
<th>diff</th>
<th>p adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>3</td>
<td>-0.064</td>
<td>0.670</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>0.030</td>
<td>0.759</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>0.094</td>
<td>0.019</td>
</tr>
</tbody>
</table>

*Note. Table compares the percent of differences in the mean values of the different pairings and their adjusted p-value for significance using Tukey HSD test.*

The extent student achievement differs as a function of the interaction between quality of collaboration and presence of administration during teacher collaboration.

Research question three investigated the extent which student achievement on a summative assessment differs as a function of the interaction between the quality of involvement by the classroom teacher in collaboration, as determined by self-evaluation using the Teacher Collaboration Assessment Rubric (TCAR) and the involvement of leadership in the collaboration process throughout the unit of study. The student assessment scores were gathered from the assessment platform, the quality of collaboration was determined by the TCAR administered via
a Google Form, and the presence of administration at collaboration during this unit was
determined by an email to each principal.

**Assumptions**

Sample sizes in the different pairings have a similar n-size. These n-sizes can be seen in
Table 3.

**Table 3**

*Student Participation by Administration Presence*

<table>
<thead>
<tr>
<th>Administration Present:TCAR Rating</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes:Avg</td>
<td>73</td>
</tr>
<tr>
<td>Yes:Avg_High</td>
<td>81</td>
</tr>
<tr>
<td>Yes:High</td>
<td>0</td>
</tr>
<tr>
<td>No:Avg</td>
<td>50</td>
</tr>
<tr>
<td>No:Avg_High</td>
<td>70</td>
</tr>
<tr>
<td>No:High</td>
<td>65</td>
</tr>
</tbody>
</table>

*Note.* Data was retrieved from Performance Matters Comparative Results.

Most pairings (administration presence X TCAR rating) produced students summative
assessment scores which suggest deviation from normality when applying the Shapiro-Wilk Test.
All pairings produced p-values smaller than 0.05. However, when considering the maximum
score of a 1.0 on the summative assessment, the small sample sizes, and the fact that all Shapiro-
Wilks tests had a W-value > 0.856, normality in the pairings exists and is parallel to the
normality in the overall scores on the summative assessment. Table 4 shows the W-value and p-
value for each pairing.

102
Table 4

Shapiro-Wilk Test for Normality on Pairings

<table>
<thead>
<tr>
<th>Pairing</th>
<th>W-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Admin.Presence:TCAR Rating)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes:Avg</td>
<td>0.937</td>
<td>0.001</td>
</tr>
<tr>
<td>Yes:Avg_High</td>
<td>0.908</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Yes:High</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>No:Avg</td>
<td>0.856</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No:Avg_High</td>
<td>0.920</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No:High</td>
<td>0.921</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Student scores on the summative assessment do not violate the assumption of homogeneity of variance as Levene’s test produced a p-value = 0.143. Homogeneity of variance exists and can be seen in Figure 4.

Figure 4

Homogeneity of Variance in RQ3 Subsets

By design of the study, there was not a connection between the participants of the five pairings created by grouping administration presence with TCAR rating. No student was enrolled in more than one pairing.
**Analysis**

Results from the analysis of variance (ANOVA) suggested there is statistically significant variance in the means of student summative assessment scores as a function of administration presence when interacting with evaluation of collaboration according to the TCAR rating \( F(4, 334) = 15.85, p < 0.001 \). Overall, evaluation of collaboration according to the TCAR rating interacting with administration presence accounted for a medium amount of variance in the mean values of student summative assessment scores with an \( \eta^2 = 0.089 \), meaning 8.9% of the variance could be explained.

Given a statistically significant omnibus test (the initial ANOVA), post-hoc comparisons were examined to determine the extent of variability in student summative assessment scores, as a function of all the interactions between Administration presence when interacting with evaluation of collaboration according to the TCAR rating. Results from the Tukey’s HSD suggested there is a significant difference in the means for student summative assessment scores in four of the pairing comparisons. When comparing the pairings of administration presence and a teacher’s self-evaluation of the quality of collaboration during the unit of study with the Cohen’s D test, the effective size in summative assessment scores can be measured. Any Cohen’s D value between 0.20 and 0.49 is considered a small effect, between 0.50 and 0.79 is considered a medium effect size and any value 0.80 and over is considered a large effect size. These values should be considered in their absolute value as a negative value only determines the second group has a larger effect than the first group listed.

In Table 5, results of the differences in mean values between all pairings, along with their adjusted p-values and the Cohne’s D effect size of the ten comparisons are listed.
### Table 5

*Differences, Significances, and Effect Sizes of Mean Values between Pairings*

<table>
<thead>
<tr>
<th>Group 1 (TCAR:Admin Presence)</th>
<th>Group 2 (TCAR:Admin Presence)</th>
<th>diff</th>
<th>p adj.</th>
<th>Cohen’s D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg:No</td>
<td>Avg:Yes</td>
<td>0.052</td>
<td>0.778</td>
<td>0.116</td>
</tr>
<tr>
<td>Avg_High:Yes</td>
<td>Avg:Yes</td>
<td>0.074</td>
<td>0.275</td>
<td>-0.265</td>
</tr>
<tr>
<td>Avg_High:No</td>
<td>Avg:Yes</td>
<td>-0.178</td>
<td>&lt;0.001</td>
<td>-1.054</td>
</tr>
<tr>
<td>High:No</td>
<td>Avg:Yes</td>
<td>0.051</td>
<td>0.736</td>
<td>0.004</td>
</tr>
<tr>
<td>Avg_High:Yes</td>
<td>Avg:No</td>
<td>0.022</td>
<td>0.993</td>
<td>-0.377</td>
</tr>
<tr>
<td>Avg_High:No</td>
<td>Avg:No</td>
<td>-0.229</td>
<td>&lt;0.001</td>
<td>-1.153</td>
</tr>
<tr>
<td>High:No</td>
<td>Avg:No</td>
<td>&lt;0.001</td>
<td>1.000</td>
<td>-0.097</td>
</tr>
<tr>
<td>Avg_High:No</td>
<td>Avg_High:Yes</td>
<td>-0.252</td>
<td>&lt;0.001</td>
<td>-0.833</td>
</tr>
<tr>
<td>High:No</td>
<td>Avg_High:Yes</td>
<td>-0.023</td>
<td>0.988</td>
<td>0.233</td>
</tr>
<tr>
<td>High:No</td>
<td>Avg_High:No</td>
<td>0.229</td>
<td>&lt;0.001</td>
<td>0.927</td>
</tr>
</tbody>
</table>

*Note.* Table compares the percent of differences in the mean values of the different pairings and their adjusted p-value for significance using Tukey HSD test and lists the Cohen’s D values of effect size for the same pairings.

The summary of mean values of student achievement on the summative assessment when pairing administration presence and teacher self-evaluation of the quality of collaboration can be viewed in Figure 5.
Discussion

As pressure for national improvement on state summative mathematics assessments continues, teachers and administration at the state and local levels continue searching for the best instructional practices to achieve this goal. This study investigated the impact on student achievement of three separate sets of factors: teacher experience, self-reported quality of teacher collaboration, and the interaction between the quality of teacher collaboration and administration involvement.

It is common for people to assume more experienced teachers are more effective instructionally and have students who will score higher on assessments. Ladd and Sorenson (2017) confirm teachers with more years of experience have higher student achievement but only when teacher quality can be controlled. While this study resulted in some significant difference between select pairings of students, these comparisons showed students the less experienced
teachers scored higher than the students with the most experienced teacher. Since student achievement did not increase as teacher experience at the middle school level increased from the least to the most experienced, there is not enough evidence to support teaching experience as a predictor of higher student achievement and perhaps teacher quality needs further investigating.

In this study, the curriculum is written and learning targets are unpacked and scaled with success criteria at the district level and provided to teachers. However, Schimmer et al. (2018) and Rinkema and Williams (2018) state these targets and success criteria must be communicated throughout instruction and formative assessments must be utilized to check for student understanding. Although a common, standards-based curriculum is provided to all teachers in this district, the presentation methods, commitment to interventions, formative assessment efforts, and content knowledge of each teacher can affect the quality of learning and account for variations in teacher quality. When teachers collaborate using the PLC model, these variations in instructional practices are minimized (DuFour et al., 2016; Hattie, 2009; Ronfeldt et al., 2015). Teachers in effective collaborations share instructional experiences, calibrate their understanding of standards and learning targets, and backward design a balanced system of assessments throughout the unit of study.

With teacher collective efficacy topping the list of factors influencing student achievement (Hattie, 2021), it would make sense for students’ achievement to increase as the quality of teacher collaboration increases. In this study, students whose teacher gave the highest rating for the quality of collaboration did score significantly higher than students whose teacher scored one rating scale lower. However, student achievement did not increase as teachers’ self-reported quality of collaboration ratings increased from smallest to largest. When teachers do not have a clear understanding of what quality collaboration entails, an inaccurate evaluation, even
with a descriptive rubric, is quite probable. Also, teachers in this district understand the expectation to engage in collaborative meetings exists and may not be completely truthful in their evaluation as to not appear negligent of meeting the requirement. When relying on self-evaluation results on the rubric for quality of collaboration, it is impossible to measure the accuracy and comparative values of the ratings. A further investigation is needed to determine other factors which explain the discrepancy.

Preston and Donohoo (2021) describe how important leadership involvement is to the efficacy of teacher collaboration. When school leaders engage with teachers during collaboration, they provide encouragement, reinforce proper PLC activities, offer models of success from other teams, give examples of effective strategies or alternative ideas, encourage reflection, and prevent teachers from straying off course. Using the same rubric, each administrator from the two schools evaluated the collaboration efforts of their teachers lower than their corresponding teachers. This discrepancy between administrations’ view and the teachers’ view of high-quality collaboration suggests teachers have a different level of expectation compared to leadership. It is possible for teachers to be more aligned with leadership views when those leaders participate in collaboration on a consistent basis. The final research question disaggregates results of the student assessment scores by building, one where administration played an active role in the teachers’ collaboration process and one where teacher met to collaborate throughout the same unit of study but lacked administrative involvement in the process, and determines if one scenario is a better predictor of increased student achievement.

As predicted, in the building where administration attended collaboration meetings, student achievement increases as the self-evaluated quality of collaboration increases. Additionally, no teachers from this building gave their collaboration activities a perfect rating.
Each member recognized there were elements of their efforts in need of improvement. Teachers have a better understanding of what collaboration should look like and filled out the TCAR with greater accuracy in their evaluation. Adversely, in the building where administration did not attend collaboration meetings or provide leadership to the team, student achievement did not increase as the quality of collaboration increased.

The results of both buildings support the findings of Goddard et al. (2015) where direct involvement of leadership during collaboration time and instruction serves as a strong predictor of student achievement. On the team where collaboration took place, but administration was not present, the team members lacked the guidance, knowledge, and models of best practice from leadership. Additionally, Preston and Donohoo (2021) state teams without administration present do not have accountability to implement instructional practices discussed within collaboration, there is not an element of mediation when team members disagree or felt unheard by others, and no one ensures each group member has a voice at the table. Collaboration activities can be rated high by team members when they do not fully understand what highly effective collaboration involves. Misguided efforts or unkept team commitments fail to impact student learning and do not result in higher assessment achievement.

**Conclusion and Recommendations**

While there was a significant difference in the comparison of student achievement by teachers’ years of experience teaching mathematics at the middle school level between certain groups, those results were not conclusive enough to determine teacher longevity is a predictor of greater student learning. The results of the study and the research of Ladd and Sorenson (2017) on controlling for teacher quality for experience to matter suggests administration does
not have to hire the most experienced teachers to get quality student learning or higher student achievement. Consequently, building leadership needs to mentor and grow the quality of instruction within each teacher and their years of experience will produce even greater results.

It is not enough to train teachers on the PLC model and expect them to carry it out perfectly. Teachers’ list of duties is long, and they can find themselves taking shortcuts when trying to accomplish everything on the list. It is imperative for teachers to practice the PLC model in environments where they can receive feedback and make improvements on their efforts. Gaining the most return on investment of time in teacher collaboration would mean educators are engaging in the best practices and students are achieving at higher rates as a result. This study indicates when teachers engage in quality collaboration practices and administration continues to be present to maintain the integrity of the PLC model, it can serve as a predictor of higher student achievement.

Based on the results of this study, districts and building leadership must schedule teacher collaboration time in a manner allowing administration presence throughout units of study to increase student achievement. Additionally, leadership must work hard to create a collaboration culture with their teams which fosters reflection, mentorship, and recognition of success. This suggestion may prove difficult in settings where large numbers of teacher collaboration groups meet at the same time each day and where the administrative team is not large enough to support required management tasks and instructional leadership duties simultaneously. Creativity in scheduling, prioritization of time, and establishment of a culture where teachers feel safe and value the efforts of collaboration will produce the success Preston and Donohoo (2021) found breeds further success in areas teachers formerly believed were beyond their capacity. As a result, students learn content and master skills allowing them to achieve higher.
Limitations and Areas for Further Study

This study collected years of teaching experience as years where mathematics was taught at the middle school level in this district. Due to this method of counting years of experience, teachers may have additional years at another grade level or in another district which were not counted but may have an impact on the quality of their teaching. When comparing years of teaching experience to student achievement, the study also failed to control of the quality of teaching which Ladd and Sorenson (2017) found to be critical their results. Had quality of teaching been controlled for in this study, the sample size may have been too small to return significant results.

As a result of teachers self-evaluating the quality of the collaboration activities, this study was limited in the accuracy of the TCAR ratings used for the study. It is difficult to know if all teachers received the same professional development in the PLC model or had a similar understanding of what activities quality collaboration should include. Perhaps the study could be improved if the TCAR rating for each teacher was created through observations of participation by the administrator, However, the quality of administrator’s understanding of the PLC model was not analyzed or controlled for in this study either.

Finally, one other limitation of the study lies in the way the TCAR rating was reduced to a single score. The TCAR is split into the areas of dialogue, action, and evaluation. An opportunity for further study would be to determine if one of those areas was more predictive of higher student achievement. This study would allow administration to know if there are parts of collaboration which generate better student results or if overall quality has a larger effect on student achievement. Repeating the study and gathering TCAR ratings through administrative observation with scores from the three components of collaboration, in addition to an overall
rating, would further develop the research to support how imperative administration’s presence and leadership is to the efficacy of teacher collaboration.
References


http://www.ascd.org/publications/educational-leadership/sept12/vol70/num01/Preventing-Feedback-Fizzle.aspx


Erkens, C., Schimmer, T., & Vagle, N. D. (2017). *Essential assessment: Six tenets for bringing hope, efficacy, and achievement to the classroom.* Solution Tree Press, the Solution Tree Assessment Center.


Lopez, N., Patrick, S., & Sturgis, C. (2017). Quality and Equity by Design: Charting the Course for the Next Phase of Competency-Based Education. CompetencyWorks. iNACOL.


https://digitalcommons.georgiasouthern.edu/nyar_savannah/2020/2020/96/


http://www.ascd.org/publications/educational_leadership/mar99/vol56/num06/Realizing_the_Promise_of_Standards-Based_Education.aspx


https://doi.org/10.14507/epaa.v18n19.2010
## Appendix A

### TEACHER COLLABORATION ASSESSMENT RUBRIC

<table>
<thead>
<tr>
<th></th>
<th>Beginning</th>
<th>Developing</th>
<th>Proficient</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dialogue</strong></td>
<td>Dialogue is often related to swapping curriculum ideas or strategies</td>
<td>Dialogue is usually related to instructional practice and student performance</td>
<td>Dialogue is focused on the structured examination and analysis of instructional practice and student work</td>
</tr>
<tr>
<td></td>
<td>Tensions exist among team members, but are not discussed or resolved</td>
<td>Professional tension exists but disagreements and/or conflicts are rare</td>
<td>Professional tensions are resolved as soon as possible</td>
</tr>
<tr>
<td></td>
<td>Unequal participation in dialogue</td>
<td>For the most part members participate equally, though some members may ‘hibernate’ or ‘dominate’</td>
<td>All members participate equally with no ‘hibernators’ or ‘dominators’</td>
</tr>
<tr>
<td><strong>Action</strong></td>
<td>Team members take minimal action as a result of group decision-making</td>
<td>The team occasionally makes decisions about what pedagogical practices they will initiate, maintain, develop and/or discontinue</td>
<td>Team regularly makes decisions about what individual and collective pedagogical practices they will initiate, maintain, develop and/or discontinue</td>
</tr>
<tr>
<td></td>
<td>Team decisions are minimally informed by group dialogue</td>
<td>Most team decisions are informed by group dialogue</td>
<td>All decisions are informed by team dialogue</td>
</tr>
<tr>
<td></td>
<td>Most team actions are unrelated to the improvement of instructional practice and the cultivation of student learning</td>
<td>Actions are generally related to instructional practice and student learning</td>
<td>Actions are directly related to the improvement of instructional practice and the cultivation of student learning</td>
</tr>
<tr>
<td><strong>Evaluation</strong></td>
<td>Team members do not share evaluative data about the merits of their instructional practices with one another</td>
<td>The team does not regularly collect and/or analyze qualitative and quantitative information about member teaching practices and student learning</td>
<td>The team regularly collects and analyzes qualitative and quantitative information about teaching practices and student learning, including peer observations</td>
</tr>
<tr>
<td></td>
<td>The team does not systematically collect or analyze information about instructional practices and student learning</td>
<td>The team may rely more on “hearsay,” “anecdotes,” or “recollections” to evaluate the merit of their practices</td>
<td>The team uses student performance data to evaluate the merit of individual and collective pedagogical practices</td>
</tr>
<tr>
<td></td>
<td>The team relies exclusively on “hearsay,” “anecdotes,” or “recollections” to form the basis of their dialogue and decision-making</td>
<td>Evaluative information is usually shared publicly and forms the basis for dialogue and decision-making</td>
<td>Evaluation data and findings are shared publicly and form the basis for team dialogue and decision-making</td>
</tr>
</tbody>
</table>

---

Appendix B

Common Summative Assessment

Grade 7
Unit 4 End of Unit Proportional Relationships and Percentages
Calculator can be used.
Name ___________________________ Pd ______

1. A sweatshirt costs $32. Today it is on sale for 25% off. What is the amount of discount? Explain or show your reasoning.
2. Diego measured the length of a pen to be 22 cm. The actual length of the pen is 23 cm. Which of these is closest to the percent error for Diego’s measurement?
   A) 4.3%
   B) 4.5%
   C) 95.7%
   D) 104.5%
3. A car is 180 inches long. A truck is 75% longer than the car. How long is the truck?
   A) 135 inches
   B) 240 inches
   C) 255 inches
   D) 315 inches
4. A circular running track is 14 mile long. Elena runs on this track, completing each lap in 120 of an hour. What is Elena’s running speed? Include units of measure. Explain or show your reasoning.
5. Today, everything at the store is on sale. The store offers a 20% discount. The regular price of a T-shirt is $18. What is the discount price? Explain or show your reasoning.
6. Today, everything at the store is on sale. The store offers a 20% discount. The discount price of a hat is $18. What is the regular price? Explain or show your reasoning.
7. Today, everything at the store is on sale. The store offers a 20% discount. The regular price of an item is x dollars. Write an expression that represents the discount price or sale price.
8. Lin’s father is paying for a $20 meal. He has a 15%-off coupon for the meal. After the discount, a 7% sales tax is applied. What does Lin’s father pay for the meal? Explain or show your reasoning.
9. Tyler’s brother works in a shoe store. He earns a commission. He makes 2.5% of the amount he sells. Last week, he sold $900 worth of shoes. How much was his commission? Explain or show your reasoning.
10. A store bought a pair of shoes for $50, and sold it for $80. What percentage was the markup? Explain or show your reasoning.
11. Steve earns $12 per hour. He is offered a raise of 5% increase per hour. After the raise, how much will Tyler’s brother make per hour? Explain or show your reasoning.
Appendix C
Scoring Guide
End of Unit Assessment
7th Grade - Unit 4 - 18 points

Standards: 7.RP.A.1, 7.RP.A.2, 7.RP.A.3

<table>
<thead>
<tr>
<th>Test Question</th>
<th>Online Question</th>
<th>Score</th>
<th>Point Distribution</th>
<th>Standard</th>
<th>DOK Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>____/1</td>
<td>1 point for correct answer (8)</td>
<td>7.RP.A.3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>____/1</td>
<td>1 point for correctly selecting A</td>
<td>7.RP.A.3</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>____/1</td>
<td>1 point for correctly selecting D</td>
<td>7.RP.A.2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>____/2</td>
<td>2 points - correct speed and unit of measure (does not have to be mph, can be minutes per mile, miles per minute, etc) 1 point - correct speed but incorrect or no unit of measure 0 points - no evidence of understanding</td>
<td>7.RP.A.1</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>____/2</td>
<td>2 points - correct discount price and work 1 point - correct discount, but don’t subtract from the original price 0 points - no evidence of understanding</td>
<td>7.RP.A.3</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>____/2</td>
<td>2 points - correct regular price and work 1 point - some correct work but not right answer common mistake could be $21.60 0 points - no evidence of understanding</td>
<td>7.RP.A.3</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>____/1</td>
<td>1 point - correct response of 0.8x, (x - 0.2x), or ( \frac{4}{5}x )</td>
<td>7.RP.A.3</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>____/4</td>
<td>4 points - correct final amount 3 points - one error in work 2 points - two errors in work 1 point - three errors in work 0 points - no evidence of understanding Note: If an error in the first step, but the math following is accurate, do not take additional points off for incorrect answers</td>
<td>7.RP.A.3</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>____/1</td>
<td>1 point - correct response of $22.50</td>
<td>7.RP.A.3</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>____/1</td>
<td>1 point - correct response of 60%</td>
<td>7.RP.A.3</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>____/2</td>
<td>2 points - correct answer and correct work 1 point - one error in work 0 points - no evidence of understanding</td>
<td>7.RP.A.3</td>
<td>2</td>
</tr>
</tbody>
</table>

Notes to tell your students prior to test:

*Type the numbers only. Do NOT type your work or labels. (Questions 1, 9, 10)*

**Teachers** - Please look over questions 1, 9, and 10 to make sure they did not type a correct version that was counted incorrect. Go to **Reports - Student Item Analysis - Look at the red answers.**