# AN IMPROVEMENT SCIENCE CASE STUDY FOR ACCELERATING RECOVERY OF DISRUPTED LEARNING USING FLEXIBLE GROUPING

by

Allison A. Jonas

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Education

Major: Educational Administration

Under the Supervision of Professor Sarah Zuckerman

Lincoln, Nebraska

May 2023

# AN IMPROVEMENT SCIENCE CASE STUDY FOR ACCELERATING RECOVERY OF DISRUPTED LEARNING USING FLEXIBLE GROUPING

Allison Amy Jonas, Ed.D.

University of Nebraska, 2023

Advisor: Sarah Zuckerman

During the COVID-19 pandemic, numerous challenges specific to educational context resulted in disrupted teaching and learning for students across the U.S. Consequently, accelerating recovery of disrupted learning has become the primary concern of every school district in the nation. Prior to the pandemic, the use of data to guide instruction was considered an important means of developing individualized learning paths; however, the mere use of data traditionally failed to provide practical and actionable guidance for teachers.

This qualitative case study explored how teachers in a rural Nebraska school used flexible grouping as a strategy to support teachers in modifying their instruction when students did not know or already knew benchmark learning objectives. The teachers used flexible grouping to address the rapidly increasing skill gap between low- and high-performing students. This approach allowed students to move between groups to receive instruction tailored to their individual needs within a homogeneous classroom.

*Keywords:* flexible grouping, differentiated instruction, professional learning

# Table Of Contents

## List of Tables



## List of Figures

# An Improvement Science Case Study of Accelerating Recovery of Disrupted Learning Using Flexible Grouping

## CHAPTER 1

## INTRODUCTION

The unprecedented COVID-19 pandemic resulted in disrupted teaching and learning for the 2019–2020 school year. This period highlighted the need for individualized student learning more than ever. Due to emergency school building closures in the spring of 2020 and the inherent challenges of then-novel remote instruction, students across the U.S. missed 25% of the 2019–2020 school year. Experts projected that students would return in the fall of 2020 with just 70% of the reading gains typically observed in a year due to the lack of targeted, systematic instruction offered in traditional school settings (Kuhfeld & Tarasawa, 2020). The study also found those impacts to be more pronounced in the primary grades due to the intense support required to teach the foundational skills of reading and the need for physical movement and social interaction, all of which are nearly impossible to recreate virtually (Hinton, 2020). Continued disruptions in the form of rapidly changing quarantine guidelines, mask mandates, social distancing requirements, and the balancing of in-person, remote, and hybrid learning environments, as well as lack of equitable access to each of these environments, all contributed to a rapidly increasing gap in terms of content knowledge between low- and high- performing students (Kuhfeld et al., 2020). This supports the argument that there is an ever-increasing need for teachers to work diligently in differentiating or tailoring instruction to meet the individual needs of students.

The lack of an organized federal response left states and districts to make decisions on where and how to deliver instruction, resulting in patchy efforts with varying degrees of success (Shafer, 2020). Nationally, achievement gaps caused by the pandemic have been measured by NWEA's MAP Growth based on reading and math assessment scores. MAP Growth is a computer adaptive test, typically administered three times a year, which measures both achievement and growth in content areas such as reading and math. It is used by more than 11 million students across the country, including all of Nebraska's 244 school districts (Hahn, 2019). Normative scores are then used to demonstrate the student's growth areas compared to the national norm. It is a typically consistent and reliable metric, given a large number of participants.

Kuhfeld et al. (2020) noted that, due to the pandemic, only three out of four students assessed in Fall 2019 were assessed again in Fall 2020), indicating that the true picture of disruptive learning might not have been fully captured. They also revealed that Fall 2020 scores were anywhere between 86–107% of typical fall scores. This suggests that gains in reading growth were primarily made by students who had been performing at or above the 66$^{th}$ percentile in Winter 2020. This further highlights the gap between high- and low-performing students. The analysis also found these gaps to be greater in the primary grades, suggesting that students in the primary grades missed out on core reading skills such as phonics, oral reading fluency, comprehension, and vocabulary development, which is bound to further impact skill acquisition if not worked upon quickly (Bielinski et al., 2020).

In Nebraska, most schools re-opened in August 2020 and have largely remained open since for both in-person instruction, as well as remote or hybrid learning. However,

disruptions continued throughout the 2020–2021 and 2021–2022 school years, as teachers' efforts to close achievement gaps were thwarted by two-week quarantines that reduced the time students and teachers spent in the classroom. Other factors that impacted learning during the pandemic include inequitable internet access, varying degrees of parental support, and dissimilar responsibilities such as caring for younger siblings. On the other hand, internal factors include the navigating of health regulations that limited student grouping strategies and peer-to-peer interactions. Fridland (2020) noted a more elusive factor—a reduction in student comprehension due to masks inhibiting multi-sensory cues, including both auditory and visual information.

**Study Context**

In our rural district of just under 1,000 students, in-person instruction was the norm for the 2020–2021 school year. However, in Fall 2020, Gothenburg Public Schools opened but maintained an elevated level of pandemic risk awareness, where masks were mandated, social distancing was required, and no student transportation or adult volunteers were allowed inside the buildings. As a district, we struggled to find substitutes to cover the classrooms of teachers who, in some cases, were quarantined for up to six weeks. Further, teachers struggled to teach classrooms that dwindled from 24 students to four due to mass quarantines Additionally, teachers struggled in classrooms trying to limit the interactions between cohorts (elementary vs. secondary) and even between students within the classrooms.

Prior to the pandemic, the decision was made in the elementary building to enact flexible grouping, a strategy that allowed students to move between groups to receive instruction tailored to their individual needs within a homogeneous classroom, as a small

part of the overall comprehensive Professional Learning Community approach. As part of this approach, each teacher team was asked to engage with four critical questions outlined by DuFour et al. (2016) that should drive the action of every collaborative team:

1) What do we want students to know and be able to do? (essential learnings)
2) How will we know if each student has learned it? (formative assessment)
3) How will we respond if students don't know it? (differentiated instruction)
4) How will we respond if students already know it? (differentiated instruction)

In response to Questions 3 and 4, the elementary school embraced the flexible grouping approach to allow teachers to maximize student learning by ensuring that the instruction was specifically aligned to the skills they were ready to develop while limiting the negative impacts of tracking.

Thus, I employed an improvement science case study approach for this study because I wanted to understand better why one teacher team that was utilizing flexible grouping was observing significantly higher growth rates on NWEA MAP growth assessments than others, and to learn from this positive outlier what they were doing differently, in hopes of helping other teacher teams develop these same skills and produce similar results. The positive outlier in this case averaged 43% more student growth than the national average for the period 2018–2022 and, therefore, provided a unique opportunity to learn what strategies and practices allowed this team to have an exceptional impact on student outcomes, which could potentially be capitalized upon and taught to other grade-level teams for enhancing student outcomes. To this end, I crafted several research questions that guided each stage of this study.

Phase 1: Understanding Flexible Grouping in Use

1) How did the teachers use flexible grouping to address the diverse academic needs of the students?
2) How did the teachers work together as a grade-level team to meet the students' needs?

Phase 2: Targeted Professional Development in Response to Findings from Phase 1

Phase 3: Unmasking Perceptions of Professional Development

1) What were the teachers' perceptions of professional development opportunities and their subsequent efforts as part of a collaborative team?

This change effort was aimed at improving students' assessment scores, using a theory of action that focused on meeting students' unique needs through differentiated instruction supported by a collaborative culture among teachers. Tomlinson's (2015) differentiated instruction theory was utilized to guide the analysis and highlight the effectiveness of specific strategies that teachers utilize to positively impact student growth. The findings from this study may be valuable to classroom teachers and administrators who seek to leverage targeted professional learning experiences to accelerate student growth by utilizing flexible grouping strategies and practices.

### *A Persistent Problem*

With the passing of the No Child Left Behind Act (2001), student proficiency and academic progress were thrust into mainstream conversation. According to the U.S. House of Representatives Committee on Education and the Workforce (2006), reporting student scores on state assessments by student sub-groups "ensures that academic progress is being made." Even with a very public focus on student performance, the

reading achievement level of U.S. students has remained flat or, for struggling readers, declined further (National Assessment of Educational Progress, 2015). The COVID-19 pandemic created a renewed focus on student learning.

Instructional disruption appears linked to lower student achievement, particularly among those students who are already at risk for school failure (Bielinski et al., 2020). Before the pandemic, little attention was given to practical solutions for teachers to better equip them to meet the continually increasing diverse needs of students which was further exacerbated by the pandemic. With primary students lacking essential foundational reading skills, the long-term impact is not yet known. Therefore, teachers must find ways to address the current disrupted learning before cumulative learning loss occurs (Andrabi et al., 2021; Kaffenberger, 2021; Mangan, 2021).

Researchers believe that the achievement gap may be a result of inequality in the opportunities to learn (Datnow & Park, 2018; Flores, 2007). However, this inequality was brought to the forefront when schools across the nation closed in March 2020 due to the COVID-19 pandemic. For the first time in recent history, schools across the U.S. were closed for more than just a summer or holiday break. Schools scrambled to provide education to students who were no longer able to attend in-person. To facilitate the quick shift to remote education, it was important to ensure that students had access to the Internet. The majority of students at the national level do have access to the Internet, given that they belong to White families with a household income of more than $75,000 (NCES, 2020). However, as stated by the NCES (2020), poverty is a huge hurdle in ensuring internet access for all students. Furthermore, current research indicated that those from economically disadvantaged families are also more likely to lack adult

academic support and more likely to live in crowded households which make online learning more challenging (Samuels, 2020). At a local level, this access issue was equally problematic. A locally administered survey in Fall 2020 revealed that 32% of Gothenburg students lacked access to high-speed internet. Although just 1.3% did not have access to the internet at all, about 30.7% lacked the internet speed that could facilitate virtual instruction. About 25% and 3% of those without access cited their location (rural community) and cost, respectively, as the primary reason why they lacked high-speed internet. This translates into three out of every ten students lacking access to virtual instruction.

The data gathered in Fall and Winter 2020 using NWEA MAP growth windows reflect that 25% of students were not yet accounted for in the data due to not being back in school by the end of the first semester of 2020 (Kuhfeld et al., 2020). Yet, even with these students missing from the dataset, the data indicated up to a 14% loss of learning in reading which disproportionately negatively affected those performing below the 66$^{th}$ percentile and all students in primary grades (Kuhfeld et al., 2020). Even more concerning is the fact that in the primary grades, critical reading skills such as phonics, oral reading fluency, comprehension, and vocabulary development further impacted skill acquisition (Bielinski et al., 2020; Sparks, 2021). The pandemic has already resulted in a 17% increase in the number of students falling “well-below” the benchmark on DIBELS, a separate nationally normed assessment that measures the reading skills of students (Freitag, 2021). Educational researchers implore that failing to recognize the importance of reading instruction designed to address specific skill acquisition will have detrimental effects on long-term student outcomes (Freitag, 2021).

This study suggests flexible grouping as a strategy to address the diverse academic needs of students by ensuring instruction is continually matched to their needs in order to maximize student growth. Flexible grouping is defined as a process of continually utilizing formative data to guide within-school sorting that results in students learning with peers of similar ability in groups that shift as skills are mastered (Kalogrides & Loeb, 2013; Riley, 2016). Flexible grouping differs from tracking in that, when students are tracked, they are placed in groups for long periods, with no way out of that course of study. With flexible grouping, teachers use formative assessment data to fluidly move students between groups to ensure that students receive instruction at the level more receptive to them in terms of learning, further ensuring that all students are appropriately challenged (Perez, 2019). Flexible grouping is effective in accelerating recovery from disrupted learning (Datnow & Park, 2018; Matthews et al., 2013; Neuman, 2016; Riley, 2016; Shafer, 2020; Rytivaara, 2011; Tomlinson, 2015).

Prior to the COVID-19 pandemic, research regarding lagging proficiency scores of American students focused heavily on ensuring equitable opportunities and outcomes for all students (Datnow & Park, 2018). However, research conducted amid the COVID-19 pandemic suggests that a focus on equity may not be enough to overcome the opportunity gap caused by the fringe, internal, and elusive factors related to online and hybrid learning (Bielinski et al., 2020; Fridland, 2020; Shafer, 2020). Instead, Thompson (2021) argued that the pandemic has "simply exposed the problems that previously existed in schools" and that three core components will have the biggest positive impact: strengthen core instruction by identifying outcomes, assess students to determine where they are at in terms of achievement level, and use that information to design instruction

specific to their current skill level (Belinski et al., 2020; DuFour & DuFour, 2013; Hattie, 2009; Marzano, 2017; Thompson, 2021). Research has shown that these rather unambiguous goals prove to be quite a difficult task when a teacher tries to juggle the varying needs of all students in a typical heterogeneous classroom (Briggs, 2020; Gallagher & Herradine, 1997; Tomlinson, 2015).

Research reveals promising results of flexible grouping as a strategy that can encourage student growth (Datnow & Park, 2018; Matthews et al., 2013; Riley, 2016; Rytivaara, 2011). As the impact of COVID-19 on student learning becomes fully understood, strategies that accelerate disrupted teaching and learning have become more imperative. Research from the pre-COVID era can point to what has been effective in the past and offer hope for strategies that may be effective now. Rollins (2014) acknowledged that there is a strong tendency for teachers to delay access to grade-level materials until all missing skills have been acquired, which results in a delay in grade-level skills acquisition and, subsequently, widens the achievement/opportunity gap. Further, in the 2018 TNTP report, *The Opportunity Myth*, researchers outlined just four key components that were missing in students who had failed to thrive after graduating from high school: access to grade-level assignments, strong instruction, deep engagement, and teachers who had high expectations. Recognizing the difficulty faced by a teacher to meet all of these needs in a classroom with students who have varying levels of understanding of a particular subject, it stands to reason that this school district would seek a strategy to allow more efficient use of teachers' time and talents by allowing them to work together to streamline the lesson planning process and meet the diverse needs of students (Perez, 2019).

Researchers have identified a process of frequent within-school sorting that results in students learning with peers of similar ability (Kalogrides & Loeb, 2013; Riley, 2016). Flexible grouping is differentiated from tracking by its use of on-going assessment to ensure that students get the right instruction. Whether grouping within an entire school, regardless of grade-level (Matthews et al., 2013; Riley, 2016) or within a grade-level (Datnow & Park, 2018; Rytivaara, 2011), the primary focus of flexible grouping is always the skills that each student is ready to develop (Datnow & Park, 2018).

**Purpose**

In order for school districts to see growth at all achievement levels, instruction must match students' current abilities (Datnow & Park, 2018; Matthews et al., 2013; Neuman, 2016; Riley, 2016; Rytivaara, 2011; Tomlinson, 2010). Riley (2016) suggested further research to recognize and affirm the teaching strategies that address the unique learning needs of students that are embedded within Fullan and Quinn's (2016) coherence framework. This framework is a learning theory that emphasizes the importance of connecting new knowledge to existing knowledge in order to improve learning. Going one step further than constructivism, the coherence framework suggests that learning is most effective when new information is integrated with prior knowledge in a meaningful way. Providing an important conceptual framework for this study, the coherence framework ascertains the importance of focusing on direction, cultivating collaborative culture, deepening learning through clear learning goals and capacity building, as well as securing both internal and external accountability while underscoring the importance of leadership among each of these components. Embracing the coherence framework and rethinking traditional instruction, flexible grouping invites teachers to homogeneously

group students based on the specific skill being worked on, while providing an opportunity for students to “flex” or move in or out of the group, based on their current level of understanding.

**Study Structure**

The structure of this study describes flexible grouping as a potential hybrid strategy to increase student growth by exploring three research questions: How do teachers use flexible grouping to address the diverse academic needs of students? How do teachers work together as a grade-level team to meet student needs? What were teachers’ perceptions of the professional development opportunities and their subsequent efforts as part of a collaborative team?

The next chapter provides a comprehensive review of the literature on achievement and opportunity gaps, as well as grouping strategies that have the potential to meet the instructional needs of all students. Additionally, a review of the improvement science framework guided the methods of this research project. Given the broad focus of many of these studies, the literature review was narrowed down to primarily those studies that included elementary students and within-school factors. Chapter 3 describes the qualitative methodology used to collect and analyze the data in a three-phase approach, while Chapter 4 explores the findings specific to the research questions. Finally, Chapter 5 presents a discussion of the findings, and Chapter 6 outlines conclusions and recommendations for action and future research.

# CHAPTER 2

# REVIEW OF LITERATURE

This literature review utilized a thematic and chronological structure (Torraco, 2016) that traced the development of ideas related to the historical foundations of impactful learning strategies and their integration to allow the use of flexible grouping to address the diverse learning needs of students in order to maximize student growth (Datnow & Park, 2018; Matthews et al., 2013; Neuman, 2016; Riley, 2016; Tomlinson, 2010, Rytivaara, 2011). Further, relevant information on improvement science and inquiry-based teacher professional learning is included within the chapter on methods.

## Conceptual Framework

My observations from our district data aroused my curiosity to want to know more about what was in my purview to maximize student growth. Beginning with broadening my understanding of achievement gaps and why they exist, I then explored strategies that have been shown to have a positive impact on student growth. When I began this literature review, the parts and pieces existed in silos—each important yet independent. It wasn't until this research project was concluded that I began to make sense of how each of the pieces fit together in an interconnected and sequential manner.

This thematic review of literature is organized first by a historical overview of achievement gaps to understand the gap in student achievement between various sub-groups of students (Coleman, 1968; Huang & Sebastian, 2015) and then an overview and exploration of three areas of research that support addressing the gap: within-school strategies, instructional grouping strategies, and foundational components of a

collaborative culture, frequently referred to as professional learning communities or PLCs.

Because the conceptual framework that shows how these thematic silos are connected beyond the thematic review was not conceived until the conclusion of this study, a visual representation of the conceptual framework can be found in the section on conclusions.

**Achievement Gaps**

This section provides a historical overview of the evolution of the achievement gap literature, including between-school achievement gaps and within-school achievement gaps, with a particular focus on within-school achievement gaps associated with poverty, as those are of most concern in Gothenburg given the district is 97% White, and 37% percent of students qualify for free and reduced-price lunch (FRPL), which is frequently used as a proxy measure of poverty (Domina et al., 2018). Understanding that a gap does indeed exist, as well as the factors that contribute to that gap, are critical in aiding the understanding of the overall conceptual framework.

An achievement gap is defined as the difference, or the gap, in student achievement between various sub-groups of students (Coleman, 1968; Huang & Sebastian, 2015). The majority of achievement gap research focuses on the overall achievement in core content areas as measured by assessments such as the National Assessment of Educational Progress (NAEP) and state accountability assessments. Researchers have consistently found achievement gaps between Black, Latino, and American Indian students and their White and Asian peers. Achievement gaps have also been identified between students with disabilities and their non-disabled peers, as well as

between students from economically disadvantaged homes and their more affluent peers (Coleman et al., 1966; Hussar et al., 2020).

### *Historical Foundations*

As part of the 1964 Civil Rights Act, Congress commissioned the now-infamous Coleman Report with the intent of establishing that school segregation caused academic disparities, to support further desegregation (Alexander & Morgan, 2016). The report suggested that factors outside of schools, including family demographic characteristics, had a greater effect on student outcomes than factors inside of schools (Coleman et al., 1966). Looking specifically at between-school achievement gaps, the Coleman Report (1966) proposed that family demographics were more impactful than any within-school factors.

The Coleman Report (1966) outlined the history of education, spanning a period of 150 years during which education reforms focused on providing free education to all students, a common curriculum, similar schools for diverse students, and locally funded financial equality. The report outlined a stark contrast between the performance of the very demographics that were initially "left out" of public education—those who were very poor, Indian, or Southern Negro. The report left readers, as well as researchers, wondering if these results are because of a lack of opportunity, lack of access, or a plurality of factors from both (Atteberry & McEachin, 2020; Alexander & Morgan, 2016; Downey & Condron, 2016; Hanushek, 2016). Coleman et al. (1966) posited that, if the impact of within-school factors, such as high-quality curriculum, per-pupil expenditure, quality of teachers, and teachers' expectations of student achievement were weak,

minoritized and students living in poverty would continue to perform lower than their counterparts.

In the same vein, Christopher Jencks published *Inequality* (1972), a body of research that was widely interpreted to have proven within-school factors as having little to no influence on student outcomes, citing familial factors as the biggest influence. Instead, a critical view of this work by Ravitch (1973) outlines a viewpoint that equality of opportunity does not produce equality of environment (i.e., income, family structure, etc.), meaning that "schools are not able to overcome genetic and environmental inequalities among children" (). Cited by both supporters and critics of educational reform, Jencks's social philosophy can be simplified by stating that school reform is not the best way to eliminate poverty—only increasing the income of poor families can do that.

Both the Coleman Report (1966) and Jencks' *Inequality* (1972) represent a viewpoint that familial factors have a more influential impact on student outcomes than within-school factors. Further, both works indicate that addressing familial challenges by only improving within-school practices is likely to yield limited results.

### *Contemporary Between-Group Achievement Gap Research*

Social constructs named through the research on capitalism by Karl Marx and Max Weber, as well as the research on race and gender by W.E.B. DuBois and Charlotte Perkins Gilman, are the historical and foundational works that developed the contemporary views on education and inequality in the U.S. (Grusky & Hill, 2017). These bodies of research and others have led to a generally accepted contemporary belief that class, race, and gender are the fundamental forms of inequality (Grusky & Hill,

2017). Despite 60 years' worth of efforts to break the correlation between race, poverty, and educational outcomes since the Coleman Report, minoritized students and students living in poverty continue to be segregated within neighborhoods and schools and consistently underperform in both reading and math on a national level (NAEP, 2020).

Hussar et al. (2020) found that the correlation between the characteristics of students' families and outcomes, such as achievement scores, remains strong. These characteristics include poverty, parental education level, and family structure; however, they are not equitably distributed across racial groups. For example, Hispanic students are more likely to live with parents who have not completed high school (23%) compared with White (3%), Black (9%), Asian (6%), Pacific Islander (10%), and Native American (11%) students (Hussar et al., 2020).

Rothstein (2004) argued that "the influence of social class characteristics is probably so powerful that schools cannot overcome it, no matter how well trained are their teachers and no matter how well designed their instructional programs and climates are" (p. 5). He goes on to outline a range of policies that could significantly reduce the achievement gap, such as reducing unemployment and increasing income among the poor, ensuring stable housing and better health care, and finally, expanding early childhood education and after-school/summer programs. All of Rothstein's (2004) recommended shifts seem to support the findings of the Coleman Report (1966) and Jencks' *Inequality* (1972).

Reardon (2011) deemed factors related to the achievement gap as "complex and interconnected." He goes on to explain a concerning feedback loop, that is, family income is the primary predictive factor of a child's academic achievement, while

educational attainment has become more predictive of adults' income; therefore, those who achieve more in school are likely to make more money and later have children who benefit from higher family income. The feedback loop creates a vicious cycle of a "more unequal and economically polarized society" than the current one. Reardon (2011) added that certain factors further complicate this feedback loop: high-income families invest more time and money in their child's early development, have greater socioeconomic and social resources that benefit their children, and the widening income gap between the 10$^{th}$ and 90$^{th}$ percentiles subsequently leads to a greater difference between school quality and opportunities since most schools are funded by taxpayer dollars that most often represent the communities in which their schools are located.

Reardon (2016) later expanded on this work in which he sought to determine if segregation exacerbated racial inequality in education. He concluded that exposure to poor schoolmates was linked to larger achievement gaps, especially in the case of Black and Hispanic students. These findings indicate that reducing school segregation may be an effective means of improving access to high-quality educational opportunities.

Contemporary researchers have moved beyond measuring the correlation between student demographic characteristics and assessment scores to understand the causes of these disparities. Ladson-Billings (2006) and Datnow and Park (2018) expressed that inequitable outcomes represent situations in which students have unequal opportunities to acquire skills. One way of understanding educational inequities is through developmental systems that conceptualize learning as something that occurs within interactions between children and their environments. Osher et al. (2020) believe that this occurs because

poverty limits positive environments and racism causes stress. Further, they summarized the challenges children face when access to opportunities is compromised:

> Children's growth is characterized by complex, dynamic transactions between nature and nurture; interpretations and internalizations of these transactions; and the variations of these transactions across time, place, and individuals. Throughout this entire process, genes are chemical "followers"—their expression at the biological level is determined by contextual influences and developmental malleability and plasticity. Human development is not predetermined, fixed, or linear; it is not prefigured in a genetic program. Rather, it is unique to each and every individual, highly responsive to environments, cultures, and relationships, continuously adapting, organizing, and reorganizing, and subject to change across the lifespan. (p.23)

Considering this through the lens of instruction, there cannot be a single product, approach, instructional methodology, or even grouping system that can meet the needs of all children. Instead, we must consider Osher et al.'s (2020) deduction that this is highly responsive to continued environmental changes.

To summarize, contemporary researchers have moved beyond measuring the correlation between student demographic characteristics and assessment scores to understand the causes of these disparities. This has propelled researchers' descriptions of these situations beyond the achievement gap to a more descriptive approach that focuses more on a child's access to high-quality educational experiences throughout their educational career. Regardless of the terminology, between-school and within-school

achievement gaps continue to exist, and schools must continue to seek practical approaches to address these discrepancies.

### *Between-School Achievement Gaps*

In addition to between-group achievement gaps, researchers have focused heavily on between-school achievement gaps, which are in some ways similar to between-group achievement gaps. Due to the grouping based on race in a racially segregated school or class, or residential segregation, between-school achievement gaps can mirror that of between-group achievement gaps. As an example, Kozol (2012) called attention to exorbitant discrepancies between affluent schools and those in poverty which certainly impacts both the groups within the school, as well as the schools themselves. Notably, researchers have identified that some of the greatest discrepancies between schools include school funding, access to high-quality instructional materials, recruitment of highly qualified teachers, and disproportionate school discipline (Darling-Hammond, 2016; Kozol, 2012; Ladson-Billings, 2006; Russell et al., 2010).

Kozol's (2012) *Savage Inequalities* tells the story of East St. Louis, a community that was at the time 98% Black, and one-third of its families lived on less than $7,500 per year. This type of poverty is not unique to this area, nor are the challenges these kinds of demographics present for a school district, for instance, how it impacts the resources available for schools. While school funding is complicated and varies by state, the most recent data from the U.S. Department of Education show that, on average, 81% of a school's budget is derived from local property taxes, leaving just 11.2% and 7.8% of the school revenue coming from state and federal aid, respectively (NCES, 2020). Because such a large portion of school funding comes from property taxes, the value of those

community assets is what determines the largest portion of potential school revenue. Simply put, communities with low-value homes and low-income residents result in low property tax revenues and, therefore, contribute less revenue for local schools.

As mentioned in *Savage Inequalities*, the access to funding is so limited that budget cuts result in class sizes of 30–35 students and consistent layoffs of teachers while simultaneously employing many "permanent substitute teachers" for less than $10,000 per year as a way to save money. Further, many teachers comment that they feel isolated from educational development and often have materials that are over 30 years old, which contributes to the challenges of providing high-quality education.

Kozol's (2012) description of East St. Louis is in direct contrast with what many would consider possible. Ladson-Billings (2006) noted there is a generalized belief that equal opportunity is the norm in schools and, therefore, the assumption is that the achievement gap is simply a result of genes, culture, or lack of effort on the part of certain groups of students. Jimenez-Castellanos (2012) analyzed this contemporary deficit ideology by evaluating funding and policy implemented as a result of the Coleman Report, even going so far as to propose that reformation agendas to date have focused on changing culture and behaviors of minority students so that "they can resemble that of affluent families" by labeling them as "disadvantaged" and specifying a need to "compensate" for "inadequacies of students" as opposed to focusing on the strengths of minority populations and building on those strengths (p.50).

Deficit ideology or viewing of others as "less than" (Jimmenez-Castellanos, 2012) creates a school environment that is uniquely predisposed to the overrepresentation of minority students with regard to student discipline (Russell et al., 2010). If a student does

not conform to the specified standards, they are less likely to be valued as having the potential to succeed (Coleman et al., 1996) and more than twice as likely to receive discipline in the form of suspension or expulsion, especially in less diverse and more affluent communities (Russell et al., 2010). These statistics paint a dire picture that describes why the achievement gap was further described as "education debt" by Ladson-Billings (2006) or "opportunity gap" by Datnow and Park (2018) in an attempt to better describe the root cause of the gap observed in standardized achievement scores.

Between-school achievement gaps have been attributed to gross discrepancies in school funding which result in limited access to high-quality instructional materials and highly qualified teachers, as well as disproportionate assignment of suspension and expulsion to minority students (Darling-Hammond, 2016; Kozol, 2012; Ladson-Billings, 2006; Russell et al., 2010). With so many large-scale factors to consider, many researchers have directed their time and effort toward exploring more tangible factors that educators themselves can directly influence on a smaller scale—within-school factors.

### *Within-School Achievement Gaps*

These gaps are caused by challenges regarding cultural proficiencies, the provision of equitable resources within the school system, and the development of teacher efficacy to ensure equal access to high-quality instruction (Leithwood, 2010; Milner, 2010; Pace, 2014). The No Child Left Behind Act of 2001 required schools to report assessment scores by student subgroups (e.g., race, poverty, disability), which highlighted local student performance for the first time with high-stakes requirements (Karen, 2005). Ever since, there has been a laser-like focus on improving outcomes for students at all levels, particularly minority students, irrespective of race, socioeconomic

status, or those who qualify for additional support through an individualized education plan (IEP). I chose to focus on within-school achievement gaps, as they are within my circle of influence. With 34% of our families earning less than a living wage, it took some latitude to assimilate conceptual parallels between the research and the existing socio-economic diversity in this community.

In the text *Start Where You Are, But Don't Stay There,* Richard Milner (2010) draws attention to the current state of affairs (i.e., the existent achievement gap) and presents some suggestions for school districts to address the within-school factors that may impact student outcomes. Milner (2010) called on teachers to quit pretending that racism doesn't exist and instead embrace the culture and unique characteristics of their students. He proposes that, in doing so, teachers will be better able to respond to cultural conflicts, build awareness of situational challenges, and shift away from low-expectations and deficit mindsets. The overarching theme of the text is a call for educators to commit to learning more about the lives and experiences of their students and to work to change policies, procedures, and practices as needed to reflect the assets and needs of the school and community (Milner, 2010; Pace, 2014).

The Coleman Report (1966) and research on achievement gaps have placed the onus for educational failure on children and families, calling attention to demographic factors beyond the control of schools. However, more contemporary research has since identified that achievement gaps may actually reflect more of a lack of opportunity than a lack of potential (Datnow & Park, 2018; Ladson-Billings, 2006). Occurring both between-schools and within-schools, the recognition of the lack of opportunity

effectively shifts the responsibility for student achievement onto teachers and schools to meet the needs of diverse students rather than as a result of periphery demographics.

The extension of the concept of "achievement gap" to "opportunity gap" forces educators to look at these at-risk categories and risk factors as opportunities to grow rather than roadblocks to growth (Datnow & Park, 2018; Flores, 2007). With the onus shifted to schools, educators are then empowered to examine strategies within their circle of influence to positively impact student outcomes. This study sought to do this by learning from teachers who consistently experience higher growth rates among students than what is expected and then creating professional learning opportunities to expand the impact.

**Opportunity Gap**

Following the Coleman Report, school effectiveness researchers began to seek evidence of the effects of schools on achievement. Since the 1970s, school and district effectiveness researchers have identified correlates of student achievement: positive climate, high expectations, clear goals, opportunities to learn, instructional leadership and alignment, progress monitoring, positive school–home relationships, organizational coherence, and a clear mission (Brookover & Lezotte, 1979; Edmonds, 1979a, 1979b; Elmore & Burney, 1997; Reynolds et al., 2015; David et al., 2000; Waters & Marzano, 2006).

Without the use of effective research-based strategies, these gaps will continue to increase, further impacting student outcomes (Brookover & Lezotte, 1979; Datnow & Park, 2018; Edmonds, 1979a; Ladson-Billings, 2006; Reardon, 2011). This section reviews the literature on efforts to close achievement gaps, particularly within-school

gaps, that provide a basis for developing an intervention in Gothenburg. To do this, we looked at the research around strategies to close within-school achievement gaps, such as goal setting, professional development, implementation of differentiated instructional strategies, as well as instructional grouping strategies such as heterogeneous grouping and flexible grouping, and differentiated instruction (Leithwood, 2010; Milner, 2010; Tomlinson, 2010).

### *Summary of the Achievement/Opportunity Gap*

Contemporary researchers have provided ample support for the idea that schools can make a difference (Coleman, 1968; Leithwood, 2010; Milner, 2010). The differences in achievement between subgroups of students both between-school and within-school are factors that can positively influence student outcomes if only educators engage in a strategic process to set goals, educate teachers, and provide quality resources and opportunities for students and teachers to thrive (Leithwood, 2010; Milner, 2010). The literature in this section informs this study by providing a blueprint for designing both the methods and construct in which this study took place.

Acknowledging that the students in this case study have demonstrated inequitable outcomes, which represents that some students have had an unequal opportunity to acquire skills, offers a new lens through which to explore the "complex and interconnected" factors of the achievement gap (Datnow & Park, 2018; Ladson-Billings, 2006; Reardon, 2011). By shifting the research focus, the onus is moved from simple acceptance of the situation to that of addressing equity of access to high-quality educational experiences and implementation of strategies that have proven effective despite social class characteristics (Reardon, 2011; Leithwood, 2010). This study sought

to learn from teachers who consistently experience higher growth rates for students than what is expected.

## Within-School Strategies that Impact Student Learning

Atteberry and McEachin (2020) proposed that student growth rates vary from school to school more than achievement levels do, which indicates that within-school strategies do impact how quickly students grow. This challenges our traditional thinking to the point of needing to consider strategies that have a high likelihood of positively impacting student learning. There is a vast body of literature that has explored specific strategies to close the achievement gap, including the importance of teacher quality, collective teacher efficacy, utilization of research-based instructional strategies, and the quality of the curriculum and/or materials (Atteberry & McEachin, 2020; Donohoo et al., 2018; Downey & Condron, 2016; Hanushek, 2016; Hattie, 2009; Oakes, 1986; Rubie-Davies, 2010; Scammacca et al., 2020; Tomlinson & Imbeau, 2010).

### *Conceptualization of Student Learning*

This study focused specifically on the reading portion of a school day in this district which has embraced explicit instruction as the primary mode of instruction. As a Marzano school district, our evaluation of a teacher's effectiveness is based heavily on their ability to set the conditions right for learning by providing clear and direct explanations of what students are expected to learn, observing their process of breaking down complex concepts into smaller, more manageable parts, examining their ability to provide examples and opportunities to practice new skills, and monitoring the type and quality of feedback the teacher provides regarding student mastery of a skill (Marzano, 2017).

While there are portions of the school day that would provide an opportunity for a researcher to explore social constructivism, or how students learn, this study was intentionally designed to explore the conditions that could be placed around student learning, to allow me to make sense of knowledge that already existed and that could be further implemented in other grade levels to maximize student growth.

The overall conceptualization of student learning in this district focuses heavily on utilizing explicit instruction, as guided by the Marzano Instructional Framework (Marzano, 2017). Specifically, teachers are taught and expected to clearly state the learning objectives and goals prior to each lesson, break down complex skills or concepts into smaller, more manageable parts, provide clear and direct examples and explanations, offer guided practice and opportunities for quality feedback, and monitor and adjust instruction based on learner progress. It is this conceptualization of learning that guided the exploration of the following within-school strategies that positively impact student learning.

### *Formative Performance-Based Assessments*

Analyzing students' knowledge requires teachers to be able to specifically evaluate what they know based on clearly defined outcomes. Black and Wiliam's (2010) review titled *Inside the Black Box: Raising Standards Through Classroom Assessment* concluded that formative assessment is a powerful tool for improving student learning. Specifically, they argue that formative assessment can help teachers identify students' strengths and weaknesses, motivate students to engage in learning by taking ownership of their progress, facilitating the development of metacognitive skills, such as self-assessment and goal-setting, and encouraging a growth mindset. They assert that these

benefits can only be realized when careful attention is given to the design and implementation of such assessments and the quality of feedback provided.

Marzano (2017) argued that performance-based assessments positively impact student learning in several ways. First, by aligning assessments with learning goals and standards, students have a clearer understanding of what is expected of them and are more likely to engage in learning. Second, performance-based assessments allow students to demonstrate their understanding of academic content in a more authentic and meaningful way, rather than simply regurgitating information for a test. Third, performance-based assessments provide teachers with valuable information about student learning, which can be used to adjust instruction and provide targeted feedback to students. Finally, Marzano suggests that performance-based assessments can also help to develop students' higher-order thinking skills and problem-solving abilities, which are important for success beyond the walls of the school. Overall, Dr. Robert Marzano sees performance-based assessments as an effective tool for improving student learning and promoting a deeper understanding of academic content.

Both Black and Wiliam's (2010) and Marzano's (2017) research indicate that assessment, both formative and performance-based, should be an ongoing and integral part of teaching and learning that informs a teacher's instruction, rather than a one-time event.

### *Data-Based Collaborative Conversations*

Data-based collaborative conversations are possible when assessments are well-aligned to essential standards, providing quality data that teachers can dissect (DuFour et al., 2017). In *Learning by Doing*, DuFour et al. (2013) guide teachers through a three-step

approach to discussing their data. First, teachers analyze the strengths and concerns that the data present by reviewing item-specific and skill-specific data. The team first considers any celebrations but also opportunities for growth. Within this phase, teachers place each student in one of three categories: below target, on target, or above target. In the second step, teachers establish goals and determine which instructional strategies they will use to meet those goals. Finally, teachers establish a plan to monitor student growth where they outline how they will scrutinize if their instructional strategies are working, how often they will monitor progress, and what data they will collect. This in-depth conversation is only possible when quality data exists for discussion, which makes creating effective assessments even more important.

### *SMART Goals*

Listed last in this category not by the level of importance but in reference to the chronological order a team is likely to follow, SMART goals are Specific, Measurable, Achievable, Relevant, and Time-bound. According to DuFour et al. (2013), SMART goals are an effective strategy for improving collaborative team effectiveness by providing a clear and focused framework for planning and assessment. Collaborative team SMART goals allow a team to extend beyond a district- or school-wide goal and drill down to the specific goals they have for themselves or the students in their classroom. By creating specific goals that are both measurable and achievable, teachers have a framework within which to focus their time and talents.

### *Teacher Quality/Individual Efficacy*

Ensuring that a teacher with strong content knowledge and effective instructional strategies, instructs students is critical (Berliner, 2005; Muijs & Reynolds, 2011; Muijs et

al., 2014; Wenglinsky, 2000). Teacher quality leads to individual efficacy by building a teacher's belief in their potential to make a difference. This parameter is so critical that the Milken Foundation boldly stated that, unless a child is taught by quality teachers, the impact of any other education input, strategy, or reform will be diminished (Wenglinsky, 2000). While it is easy to define a quality teacher as someone who helps students to learn, honing in on what characterizes a teacher's quality has been challenging for scholars. Berliner (2005) identified that testing for teacher quality was a near impossibility. Wenglinsky (2000) identified a good teacher as one who helps students learn and a great teacher as one who helps students learn how to learn by engaging them in higher-order thinking skills. By extension, Chetty et al. (2012) underscored the importance of teacher quality by proposing that the effects of a great or poor teacher persist well into adulthood (Chetty et al., 2012; Dononoo et al., 2018; Eels, 2011; Hattie, 2009).

According to Wenglinsky (2000), majoring or minoring in the subject taught is the only input factor associated with improved academic performance. They add that classroom practices such as engaging in hands-on activities, teaching higher-order thinking skills, and the methods teachers utilize to assess student progress positively impact student learning. More importantly, they state that professional development in these areas further improves student learning. This provides a lens of optimism that both initial teacher training and ongoing professional development can positively impact student outcomes.

Teacher quality is variable. Teachers teaching outside of their endorsed areas, as well as a lack of ongoing professional training, can significantly impact the effectiveness of a teacher (Wenglinsky, 2000). Within the classroom, variability of teacher quality can

be impacted by the student's perception of teaching quality, teacher's expectations, teacher's conceptions of the instructional core, classroom climate, clear articulation of success criteria, and fostering effort and engagement of all students (Hattie, 2009).

Reynolds and colleagues (2015) found that the differences in teacher quality are often so pronounced that the magnitude of variation between classrooms is twice that of school variance. This finding echoes that of Wenglinsky's (2000) and Hattie's (2009) that teacher quality and classroom practices such as engaging in hands-on activities, teaching higher-order thinking skills, and the methods teachers utilize to assess student progress can and do have a positive impact on student learning.

### *Collective Teacher Efficacy*

In addition to teacher quality and individual efficacy, collective teacher efficacy is an important factor that influences the overall impact of a teacher team on a group of students. Collective teacher efficacy is defined as the collective belief of teachers in their ability to positively affect students (Hattie, 2009). Studies show that collective efficacy results when school staff believes they can collectively accomplish great things; essentially when teachers believe they can make a difference, they are more likely to do so (Hattie, 2009). Where collective teacher efficacy is present, student outcomes are improved significantly (Donohoo et al., 2018; Hattie, 2009).

According to Hattie (2009), collective efficacy focuses on a teacher team's ability to set high expectations for student achievement, adopt effective instructional strategies, provide students with high-quality feedback, and collaborate to improve instruction. While these conditions might certainly impact culture, the outcomes are focused more on specific results. Each of these strategies has an implicit undertone that teachers and

administrators already know how to do these things. More accurately, when a teacher team has not yet developed these skills, we often observe student outcomes that fall below what is expected.

Collective teacher efficacy has been studied extensively. Hattie (2009) synthesized 800 meta-analyses that included over 50,000 studies; these studies represented millions of student experiences regarding 252 factors that influence student learning. His analysis determined that teachers are among the most powerful factors that influence student learning. Donohoo et al. (2018) further explored Hattie's (2009) claim and found the following: effect size = 1.57; hinge point = .4. This staggering effect size indicates that collective teacher efficacy is three times as impactful as classroom management (effect size = .52). This indicates that fostering "a group's shared belief in its conjoint capability to organize and execute the courses of action required to produce given levels of attainment" is essential for schools looking to improve student outcomes (Bandura, 1997, p. 477).

Leithwood (2008) found that schools with higher collective efficacy tend to have higher levels of student achievement. He concluded that, by fostering a sense of shared responsibility and collective action among teachers and administrators, schools can create a more supportive and empowering learning environment that benefits both teachers and students. To develop this capacity, districts must consider what knowledge and skills must be cultivated. The remainder of this section explores professional development to support teachers and administrators in implementing research-based instructional strategies and ensuring appropriate usage of High-Quality Instructional Materials (HQIM) to impact teacher will and develop teacher capacity.

### *Professional Development to Support Teachers and Administrators*

Professional development is an integral component of both individual and collective efficacy, not just for teachers but for administrators as well. Identified as a key factor in school change efforts (Moursehd et al., 2010; Muijs et al., 2014), professional development ensures those with the greatest potential impact on students have the knowledge and skills necessary to make that impact.

Education literature consistently demonstrates that school improvement efforts tend to follow pendulum swings. With each swing, programs are implemented quickly at a large scale, with the idea that problems can be fixed later. Ultimately, this strategy has failed repeatedly because educators typically do not know how to execute innovative ideas and lack individual expertise and organizational capacity to support changes at scale; further, effective approaches to change are often ignored by policymakers, which makes innovation even more challenging (Bryk et al., 2017).

As a critical realist, Bryk (2020) indicated that initiatives tend to lead districts to embrace good ideas, but districts take them up in superficial ways. This often leads to frantic implementation, which ends with minimal improvement in student outcomes. This study focuses on instructional grouping strategies and draws on three bodies of conceptual understanding to create an intervention aimed at increasing the use and quality of flexible grouping strategies. This study was guided by Tomlinson's (2010) model of differentiated instruction, Bryk et al.'s (2015) model of improvement science, and Timperley's (2008) model of inquiry-based teacher professional learning. This framework addresses the specific strategies needed for teachers to impact student learning through differentiated instruction while also addressing the need for professional

development focused on inquiry. Seeking to reverse the pitfalls of what Bryk et al. (2017) noted as “going fast and learning slow”, this study seeks to learn fast and implement well to ensure there is adequate time to develop the individual expertise and organizational capacity that Bryk et al. (2017) considered the biggest hurdle in developing change at scale.

### *Inquiry-Based Teacher Professional Learning*

Literature notes the importance of impactful professional development as part of the PDSA process (Bryk et al., 2017; Wilcox et al., 2017). In inquiry-based professional learning, teachers engage in a non-sequential cycle of learning rather than a sequential process so teachers can ideally observe positive outcomes of their efforts, which increases continued motivation and progress (Timerley & Phillips, 2003).

More specifically, Timperley and Phillips (2003) outlined a three-pronged, non-sequential approach to professional learning. First, professional learning focuses on confronting current teacher beliefs about student learning so teachers can set higher goals and expectations for their students. Second, professional learning focuses on building teacher self-efficacy by acknowledging the relationship between teacher actions and student outcomes. Third, a more traditional approach to professional learning is where teachers learn the theory of the new skill and develop their own skills. Teachers and districts move between these three prongs as needed when recognizing goals to be monitored or adjusted, teacher beliefs to be confronted, or new skills to be developed.

Timperley et al. (2008) suggested that a three-pronged approach to the inquiry had a greater impact on changing teacher practice than any other traditional professional development approach. By combining the strategies and lessons learned with flexible

grouping and implementing an inquiry approach to professional development, I hope to contribute to the field of education by identifying strategies in practice and outlining a successful implementation of such strategies on a larger scale.

Timperley and Phillips (2003) recommended self-regulated learning, where teachers intentionally set high expectations for students, monitor their progress, and adjust their instruction when student performance does not match the expected outcome, which nearly mirrors the PDSA process outlined by Bryk et al. (2017). The improvement science framework was specifically explored to aid in addressing the research–practice gap that exists in schools and how professional development as a result of this study can assist teachers in bridging this gap.

In addition, the literature highlights the importance of communication and strategic planning in the initial phase of the PDSA cycle (Bernhardt, 2018; Timerley & Phillips, 2003; Wilcox et al., 2017). Timperley and Phillips (2003) stressed the importance of helping teachers see the explicit connection between professional learning and increased student outcomes. Providing an iterative opportunity to confront teacher beliefs about student learning, acknowledging the relationship between teacher and student outcomes, and diving into theory and action will support districts in ensuring lasting change (Timperley and Phillips, 2003).

Knapp et al. (2007) emphasized the phrase “data-informed decision making,” which places the emphasis back on utilizing data to inform how instruction is developed. With this in mind, literature on professional development helps us further understand what steps need to be taken to maximize both individual and collective efficacy through professional learning experiences.

### *Research-Based Instructional Strategies*

In addition to teacher quality, individually and collectively, the strategies teachers use to meet student needs have an impact on student growth and achievement. Of the 256 factors that Hattie (2009) found to impact student learning, the top five included collective teacher efficacy (effect size = 1.57), self-reported grades (1.33), teacher estimates of achievement (1.29), cognitive task analysis (1.29), and response to intervention (1.29). Each of these directly reflects the importance of quality teachers and instructional strategies. In addition, Hattie (2009) also identified specific teaching approaches associated with student learning: setting learning intentions, providing challenging tasks, providing multiple opportunities for practice, knowing when teaching and learning goals have been met, understanding the role of teaching strategies, planning and talking about teaching, and seeking feedback on instruction.

Teachers and administrators recognize the importance of instructional strategies. However, sifting through those strategies can be a challenge for busy teachers. Carpenter (2000) counted nearly 400 "good ideas" that had been published in the previous decade in the widely read practitioner journal *Phi Delta Kappan*, including the Madeline Hunter method, whole language, performance assessments, assertive discipline, cooperative learning, block scheduling, outcomes-based education, national standards, looping, constructivism, full inclusion, interdisciplinary teaching, detracking, writing to read, and character education. This extremely long list of "ideas" that demonstrates the level at which educators are seeking answers to address the instructional challenges in their classrooms is a good reminder that not all ideas are grounded in research.

Professional development must bring attention to research-based instructional strategies that have a proven impact on student learning (Hattie, 2009; Wenglinsky, 2000). By building teacher quality, individually and collectively, as well as developing the tools teachers utilize for instruction in the form of research-based instructional strategies, schools can focus on a more tangible aspect of within-school variations—ensuring students all have access to high-quality instructional materials.

### *High-quality Instructional Materials*

In addition to quality teachers and instructional strategies, evidence suggests that quality instructional materials and curriculum also positively impact student learning, especially among disadvantaged student populations (Chingos & Whitehurst, 2012; Kane et al., 2016; Opfer et al., 2016). Interestingly enough, Chingos and Whitehurst (2012) found that quality instructional material impacted the teacher's instructional choices as much as it influenced learning directly from the materials, meaning that the materials served a two-fold purpose of positively impacting the learner, as well as the teacher. This is supported by Erberber et al. (2015) and Kane (2016), who found that the quality of the instructional materials themselves is equally as important as the quality of the teacher.

Although it is difficult to separate the teacher from the materials for studying the impact, Kane et al. (2016) attempted to do so. Their study found that teachers who utilized HQIM for teaching math resulted in student achievement gains of 3.6 percentile points, which is more than the improvement attributed to the increase of a teacher's effectiveness over the course of their first three years of teaching. These findings are in agreement with those of Jackson and Makarin's (2017), who found that HQIM had the

same statistical effect on learning as a teacher above the 80th percentile with regard to teacher quality.

Supporting the importance of HQIM one step further is the research of Boster et al. (2015) which found that improving the quality of instructional materials is 40 times more effective than class-size reduction. When we look at the cumulative impact of this research, the implication is that an average teacher using HQIM would be likely to see greater gains than a good teacher using poor materials. Therefore, it is concerning that a 2017 RAND analysis found that 96% of teachers use Google to find lessons and materials, while nearly 75% of teachers use Pinterest for the same purpose, leading to inconsistent quality of instruction overall (Opfer et al., 2016). Perhaps most concerning is Opfer et al.'s (2016) finding that teachers working in schools with a high percentage of free and reduced lunch are turning to Google and Pinterest more frequently than their more affluent peers and reflect lower quality than what was available to them in their classroom already (TNTP, 2018). Given the findings of Chingos and Whitehurst (2012), Kane et al. (2016), and Opfer et al. (2016), HQIM are as important as a quality teacher, meaning that HQIM must be considered part of the puzzle when looking at within-school strategies that impact student learning.

### *Summary of Within-school Strategies that Impact Student Learning*

Many factors can positively impact student learning within a school. Individual and collective teacher quality, as well as the strategies and materials that teachers use to meet student needs, all have the potential to positively impact student growth and achievement (Berliner, 2005; Chingos & Whitehurst, 2012; Donohoo et al., 2018; Hattie, 2009; Kane et al., 2016; Opfer et al., 2016; Muijs & Reynolds, 2011; Muijs et al., 2014;

Wenglinsky, 2000). With this research indicating that within-school factors can positively impact student outcomes, the importance of the quality of the teacher, quality of the curriculum, research-based instructional strategies, and collective teacher efficacy becomes vital if we are to leverage school resources to positively impact student learning (Atteberry & McEachin, 2020; Donohoo et al., 2018; Downey & Condron, 2016; Hanushek, 2016; Hattie, 2009; Oakes, 1986; Rubie-Davies, 2010; Scammacca et al., 2020; Tomlinson & Imbeau, 2010). With these tools in the proverbial teacher toolbox, we can set aside the mindset that schools cannot make a difference (Atteberry & McEachin, 2020; Downey & Condron, 2016; Hattie, 2009).

**Implications of Research at Local Level**

Murphy (2010) argued that, in order to actually close achievement gaps, we must accelerate the rate of learning of targeted student groups and that change strategies must disproportionately advantage those on the wrong side of the achievement gap. Essentially, "disadvantaged students cannot catch up to their initially higher scoring peers by making the same progress as those peers" (Ding & Davison, 2005, p. 94), and "as long as the same level of improvement occurs, the gap will not close" (Shannon & Bylsma, 2002, p. 48). While this is indeed a pragmatic approach, explaining that a child, who is at benchmark or above, will be systematically disadvantaged to provide an opportunity for others to catch up is not likely to be considered acceptable to any parent or teacher. Murphy's (2010) interpretation of equity portrays a learning environment in which some get what they need to maximize their personal success (those on the wrong end of the achievement gap) while others do not (those on the right side).

It is the belief of Gothenburg Public Schools (and mine) that every person is worthy and valued in our school community (Gothenburg Public Schools, 2021). Specifically, in our strategic plan, we claim that we will not tolerate any behavior that undermines the mission and beliefs of the district. That being said, systematically disadvantaging any student demographic would be considered highly unacceptable. It is because of this moral compass that this research project exists.

Research tells us the importance of teacher quality, the usage of effective instructional strategies, effective classroom management, HQIM, and collective teacher efficacy in improving student outcomes (i.e. Atteberry & McEachin, 2020; Donohoo et al., 2018; Downey & Condron, 2016; Hanushek, 2016; Hattie, 2009; Oakes, 1986; Rubie-Davies, 2010; Scammacca et al., 2020; Tomlinson & Imbeau, 2010). My ultimate responsibility as an educational leader is to continually seek ways to mitigate these potential roadblocks. We must ensure every student, not just a select few on the right or wrong side of the gap, has an opportunity to become the best possible version of themselves. It is because of this belief, as well as a logical conclusion that if we can meet students where they are and challenge them appropriately, they will experience success, that I selected instructional grouping strategies as the focus of this research project.

**Instructional Grouping Strategies**

There are two common classroom composition strategies: tracking, otherwise known as homogeneous grouping, and mixed-ability classrooms, otherwise known as heterogeneous grouping. In addition to reviewing the literature on these strategies, this section also explores a third hybrid model called flexible grouping.

### *Tracking and Ability Grouping*

Tracking and ability grouping have a long history in the U.S. (Alexander et al., 1978; Loveless, 1998, 2013; Lucas & Gamoran, 2002; Oakes, 2005). Often used interchangeably, tracking typically references junior high or high-school students being placed on a path to learning, between classes, that carries them to graduation, whereas ability grouping takes place within classes, typically in an elementary setting (Loveless, 2013). These two homogeneous grouping strategies were born from an increased level of accountability and a call for educational reform dating back to the early 1950s when, after the Cold War, concerns about American students competing academically at a global level were brought to the forefront (Nesmith, 2018). At that time, tracking was utilized to ensure students were ready for their college or career upon graduation from high school (Burris, 2014).

The logic behind tracking and ability grouping is to allow teachers to focus on learners deemed to have similar abilities (Loveless, 1998, 2013; Oakes, 2005). However, meeting the needs of students in this manner has shown uneven outcomes. Hattie's (2009) analysis included more than 300 studies of tracking, across a wide variety of school settings, in most curriculum areas, and across all grade levels, and using most major achievement outcomes. He found a small average effect size for tracking of .12, significantly less than the average .4 which equates to a full year of learning, revealing that tracking has minimal effects on learning outcomes. At the same time, qualitative studies in particular have identified profound negative equity effects of tracking, particularly for lower-ability students (Hattie, 2009; Oakes, 2005).

Foundational studies on tracking from the 1980s and 1990s paint a mixed picture. Recognizing the methodological limitations of such research conducted prior to the advent of modern computing, these mixed findings on tracking include a meta-analysis by Kulik and Kulik (1992), who found that students in high-ability groups received enriched instruction in honors classes, thus producing large achievement gains. However, students in average and below-average groups showed near-zero effects. In a later meta-analysis, the authors reviewed 11 studies on within-class ability grouping and found a mean effect size of 0.25 (Kulik & Kulik, 1992). Per Hattie's (2009) explanation, this would not be considered sizable and still falls well below the average value of .4. However, that effect size was influenced by the larger effect sizes found for students in higher ability groups (0.30), compared to those in the medium and low groups (0.18 and 0.16, respectively) (Kulik & Kulik, 1992).

Much like Kulik and Kulik's (1992) meta-analysis, findings on tracking and ability grouping in the last 40 years have been mixed. Studies with negative findings include a foundational study by Oakes (1986), which surmised that ability grouping was just another way to uphold the distribution of power and privilege. Oakes et al. (1990) found that tracking and ability grouping contributed to inequitable opportunities for learning in math and science, rooted in race, social class, and gender. They found a significant difference in terms of how the curriculum was implemented across tracks and that students were not exposed to the same curriculum (Oakes, 1990). Additionally, Oakes and Guiton (1995) concluded that, when students are tracked, low-income and minority students are concentrated in low-ability and non-college-bound classes, while

their more affluent and Caucasian counterparts are more likely to be placed in high-ability and college-bound classes.

Similarly, Sorensen (1987) claimed that unequal distribution of students in high-level classes is not just a result of a student's potential success but also a function of structural limits such as available slots and the desires of other students, which is further complicated by widely held beliefs that few low-income or minority students are capable or interested in rigorous academic work. Compounding this belief is a later conclusion by Loveless (1999) that revealed teachers strongly believe tracking places a student at a particular level for their entire educational career, consequently impacting their potential and outcomes.

Further, Fu and Mehta (2018), using data from the late 1990s, found that, if tracking were banned, peer composition would shift significantly and, in doing so, could increase achievement for students in low-ability tracks and reduce achievement for those in high-ability tracks, reflecting Hattie's (2009) finding that peer influences have a meaningful, positive effect on learning, particularly for those in the low-ability tracks (Argys et al., 1996; Hattie, 2009).

Despite the negative outcomes identified in this early research, other studies have found positive impacts of tracking and ability grouping. For example, a study by Argys et al. (1996) revealed that overall achievement was approximately two percentage points higher in schools with tracking than in detracked schools (Loveless, 1999).

More recent studies also demonstrated mixed impacts of tracking. For example, Mickelson et al. (2013) contended that, when students are grouped based solely on skill, they are actually being grouped by race and economics. Additionally, Rubie-Davies

(2010) studied teacher beliefs on whether students' ability to learn contributes to unequal expectations for tracked classrooms and encourages more veteran teachers to choose to teach the higher tracks, further resulting in unequal access to quality instruction. The critics of tracking and ability grouping believe strongly that, when students are sorted by ability, student achievement is negatively impacted, and segregation increases (Kalogrides & Loeb, 2013; Mickelson et al., 2013; Rubin, 2006).

However, Brulles et al. (2010) studied the percent change of pre- versus post-assessment scores on a standardized math assessment. Brulles (2010) noted that student growth from the pre- to post-assessment nearly doubled when students were in homogeneous groups across all demographics, without exception. Additionally, Duflo et al. (2011) studied 10,000 students and found that students who were in tracked settings showed a significant increase in testing scores, regardless of high, middle, or low tracks. Matthews et al. (2013) studied over 200 students and found that both gifted and non-gifted students benefit equally from ability grouping.

Although the research on tracking and ability grouping remains mixed, these instructional grouping strategies continue to enjoy wide support from stakeholders. Teachers perceive tracking as a way to provide students with instruction that matches their ability level, which makes instructional planning more efficient and effective (Ansalone, 2010; Kim, 2012). On the other hand, Biafora and Ansalon (2008) stated that nearly all of the elementary teachers in their study supported some form of tracking in their classrooms due to being overwhelmed by the task of teaching large classes of diverse learners. Nesmith (2018) found that teachers were quick to point out the merits of ability grouping as a means of meeting the instructional and curricular needs of all

students while decreasing the amount of planning and preparation necessary. Similarly, Ansalone (2010) identified that teachers feel tracking allows them the chance to enrich or remediate the curriculum based on the ability of the students.

Teacher perceptions also acknowledge the potential downsides of tracking and ability grouping. McKown and Weinstein (2008) noted that even when teachers had a positive viewpoint of ability grouping, they are equally concerned for students in the low-ability group. Despite this, Ansalone (2010) found that while teachers acknowledge the negative impacts of tracking and ability grouping, such as unequal access to and quality staff, they viewed ability grouping as necessary in order to manage the instructional planning aspect of teaching.

**Summary.** While the logic of tracking and ability grouping is to provide equitable learning opportunities—where everyone gets what they need—in reality, it appears to be more complicated. The mixed findings on tracking may be due to the differences in how schools implement and interact with these policies (Gamoran & Hallinan, 1995). For example, Kalogrides and Loeb (2013) suggest that tracking itself is not flawed, but rather it is the tendency to put lower qualified and lesser experienced teachers with students in lower tracks that limit the likelihood of success. Additionally, Gamoran and Hallinan (1995) proposed that mixed findings occur in some instances because each ability group covers similar material, moving as quickly as they can, whereas others cover different material. They also proposed that high-ability classes tend to have a better instructional climate than low-ability classes, which impacts the overall results as well. Ultimately, since most critics of tracking focus on the poor learning

conditions for lower-level classes (Oakes, 1985), Gamoran and Hallinan (1995) analyze conditions in which learning opportunities can be maximized.

If educational leaders can accept that every classroom needs highly qualified teachers who use effective instructional strategies and classroom management, as well as HQIM, the major "negatives" of tracking can be mitigated (Gamoran & Hallinan, 1995). However, without these mitigating factors, tracking will only increase inequity and, ultimately, the opportunity and achievement gap for those students most at risk (Oakes, 2005, 1990, 1986).

### *Heterogeneous Grouping*

In contrast to tracking and ability grouping, heterogeneous classrooms use only age to separate students into grade-level classes to provide the same education for all students (Oakes, 2005). Seen as a response to tracked classrooms where students have similar abilities, heterogeneous classrooms or mixed-ability classrooms ensure learners of all levels receive the same instruction. Supporters of the heterogeneous classroom grouping model argue that this type of arrangement is the only way to truly ensure the achievement gap does not widen but instead decrease, by providing all students with equal access to the school's best teachers and curriculum (Anselone, 2010; Burris & Garrity, 2008; Collins, 2013; Kalogrides & Loeb, 2013; Loveless, 2013; Oakes, 2005). For example, as a staunch supporter of this model of education, Loveless (2013) proposed that the heterogeneous grouping of students is "the only way" to ensure equal learning opportunities for students of diverse races, ethnicity, and socioeconomic status.

Unlike ability grouping, which assumes that ability is fixed, proponents of homogenous grouping assume student performance is malleable. For example, Hart et al.

(2004) proposed that a significant benefit to mixed classrooms is the opportunity to accept the presumption that "current patterns of achievement reflect stable differences in young people's potential" (p.9) is fundamentally incorrect and instead that learning ebbs and flows over time and, therefore, should not be used to determine permanent or even semi-permanent learning paths (Roberts, 2016). In the simplest form, heterogeneous classrooms are any class with students whose current level of understanding is not the same (Bailey & Bridges, 2016). However, Bailey and Bridges (2016) argued that how a teacher decides to approach learning in this classroom can vary significantly from one classroom to the other. Despite the differences in the students' current level of knowledge, a teacher may attempt to provide the same teaching to all students, create workstations within the classroom based on ability, create individualized learning plans, or even intentionally place students of different levels together for collaborative work. All these approaches are considered to be heterogeneous grouping.

Even though these examples essentially take a large heterogeneous group to make several homogeneous groups, this is still considered mixed-ability grouping in large part because all students, in theory, still have access to the same teacher, same materials, and same instructional pedagogy. However, this may not always be the case. Strikingly, Oakes (2005) found that the quality and effectiveness of these different heterogeneous grouping strategies varied based on whether the students were predominantly White versus non-White or poor versus non-poor, which brings back into question teachers' perceptions of students' abilities based on their characteristics, as well as the out-of-school factors that may impact student achievement. Similarly, opponents of heterogeneous classrooms argue that when everyone gets the same thing, growth is

limited for all students (Datnow & Park, 2018; Kulik & Kulik, 1992; Ladson-Billings, 2006). However, within heterogeneous classrooms, grouping is encouraged to better meet the needs of students (Bailey & Bridges, 2016; Johnson et al., 2000). Slavin's (1991, 2011) research found that grouping within a classroom promotes enhanced learning, engagement, and accountability.

The challenge in monitoring the effectiveness of such strategies is compounded by the fact that the outcomes of these small groups vary significantly within the group. For example, a small group may consist of homogeneous or heterogeneous students, have specified goals or no goals at all, or be impacted by the facilitation skills of the teacher or para leading the group (Azmitia,1988; Lou et al., 1996; Murphy et al., 2017; Saleh et al., 2005). Despite this, research suggests the benefits of small-group versus whole-class instruction (Murphy et al., 2017). In fact, meta-analyses conducted by Kulik (1992), Lou et al. (1996), and Slavin (1987) clearly show that small groups promote student learning more than whole-group instruction, as reported by their higher achievement levels in both reading and math when students participated in small groups (ES = +.34 and +.25, respectively). Even though these effect sizes are considered statistically significant, compared to Hattie's (2009) scale, they can be considered moderately effective.

In a larger meta-analysis of 51 studies, Lou et al. (1996) found a slightly lower impact of small groups, with an effect size of +.17. In this analysis, the effect size of grouping in math and science was higher (ES = +.20) than in English Language Arts (ES = +.13). Additionally, Lou et al. (1996) identified that low-ability and high-ability students benefited significantly more from the small-group instruction than average-ability students did (ES = +.28 and +.37, respectively, versus +.19). Additionally, their

analysis found that small groups of 3–4 students were significantly more impactful than larger small groups of 5–7 students (ES= +.22 versus -.02).

Whether small groups should be heterogeneous or homogenous remains a controversial topic among researchers and teachers alike (Murphy et al., 2017). While there remains limited research on how the composition of grouping affects achievement, Lou et al. (1996) found that homogeneous grouping created effect sizes of +.12 more than that of the heterogeneous group. However, they found that this benefit was not consistent across all demographics. For high-ability learners, the group composition did not matter; they benefitted equally from both heterogeneous small groups and homogeneous small groups. In contrast, low-ability students benefitted more from heterogeneous groups, and average-ability students benefitted more from homogeneous groups (ES = +.12).

In a similar meta-analysis, Saleh et al. (2005) considered social interactions and their impacts on student motivation. They found that individual student answers were observed more frequently in heterogeneous groups than co-constructed responses, but the reverse was true for homogeneous groups, where co-constructed responses were observed more frequently. This indicates that the composition of the group impacts student achievement, as indicated previously, but also impacts their social interactions. Murphy et al. (2017) drew on Piaget's (1932) theory of learning to explain that social interactions with peers force students to recognize gaps in their own understanding, address misconceptions, and think more complexly. Additionally, Vygotsky's (1978) theory of learning highlights a student's zone of proximal development (ZPD), or the point at which a student can learn with guidance, as the point at which students work with other students of higher ability and begin to internalize the higher-level skills so they can

complete tasks independently. These two theories of learning provide insight as to why high-ability students benefit more from homogeneous groups and low-ability students benefit more from heterogeneous groups.

**Summary.** For teachers, the inconclusive findings on student grouping do not translate well to a classroom. Where does an educator place the most emphasis? Do you utilize mixed groups for the benefit of your low-ability and high-ability students, or do you utilize homogeneous groups for the benefit of your average-ability and high-ability learners? In either situation, a group of students does not get what they need to be successful, but what is it exactly that helps them to become successful in these particular groupings? This pull between tracking and heterogeneous classrooms caused the district to begin looking for other options.

### *Flexible Grouping and Differentiated Instruction*

Accepting that homogeneous instruction and heterogeneous grouping are effective for only some students and that for some students, the district sought a hybrid solution, one that would draw on the biggest benefits of heterogeneous grouping while minimizing the negative impacts of homogeneous groups. As a result, the concept of "flexible grouping" came into the picture.

Flexible grouping and differentiated instructional approaches reflect attempts to meet student needs while reducing the potential of segregation within schools. Both strategies rely on the use of small groups of students, rather than whole-class instruction. Hattie (2009) found small group work to have an effect size of .47, slightly more than the effect size of .4 which represents a full year's worth of growth. However, response to intervention (RTI), reflects an effect size of a whopping 1.29. Defined as a process that

"provides early, systematic assistance to children who are struggling in one or many areas of their learning", RTI is essentially a mastery-based learning approach that focuses on ensuring that instruction is aligned to student skill level and is mastered before moving on to the next level of skill (Hattie, 2009). Radenich and McKay (1995) defined this concept of considering the strengths and weaknesses of students in a grouping approach as "flexible grouping". According to this early definition, the groups are created and adjusted as the needs of the students change, which allows for increased flexibility that avoids the static nature of tracking.

***Differentiated Instruction***

Like a thread woven through the critical components of differentiated instruction, flexible grouping typically does not occur in isolation; instead, flexible grouping can be seen among all components of differentiated instruction (Gorman, 2011). Defined as a teacher's reaction to student learning in which teachers distinguish "what students learn (content), how they learn it (process), and how they measure what they have learned (product) and how students respond (affect)", differentiated instruction is an instructional approach aimed at meeting the individualized needs of students (Gorman, 2011; Tomlinson, 2008).

The flexible grouping component of differentiated instruction seeks to match students with challenging and specific instruction and learning opportunities at their individual levels within a homogeneous setting (Tomlinson & Imbeau, 2010; Tomlinson, 2018). More specifically, regarding the content component of differentiated instruction, teachers must design instruction appropriate to students' stages of development, learning styles, strengths, and needs (Tomlinson & Imbeau, 2010). A commitment to content

means that frequent assessments must take place to ensure teachers are aware of the stage of student development. Not outlining specific strategies for doing so, differentiated instruction approaches this process from a commitment to a "do whatever it takes" attitude to ensure instruction is designed at the appropriate level (Tomlinson & Imbeau, 2010).

The second phase of differentiated instruction is the processing component, which implies that teachers and administrators must embrace that students approach learning from a perspective of multiple intelligences, including linguistic, logical-mathematical, musical, special, bodily-kinesthetic (Gardner, 1998; Tomlinson & Imbeau, 2010).

Thirdly, with differentiated instruction, teachers and administrators must commit to utilizing appropriate services or resources to meet varied learning needs and adjusting instruction to accommodate the learning differences of students (Tomlinson & Imbeau, 2010). This approach requires a commitment to the idea that students have different needs regarding the specific tasks that should be assigned to individual students and that students have different communication styles and response modes that must be embraced (Tomlinson & Imbeau, 2010).

Finally, the fourth component is a recognition by teachers that student emotions and feelings about learning impact both the process and the product of learning—this is known as affect (Tomlinson & Imbeau, 2010). Tomlinson (2015) states that the instruction-centered classroom is also tightly aligned with the identified essential knowledge but takes this to an instructional level in which "formative assessments are utilized to plan instruction targeted to students' varied next steps in learning" (p. 207).

Flexible grouping requires groups to be created and adjusted as the needs of the students change, which allows for increased flexibility that avoids the static nature of tracking (Radenich & McKay, 1995). As seen through the lens of differentiated instruction, the content is driven by what students know, which is determined by the product (assessments). This information is influenced by the process of instruction, which is determined when a teacher is well in tune with the students' strengths. Flexible grouping is not something "new" but rather a hybrid approach that allows groups to be fluid based on student mastery of skills. It is through the differentiated instruction approach that we can see what flexible grouping can look like, which provides a useful guiding lens to apply in this study.

**Summary.** Flexible grouping provides a hybrid approach to instruction, which allows teachers to see both the quantitative and qualitative benefits of homogeneously grouping students in a fluid manner. When students are placed in groups based on ability, student achievement is shown to significantly increase (Matthews et al., 2013; Brulles et al., 2010). Yet, when students don't have an opportunity to flex, or move, between groups they may be denied equal access to teachers that have both experience and skill in teaching students with the highest needs (Kalogrides & Loeb, 2013). By providing an opportunity for students to move in and out of groups frequently, a district ensures consistent, grade-level instruction that promotes the frequent use of data to ensure students are in the right classrooms at the right time. Embracing a flexible grouping model allows teachers to have the best of both worlds. Through this approach, students experience multiple teachers as they ebb and flow through lessons, as well as multiple

peer groups, which further ensures equitable access to quality instruction and material,s as well as instruction that is specifically aligned to their skillset.

**Collaborative Culture**

Professional learning communities (PLCs) are widely known as a mechanism to provide educators with a structured, collaborative approach to improving teaching and learning. I have intentionally titled this section "Collaborative Culture" to emphasize the skills and strategies teachers must focus on to be effective collaborators, rather than a highly commercialized acronym that often has different connotations for each educator.

An effective collaborative culture focuses on providing educators with time to work together to share ideas, strategies, and best practices (DuFour et al., 2013). A truly collaborative culture focuses on student learning as the ultimate goal, decision-making based on student data, and professional growth of educators through inquiry-based action research by continually seeking answers to four guiding questions:

- What do we want students to know and be able to do?
- How will we know they know it?
- What will we do if they don't know it?
- What will we do if they already know it?

These four questions engage teachers first with the curriculum by asking them to identify exactly what they want students to know and be able to do. Marzano (2017) called this a "guaranteed and viable curriculum"; Hattie (2009) described this in *Visible Learning* as "appropriate, specific, and challenging goals"; DuFour et al. (2013) described "what we want students to know and be able to do" as "essential standards".

The second question is primarily focused on assessment. The literature is saturated with researchers affirming the importance of assessment being both formative and performance-based (Black & Wiliam, 2010; DuFour et al., 2013; Hattie, 2009; Marzano, 2017), which means to answer this question, teachers who are strong in collaborative culture utilize measures to guide their discussions.

The last two questions can be grouped together. Both these questions focus on adjusting instruction to meet students at their specific levels. Covered extensively within differentiated instruction, this approach ensures teachers align instruction with what students are ready to learn (Tomlinson, 2008).

Taken together, these four questions guide the collaborative work that effective PLCs engage in and ensure a data-driven approach to improving teaching and learning (DuFour et al., 2013).

### *Essential Standards*

To answer the first of the four guiding questions, DuFour et al. (2013) outlined the importance of clarity when telling students what we want them to know and be able to do. Marzano (2017) claimed that in a typical K-12 school system, there are 3,500 benchmarks, 9,000 hours of instruction available, and 15,500 hours of instruction needed to cover the benchmarks. This equates to 22 years of instruction, which means that pairing down the content to only what is most essential is the only way to ensure that the most important skills are mastered (DuFour et al., 2013).

By definition, essential standards are what teachers can guarantee all students will know and be able to do by the end of the year (DuFour et al., 2013). These essential standards then guide what teachers will assess and how they will provide appropriate

differentiation for those who have not yet mastered the skills or how they will adjust for those who are ready for extension. DuFour et al. (2013) guide teachers through a process of evaluating standards through a lens of endurance, leverage, readiness, and success. Endurance captures the value of a specific skill beyond a single grade level, while leverage evaluates whether a specific skill will be utilized in multiple content areas. Readiness is the process of determining if the skills are necessary for success at the next grade level, and success focuses on the frequency at which a skill is assessed at a statewide summative level. When taken together, these four evaluation criteria help teachers identify a well-balanced set of essential standards for teachers to focus on with students. Once these essential standards are identified, teacher teams can begin to determine how they will evaluate the second question: "How will we know if they know it?"

**Gap in the Literature**

Despite the studies that recognize achievement gaps or provide general strategies for addressing the achievement gaps, few studies exist that provide explicit examples or instructions for practitioners. Therefore, the purpose of the first phase of this study is to examine how teachers see more than expected growth plans for all levels of instruction (Booher-Jennings, 2005; Gillborn & Youdell, 1999).

Though research on the effectiveness of grouping, in general, is abundant (Datnow & Park, 2018; Matthews et al., 2013; Neuman, 2016; Riley, 2016; Tomlinson, 2010, Rytivaara, 2011), a notable deficiency is found on the topic of how teachers actually utilize this strategy to impact student growth for all students. Additionally, nearly all studies call for additional professional development but do not typically follow up to

see how professional development impacts the teacher's will and capacity to implement such strategies. Researchers tend to focus on proving the effectiveness of the process with a suggestion for future research revolving around the need for professional development (Blecker & Boakes, 2010; Brulles et al., 2010; Matthews et al., 2013). However, these suggestions are vague regarding the actual strategies teachers could utilize to make flexible grouping effective. Specifically, Brulles et al. (2010) suggested that further research should focus on training teachers in the differentiation of curriculum, diversified instructional techniques, and clustering of students utilizing assessment data. Furthermore, Blecker and Boakes (2010) suggested that training should focus on the development of diverse instruction, utilizing interest centers, implementing tiered lessons, and employing performance-based assessment strategies.

With national statistics flatlined, schools must pull themselves inward and reflect on the practices within their control to further evaluate and adjust based on their specific needs (Ladson-Billings, 2006). Carpenter (2000) noted that the lack of progress can be attributed to the practice–research and research–practice gaps. He identified flawed study designs that involve single-variable studies as "ultimately useless" in a complex educational setting. He further pointed to the complexity of within-school problems and the short duration of studies as reasons why silver bullets never really pan out.

Carpenter's (2000) proposed solution includes rewriting teacher education to create better-prepared teachers, giving schools the leverage to fire underperforming teachers, and decreasing teacher workload by creating smaller class sizes and fewer classes to prepare for. Recognizing that this kind of reform is unlikely to happen, he encouraged readers to refocus on providing teachers time to learn, think, and reflect.

Therefore, this study reflects on several theories and concepts that contribute to a framework of differentiated instruction supported by locally driven inquiry using improvement science and inquiry-based professional learning to positively impact growth for all students.

**Summary**

The existence of achievement gaps is widely acknowledged, as is the plurality of reasons why they exist. Three instructional grouping approaches were reviewed: heterogeneous classrooms, homogeneous classrooms, and a hybrid approach called flexible grouping. While heterogeneous classrooms focus on equality, making sure every student receives the same opportunities regardless of need, homogeneous classrooms focus on equitable learning opportunities, where every student gets what they need even if that means an unequal distribution of human or financial capital. Acknowledging that a heterogeneous approach does not allow a teacher to address the diverse needs of students and that tracking causes unintended segregation, researchers and practitioners alike continue to explore a more hybrid approach that maximizes student growth while eliminating segregation. Following the recommendation of Datnow and Park (2017), who noted that student grouping should be "flexible", a deeper dive into the heterogeneous approach of flexible grouping was explored.

In an effort to extend a positive case of flexible grouping, the improvement science framework was explored to address the research–practice gap that exists in schools and how professional development created as a result of this study can assist teachers in bridging this gap. Utilizing the Plan-Do-Study-Act cycle of professional learning (Bryk, 2020; Carnegie, 2020), this research project focused on developing

teachers' will and capacity to implement all aspects of flexible grouping: utilizing data to create instruction specifically aligned to student skill level, and then creating and utilizing formative assessments to monitor student progress to reflect on instruction and shift students between groups based on their current skill level to ensure instruction is well-aligned to their needs to maximize student growth.

# CHAPTER 3

# METHODOLOGY

The purpose of this qualitative, improvement science case study was to explore how teachers utilize flexible grouping as a means to address the diverse academic needs of students to develop professional learning experiences that can be utilized to impact an entire building. Case study research methodology was selected to allow for an in-depth description and analysis of a bounded system (i.e., a single grade level) utilizing multiple forms of data (Merriam & Tisdell, 2016). To increase what could be learned from a single case, the study began with a focus on a positive outlier (Bryk et al., 2015; Stake, 1995), that is, the teacher team that consistently shows the highest percentage of growth above the norm on NWEA MAP Growth was selected for the first phase of this study.

This study utilized an improvement science approach (Bryk et al., 2015) to construct a three-phase design. Phase 1 analyzed a high-performing teacher team's use of flexible grouping, through interviews, classroom observations, and reviews of lesson plans. Phase 2 used the findings from the first phase to create professional development opportunities aimed at building teachers' capacity for employing flexible grouping strategies, and Phase 3 examined teachers' perceptions of the impact these professional development opportunities had on their collaborative team.

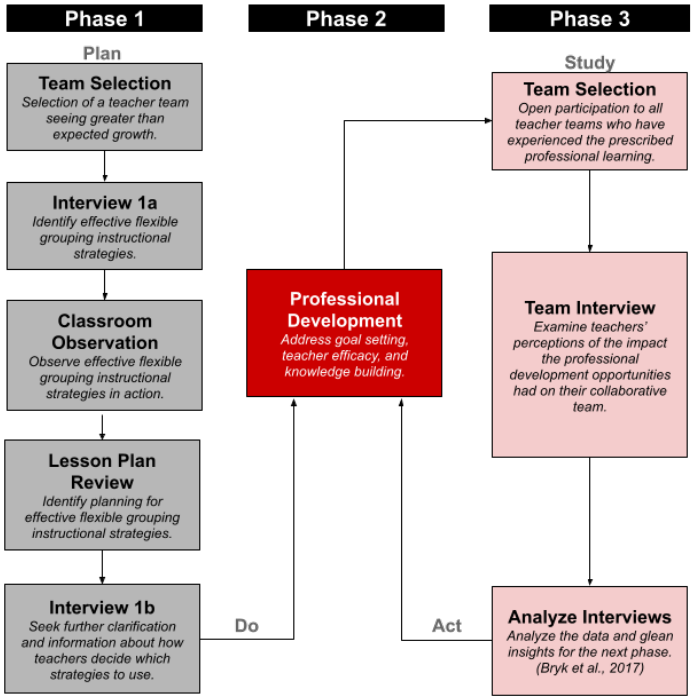
The three-phase approach of this study was designed thoughtfully to answer the research questions of this study: How do teachers use flexible grouping to address the diverse academic needs of students? How do teachers work together as a grade-level team to meet student needs? Throughout, I sought to make sense of what could be learned from a positive case, create professional learning opportunities based on what I learned,

and reflect on those opportunities to determine perceptual impact. In the final phase, I explored teachers' perceptions of the professional development opportunities and their subsequent efforts as part of a collaborative team. This case study methodology was aligned with a general theoretical construct of prior scholarly work involving differentiated instruction as a means of maximizing student growth (Datnow & Park, 2018; Matthews et al., 2013; Neuman, 2016; Riley, 2016; Schwandt & Gates, 2018; Tomlinson, 2010, Rytivaara, 2011) and the improvement science framework that guided the inquiry-based teacher professional learning outlined in Phase 2.

As addressed within the literature review, little evidence is offered concerning how teachers actually implement specific strategies that are effective. Therefore, this study sought to fill the gap in the literature and increase the understanding of effective strategies teachers utilize for differentiated instruction, as well as how and why they utilize those strategies when viewed through the lens of the improvement science framework.

**Figure 1**

*Data Collection Phases*




## Research Questions

Phase 1: Understanding Flexible Grouping in Use

1) How did teachers use flexible grouping to address the diverse academic needs of students?
2) How did teachers work together as a grade-level team to meet student needs?

Phase 2: Targeted Professional Development in Response to Phase 1 Findings

Phase 3: Unmasking Perceptions of Professional Development

1) What were teachers' perceptions of the professional development opportunities and their subsequent efforts as part of a collaborative team?

**Table 1**

*Data Collection and Analysis Strategies*

| Research Phase | Research Question | Data Sources |
|---|---|---|
| Phase 1:<br>Understanding Flexible Grouping in Use | How did teachers use flexible grouping to address the diverse academic needs of students?<br><br>How did teachers work together as a grade-level team to meet student needs? | Individual interviews with the positive outlier<br>Classroom observations<br>Review of lesson plans |
| Phase 2:<br>Targeted Professional Development in Response to Phase 1 Findings | | Improvement science |
| Phase 3:<br>Unmasking Perceptions of Professional Development | What were teachers' perceptions of the professional development opportunities and their subsequent efforts as part of a collaborative team? | Team interviews<br>Field notes |

## Researcher Positionality

Four main components have impacted my positionality and potential implicit biases I may hold on this particular research project. My own experience with education, my experience as an educator, my role as part of an administrative team, and my epistemology and ontological orientation as a researcher, all either impact or inform my positionality. To ensure what Yin (2018) outlines as internal validity, I worked diligently to create a methodology that included triangulated data from interviews, observations, and lesson plans to ensure my own perceptions were inconsequential when designing the second and third phases. The questions asked in the third phase were reflective in nature and designed to minimize my own bias through the timing of the interviews (prior to my transition to the role of superintendent) and were team-focused to ensure I could have the

teams reflect upon my findings. It is my belief that through exploring my positionality, I can be intentional in identifying a methodology that is as free from any biases or judgments as possible.

**An Early Education.** When I was in Grade 3, my parents divorced. Back then, divorce was still uncommon. I was one of two families in my grade of 100 students who suddenly did not have two parents living at home. Seemingly overnight, our family was thrust from the middle-class to grappling with the effects of poverty. I remember getting special lunch tickets that told everyone just how poor we were, wearing an outfit I had gotten at a garage sale only to have another student comment that it used to be hers, and watching my mom struggle with taking college classes while working two jobs and raising two children.

We were fortunate; my grandfather was a rancher and provided us with beef. There were many nights we had hamburgers because that was all we had. One night in particular, my mother and I ran to the grocery store in need of hamburger buns. We looked in the car to find enough change to pay for them. As we were looking, a gentleman came by and asked, “What are you doing?” My mother responded, “Finding some change for buns.” A more observant person would have noted that as a newly divorced mother of two, he should refrain from further comment. Unfortunately, his response was, “I’ve got some buns right here!” As he grabbed his behind, laughing at his own joke, the reality of our situation shook me to my core.

This experience would eventually become my “why” as an educator. The experience was not unique or even particularly traumatic in the big scheme of things, but it did help me better understand part of what other students may experience outside the

walls of the school. Most recently, 114 million people lost their jobs as a result of the COVID-19 pandemic; students from such families might be dealing with food scarcity or housing instability. Further, an increase in domestic violence during this period complicated matters for students.

After the abovementioned unfortunate encounter, I did my best to avoid being noticed by anyone for the next several years. I did just enough to get by, but never more. Not until my mother graduated with her master's degree in School Counseling did I realize just how important education is. At that moment, I realized what George W. Bush would profess years later with regard to No Child Left Behind: "Education is the key to opportunity. It's a ticket out of poverty." It was the culmination of these experiences that provided the foundation for my entire educational philosophy. I knew that if education could provide a way out of poverty for our family, then as a teacher, I could help other students find that same opportunity.

***Educator in the Making***

Once I became a classroom teacher and noticed the impact I had on my students to help them grow, my desire to reach every student grew stronger. I became an administrator to have an even greater impact, to give a voice to those who found themselves in a situation they felt they could not control and were unsure how they would ever thrive as adults. As I spent time in classrooms, I was shocked that not all classrooms looked like mine. What I learned is that, by the nature of the profession, teachers have traditionally held a very narrow perspective. Their classroom was often the only classroom they had experience with and, therefore, had not had an opportunity to see new

or different ways of teaching, so they'd keep doing their own thing, often missing out on an opportunity to collaborate and learn from their colleagues.

### *Administrative Aspirations*

In my early years as an administrator, I came to understand that there were many different kinds of teachers in our building, each of them doing the best they could, albeit some using more effective research-based strategies than others. To help build capacity, our administrative team made some radical changes, including encouraging peer observations and weekly professional development time focused on building collaborative teams and utilizing assessment data.

Through our journey, we discovered the work of DuFour and DuFour (2008). Of particular interest in our elementary school has been the concept of flexible grouping. DuFour and DuFour (2008) outlined flexible grouping as the process in which teachers work together, utilizing both formative and summative data to create fluid groups of students based on the specific skills they are ready to develop. At this point, I was between my third and fourth year of being an administrator and had been allowed to serve in a district-wide role as the Director of Teaching and Learning. In a new position that we were creating while we went, our superintendent, a dear mentor and friend, continually prompted me to move the needle for our teachers and students by learning everything I could about professional learning communities (PLCs) and collaborating with our administrative team and teacher leaders about what this opportunity could and should look like for our community. To call this a passion project would be an understatement. I could not afford to let our staff or students down with the poor implementation of such an essential process.

Further challenging this research project was an additional opportunity for me to transition to the role of the superintendent within this district I am so passionate about. Ideal timing would have allowed for this project to be completed prior to this transition so that I could be assured the data collected was free from responses that were influenced by positional authority. Instead, as a concession to the changes that were coming, the timeline of this project was condensed to ensure all data collection was completed before this transition took place. Analysis, discussion of results, and recommendations were all developed after I transitioned to my new role of Superintendent. This scenario, while not ideal, was the next best scenario given the timing of my transition.

### *Epistemological and Ontological Orientation*

My prior personal and professional experiences have cultivated my interest in ensuring that the individual needs of students are met to the greatest extent possible. I selected a single case study methodology to allow for an in-depth description and analysis of a bounded system (i.e., a single grade level) utilizing multiple forms of data (Merriam & Tisdell, 2016). Through a theoretical construct and post-positivist perspective, I seek to reveal knowledge by exploring empirical evidence and how that evidence addresses the complex social phenomena of meeting diverse academic needs in the classroom (Lincoln et al., 2011; Scotland, 2012; Schwandt & Gates, 2018; Yin, 2014). This post-positivist perspective allows me to recognize that the variables of this study will likely never be fully understood.

Through a post-positivist lens, I seek to reveal what effective flexible grouping strategies are currently taking place, so I can learn from those effective practices (Creswell & Creswell, 2018). Flexible grouping is a case that Yin (2018) defines as “a

contemporary phenomenon within its real life context" (p.15). Merriam's hybrid approach to collecting data for a case study allowed me to conduct a literature review and then construct a theoretical framework based on that literature review. I utilized Yin's (2018) approach which provides room for the use of both qualitative data and quantitative data. As I gathered data, I realized that hidden variables such as mental illness, addictions, or family situations will always impact the outcome and can never be fully accounted for in a research project. Therefore, as I began to analyze the data, Merriam's epistemological purpose of predicting, controlling, and generalizing was the lens that most closely aligned with my research questions. My post-positivist worldview guided my process and interpretation of data primarily because I was not looking to construct new knowledge but rather looking to "achieve an understanding of how people make sense out of their lives, delineate the process (rather than the outcome or product) of meaning-making, and describe how people interpret what they experience" (Merriam, 2016, p.15).

Through the research process, I drew upon Yin's (2018) work on qualitative and case study research to enhance internal validity. He defined internal validity as "the extent to which you can be confident that the case study's results are true, and not influenced by alternative explanations, or confounding variables.(p.44)" I worked to enhance internal validity through the triangulation of data, ensured reliability by disclosing my researcher's position, and enhanced external validity by utilizing a thick description of the setting. It was my goal to utilize the findings of "how" teachers utilize flexible grouping and how teachers work together as a grade-level team to impact student

growth to communicate this knowledge internally through professional development and with other scholars and practitioners in similar settings (Yin, 2018).

My experiences as a child, teacher, building principal, and district administrator make me particularly passionate about the results of this study. These same experiences have uniquely informed my interests and inspired assumptions about what constitutes an effective educator. For instance, the teachers who are most resistant to new ideas tend to show the lowest student growth rates. This potential personal bias of how I perceive these resistant teachers merits investigation, and I forced myself to be open to other possibilities and explanations; in fact, I inherently sought them. For example. the collaborative culture was not something I had expected to find. Additionally, I would have never expected the importance of formative assessment in the process of flexible grouping. Finally, I explored the challenges of flexible grouping through intentional question design and was open to the idea that flexible grouping may not be best suited in all situations or for all teachers or all students. Underpinning all research was a foundational recognition that education is a human business, which means the outcome is not predictable or certain.

By incorporating this interpretative paradigm, I affirmed that meaning is not discovered but rather constructed through the interaction between consciousness and the world (Scotland, 2012). In this particular study, the phenomenon being explored was how teachers implement specific flexible grouping strategies to ensure student growth. Through this perspective, I sought to reveal how teachers address the complex social phenomena of meeting diverse academic needs in the classroom and what can be done to expand upon a positive case (Schwandt & Gates, 2018; Yin, 2014).

**Improvement Science Case Study Design**

For this research opportunity, an embedded case study design was carefully chosen to capitalize on a positive case in order to assist with implementation on a broader scale through professional development opportunities (Creswell & Creswell, 2018). Because this team is embedded within the elementary school that will serve as the expanded case for Phase 3, this design was an appropriate selection for the primary goal of this research endeavor. For selection, the teacher team consistently showing student scores with the most growth on the NWEA MAP Growth assessment was chosen for Phase 1, which served as the basis for the creation of the professional development experiences utilized in Phase 2 and reviewed in Phase 3.

Yin (2018) defined a case as "a contemporary phenomenon within its real-life context (p.4)." For this study, the contemporary phenomenon is why one grade-level team consistently sees more growth than other teams. The real-life context for this study is Dudley Elementary School in Gothenburg, Nebraska. For the purpose of this study, the smallest bounded system available was a grade level consisting of three classroom teachers (Merriam & Tisdell, 2016). In this district, a grade-level team is considered a bounded system because they have common planning each day, are structured around student needs, work together to offer differentiated instruction, and work collaboratively to meet the needs of all students in that grade level, utilizing a common curriculum. I chose to look at an entire grade-level team to see how a unit worked together and where their strategies differed.

Merriam and Tisdell (2016) explain the purpose of a case study as the search for meaning and understanding utilizing the self as the primary instrument of data collection

and analysis, beginning with an inductive investigative strategy with a richly descriptive end product that aimed to understand one thing well. This purpose aligned well with this study in that I sought to understand the strategies teachers were utilizing, as well as how they interacted with one another to meet the diverse needs of their students. While I was not looking to construct new knowledge, Merriam's (1998) statement that "reality is constructed by individuals interacting with their social worlds" (p.6) is an excellent reflection of the research questions that guided this project.

The case study design's hallmark use of data source triangulation allowed for the collection of both qualitative and quantitative data (Yin, 2018). Quantitative data reviewed for this research project included an analysis of MAP Growth data from the Winter 2019, 2020, and 2021 assessment windows to determine the embedded teacher team with the highest growth rates. The remainder of the study relied on qualitative approaches of interviews, observations, and document review. This important use of triangulation of data increases the validity of this study and allows scholars to more confidently "follow the derivation of any evidence, ranging from initial research questions to ultimate case study conclusions;" (Yin, 2002, p. 83).

### *Improvement Science*

This study utilizes improvement science to guide the design of professional development meant to increase teachers' understanding and use of flexible grouping strategies, with a particular emphasis on Plan-Do-Study-Act (PDSA) cycles. Improvement science offers strategies for teachers and school leaders to develop teacher will and capacity while rejecting the "solutionists" described by Carpenter (2000) in the failure of the silver bullet, technical solutions.

In *Learning to Improve* (Bryk et al., 2015), educators are introduced to a new idea, one in which the responsibility for inquiry and improvement moves from university researchers and policymakers to practitioners. Underscoring improvement science is the notion of learning fast and implementing well described by Bryk et al. (2015) as a way to combat the tendency to accept solutions without proper training to implement them.

According to Bryk et al. (2017), improvement science aims to make the work problem-specific and user-centered by asking "What specifically is the problem we are trying to solve?" This allows teams to consider the variation that exists within the core of the problem. Improvement science requires teams to examine how local practices have influenced current outcomes and define exactly what the team hopes to improve by determining measurement metrics in advance. Then, teams can engage in Plan-Do-Study-Act (PDSA) cycles to learn fast, fail fast, and improve quickly. Perhaps, most importantly, Bryk et al. (2017) encourage teams to embrace the wisdom of others.

Improvement science has been used to establish a developmental continuum for reliable change on a large scale by examining the specific problem and determining how it will be measured along the way. By engaging in rapid cycles of Plan-Do-Study-Act, teams can learn fast and implement well.

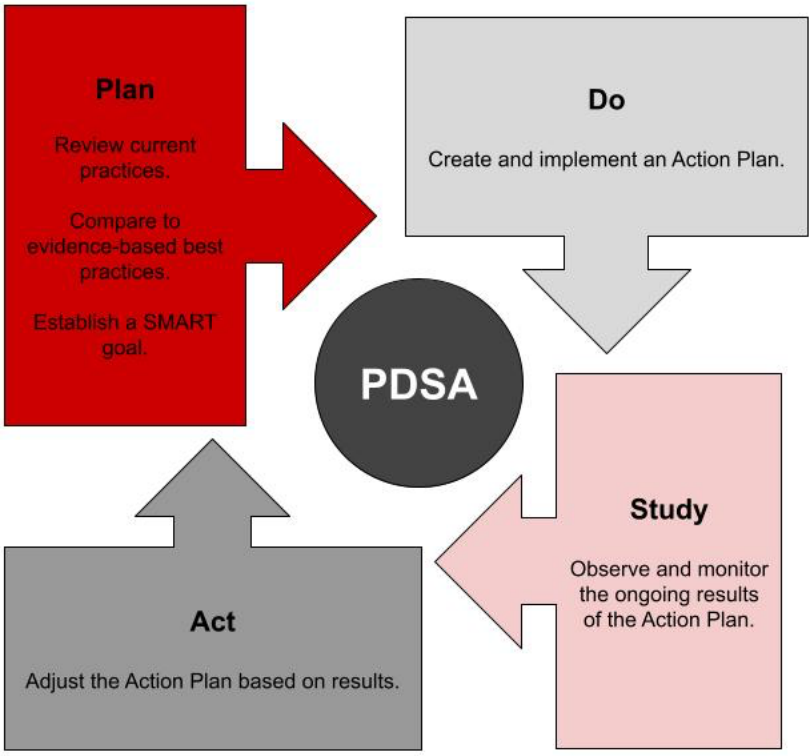
### *Plan-Do-Study-Act*

When paired down, the improvement science framework contains one component specifically that I felt would have the biggest impact on the design of this case study—the Plan-Do-Study-Act cycle. Not too long ago, the case study district suffered from what Bryk (2017) described as "solutionitis", that is, the process of continually implementing the next silver bullet without consideration of the need for training for systemic planning.

In hopes of making the planning process more systematic, I chose to focus specifically on the use of PDSA cycles as a way for the district to move toward systemwide implementation through inquiry cycles that focus on rapid learning and effective implementation. To implement well, we must first understand each phase and its overall purpose in the improvement science process.

**Figure 2**

*Plan-Do-Study-Act Cycle (Adapted from Langley et. al., 2009)*




**Plan.** In the first phase of the Plan-Do-Study-Act (PDSA) cycle, practitioners plan by reviewing available data, setting measurable goals, establishing criteria for progress evaluation, planning professional development, and creating processes to support change (Tichnor-Wagner, 2017; Woulfin, 2015). Due to the complexities of any new implementation, Woulfin (2015) encourages district leaders to reflect and prioritize

key components and ensure those components are systemic before considering supplemental foci (Woulfin, 2015).

During this phase, how leaders communicate key messages frequently and consistently is important (Wouflin, 2015). Wilcox et al. (2017) defined the "how" as "opportunity discourse", that is, the ability to portray a process more in terms of opportunities and solutions than fixing problems. By using the opportunity discourse, leaders can frame change through a positive, rather than a negative lens, to establish a mindset of innovation as opposed to a fix-it-and-forget-it mindset (Wilcox et al., 2017). The overlying theme across the literature is that districts must communicate clearly first the purpose, then the plan, and finally the results; the results must be shared if districts want the information to be utilized (Bernhardt, 2018).

**Do.** This phase focuses on carrying out the change and collecting data to document how the change was implemented (Bryk et al., 2017). This is the phase in which professional learning occurs and changes in practice take place (Bryk, 2020).

Pulling from Bryk et al.'s (2017) earlier reference that networked communities are an important component of improvement science as they allow teachers to "learn fast and implement well", one would be remiss not to include a reference to Solution Tree's *Learning by Doing: A Handbook for Professional Learning Communities at Work* (DuFour et al., 2016). Professional learning communities (PLCs) are a critical component in this phase. It is within these teams that educators carry out the change collaboratively (Bryk et al, 2017; DuFour et al., 2019).

**Study.** Perhaps the most important piece of the PDSA cycle is the results. Teacher teams must share and review results if the information is to be used for

implementing change (Bernhardt, 2018). When teachers share proven strategies, it decreases the negative impact of changes on others and supports teachers' efforts, which further develops and supports their will. The study phase of the PDSA cycle is where teachers analyze the data and begin to glean some insights (Bryk et al., 2017). In this phase, we must be careful not to leave teachers to their own devices but to support them as they look through the data (Bryk, 2020).

Administrators must support a fundamental change in which they transition from "what does your gut say" to "what does the data tell us?" (Bryk, 2020; Bernhardt, 2018). Additionally, they must support teachers as facilitators, provide data in usable forms, and scaffold the processes to support the varied needs of teachers and schools (Bernhard, 2018; Wilcox & Zuckerman, 2019).

**Act.** The final phase of the PDSA cycle is to act. Based on what is learned, it is important to explore if the idea will be abandoned, adjusted, or expanded (Bryk et al., 2017). As Bryk states in *Learning to Improve* (2017) and later reiterates in *Improvement in Action* (2020), there is no shortage of reform ideas in education…and, as he states, developing a practice-based know-how is the objective of improvement research. To round out the PDSA process, information is collected to monitor the degree to which the goal has been met and plans are made for the next steps to adjust for the amount of progress made while accounting for changes that still need to take place (Bryk, 2020). It is here within this context of "Action" that lessons learned become better ideas along an iterative learning journey (Bryk, 2020).

### *Will and Capacity for PDSA Cycles*

When districts consider PDSA cycles, they must consider that will and capacity are much like the chicken and the egg, that is, it is unclear which comes first and whether it even matters. Utilizing the communication tools outlined previously, districts can develop teacher will, but without capacity, they cannot continue making progress toward their goals (Bryk et al., 2015). Will is measured as a commitment to engage in district-wide and school-wide improvement efforts (Wilcox & Zuckerman, 2019). Equally as important, teacher capacity is defined by Tichnor-Wenger et al. (2017) as the "knowledge, skills, organizational routines, resources, and personnel available to support implementation" (p.8). It is important to note, however, that developing will and capacity is not as straightforward as it might seem. One must consider the specific context of the situation, as well as the complexities to achieve the desired outcomes (Honig, 2009).

When it comes to the process as a whole, the lynchpin within the PDSA process is districts being able to share positive results. Once districts share proven strategies, it eliminates the potential for negativity to impact others and supports teachers' efforts, which further develops and supports their will (Bryk, 2020). Therefore, communication of results is a critical component of developing will within the larger context of systemic change.

### *Validity*

Although the term "validity" is somewhat in flux, the ultimate goal of producing valid and reliable knowledge ethically is to ensure research results could be trusted (Merriam & Tisdell, 2016).

In the initial planning phase, I established criteria for the teams to be considered. Because all non-tenured staff are evaluated by me, they were not considered for this study for fear that their interviews could contaminate the data because they might have a reason to not be truthful. Additionally, they may have wanted to say the right thing as opposed to sharing what they actually do, which would have skewed the data. In addition, MAP Growth data from the Winter 2019, 2020, and 2021 terms were reviewed to narrow the pool of candidates to only those demonstrating student growth above the national norm. MAP Growth scores are reported on an RIT scale (Rasch unIT scale) that informs teachers and administrators about the specific instruction students are ready for as per the areas. By utilizing this particular unit of measurement, I was able to evaluate the teams based on how far above the norm their students were performing. This ensured that only those teams with the highest growth rates would impact the results of this study.

Prior to beginning research, I developed the interview protocols. Upon completion of the drafts, I consulted with Dr. Sarah Zuckerman, a professor who specializes in case study methodology, to ensure adequate background information was sought and that all questions were open-ended without leading the participant toward a particular answer.

Finally, interviews took place in a location inside the school, and the time was chosen by each teacher. This was done to ensure the participant felt at ease and that confidentiality could be maintained to the greatest extent possible (Elwood & Martin, 2000). During the interviews, an audio recording was used to ensure that the data were recorded accurately. After the interviews, the files were stored securely and uploaded to NVivo for analysis. This chain of evidence ensured the initial evidence and the circumstances in which the evidence was collected remained authentic.

Most importantly, as a case study, triangulation using multiple sources of data collected through interview data, including classroom observations, lesson plans, and student growth analysis, was utilized to enhance the validity of this study (Meriam & Tisdell, 2016). This process was informed by a post-positivist perspective rooted in the belief that nature can never be fully understood and that hidden variables in this study may never be fully recognized (Denzin & Lincoln, 2018). Ultimately, this empirical inquiry sought to understand a real-world case involving important contextual conditions pertinent to this particular phenomenon (Yin, 2014).

Because of the constructs put into place that were designed to enhance validity, I believe the results of this study have produced both valid and reliable knowledge in an ethical matter that can be trusted and fall within the scope and context of this study (Merriam & Tisdell, 2016).

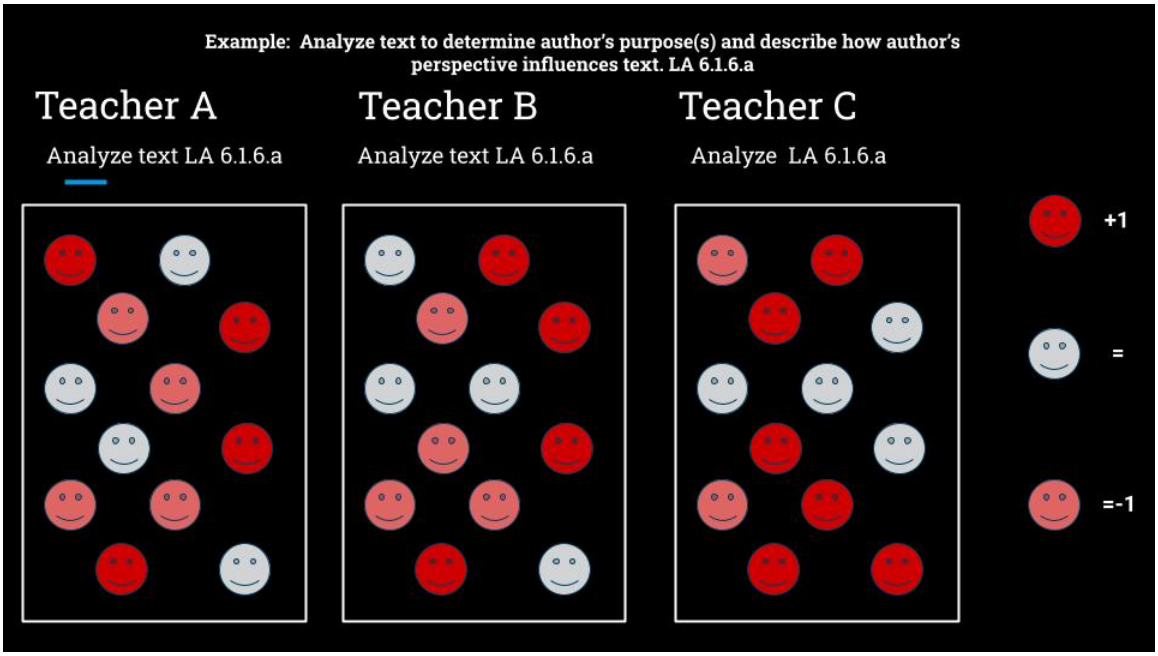
**Pilot Phase**

To develop professional knowledge and academic capacity on this topic, I engaged in an informal pilot phase that was developed in part through a research methods class and, therefore, was not conducted under the guidelines of the Internal Review Board (IRB) unlike the formal Phases 1, 2, and 3. As the Director of Teaching and Learning in the district, I sought volunteers to commit to learning about and implementing the following concepts for one year: 1) weekly team meetings to discuss lesson planning, 2) review of student data, 3) evaluate instructional changes, and 4) identify additional supports for students. Additionally, teams had to be willing to shift students among their grade-level classrooms when they mastered each identified skill or unit and not by a more traditional shift at the end of a quarter or semester.

In a school with three classrooms per grade level, a traditional mixed-ability classroom is composed of students who are ready to learn beyond grade-level content, ready to learn on-grade-level content, as well as who need support to build on-grade-level content. Figure 3 below shows what a heterogeneous classroom using only age to separate students into grade-level classes to provide the same education for all students (Oakes, 2005) looks like, as well as an example of what a standard sixth-grade classroom might be working on.

**Figure 3**

*Heterogeneous Classroom Model*

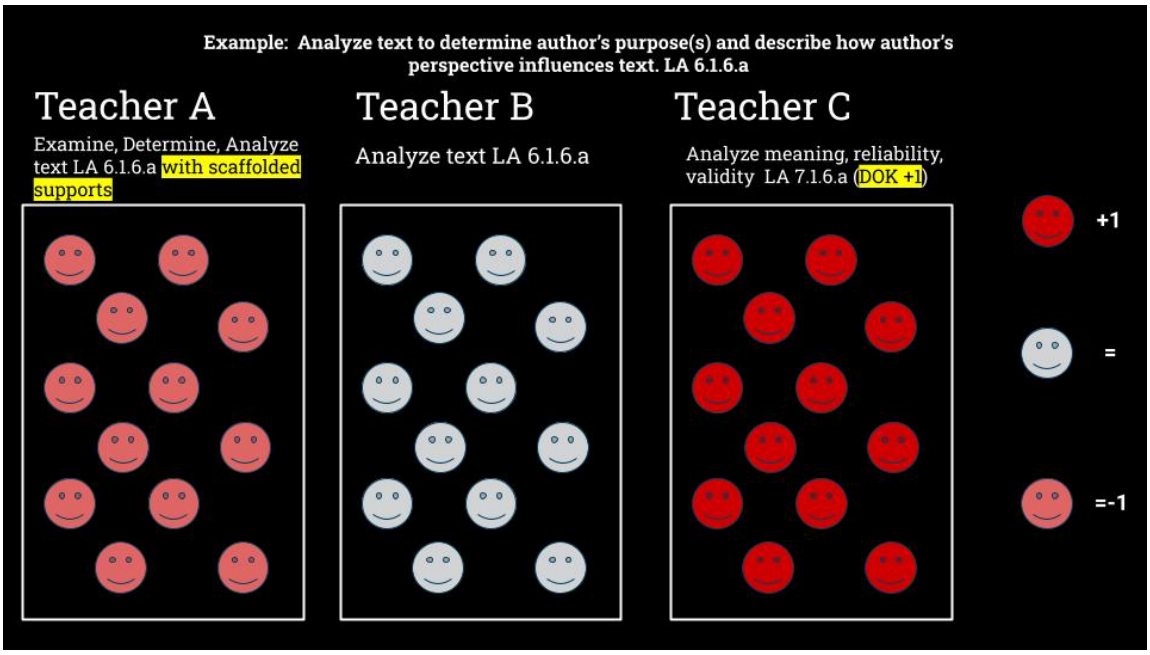



In contrast, the flexible grouping model originally explored in the pilot phase sought to find a solution that would draw on the greatest benefits of heterogeneous grouping while minimizing the negative impacts of homogeneous groups by creating an environment where students could fluidly move in and out of classrooms within that grade level as they mastered each skill, in order to avoid the static nature of tracking

(Radenich & McKay, 1995). Figure 4 below shows what a homogeneous classroom using flexible grouping as a component of differentiated instruction looks like as the classroom teacher seeks to match students with challenging and specific instruction and learning opportunities at their individual levels. In practice, the students in each of these groupings would adjust frequently based on the skill level demonstrated by students on a daily or weekly basis.

**Figure 4**

*Homogeneous Flexible Grouping Classroom Model*




For the homogeneous flexible grouping model to work, each classroom must focus on the same skill but at different levels. Figure 4 shows that Teacher A is teaching missing skills along with the grade-level standard, Teacher B is focusing on the grade-level standard, and Teacher C has moved on to the same skill but at the Grade 7 level. With each teacher focusing on the same skill, students can move fluidly as they master skills without fear of missing critical instruction. This is markedly different from tracking

or ability grouping, where students are "stuck" in these groups without the opportunity to master the skill and be further challenged. In this example, students are working with the same curriculum, materials, and standards, while teachers differentiate their instructional strategies and learning goal depth of knowledge to reflect the skills students are ready to develop. This process ensures students can easily move between groups once they have mastered the skill.

After reviewing the required commitments in the informal pilot phase, four out of nine grade levels were willing to commit to weekly meetings focusing on the required criteria and to shift students as they mastered each skill. Using the Spring 2019 NWEA MAP Growth scores, I calculated the percentage that each grade level was above or below the grade-level norm. Because the RIT scores increase with each grade level, calculating this as a percentage gave the fairest representation of how much more or less growth had taken place in each grade level. The results indicated that the classrooms that were elected to be part of the pilot had significantly higher growth rates than the classrooms that elected not to participate. Figure 5 below shows each grade level—they are placed randomly to protect the identity of the grade-level teachers—and the percent above or below the national growth norm that the students grew that year. The groups who elected to participate in the pilot are marked with a star.

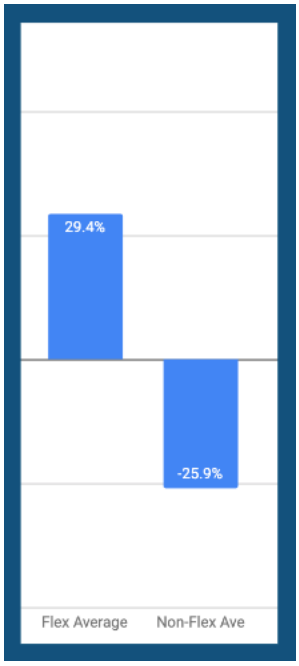
**Figure 5**

*Percent Above or Below the Fall to Spring Growth Norm on NWEA MAP Growth*


When those averages were compiled into just two sections, either “Flexibly Grouped” or “Not Flexibly Grouped”, the difference was even more profound. Figure 6 shows that the number of students in a classroom utilizing flexible grouping as part of differentiated instruction grew nearly 30% more than the national average. On the other hand, those in a classroom that had not engaged in flexible grouping grew 25.9% less than the national average.

**Figure 6**

*Flexibly Grouped Classroom Growth versus Non-Flexibly Grouped Classroom Growth*




I collected this data informally to develop professional knowledge and academic capacity on this topic. It was developed in part during my research methods class and, therefore, was not conducted in accordance with the guidelines of the Internal Review Board (IRB).

Administrators and teachers must know how teachers experiencing success with this strategy actually implement flexible grouping and what can be learned from those teachers in order to impact a larger number of teachers and students. Therefore, this study aimed to understand how teachers utilize flexible grouping to address the diverse academic needs of students and how those teachers work together as a grade-level team to impact student growth. As an extension of this learning, the purpose was to observe how targeted professional learning in these areas impacts teachers' perceptions and usage of flexible grouping as a strategy to differentiate instruction. The overall goal of this

change effort was to improve student assessment scores using a theory of action that focused on meeting students' unique needs through differentiated instruction supported by a collaborative culture among teachers.

## Phase 1: Understanding Flexible Grouping in Use

### *Data Collection*

Phase 1 data collection served as the foundation for this study. Thus, I needed to ensure the validity of the data informing the subsequent phases of this study. Careful attention was given to team selection in this phase. By first selecting a team whose students consistently demonstrated more growth than what was expected, the information learned from this team was used to inform the professional development creation in Phase 2 that would impact every other teacher team in the school. To further increase the validity of these results, triangulation using multiple sources of data—an initial interview (Interview 1a), classroom observations, and a review of teacher lesson plans—was performed to explore Phase 1 research questions of "How do teachers use flexible grouping to address diverse academic needs of students?" and "How do teachers work together as a grade-level team to meet student needs?".

### *Team Selection*

The first step in Phase 1 began with identifying the teacher team with the highest MAP Growth scores. To do this, I conducted a review of MAP Growth data from the Winter 2019, 2020, and 2021 assessment windows to narrow the pool of candidates to only those who consistently demonstrated student growth above the national norm. The national norm was selected as the comparative point for analysis because MAP Growth norms are designed for the explicit purpose of allowing educators "to compare

achievement status—and changes in achievement status (growth)—to students' performance in the same grade at a comparable stage of the school year or across two test events within or across school years" (NWEA, 2020, p.1). More importantly, these norms can be considered reliable and valid because "records are sampled from between 3.6 and 5.5 million test scores from 500,000 to 700,000 students attending over 24,500 public schools in 5,800 districts spread across all 50 states" (NWEA, 2020, p.2).

The purpose of narrowing the pool of candidates was to ensure I could understand a positive case where real-world contextual conditions yielded positive results (Yin, 2014). From there, the teams were evaluated based on how far above the norm their students were performing. This ensured that only the team with the highest growth scores would impact the professional learning opportunities developed in Phase 2.

Aggregate grade-level reports were used. These reports only identify the grade level and do not include teacher or student names, ensuring confidentiality. These reports show how much more growth students demonstrated compared to what was expected (the norm) from Fall to Winter in each of the years reviewed. This review of initial information collected from the MAP Growth system yielded three possible teams comprising three teachers each. I approached the team with the highest growth grades first by meeting with them in person and outlining the goals of this research project. After assuring them that the study was non-evaluative and would help develop professional development opportunities for their colleagues to expand upon the significant achievements happening within their classrooms, this team agreed to complete two one-on-one interviews with me, be observed in the classroom, and make lesson plans available for analysis.

**Interview 1a.** The purpose of the first interview was to gain additional insight as to how teachers meet the diverse academic needs of students. The interviews took approximately 30–45 minutes and were conducted in a place and at a time both selected by the participants, in order to ensure that the participant felt at ease, as well as the confidentiality of the participants.

During the initial interview, a semi-structured interview protocol allowed for comparability between classrooms while still allowing the flexibility to ask additional questions when the research phenomena were not initially revealed. Questions focused on participants' background information, such as "How long have you been teaching?" and "What other teaching experiences have you had?" These questions provided an opportunity to examine how these experiences impacted the participant's interaction with flexible grouping strategies and to understand each teacher's overall philosophy of teaching.

Questions about flexible grouping sought to identify the participants' personal meaning of flexible grouping and how the team member utilizes flexible grouping strategies within their classroom. These questions utilized Tomlinson's (2015) four classroom elements of differentiated instruction (content, process, product, and affect) to ensure the basic components of differentiated instruction were embedded in the interview protocol. Each of these questions was designed to help each participant describe the process of flexible grouping from their perspective. Questions such as "Suppose it was my first day as a teacher on your grade level team, how would you explain flexible grouping?" allowed each teacher to articulate their personal meaning. Then, follow-up questions such as "How often do you adjust your groups?" and "What criteria do you

look for?" provided an opportunity for expanded information on how teachers group students. Each team member was asked what their typical team meeting looks like and to expand on the goals of those conversations. Additionally, the teachers were asked to elaborate on how they plan for students at different levels. Specific questions about flexible grouping were designed to understand the strategies actually utilized.

***Classroom Observation***

Each participant selected a class period for the observation that they felt would best demonstrate flexible grouping strategies. The purpose of the observation was to expand upon each teacher's personal meaning of flexible grouping by observing first-hand how they utilized flexible grouping strategies within their classroom. Each class period observed was 90-min long. In these periods, each classroom had been homogeneously grouped based on MAP Growth data, DIBELS data, and teacher observation for the whole-group instruction portion of each lesson. Further, each classroom developed even more specific homogenous small groups based on specific skills students was ready to develop. This "group within a group" concept acknowledges the diverse needs of learners even within a single classroom. It is critical to note that what makes these flexible grouping classrooms different from a "tracking" system where students are placed on a path for learning that does not change (Loveless, 2013) is that these classrooms review formative data on a daily basis to determine if students need to move within the classroom groups or between classrooms (Radenich & McKay, 1995).

The observation protocol was not shared with them in advance, as this had the potential to alter how they typically teach, consequently skewing the observation data. It is important to note that I did not include any personally identifying student details; I

recorded student–teacher interactions using codes such as T and S1, S2, and S3. Additionally, for classrooms that were utilizing a paraprofessional during the observation, codes such as P1, P2, etc., were utilized to ensure there were not any personally identifying attributes disclosed in the observation protocol.

The literature review highlighted many strategies that could be observed in a classroom with highly effective instruction. To streamline the data collection process, I condensed Tomlinson's (2015) four classroom elements of differentiated instruction, as well as instructional differentiation strategies shared by DuFour and Schwartz (1990) and Tomlinson and Imbeau (2010) into three broad topics: 1) differentiation of curriculum, 2) diversified instructional techniques, and 3) assessment practices. Each is described in the following section.

In addition to these three broad topics, attention was given to the more subtle factors within the classroom. Drawing upon Merriam's (1998) guidance, the observation protocol called for anecdotal evidence of the environmental and physical setting, such as seating options and available resources. Additionally, I sought to capture rich detail of what was happening with the participants in terms of their roles and medium of interaction, as well as the activities and interactions that took place throughout the lesson. Paying acute attention to the sequence of activities, how students and teachers interacted with one another, how events were connected or interrelated, and directional conversation were of particular interest.

Finally, I also aimed to see what didn't happen: When instruction didn't go as planned, how did the instructor adjust? Was there nonverbal communication taking place and what was the intent? How did my presence affect the classroom dynamic (Merriam,

1996)? While not as explicit as the instructional strategies identified through the literature review, data were collected on all of these more subtle factors to help develop a more rich, thick description of what was taking place in the classroom.

To ensure field notes yielded organized information, an observation protocol was created to categorize observations in real-time as I observed and recorded observations in the field.

**Differentiation of Curriculum.** Each of these three broad topics embodied multiple categories and examples of strategies that guided me in capturing a comprehensive record of classroom events. In the first category, "Differentiation of Curriculum", I was specifically looking to see how the teacher identified what students should know and be able to do by the end of the lesson (DuFour & Schwartz, 1990; Marzano, 2017; Tomlinson, 2010; Tomlinson, 2015), how the teacher ensured the instruction was appropriately challenging (Blecker & Boakes, 2010; DuFour et al., 2013; Tomlinson & Imbeau, 2010; Tomlinson, 2015), and how the teacher-directed support staff in a manner that allowed specific and intentional practice for students (DuFour et al., 2013; Tomlinson & Imbeau, 2010; Tomlinson, 2015).

**Diversified Instructional Techniques.** Looking more at how the teacher planned for and exercised student engagement, the "Diversified Instructional Techniques" section specifically looked at how students were engaged in collaborative learning (Marzano, 2017; Tomlinson & Moon, 2013), what opportunities existed for students to learn in interest centers (Blecker & Boakes, 2010; Brulles et al., 2010; Tomlinson & Imbeau, 2010; Tomlinson, 2015), and how instruction was adjusted when students either already

knew or didn't know the desired skill (Brulles et al., 2010; DuFour et al., 2013; Tomlinson, 2018; Tomlinson, 2015).

**Assessment.** Finally, the observation protocol called for evidence of both formative and summative checks for learning, as well as how performance-based assessments were utilized within the observation period (Blecker & Boakes, 2010; Marzano, 2017; Tomlinson & Moon, 2013).

### *Lesson Plan Review*

Just as the classroom observation validated the information in Interview 1a, the lesson plan review was designed to validate the classroom observation by checking to see that not only teachers were implementing flexible grouping strategies within the classroom but that they were planning for them in advance. After the conclusion of the classroom observation, the teacher shared the lesson plans for that day's lesson, and I then reviewed the lesson plans. The primary purpose of this review was to identify innovative ways in which the teacher planned for the three broad topics utilized in the observation protocol: 1) differentiation of curriculum, 2) diversified instructional techniques, and 3) assessment practices (Blecker & Boakes, 2010; Brulles et al., 2010; DuFour et al., 2013; DuFour & Schwartz, 1990; Marzano, 2017; Tomlinson & Imbeau, 2010; Tomlinson & Moon, 2013; Tomlinson, 2015). To accomplish this task, NVivo was utilized for coding each of these categories.

**Interview 1b.** The design of Phase 1 called for data collection to conclude with a follow-up interview with each team member. Because of scheduling challenges, this worked for two of three participants; the third participant asked to combine both Interviews 1a and 1b, which occurred after the classroom observation. This combined

interview, as well as the two follow-up interviews, were less structured and focused specifically on the instructional strategies I observed in the classroom. While deviating from the specified plan is not an ideal situation for any researcher, flexibility in this situation was important in order to explore this particular team's approach. Because the primary purpose of Interview 1b was to seek clarification, I was able to seamlessly combine both Interview 1a and 1b for one of the three participants into a single session without compromising the integrity or credibility of this study.

More specifically, the second interview intended to seek further clarification and information about how teachers decide which strategies to use and what information they base those decisions on. To do so, I drew directly from interview notes to promote a more open-ended conversation based specifically on what was observed in the classroom. Interview questions such as how teachers determined who would work together, whether or not groups had adjusted since that day, and what criteria were utilized to make that decision, encouraged the teachers to expand upon the strategies I observed when I was in the classroom.

Additionally, the teachers were asked to speak about not just how their instructional block was arranged but also why they arranged it in that manner, and even though the second interview took place several days after the classroom observation, the teachers were asked to recall how many students were proficient at the end of the lesson and whether or not they shared that information with their team. Finally, teachers were asked to expand upon the planning they utilized to ensure their level of instruction was aligned with student needs and how that differs from the instruction taking place in the

other classrooms. This less structured approach ensured each participant had an opportunity to fully explore their situation during the discussion with me.

**Data Analysis**

An initial analysis of the Phase 1 data began by utilizing a deductive coding strategy. This approach was selected due to the abundance of research on effective instructional strategies and allowed me to generalize my findings quantitatively based upon the categories explored through the literature review (i.e., how often was the differentiation of curriculum, diversified instructional techniques, or assessment referenced?). To accomplish this deductive coding task, transcripts from Interviews 1a and 1b, field notes from the classroom observations, and anecdotal notes from the lesson plans were all coded using QSR International's NVivo 12 for Mac software and a set of codes aggregated from the literature review.

As noted in the literature review, there was no shortage of literature that identified effective strategies even though specific examples of what these strategies looked like in practice were sparse. Once I reached a point of saturation where no new ideas could be discovered within the literature (Merriam, 1998), six broad strategy categories were identified as codes to assist me in identifying innovative ways in which the teachers implemented these strategies: 1) clustering of students utilizing assessment data, 2) differentiation of curriculum, 3) diversified instructional techniques, 4) utilizing interest centers, 5) implementing tiered lessons, and 6) employing performance-based assessment strategies.

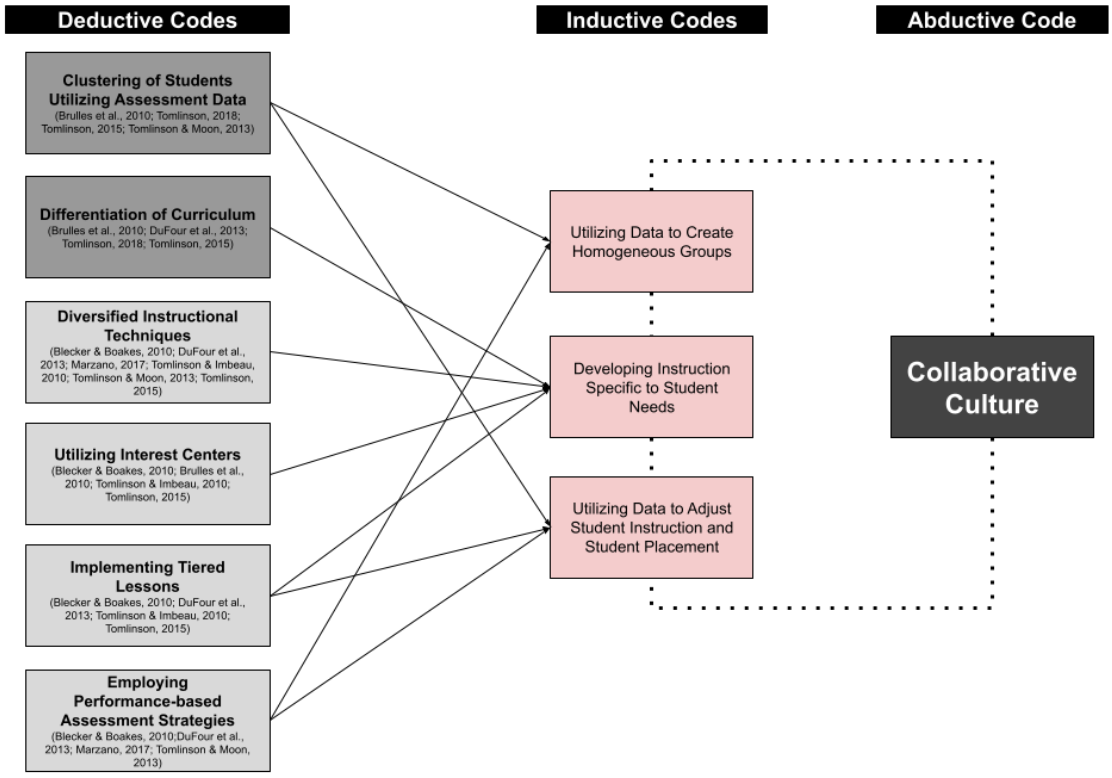
Once the interviews, classroom observations, and lesson plans were coded based on these categories, I began to sort the descriptive data so that the materials could be

physically separated from other data (Maxwell, 2013; Merriam & Tisdell, 2016). These six broad categories were eventually narrowed down to three inductive codes: 1) utilizing data to create homogeneous groups, 2) developing instruction specific to student needs, and 3) utilizing data to adjust instruction and student placement. In addition to these three inductive codes, a fourth was added as a result of the sensemaking process when I realized that often these teachers referenced their "team" during the interviews. The additional code was named "collaborative culture" to reflect the regular and deliberate approach this teacher team took on working together to improve student growth (DuFour et al., 2013).

As I analyzed the data, I found myself straddling between Merriam's (1998) constructivist lens, through which I was looking to make sense of a teacher's practice and the construction of meaning that occurs, but ultimately, my post-positivist worldview guided my interpretation of that data primarily because I was not looking to construct new knowledge but rather to reveal the meaning of the knowledge that already existed (Yin, 2018). Through this coding process, I worked to enhance internal validity through the triangulation of data, ensuring reliability by disclosing my researcher's position and enhancing external validity by utilizing a thick description of the setting. This analytic approach allowed me to utilize my findings of how teachers utilize flexible grouping and what flexible grouping means to teachers, to communicate this knowledge internally through professional development and with other scholars and practitioners (Yin, 2018). Ultimately, the broad deductive codes and narrowed inductive codes have a significant impact on the application of this research and are therefore described in more detail below; the impact of this research will be further explored in subsequent chapters.

**Figure 7**

*Coding Process*




### *Deductive Codes*

**Clustering Students Based on Assessment Data.** According to many researchers, clustering students based on current assessment information is essential in ensuring they receive instruction at the appropriate level (Burlles et al., 2010; Tomlinson, 2018; Tomlinson, 2015; Tomlinson & Moon, 2013). As an operational definition, I looked specifically for instances where student groups were adjusted based on formative or summative information that the teachers had gathered. For example, each of the three teachers had a clipboard system for tracking skills that students could demonstrate. When a student successfully completed a skill such as blending a digraph, a checkmark was added next to their name. Those that did not have a checkmark were placed in a skill-

specific small group that day, which focused on blending that digraph. By utilizing this in-the-moment formative data collection, the teachers who were observed had truly flexible small groups that focused specifically on a missing skill or extended upon a skill that students had already mastered. Beginning with the end in mind, this deductive code connected directly with Tomlinson's (2015) "affect" category, in which formative assessments are utilized to plan instruction in alignment with the skills students are ready to develop.

**Differentiation of Curriculum.** Shifting toward Tomlinson's (2015) category of "content", this code draws upon the idea of clustering students based on assessment data. The implication of clustered groups is that each group receives instruction adjusted either up or down to meet their specific instructional needs through either support, reinforcement, or extension of the content (Brulles et al., 2010; DuFour et al., 2013; Tomlinson, 2018; Tomlinson, 2015). My operational definition of this category was to see if the participants utilized different depths of knowledge to guide the skills that students were working to develop. To standardize this reflection, I utilized Bloom's (1956) Taxonomy to reflect upon whether materials were differentiated. Specifically, I was interested in whether teachers were supporting instruction of grade-level acquisition of skills by scaffolding information and if they were pushing students beyond grade-level skills when they were ready. Ultimately, I was looking for ways the core curriculum was adjusted to be more specific to the needs of students in each classroom and each small group.

A prime example of this strategy was that in one classroom, the core curriculum was being utilized to teach the skill of identifying the main idea and details of a text. In

the classroom with students identified as “beyond” (i.e., working “beyond” the grade level skill), students were expanding upon this skill by seeking evidence within the text that supported the main idea and details. Finally, the classroom with students identified as “approaching” (i.e., approaching the grade level skill) received scaffolded instruction provided in ther form of details and were working together as a class to determine only the main idea based on the details provided. Each classroom was working on the same skill with the same core materials and content but were developing their skills at much different levels.

**Diversified Instructional Techniques.** Taking differentiation of curriculum to a more practical level, diversified instructional techniques bring Tomlinson’s (2015) category of “process” to the forefront. Because Tomlinson’s (2015) phase of “process” recognizes each student approach’s learning differently, this code specifically reflected on the way teachers adjusted how they actually taught the curriculum. Strategies that I looked for had specifically been called out in research as those shown to increase student engagement in a homogeneous setting (Blecker & Boakes, 2010; DuFour et al., 2013; Marzano, 2017; Tomlinson, 2015; Tomlinson & Imbeau, 2010; Tomlinson & Moon, 2013). Operationally, I was looking for strategies such as gradual release, where teachers work with the whole class, then have students work in groups or pairs, and then finally on their own are a common way for teachers to differentiate their instructional approach to the content. The teachers who participated in the Phase 1 interview indicated that the beyond-benchmark group would often work independently, pulling in enrichment activities that hit on the next grade level standard and require deeper thinking, whereas the “approaching-benchmark group” would do a lot more hands-on activities, working in

small groups with an adult leader to accomplish the grade-level task. Additionally of interest was the amount of wait time given to students. Examples where teachers adjusted wait time to be longer for the approaching-benchmark group or prompted students to dig deeper for the beyond-benchmark group also were coded to this category. Ultimately, I was looking for instances where a teacher adjusted how they were teaching based on the students they were working with.

**Utilizing Interest Centers.** Tomlinson and Imbeau (2010) suggest that interest centers, also a "product", are a way to encourage learning through highly engaging content geared specifically to student interests while providing students with the opportunity to learn and practice essential skills. Blecker and Boakes (2010) surmise that this strategy is critical in ensuring the growth of students even though many teachers are ill-prepared to implement them. Observing this in practice was done on a more limited scale, as these tended to exist in the context of small groups that consisted of approximately one-third of the 90-min instructional block. That being said, these were still implemented consistently during this time across all three classrooms.

In the beyond-benchmark classroom, students had two centers with pre-determined work where they worked independently on high-interest content aligned with the day's instruction while their third station was where they worked directly with the classroom teacher on extending or reinforcing skills specific to them. For instance, the day's lesson was about planets, so activities revolved around the topic of planets and provided supplemental information not found in the text. For those that needed support finding information within the text, the teacher guided them through that process during that station. This content was pulled from various sources beyond the district-approved

core materials. Another example observed was that of folders of activities preplanned by teachers for paraprofessionals to work one-on-one with students throughout work time.

**Implementing Tiered Lessons.** Blecker and Boakes (2010), DuFour et al. (2013), Tomlinson (2015), and Tomlinson and Imbeau (2010), all agree that tiered lessons are a critical component of flexible grouping. Tomlinson's (1999) early work identified tiered lessons as "the meat and potatoes of differentiated instruction" Additionally, she notes differentiated instruction is a strategy that addresses a particular standard, key concept, and generalization, but allows several pathways for students to arrive at an understanding of these components based on their interests, readiness, or learning profiles." Operationally, I was looking for ways in which the teacher scaffolded instruction or added to instruction concerning a grade-level skill. Tiered lessons were observed throughout each classroom, but examples in this category were typically coded to multiple categories at a time. Differentiation of curriculum, diversified instructional techniques, and utilizing interest centers all lead to an overarching process of developing instruction specific to student needs (Tomlinson, 2015). The one thing that sets this category apart from the others is how each teacher ensured that grade-level content was being taught as a supplement to grade-level content without supplanting grade-level content.

All three teachers indicated that their core instruction was directly out of the core text and that any tiering of instruction took place beyond that core instruction. The importance of students being able to communicate about "what they're learning" with other students in the grade regardless of which classroom they were in was reiterated multiple times by all three teachers, meaning it was critically important to them that

students knew they were all working on the same skill, while teachers knew they were working on those skills at different levels. Pacing, approach, scaffolding, and depth could all be adjusted for tiered lessons, but overall, the tiered lessons were designed beyond the core materials utilizing the previously mentioned strategies for accomplishing this task.

**Employing Performance-based Assessment Strategies.** Finally, beyond the instruction and adjusting curriculum, DuFour et al. (2013) implored teacher teams to inform their collective practice by utilizing both formative and summative performance-based assessment procedures to guide interventions, inform next steps, assess team member strengths and weaknesses, and develop better strategies for meeting the needs of those students. Defined as the combination of process and product by Tomlinson (2015), the "affect" of student learning is evidenced when they can demonstrate what they know. Tomlinson and Moon (2013) expanded upon this use of after-learning assessments by seeking the use of pre-assessments to inform instruction before any instruction actually takes place. To ensure consistency in assessments, Marzano (2017) suggests the use of rubrics based on specific criteria that can aid an instructional team in utilizing assessment results to more accurately inform instruction.

I looked for formative assessment opportunities in observations and lesson plans, such as bell ringers and exit tickets, as well as less structured checks for learning, such as "thumbs up, thumbs down" and "scale of one to three" as indicators that teachers were gathering information to guide instruction. In the instances I observed, teachers were using this information to guide both what and how they taught the next "step." Additionally, in the follow-up interview, information was sought specifically about what

teachers did with this information and how students eventually demonstrated their learning in a summative manner.

### *Inductive Codes*

Ultimately, the six deductive codes helped me begin to make sense of what I was observing within the classroom and planning process, but sifting through six categories began to get overwhelming and often redundant. As I began to look deeper at the analysis of coding, I decided it would aid the overall understanding of the data if a second phase of coding was conducted. The second phase of coding focused on inductive codes that were developed by looking at which deductive codes were often categorized into more than one code. Then I tried to define what those broader categories were representing. A visual of this process can be viewed in Figure 7.

I assembled three inductive coding themes with the intent of incorporating all six deductive codes into three inductive codes: utilizing data to create homogeneous groups, developing instruction specific to student needs based on student data, and utilizing data to adjust instruction and student placement. Intent that categories should be mutually exclusive with units of data fitting into only one category, Merriam and Tisdell's (2016) work influenced this process because many of the six deductive categories were similar, often resulting in multiple codes to a single unit. Therefore, condensing these categories assisted with the overall sensemaking process.

**Utilizing Data to Create Homogeneous Groups.** Commonly coded together, "clustering of students utilizing assessment data" and "employing performance-based assessment strategies" indicated a need for these two categories to be grouped together. This conceptually makes sense because a teacher can't effectively group students,

according to assessment data, without having quality assessment data, to begin with. Therefore, these categories were combined into a single category of "utilizing data to create homogeneous groups." As with each of the inductive codes described here, the conceptual definitions of these groups did not change, nor did the content coded to these categories; only the overarching code connected to them was changed.

*Developing instruction specific to student needs.* Categories pertaining to curriculum and instruction were often coded together. It was sometimes difficult to determine if a technique was what led to the interest center or if the interest center was the definition of the technique. Similarly, tiered lessons looked a lot like differentiated curricula and vice versa. As these codes became less and less defined, the need for a code to encapsulate their theme became more and more apparent. In reflecting upon the four deductive codes of "differentiation of curriculum", "diversified instructional techniques", "utilizing interest centers", and "implementing tiered lessons", an inductive code of "developing instruction specific to student needs" was developed.

**Utilizing Data to Adjust Instruction and Student Placement.** This category was a result of three different codes often being coded together though not exactly similar. "Clustering of students utilizing assessment data", "implementing tiered lessons", and "employing performance-based assessment strategies" often occurred together even though they describe vastly different phases of instruction (content, process, and affect). What is unique about these three codes is that they rarely occurred in isolation. The thread that wove the three together was the use of data. Based on this, a new inductive code was developed to represent the instances where these three codes occurred together—utilizing data to adjust student instruction and student placement.

### *Abductive Code*

Even after I created three inductive codes to assist with sensemaking, there was a recurring theme that continued to appear in the interviews that did not fit nicely into any of the deductive or inductive codes. Instead, in a key moment of discovery, my work to this point led to a cumulative discovery in the overall process of abductive logic (Vila-Henninger et al., 2022). This discovery was that the teachers were saying more about how they worked together as a team than about the specific strategies they used to be successful.

**Collaborative Culture.** Revealed through the reanalysis of the inductive coding process was this identified area that I was not initially searching for—collaborative culture. Though a relatively new and underdeveloped coding process (Vila-Henninger et al., 2022), this abductive theme was derived from interviews that reflected this team's willingness to work together, support one another, and focus on student learning, which was not a deductive code I initially highlighted in the data collection protocols. Even though research supports the importance of collaboration as a key component in student success, I did not initially consider it a "strategy" because I was more focused on tangible skills that could be translated to professional development as opposed to softer skills such as working well as a team which are harder to grasp from a professional development standpoint (DuFour et al., 2013; Timperley, 2008).

In visiting with the participants, it became very clear that the collaborative culture was something that supported this team in being a positive case. All three teachers cited examples of when they were ill and team members stepped in to cut out lamination, cover lunches or recesses, or cover classes without any expectation. When asked what they

each bring to the team, they were each able to identify the strengths of their team members long before they could identify what they themselves bring to the table. When pushed on this, they were able to identify that they all bring ideas, are willing to do whatever it takes, and that there is no "my students" on this team, but rather a collective group of "our students", free of competition between teachers but full of support for one another, drawing upon each of their strengths. So while this wasn't an initial deductive code called out specifically in the protocols, it became apparent through the reanalysis of the inductive codes that an additional abductive code was needed to tie together the more tangible strategies I was seeking to uncover initially. Again, a visual of this process can be observed in Figure 7.

**Phase 2: Targeted Professional Development in Response to Phase 1 Findings**

The three-phase approach to this research project was intentionally designed to guide me in gathering information in Phase 1 that could be learned from a positive case which then could inform the professional learning opportunities provided in Phase 2. This design allowed for three one-hour learning opportunities in the Fall as teacher teams were focusing their work for the year. The hour-long sessions were embedded in the overall professional learning community approach for the district where these ideas were first introduced to the whole group, but then the administrative team followed up with teacher teams in the subsequent weeks, providing additional support, inquiry, and direction. With this phased approach and gradual release of responsibility, I was able to activate a Plan-Do-Study-Act cycle (see Figure 2) to review needs, develop support, and determine the next steps. Plan-Do-Study-Act cycles assist schools in focusing their school improvement efforts utilizing an improvement science framework that calls for strategic identification

of the challenge, as well as how to address it, such as scaling up flexible grouping strategies across grade-level teams to increase effective instruction (Bernhardt, 2018; Langley et al., 2009; LeMahieu et al., 2017).

Timperley's (2008) inquiry-based professional learning model, which is grounded in the constructivist theory, emphasizes the active construction of knowledge by learning through exploration, reflection, and interaction. Different from Yin's (2018) post-positivist analytic approach, where I was not seeking to construct new knowledge but to make sense of knowledge that already existed, Timperley's (2008) learning approach engages teachers in active, self-directed inquiry and reflection, using which new knowledge can be constructed. This was a good fit for this particular project since the time available with staff was limited to one-hour sessions.

Because the concepts of adjusting instruction and developing a collaborative culture are very different, it was important to determine a research-based approach to guide the development and implementation of these professional development opportunities. The planning phase of the professional development opportunities drew from the literature review and Phase 1 data analysis of this study. Information gathered in these areas fully prepared and informed the action plan developed for the "Do" phase, or Phase 2.

To ensure that the professional development was user-centered, careful attention was given to ensure teachers were encouraged at every opportunity to utilize professional discretion in meeting the defined goal of improving student outcomes (Bryk et al., 2015; Wilcox et al., 2017). Further, careful attention was given to ensure teachers could develop a shared understanding of what needed to be accomplished and what their role

was in attaining that goal. The challenge was taking all of the information from Phase 1 to inform Phase 2 with these goals in mind while not overwhelming the process or the teachers.

Teacher interviews, classroom observations, and the lesson plan review in Phase 1 provided ample evidence that effective practice could be capitalized on with additional learning opportunities for staff in the area of adjusting instruction to meet the needs of diverse learners. Abductively, it was revealed that the skill that appeared to impact this process the most was that of developing a collaborative culture. Knowing these two concepts are not explicitly connected, the hope of ensuring teachers could move toward the more explicit connection of improving student outcomes meant it was essential for teachers to be able to focus on specific goals and to see their role in the attainment of that goal in order to bolster the collaborative culture of each team (Fullan & Quinn, 2016).

Facilitation of this process proceeded in the form of three one-hour sessions that occurred in the Fall as teacher teams were focusing their professional learning community's (PLC's) work for the year. Professional development focused on developing a collaborative culture, identifying and refining specific and shared goals, and adjusting instruction. Rooted in improvement science, I hoped that these three opportunities would pull teams together toward establishing a collaborative culture and instructional strategies that would lead to improved student outcomes.

***Developing Collaborative Culture***

Fullan and Quinn (2016) defined collaborative cultures as a team where "leaders establish a non-judgmental culture of growth where it is okay to make mistakes as long as you are working on the goals and learning from your action (p.41)." Collaborative

cultures are one of four aspects of coherence, which they defined as a "shared depth of understanding about the nature of the work" and conceptualized as "a four-pronged wheel that encompasses a focusing direction, which builds collective purpose; cultivating collaborative cultures, which develops capacity; deepening learning, which accelerates improvement and innovation; and securing accountability based on capacity built from the inside out" (p.32). This conceptualization builds on the idea of DuFour et al. (2013), who defined collaborative culture as

> an ongoing process in which educators work collaboratively in recurring cycles of collective inquiry and action research to achieve better results for the students they serve. Professional learning communities operate under the assumption that the key to improved learning for students is continuous job-embedded learning for educators. (p.32)

Essentially, teachers must be willing to commit to behaviors that can help the school or classroom to achieve their agreed-upon purpose.

In Phase 1, participants identified student learning as their common purpose. Often reiterating that their approach was a team effort and never once taking sole credit, the teacher team focused on solution-based approaches to learning and exhibited a "whatever it takes" attitude repeatedly. Of utmost interest was the continued reference to a team approach that never ends, solidified by absolute trust in one another to get the job done. How does one teach this intangible quality?

In an effort to create tangible steps that would help build this collaborative culture, I went back to the literature review to determine which skills and strategies can help support this work and developed the first professional development session. This

session reflects the work of many researchers who have identified the importance of clarifying both the mission and vision when working to develop staff (Bernhardt, 2013; Bryk, 2020; Bryk et al., 2017; DuFour et al., 2013). Because of this, the first session for teachers began with a reminder of the school's newly adopted and collaboratively developed mission statement:

> Gothenburg Public Schools is the cornerstone of a proud, passionate, and progressive community where students and staff thrive in a positive and innovative environment. We ensure the growth of all individuals by inspiring them to own and maximize their potential.

Teachers discussed in their collaborative teams what parts of this mission statement spoke to them and where they saw themselves as part of this mission. Teacher responses included focusing on growth, being creative with solutions, and ensuring that "all means all", meaning all children, not just those at benchmark, have opportunities to succeed. The staff was reminded of the district's strategic plan and the day's objective of being able to define each team's most important "next step" that was connected to two strategic plan strategies: we will expand educational opportunities to allow students to maximize their potential, and we will collaborate to maximize student growth.

Teachers were reminded that collaborative work is a broad process. In this district, the administrative team has determined that collaborative work must connect to one of four guiding questions and evaluates this commitment based on weekly PLC exit tickets submitted by each team. The four guiding questions of our PLCs include (DuFour et al., 2013) the following:

- What do we want students to know and be able to do?

- How will we know they know it?
- What will we do if they don't know it?
- What will we do if they already know it?

The identification of this focused approach was essential because, at the beginning of the implementation of PLCs, several teams would often ask if they could work on something specific and unrelated to collaboration. Because of this identified focus, administrators can respond consistently to these requests with the question, "which of the four guiding questions does this connect to?" Once teams had a shared understanding of the expectation to work collaboratively and what these concepts entailed, their next phase of learning was supported by a stronger foundation.

Professional learning communities (PLCs) utilize a common language that identifies what DuFour et al. (2013) identify as "critical issues." These critical issues encompass a wide range of tasks from ensuring teams have established norms, goals, and a set of essential skills to teams having developed frequent common formative assessments that help them determine each student's mastery of essential skills.

Foundational work conducted by the school district prior to this research project ensured each team had foundational information regarding each of these critical tasks. However, teams arrived at this August professional development opportunity with varying degrees of completion. Although, each team had at least started to address each of these critical issues within their team despite not much of the work yet being complete, which was why team evaluation was needed.

To support the development of a collaborative culture among teams and assist them in focusing on their "next step" for the upcoming school year, the first professional

development activity required collaborative teams to appoint three specific roles to cultivate deeper discussion: 1) a facilitator—someone responsible for beginning the discussion with a question and for continuing conversation in times of a lull with inviting questions; 2) a scribe—someone responsible for recording points of discussion specifically making note of big ideas and questions that arose; 3) a publicist—someone responsible for summarizing the group's discussion at the end and reporting to the whole group.

Once the roles were established, the teams then utilized a forced ranking discussion on each crucial issue. The facilitator read the statement and counted down from 3 to 1. At that time, each team member would raise their hand with either five fingers to indicate they felt their team had completed this task, four fingers to indicate they felt their team was over halfway to completing this task, two fingers to indicate they felt their team was just beginning this task, or one finger to indicate their team had not yet started this task. The facilitator would then ask those who differed by more than two fingers, "Tell me why you feel that way." This process would continue until all team members could come to a consensus on where the team was at with regard to each critical issue. The teams then engaged in conversation about what needed to be prioritized for this school year.

At the end of this session, the teachers were brought back to the objective for the day, that is, they could identify their team's most important "next step". After the teams quickly assessed their level of proficiency as either beginner, proficient, or advanced, it was determined that the teams were in different places with regard to being able to set Specific, Measurable, Achievable, Relevant, and Time-bound (SMART) goals. I agreed

at that time to develop a differentiated curriculum for teachers the following week that would provide tiered lessons to support teams in developing SMART goals for their collaborative teams.

### *Developing SMART Goals*

Phase 1 revealed that the studied team continually focused on a single goal—improving student learning. The team members were not only in agreement with this goal, but they were committed to this goal in all facets of their day, from meeting daily as a team to share ideas on meeting their goal, to helping team members out when they were ill or absent to ensure the overall goal of student learning could still be met.

Therefore, each professional learning community was asked to develop a SMART goal for the 2021–2022 school year. To differentiate this experience for each PLC, teams were asked to self-identify as beginning, proficient, and advanced by reflecting on whether they were ready to describe a SMART goal and draft one as a team (beginning), construct a SMART goal (proficient), or analyze each other's SMART goals and revise their own based on feedback (advanced). Out of a total 22 teams, seven self-identified as beginners, 10 self-identified as proficient, and only five teams self-identified as advanced. It was interesting to note that the team who participated in Phase 1 self-identified as advanced. These numbers indicated the importance of differentiating professional development in order to best support teacher teams.

The "Developing SMART Goals" session took place in early September when teacher teams were very busy. To help develop will, the lesson activities began with establishing why goals are important. The advanced group created their own list, while the proficient and beginner groups looked for evidence in a shared text, "Why do we need

SMART Goals?" (DuFour et al., 2013). Once the importance of goals was established, the teams began the work of developing capacity by learning which critical issues, identified the week prior, would have the greatest impact on student outcomes. Utilizing the SMART Goal Discussion form, the advanced teams received a quick overview but then progressed independently, as a collaborative team, through the phases. These phases included having teams look at the critical issues they marked as a "2" or "4" (just beginning or over halfway) on the Critical Issues for Team Consideration worksheet. The collaborative team was then prompted to discuss the question "Which of these will have the biggest impact on student learning?" The administrative facilitator then asked the teams "Are there any items that you have marked as a "2" or "4" that has to be done before you can complete the task(s) you listed in #1?" These items became a priority. Equally important though was identifying which critical issues could "wait". The teams came to a consensus and placed these items on their agendas for future consideration. The purpose of this elimination step was to essentially "give permission" for teams to not try and do it all but rather be focused on what would have the biggest impact. Once this activity was complete, teacher teams were asked "What topics have you identified as "first steps?"

Once this baseline information was identified, the teacher teams discussed "What strategies could we use to achieve our goal?", "What will be the indicators of attainment? (i.e. how will we know we've achieved our goal)", "What resources are needed to achieve this strategy?", "What roadblocks do you anticipate?", and finally, "How might you pre-plan to navigate those?" This discussion protocol was intentionally designed to walk teachers through each phase of the SMART goal process without being

overwhelming. In the end, each team ended with a SMART goal to guide their work for the 2021–2022 school year even though the steps they took to get to that point looked a little different based on their self-identified groups.

The importance of differentiation in professional development became very apparent early on. For example, teams who self-identified as proficient required more scaffolding. As instruction began, there were a lot of questions that indicated the facilitator would need to scaffold the instruction so as not to overwhelm the teacher teams with too much information all at once. Adjusting while instructing, the facilitator ensured the teams received instruction on each step, worked on each step, and then came back together for further instruction on the next step. Proficient and advanced teams then translated the green items to "first steps" on their discussion document and the yellow items to the "goal" section of their document. The beginning teams completed this same process utilizing a high-quality pre-written SMART goal as an example to first see what the essential components of this process should look like and then used pre-developed sentence frames to develop their own.

Additionally, the discussion questions were utilized differently by each team. The advanced group utilized these completed answers as an opportunity to look for missing pieces or provide new information to the team whose goal they were analyzing. The proficient group utilized these as team discussion. The beginning group completed these as a whole group using the example SMART goal.

Once this process was completed, both the advanced and proficient groups transferred their information to a common SMART goal template that was placed in their team's shared drives. The beginning group then met a second time the following week,

working through the process, in the same manner the proficient groups had the week prior. Ultimately, all staff completed high-quality SMART goals that included all essential components and were tied to critical issues identified by each team.

### *Adjusting Instruction*

Phase 1 identified that a team with MAP Growth scores above the national norm adjusted instruction along multiple dimensions, including utilizing data to create homogeneous groups, developing instruction specific to student needs, and utilizing data to adjust student instruction and student placement. However, through the abductive coding process, collaborative culture was identified as an overarching component of this team's above-average student growth. According to Zuckerman et al. (2018), in order for a teacher's classroom behaviors to actually change, a process must be in place for collective goal setting, instructional feedback, collective and guided learning in professional learning communities, and trusting relationships. Drawing from Phase 1 findings, I referenced the team's continual ability to adjust instruction up and down based on the skills a student was ready to develop. This teacher team would plan their instruction for benchmark students together and then provide each other with ideas on how to scale the skills up or down for each day's lesson and, therefore, knew what was happening in all three classrooms each day. This observed process reiterates the importance of the findings of Zuckerman et al. (2018) by outlining these four components that were observed and resulted in improved student outcomes.

Recognizing that adjusting instruction is a skill that needs to be practiced and collaborated upon, the final session, was developed specifically for this purpose. In this session, the teacher teams reviewed their current MAP Growth data with me, identifying

reports that could be utilized to guide the grouping of students and those that call attention to specific skills that a student is ready to develop. Then, team meeting notes were reviewed as one way for teams to be intentional about adjusting their instruction.

Teachers in this school have been utilizing a team meeting template for weekly team meetings for the past four years. Recently, I noticed a tendency to "skip" the tougher parts of these conversations, such as those dealing with adjusting instruction. Therefore, teachers were provided an opportunity to reflect on this process by examining meeting notes that had provided evidence of these tougher discussions and discussing as a team "what do we notice about these notes that is similar to ours?" and "what do we notice about these notes that is different from ours?" Finally, teacher teams also answered, "what practices do we observe here that could positively impact our conversation as a team?" The notes these teams collected remained with each team so they could reference them during subsequent meetings.

Once the teams had identified the importance of having discussions about adjusting their instruction, they identified an objective for the next week's lesson of their choice in either reading or math and worked through the process of creating tiered objectives and activities. Beginning with the end in mind, the teachers first identified what they wanted benchmark students to know and be able to do by the end of the lesson. Then, utilizing Bloom's Taxonomy (1956), they adjusted the objective at least one level up or down, focusing specifically on the verbs utilized in the objective. Those verbs were then used to establish the performance-based assessment that would indicate to the teacher the student's level of mastery.

This activity ensured that the teachers had the opportunity to practice with their own set of information, something they'd be using the very next week, and provided scaffolded instructions on how to accomplish this goal. Additionally, teachers were provided with examples of when and how to have this discussion the next week during their team meeting time with the hopes that, by assisting teachers in developing this skill, this could become part of their weekly routines. Feedback was provided each week through comments left in team meeting notes by the building principals and myself.

### *Summary of Phase 2: Professional Development*

All staff at Gothenburg Public Schools participated in three one-hour professional development sessions that focused on three powerful practices identified in Phase 1 through teacher interviews, classroom observations, and lesson plan review. These powerful practices included developing a collaborative culture, focusing on specific goals, and learning how to adjust instruction. Professional development focused on building both will and capacity among teacher teams in each of these areas. Even though all teachers participated in the professional development learning opportunities, no teacher-specific data was collected from teachers during Phase 2 though these experiences impacted the data collected in Phase 3 where teachers' perceptions of the impact professional development opportunities had on their collaborative team was investigated.

## Phase 3: Unmasking Perceptions of Professional Development

Following the three professional development sessions, the third phase of this study focused on understanding the teachers' perceptions of the impact that professional development opportunities had on their collaborative team. Because MAP Growth data

later revealed the quantitative impact of this research project on student growth, the third phase of this study focused more specifically on whether or not teachers recognized a difference in how they conducted business as a collaborative team as opposed to what they had experienced in years prior. Aligned with the Plan-Do-Study-Act cycle, this phase focused on the "study" portion of this approach. Bryk et al. (2017) emphasized the importance of taking time to study the effects of the action plan by taking time to analyze the data, observe what happened and compare it with what I thought would happen, as well as glean insights for the next phase. Following the professional development phase, Phase 3 of this study was constructed to fulfil this demand to analyze the impact of the professional development opportunities afforded to teachers in order to inform professional development approaches for the district moving forward.

***Team Selection***

I asked all grade-level teams to participate in Phase 3 interviews. These individuals were recruited in person so I could fully explain the purpose of this phase of research, as well as answer any questions the team had. All teachers, including those who had already participated in the study, were eligible to participate, including non-tenured staff. Initially, non-tenured teachers were not considered eligible for this study since I directly evaluate them, which could have contaminated the data because they might have had a reason to not be truthful. In Phase 1, I was worried they might want to say the right thing as opposed to sharing what they actually do. However, Phase 3 data collection was geared more toward exploring the impact of the professional development experiences, as well as exploring collaborative conduct between team members, which meant it was essential that the whole team be present for the interview.

### *Data Collection*

During the third phase of the project, the teachers were asked to complete a single 30-45 min team interview with me. The interviews took place in one of the team member's classrooms during school hours to ensure that the participants felt at ease. I used a semi-structured interview protocol to allow for comparability between teams while still allowing the flexibility to ask additional questions if information regarding the research phenomena was not initially revealed (Merriam & Tisdell, 2016).

Questions for team members focused on identifying the participants' meaning of flexible grouping and how the teams utilize flexible grouping strategies within their classroom. Furthermore, I sought to understand each team's philosophy of collaboration by asking questions such as "What does each team member bring to the table?", "How do you utilize team planning time?", "What are each team member's roles from week to week?" and their perception of what flexible grouping is by asking them to describe the process. These questions provided an opportunity to examine how these experiences impacted the participants' interaction with flexible grouping strategies, as well as explore the collaborative culture of each team. Furthermore, these questions provided additional insight into how teachers meet the diverse academic needs of students.

Of primary importance in this interview was identifying any impact the Fall professional development had on their use of flexible grouping strategies in the classroom, as well as how their collaborative culture had been impacted.

### *Data Analysis*

Much thought was put into the design of this phase so as not to overwhelm the project with data and render it useless. A single design decision that was made was to

interview the teams as a whole, rather than individually. I hoped to be able to observe collaborative culture, or lack thereof, while the teams were together. Data were collected through the interview protocol and analyzed by uploading to the NVivo database, transcribing and de-identifying the information, and then coding the data using the inductive and abductive codes from Phase 1. While Phase 1 focused primarily on identifying how teachers use flexible grouping to address the diverse academic needs of students and how teachers work together as a grade-level team to meet student needs, Phase 3 shifted that focus to a reflection of the professional development that was implemented as a result of Phase 1. Because of this, additional questions were utilized to help answer the Phase 3 research question, "What are teachers' perceptions of the impact professional development opportunities had on their collaborative team?"

Drawing upon the organization of the professional development, three initial codes were utilized: 1) collaborative culture, 2) focusing on specific goals, and 3) adjusting instruction. Each transcript was read through and then coded using these codes that reflected the design of this study and would aid in answering the Phase 3 research question. Once all the transcripts were coded, I looked for the identification of themes and patterns in the data.

The collaborative culture code identified strengths and weaknesses in each team's ability to establish a non-judgmental culture of growth where it is okay to make mistakes as long as they are working on their specified goals and learning from their actions (Fullan & Quinn, 2016). I was specifically looking for evidence that there is an ongoing process through which educators collaborate in recurring cycles of collective inquiry and action research to achieve better results for the students they serve (DuFour et al., 2013).

As part of the collaborative culture coding process, I made field notes regarding the team's interactions with one another. Was a single person doing all the talking? Does everyone get credit for their contributions to the team? Does everyone contribute to the team? Do certain team members get interrupted when they're talking? These reflective questions were important for me to reflect on, specifically to call attention to the less-tangible factor of collaborative culture.

For focusing on specific goals, questions included references to the team's SMART goal, as well as how that goal guided their work for the year (Zuckerman et al., 2018). The questions encouraged teacher teams to identify what their goal was and articulate specific steps they had taken to achieve that goal. I looked for evidence that their goal had served as a guide for their weekly professional learning community (PLC) and whether or not the SMART goal was collectively recognized by the teacher team member as their focus.

Finally, the adjusting instruction code sought to reflect back upon the inductive codes outlined in Phase 1. Within this single code, I looked for evidence that teachers were utilizing data to create homogeneous groups, developing instruction specific to student needs, and were utilizing data to adjust student instruction and student placement. This code provided feedback on whether changes were made to how a team was using data to adjust their instruction based on the professional development that was provided in the Fall.

Throughout the coding process, notes were kept to record themes as they emerged. These themes and coding analytics are presented in Chapters 4 and 5. Additionally, because I transitioned into a new role, the timeline of this research was such

that also allowed MAP Growth data to be reviewed in addition to the quantitative data regarding this research project. Lastly, the third phase of this study was solely focused on whether or not teachers recognized a difference in how they conducted business as a collaborative team as opposed to what they had experienced years prior.

**Phase 4: Acting on What I Learned**

The fourth and final phase of this research project is the "Act" phase. This phase is expanded upon in Chapter 5 titled "Discussion" and Chapter 6 titled "Conclusions". In this section, I detail what to do next based on what I learned through the research and coding process (Bryk et al., 2017). It was anticipated that additional support would be required for each team in order to see the kind of MAP Growth rates the Phase 1 team consistently demonstrates, simply due to the time constraints of this research project. As I coded the final interviews of this study, I specifically sought to determine what adjustments must be made to increase the impact of our positive case on the school as a whole (Bryk et al., 2017).

The leadership structure of Gothenburg Public Schools is such that no decision such as this can be made solely by one person. Instead, this information was taken to our leadership team, a team of administrators and counselors who work together to determine the best way to move forward on any topic of importance. Many changes took place during the year of writing up this final report that most certainly impacted this process, including the appointment of a new Superintendent, a new Director of Teaching and Learning, and a new Elementary Principal. Education is a forever-evolving field; thus, conducting a long-term case study that is also valid and reliable becomes very difficult.

My original goal was to determine how teachers use flexible grouping to address the diverse academic needs of students and how they work together as a grade-level team to meet student needs. I was fortunate in that I was able to accomplish this goal in the first phase of this study and uncovered a striking reality in the process—collaborative cultures appeared to be the most impactful component, more so than any tangible instructional strategy. With the encouragement of Dr. Sarah Zuckerman, I expanded Phases 2 and 3, which shifted my research focus. I wanted to capitalize on the professional development needed to develop these collaborative cultures, as well as reflect on whether those changes made a difference.

Admittedly, an analysis of just the team interviews would not have accomplished this. Therefore, I compared my findings of the coding process with an increase or decrease in MAP Growth scores from Spring 2019, 2021, and 2022 to observe if there was a correlation between team interactions and student growth compared to the norm. These comparisons are reflected upon in greater detail in the following chapters.

### *Community Context*

The context of this project is essential to understand, as it likely impacted the findings. The school district evaluated in this study is not what would be considered typical. This particular elementary school is a rural school in central Nebraska located in a community of 3,500 people. This is the only public elementary school in town and is attached to the Junior/Senior High building. The district serves 820 students, and 440 of those are elementary students. In 2019, the elementary school was categorized as “Great” by the Nebraska Department of Education's AQuESTT school rating system.

The student population is significantly less racially diverse than the surrounding communities. Over 97% of students identify as Caucasian compared to 59% in Dawson County. The median household income is $81,029 compared to $62,540 in Dawson County. The percentage of people living in poverty is 7.8% compared to 13% in Dawson County. About 37% of students qualify for free and reduced lunch. The graduation rate is 95%, well above the state average of 88%. The district reported zero students pursuing English as a Learned-Language (ELL) in 2019. Proficiency rates on the 2019 state assessments for Grades 3–8 were at the state average in English language arts (52%), slightly below in math (51% vs. 52%), and significantly higher in science (73% vs. 66%).

**Table 2**

*Student Demographics 2018–2019, Gothenburg Public Schools*

| | Gothenburg | Dawson | Nebraska |
|---|---|---|---|
| Population | 3,500 | 23,804 | 1,904,760 |
| Total enrollment | 820 | 5,172 | 325,984 |
| Students identifying as Caucasian | 97% | 59% | 66% |
| Median household income | $81,029 | $62,540 | $75,123 |
| Living in poverty | 7.8% | 13% | 11.6% |
| Free and reduced lunch | 37% | 57% | 45% |
| English as a learned-language (ELL) | 0% | 35% | 7% |
| English language arts proficiency rate | 52% | | 52% |
| Math proficiency rate | 51% | | 52% |
| Science proficiency rate | 73% | | 66% |

*Source: Gothenburg-Dawson County Community Well-Being Needs Assessment – January 2020*

Clearly, these statistics outline a different context than what would be considered typical in a public school setting. With this context in mind, the validity and reliability of

this study's results can be appropriately considered as one reads the final chapters of this study.

# CHAPTER 4

# FINDINGS

This case study was designed to understand, during Phase 1, how a high-performing teacher team utilized flexible grouping to address the diverse academic needs of students and how those teachers work together as a grade-level team to impact student growth in order to spread those practices to other grade-level teams. As an extension of what was learned in Phase 1, Phase 2 focused solely on developing quality professional development experiences that extended what was learned in Phase 1. Lastly, the third and final phase of this research project observed how targeted professional learning impacted teachers' perceptions and usage of flexible grouping as a strategy to differentiate instruction.

Relying on traditional instructional practices risks increasing the gap between low- and high-performing students during a period when accelerating recovery from disrupted learning should be of primary concern for every educator.

While there were no research questions in Phase 2 due to this being an "action" phase focused on extending what was learned in Phase 1, the research questions that served as the primary guide for this case study were as follows:

<u>Phase 1:</u>

1) How do teachers use flexible grouping to address the diverse academic needs of students?
2) How do teachers work together as a grade-level team to meet student needs?

<u>Phase 3:</u>

1) What are teachers' perceptions of the impact that professional development opportunities had on their collaborative team?

**Table 3**

*Data Collection and Analysis Strategies*

| Research Phase | Research Question | Data Sources | Findings |
|---|---|---|---|
| Phase 1<br>Understanding flexible grouping in use | How did teachers use flexible grouping to address the diverse academic needs of students?<br><br>How did teachers work together as a grade-level team to meet student needs? | • Individual interviews with the positive outlier<br>• Classroom observations<br>• Review of lesson plans | • Using data to create and adjust homogeneous groups<br>• Developing instruction specific to student needs<br>• Ensuring flexibility in student grouping<br>• Collaborative culture |
| Phase 2<br>Targeted professional development in response to Phase 1 findings | | Improvement science | Collaborative culture<br>SMART goals<br>Adjusting instruction |
| Phase 3<br>Unmasking perceptions of professional development | What were teachers' perceptions of the professional development opportunities and their subsequent efforts as part of a collaborative team? | Team interviews<br>Field notes | • Performance-based assessments<br>• Data-focused conversations |

Chapter 4 outlines the findings of Phase 1, followed by a discussion of how those findings influenced professional development during Phase 2. Finally, the findings of Phase 3 are shared. Chapter 5 then outlines a discussion related to the literature, as well as unexpected findings, followed by conclusions in Chapter 6, along with implications of this research and recommendations for further research.

**Phase 1: Learning from a Positive Outlier**

The purpose of identifying a positive outlier case for this phase of the study was to maximize what could be learned from a high-performing teacher team. To identify a positive outlier case, I looked for grade-level teams with growth norms that were greater than those posted by NWEA MAP Growth as the "norm". Because increased student growth is the desirable outcome of this project, better understanding this positive case where real-world contextual conditions were yielding positive results was a critical step in this process (Yin, 2014). The MAP Growth data for Winter 2018, 2019, and 2020 showed that the team selected had consistently demonstrated student growth rates that were, on average, 40% higher than the national norm and double that of the next closest average within the same school. Typically, growth data recorded during spring are reviewed for the most consistent picture of growth, but due to the COVID-19 pandemic, data for Spring 2020 was not available; thus, data for the Winter semester was collected to ensure the most consistent data from each grade level.

The initial interview was designed to help me further understand how this positive outlier defined flexible grouping, as well as what strategies were utilized within the classroom to meet the diverse needs of students. To confirm what was shared in the interview, I also observed each classroom and reviewed lesson plans to triangulate the strategies and ideas expressed throughout the first phase: 1) utilizing data to create and adjust homogeneous groups, 2) developing instruction specific to student needs a coordinated effort, and 3) collaborative culture.

### *Putting the "Flex" in Flexible Grouping*

Several themes emerged during the interviews that are important to expand upon here. The teacher team that participated in Phase 1 outlined several instances where data was used not only to create homogeneous groups but also to adjust those groups frequently. Additionally, the teacher team outlined various ways in which they customized instruction specific to the individual needs of students while maintaining consistency among each of the three classrooms. Below are detailed findings among each of these themes.

**Theme #1: Using Data to Create and Adjust Homogeneous Groups**. All three teachers indicated that data was the basis of all grouping. After each benchmarking period, the members of the teacher team sit down to analyze the data. They reported that they sort the students first by their RIT score, the score provided by MAP Growth, which is an estimation of the student's instructional level, and then determine if additional shifts are needed; this formal process happens three times per year. Once students are placed among the beyond-, benchmark, and approaching- groups, students receive instruction geared specifically to the skills they are ready to develop.

The teacher team reported that students moved between groups much more frequently than in the three benchmarking periods noted here. For example, when asked how often they adjust their groups, all three teachers responded with a variation of "when the student has mastered the skill we're working on." Members of this teacher team noted that they talked about formative assessment student data every day. For example, one teacher noted that she would often go out to recess to discuss exit tickets with her colleagues even when it was not her duty day. Another teacher had a clipboard she

utilized as a “skills checklist” so she could quickly share with the student’s homeroom teacher what skills they had recently mastered.

One surprising revelation of this group was that they all indicated, unprompted, that they would “be lost” without their common planning period and professional learning community team time. One teacher captured the flexibility of flexible grouping in this grade level as follows:

> We are talking about students every day in our world. We talk about our kids, how one is doing well or not, and we get the other teachers input because some part of the day they see them. We know a lot of times a kid may be there [in a certain class] the whole year, but they’ve always been discussed. It’s never just a ‘we don’t discuss them anymore’, but that the door is always swinging wide open to other kids coming in and out depending on their needs. (ER, Interview 1)

This description outlines how this teacher team feels a sense of interdependency as a team and that they function so closely as an integrated unit that they cannot imagine functioning in a silo.

Another team member agreed, noting that “we have to make sure that we’re not just sticking kids where their data says they need to be and leaving them there even if they need to be moved.” Further, the third team member agreed that “we just start piecing it all together like a puzzle, like who's going to be best at a certain spot who may not work there. And if there's like a red flag, like, oh, my goodness, this kid really needs some extra support then we make sure they get it.” All three quotes show that this teacher team understands flexible grouping as an ongoing process of reviewing student progress, rather than one that relies only on benchmark assessments. Most importantly, these three

teachers indicated a frequent review of all data available to continually review and adjust student placement.

This particular teacher team outlined a process of sorting students into three groups that they reference as approaching-benchmark, on-level, and beyond. These words were used consistently by the teacher team, and not once were the students or classrooms referenced as low or high, although middle was synonymous with on-level among all three teachers. This verbiage is reflective of a teacher team that is dedicated to utilizing data to create and adjust instructional groups but also an underscored commitment to ensuring instruction aligns with each group.

**Theme #2: Developing Instruction Specific to Student Needs**. In addition to using their planning time to group and regroup students, the positive outlier team members reported many strategies that they feel have been effective in expanding student growth. All three teachers indicated that, through flexible grouping, they can be specific about the skills that students are ready to develop. One teacher indicated that "our students would not have the growth that they do if we were not flexibly grouping them." When asked why, the teacher shared that she would not be able to make sure her instruction "met them right where they are at." Meaning, the teacher can develop instruction specific to each student's needs because the teacher team has already looked at data and determined what each student is ready to learn. Then, because the groups are already homogeneous, the teachers can instruct on skills specific to that group of students as opposed to a more broad heterogeneous classroom that would include students that were beyond-, -at, and approaching-benchmark.

Diving deeper into this idea that the teachers were using specific strategies to ensure high student growth, the teacher team reported and was observed using several strategies to differentiate instruction based on the individual needs of students.

To develop instructional content, this teacher team reported being intentional about making sure the instruction was aligned with the skills their students were ready to develop. They reported that they did this by utilizing the learning continuum within MAP Growth to dive down into specific skills that the students were ready to develop, as well as by utilizing assessments they had created as a team to reflect the skills they were working to develop. The teacher team described the ReadyGEN materials the school had adopted as the starting point for what students needed to know. However, they further articulated the importance of truly knowing the students, which they described as something they could accomplish through observation of individuals during small group or one-on-one time with students.

The teacher team articulated the importance of each team member teaching the same standard and story during whole group instruction as a means of ensuring groups could continue to be flexible. This process was observed in the classroom and affirmed within the lesson plans. Even though I observed each individual classroom on different days, field notes show that each classroom began with direct instruction, was followed by student practice opportunities, and then ended with small groups. Teachers indicated that, by ensuring each teacher was on the same standard and story, they could talk about reading across other classes such as math, writing, or homeroom time because they all had the same frame of reference. For example, while they read their specific trade book for that unit, all of the students could reference the same setting, characters, plot, conflict,

resolution, and story elements. Teachers could then say things such as “remember how we were comparing and contrasting the characters in our story? Today in science, we will be comparing and contrasting a solid and a liquid. Who can tell me what keywords we might be looking for?” Then, because students were in homerooms instead of their flex groups for science, the teachers could just reinforce the idea that “they are all learning the same things” even though how those students are supported looks completely different in each classroom.

Each teacher was observed following direct instruction with some kind of student activity where students worked on practicing the skill of the lesson. However, each teacher described differentiating this activity. For example, the teacher who had the beyond-benchmark students shared how she would have students take the story and independently look for evidence in their text because when practicing new skills, they struggled to work collaboratively. In contrast, the on-level teacher strategically partnered students up with someone of similar skill level within that classroom because she has observed this keeps her students more engaged. In comparison, the teacher with the students who were approaching-benchmark did this in a whole group setting, with students seated on the floor up at the front of the room with the book displayed on the smartboard, and they could annotate the text where everyone could see it because she felt her students who were well-below-benchmark needed the opportunity to develop their skills in this setting before moving on to independent practice.

I observed and verified within the lesson plans that all three classrooms were working on finding evidence in the text and were asking the same questions, but how they arrived at the answers differed because these teachers had developed instructional

strategies that were specific to the needs of the students within their flexibly grouped classroom. The approach that was utilized may not have provided equal opportunities practice for each student, but it was critical in allowing these teachers to accomplish the same task in a similar amount of time, which allowed for all three classrooms to then have enough time for the next part of their reading block, which was small groups.

The teachers described small group time as the period during which they drill down to learn about what each student knew and could do. The teachers described a process where they worked on specific skills and then utilized informal assessments to give them feedback on how to adjust their instruction later that day during homeroom time and the next day during reading. I observed that each teacher had three groups within their flexibly grouped classroom that were also leveled based on skill. Each group would rotate after approximately 15 minutes, completing three activities over the course of the small group time.

Consistently, I observed that each room had a small group that was working directly with a teacher and a group that worked independently. Each of these groups was planned for within the lesson planning documents. The on-level and approaching-benchmark rooms had a group that worked with a paraprofessional, whereas the beyond-benchmark classroom did not have a paraprofessional but instead had a group project for the third rotation. The teacher working with the beyond-benchmark students noted that working together in a group was a challenge for students who “were used to knowing more than others”; so this was an intentional strategy to help support them in developing this skill.

During classroom observations, the teacher-led rotation included reading out loud so the teacher could hear fluency and problem-solve specific sound patterns with individual students while utilizing leveled texts. In the beyond-benchmark classroom, students used leveled texts that had a reading level of the next grade level, whereas the on-level classroom used a leveled text that was right at that grade level. The approaching-benchmark classroom did not use a leveled text but instead participated in an activity where students would roll a dice with the sound patterns of the week and they would have to read the sound pattern out loud, find it on their bingo chart, spell it out loud, reread it, and then use a highlighter to color the square. Each time a student went, they all found the sound pattern on their own page, which ensured that no one got bored waiting for their turn. Then, once this small group rotated to work with a para, they coded and read a leveled text as a means of extending their learning with another adult. All three classrooms were observed to be working on the same sound patterns, but the beyond-benchmark classroom paired that with higher-level text, the on-level classroom affirmed these skills with on-level text, and the approaching-benchmark level classroom worked to reinforce the skill.

While the teacher was with the small group who had direct teacher interaction, I observed each teacher intentionally recording student information. On their clipboards, they had each student's name down the left side and the spelling patterns across the top. All three teachers were reflecting on each student's ability to fluently read the sound patterns even though the context of those sound patterns was different. These assessments were what the teachers described as making up the content of many of their conversations as a team. During observations, it appeared that classroom assessments were given so

informally, utilizing a clipboard and quick notes, that students often did not even realize they were being assessed. Instead, students thought they were simply reading or playing a game of bingo where they found specific phonemic patterns. It is important to note here that this is a single observation as a snapshot and certainly not intended to capture the full extent of what an elementary reading program looks like, as this observation reflects a single lesson out of one unit and is intended to capture the specificity these teachers focus on to develop important skills within their classrooms.

When I observed the approaching-benchmark classroom, she mentioned a paraprofessional working independently with students one-on-one throughout the whole group part of the lesson. In the follow-up interview with this teacher, I asked for more specifics as to what the paraprofessional was working on with the students. This teacher shared that rather than having a para sit in the back of the room or assist one or two students within the classroom with staying on task, she chooses to have them sit one-on-one with the students out in the hallway and work on specific skills they need to master, as she feels this is a more effective and efficient use of additional adult help within her classroom.

The teacher creates a folder every day that has a specific activity in it. On the front of the folder are the names of students that have not mastered the skill yet, and on the inside of the folder are instructions for the para. The para then works with the student and makes notes about the child's progress, and then the teacher uses that information to adjust the folder and the skill for the next day. I asked in the follow-up interview offhandedly where this teacher found the time to prepare for the para, to which the teacher replied that "when it is important, you make time."

All teachers indicated specific strategies that are not new to educators as being effective tools. In fact, the literature review has outlined these in detail. Strategies such as tiered lessons, interest centers, and diversified instructional techniques were not in short supply, as seen by the examples provided here.

**Theme #3: Ensuring Flexibility in Student Grouping.** The lynchpin of the entire process, the teacher team from Phase 1, provided some excellent examples of how they develop instruction specific to student needs while utilizing data to not only create but adjust homogeneous groups. Ensuring they all teach the same skill, utilize the same trade book, and develop small group activities specific to the skills their students need are all highly regarded strategies by this team. Having an environment where students can fluidly move from one classroom to another does not happen by accident. If these teachers were differentiating by the amount of time they provide students to do the work, the beyond-benchmark classroom would be further along in the materials than the on-level or approaching-benchmark classrooms. Therefore, when a student would be ready to move, it would be quite difficult, as they would miss whatever instruction had happened when they were not in that room. Instead, by aligning daily instruction and either diving deeper with skill by going to the next grade-level skill or supporting with scaffolded strategies such as whole group annotation of a text, each classroom remained on the same lesson of the same story on the same day, which allowed for fluid movement between classrooms when students were ready. This consistency is the lynchpin of the entire flexible grouping process. Without it, classrooms begin to reflect more “tracking” characteristics where students are “stuck”, rather than “flexible” characteristics where

students can move freely from one classroom to the next without loss of instruction or content.

### *A "Surprise" Finding: Collaborative Culture*

I had a preconceived notion that this positive outlier would reveal specific strategies that were critical to the process of flexible grouping. Specifically, I thought that this team would share strategies that could be taught to other teachers to ensure they, too, experience the kind of success with student growth that this grade level was seeing. This team articulated their process well. However, while coding the interviews, observation notes, and lesson plans, I began to synthesize the information and realized that many times, there was something important that needed to be marked but neither had a deductive code from the initial literature review nor did it fit into the inductive codes developed from my data. Instead, the three inductive codes seemed to be interdependent, connected by a culture of working together to maximize student growth of every student in this grade level.

It was the comment initially noted in this section that I kept coming back to—"we would not survive without our common planning time and PLC time." Each teacher reflected on this to some effect. I wondered if every teacher team in this school would say the same thing. While reflecting on this, I had an epiphany of sorts. When Phase 1 took place, the District was in the second year of creating a culture of professional learning communities (PLC). While working with a group of new teachers, I shared with the cohort that PLC is defined as "a systematic process in which teachers work together interdependently in order to impact their classroom practice in ways that will lead to

better results for their students, for their team, and for their school" (All Things PLC, 2007).

It was in that moment that I began to wonder if collaboration was indeed the most important component of this team's success. Going back through the interview transcripts, I created a new, abductive code—collaborative culture. Coding anything that fell into the above definition, I uncovered a strong theme among all three teachers—they liked each other, trusted one another, and didn't think they could do their jobs without one another.

One teacher commented that it "isn't just one person moves the whole roost, we do a good job and try really hard to be accepting of new ideas and different things because when you work as a team, that's what it should be." The teachers shared how they always talk about how the students are doing. They talk about them while they are out at recess, making copies, and lining up for lunch and expressed that this was really important because they wanted to truly know their students. One of the teachers even commented that this constant communication is why their team is so successful, but she was quick to point out that it isn't a "huge professional sit down, but a 'hey I'm really worried about this kid. What are you seeing?'" conversation. All three teachers affirmed the idea that constant communication was a key component. However, I picked up on the second part of that quote—"what are you seeing?" This tidbit indicated the teacher was genuinely looking for feedback. All three teachers expressed that they all bounce ideas off each other about what is working and what they need to adjust, and expanded that they are always open to trying new things. One teacher captured this idea by saying "if something is working well for one of my teammates, I need to try that!"

Another teacher went on to talk about how they all "pull their own weight" and have strengths they bring to the table. Interestingly, all three team members were quick to identify the strengths of their teammates but struggled to identify their own strengths, indicating a true interdependency among teammates. These teammates stated that trust is the core component of their team and shared examples of how that trust was built. One teacher shared a story about how she had a lot to do one day, both personally and professionally. Her team helped her out with work on the professional front since they knew she had had a rough day. Her colleague recalled this same situation and indicated that when she had missed some work recently for personal reasons, her team had seamlessly fulfilled all of her planning responsibilities. This teacher indicated that they "don't expect anything in return. You care about your teammates and do what you need to do to help support them especially when they're hurting." While this trust has been built over time, the team was still able to affirm the foundation of their relationship—trust. One teammate summarized it succinctly by stating "I feel safe with my teammates. I don't feel like there is any judgment. This is a completely safe zone where it is okay to not know something."

When asked what motivates this team to spend so much time talking about kids, responses consistently revealed a theme of not wanting to let their team or their students down. One teacher shared that "You want to do better. You want to show up and be there for your team. You want to give everything…because you don't want to let them down."

An additional literature review on this topic of developing collaborative culture summarizes what I have just described here as collective teacher efficacy. Hattie (2009) defines this as a collective belief of teachers in their ability to positively affect students.

This teacher team is committed to making a difference in the lives of students regardless of time, vulnerability, or pride. With an effect size that is three times the hinge point where a strategy is recognized as effective, it is no wonder this team is experiencing shared success. However, the challenge was to design a process that would support other teams in developing the type of interdependency described by this positive outlier team.

Phase 1 provided some great insights, using which I was able to develop professional learning experiences for Phase 2. Teachers shared and were observed utilizing data to create and adjust homogeneous groups while also developing instruction specific to student needs. This positive outliner team was able to do this by employing a coordinated effort, where each classroom was teaching the same standard and skill, albeit at different depths of knowledge. Ultimately, it was the collaborative culture shared among this group that most impacted the development of Phase 2 learning opportunities. Each of the skills, strategies, and approaches outlined by this teacher team was done so in such a way that collaboration was not something they did but who they were, and it dwarfed anything else I observed or learned from this teacher team. It was with this knowledge that I established a plan to guide other teacher teams to develop collaborative culture, create common goals, and differentiate instruction.

**Phase 2: Professional Development**

In the second stage of the project, three one-hour professional learning opportunities were developed for the fall as teacher teams focused on their work for the year. While there are no findings from this particular phase, it is important to note that this phase, described in detail within the section titled “Methods”, was developed and informed by the findings from Phase 1. Fullan and Quinn (2016) defined collaborative

culture as a team where “leaders establish a non-judgmental culture of growth where it is okay to make mistakes as long as you are working on the goals and learning from your action.” With this in mind, establishing a common goal seemed to make the most sense as a “first step” which the teacher teams completed to zero in on their focus for the year.

All teacher teams in this district completed this self-assessment which helped them identify their team’s “next step”. Developing a SMART goal is the process used to ensure that each team is “working on a common goal” (Fullan & Quinn, 2016). Teaching teachers how to develop a goal that was specific, measurable, attainable, realistic, and timely was new for this district and, therefore, impacted the findings in Phase 3. The teacher teams in this study were able to determine what level of support they needed in developing their first SMART goal and attended sessions based on this insight. Finally, teacher teams were coached on how to adjust instruction based on student data by working through the instructional planning process with their own resources.

In this study, professional development focused on building both will and capacity among teacher teams in the areas of developing collaborative culture, focusing on specific goals, and adjusting instruction. Even though all teachers participated in the professional development learning opportunities, no teacher-specific data was collected during Phase 2 though these experiences impacted the data collected in Phase 3 during which teachers’ perceptions of the impact of professional development were investigated.

**Phase 3: Understanding the Impact**

The third phase of this study focused on understanding teachers’ perceptions of the impact professional development opportunities had on their collaborative team. Initially, it was thought that the same MAP Growth data that was used to identify a

positive outlier would not immediately be available for review. However, I ended up transitioning into the role of Superintendent and needed some additional time to compile the findings and recommendations. During this additional time, the data became available and, therefore, can be used in addition to the interviews to help scholars understand the impact of this research project.

For the third phase of this study, teams were interviewed as a whole so I could observe collaborative culture while the teams were together. Using the inductive codes from Phase 1, each transcript was coded for the identification of themes and patterns in the data.

All seven K-6 teams were interviewed for Phase 3. It is important to note that the interviews took place at the end of the school year as teams were wrapping up the year. The timing of these interviews could have potentially impacted the results of this study given the teachers could have felt rushed and tired at the end of the school year. However, I observed during interviews that teachers seemed at ease and were open to sharing their experiences through the interview process. Because I was transitioning into a new role within the district, I did not want to wait until the fall for fear that my transition into the new role could impact the findings, as teachers might possibly feel more obligated to share what they thought I wanted to hear as opposed to their true experiences. For the purpose of consistency, the themes are reflected here within the same categories used for professional development utilized in Phase 2: developing collaborative culture, developing SMART goals, and adjusting instruction. To provide additional details, performance-based assessments, data-focused conversations, truly

flexible groups, strengths-based teaching, and whether teams met their SMART goals are represented visually in the chart below.

**Table 4**

*Common Themes Among Teams Based on Professional Development Sessions*

| | Professional Development Sessions | | |
|---|---|---|---|
| **Growth** | **Collaborative Culture** (evidence reflective of Fullan & Quin (2016)) | **SMART Goal** (primary focus of team) | **Adjusting Instruction** (could articulate more than pacing OR decreasing DOK) |
| 1st | X | X | X |
| 2nd | | | |
| 3rd | X | X | |
| 4th | | X | X |
| 5th | X | X | X |
| 6th | | X | X |
| 7th | | X | |

***Collaborative Culture***

The teacher teams described varying degrees of what they interpret as the practice of collaboration, even though they had all experienced the same professional development outlining what this relationship should look like. One team indicated that they "are always working together and sharing ideas" and that "they could not survive without their common planning time." While others described the collaboration as "throwing out things they've made and others can use it if they want to." This contrast tells me that if the district wishes to develop a consistent operating definition of collaboration, additional learning needs to take place. Teacher teams that reported they were meeting consistently also indicated increased trust among team members and could articulate what other teachers were doing within their classroom. These teacher teams gave examples of what they had learned from their colleagues and how they had adjusted

their own instruction because of that. One teacher shared that she doesn't know everything, that there is no way for her to know everything, and that is why their team has to work together. This openness to admit "not knowing everything" indicates a level of trust she has in her colleagues to help her grow to meet the needs of her students.

Four teams indicated that the templates they used for team meetings had helped organize their thoughts and support additional discussion about students, particularly the portion that aided discussion on students needing support. Of the three teams whose students showed the most growth from fall to spring, two of those three teams shared their instructional planning workload. They reported that they met consistently, talked about student data, and recognized what each teacher brought to the planning process.

The themes presented in this section were not exactly what I had expected to find, though. The team that had the second highest growth rates in the school was also the team that indicated they "had not met as a team in two months." Although the fully interdependent team had observed more growth, this team, who hadn't met in two months, was second among all seven teams in terms of student growth.

This was also the team that jokingly shared that during that week's PLC time, when teachers were given plan time in their rooms instead of team time, one of their teammates had not read the email. When one particular teacher on their team heard that they could work in their room, their response to the team was "well, then why am I even in here?" and the teacher left. It was also this team that indicated it had "never occurred to them to share the responsibility of planning for a lesson." This finding is unexpected, but perhaps it tells us that relationships are necessary but not sufficient on their own.

Further analysis of the data revealed that this team, although they were not meeting frequently, had a strong commitment to collecting and analyzing student data. This team would take that information and adjust their lesson for the next day or use flex time to support the missing skill. Students were not flexing frequently, but they were getting instruction specific to their skill level consistently, which indicates that the team does not yet have the skills needed to talk about student performance within the context of shifting rooms but can differentiate their own instruction within their own classroom. This team was operating more in silos than a melting pot, but their system was working for them. So while collaborative culture emerged in Phase 1 as a surprise, it was not a consistent finding among the top three performers. However, this particular group of teachers, as shown here, was able to articulate the instructional strategies they used, as well as the data they collected regarding student skills, which indicates that teaching strategies are perhaps also necessary but not sufficient on their own.

Therefore, it is possible that what truly distinguished the highest-performing teams was both a collaborative culture and commitment to utilizing data to create instruction specific to student needs and not one or the other.

***Common Goals***

The second phase of professional development involved assisting teacher teams with developing common goals. Six of the seven teams identified that the SMART goal that was developed in Phase 2 was the basis for their professional learning community (PLC) time each week. These teams identified that this goal kept them grounded and focused more so than they had been previously, although only two teams actually met

their goal. Specifically, knowing what the essential standards are for each grade level and staying focused on student data were areas mentioned by these teams.

Each team was asked if they had met their goal, and if not, what roadblocks they had experienced. The two teacher teams who had SMART goals that focused on teacher tasks such as identifying essential learnings and developing assessments had met their goals for the year. Five teacher teams identified goals based on student growth or achievement. None of these teams met their goals and indicated that they had set their goals too high. Interestingly, two of these teams blamed the students for not meeting the goal, and one indicated they had set the goal too high. Three of these teams indicated that the "core was too difficult" for their students and that they had modified the materials to make them easier for students to be successful; so while students had As in the classrooms, they were not meeting growth or achievement norms on MAP Growth. Considering that a student is unlikely to meet the benchmark if the instruction they are receiving is not at the benchmark, this is a significant concern. As this district has transitioned to a new Director of Teaching and Learning, this finding is certainly one that must be shared. Specifically, it is important to understand what support can be provided to these teachers to help them understand that all students need access to grade-level content. More specifically, it is important to devise strategies to support teachers in making grade-level content more accessible for students.

### *Adjusting Instruction*

The crux of the process, adjusting instruction to meet the diverse academic needs of students, proved to be more challenging than anticipated. Even after providing professional development on how to adjust instruction, there were three teams, in

addition to the positive outlier team in Phase 1, that was able to articulate what this process looked like. They could articulate that they were all working on the same skill, at different levels, so students could move fluidly as needed. Of these three teams, only one of them was in the top three teams in terms of MAP Growth.

Four of seven teams identified pacing (going slower for the 'low' group and going faster for the 'high' group) as their primary means of adjusting their instruction. While some could identify other strategies such as scaffolding and partner work as important strategies, all four of these teams recognized they were shifting students only after benchmark data had been collected. So while instruction was being adjusted within these classrooms, students were not freely moving among classrooms as had been indicated by the positive outlier in Phase 1.

**Summary**

With so much variation among these teams with regard to a collaborative culture, SMART goals, and adjusting instruction, it was important to find the things the strongest performers had in common. This proved to be more difficult than expected because the only two things the top three performing teams had in common were their utilization of performance-based assessments and data-focused conversations, neither of which were the focus of the three professional development sessions provided in Phase 2. This evidence indicates that neither collaborative culture nor a commitment to utilizing data is sufficient on its own to impact student growth. Further, it is possible that what truly distinguished the highest-performing teams was both a collaborative culture and commitment to utilizing data to create instruction specific to student needs and not one or the other.

The following chapter summarizes this research project and outlines important conclusions drawn from the data presented in this chapter.

# CHAPTER 5

# DISCUSSION

This case study was designed to learn from a positive outlier—a teacher team that consistently produced growth rates far beyond that of their peers. By creating a professional development experience based on what I learned in Phase 1, the intent was to extend the practices being displayed by the positive outlier teacher team to other grade levels within the school. This professional development focused on developing collaborative culture, the use of a SMART goal, and adjusting instruction based on student mastery. Following the professional development, all teacher teams were interviewed, of which only two teams articulated all three themes that were focused on in the professional development. Of note, these two teams ranked first and fifth (out of seven) teacher teams with regard to student growth. This suggests that there are other factors that impact student growth.

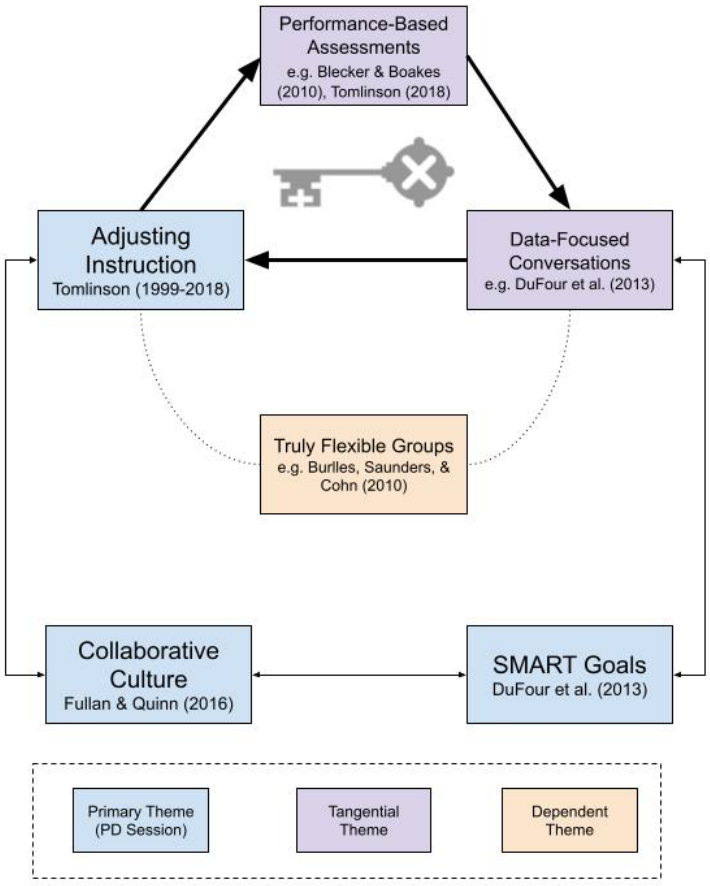
In addition to the three professional development themes, four extended themes were identified in the coding process: 1) usage of performance-based assessments, 2) having data-focused conversations during team meetings, 3) utilizing truly flexible groups that were adjusted based on skill and not a timeline, and 4) assigning teachers based on their specific strengths. Additionally, I made note of those who did and did not meet their SMART goal for the year.

This section is organized around key themes identified in Phase 1 and Phase 3. The themes are related to previous literature by implications of this research and recommendations for further research within the context of the research questions that guided this study. This discussion is followed by the conclusion and implications for

practice and future inquiry. Figure 8 helps to understand how each of these themes is interconnected in the findings and discussion. The figure depicts the preliminary themes derived from what I learned in Phase 1 (Adjusting Instruction, SMART Goals, and Collaborative Culture) and how these intertwine with tangential themes that were revealed in the interview data during Phase 3. Shown in the middle, as a completely dependent factor, truly flexible grouping can be achieved but only when all the other factors are in place as well.

**Figure 8**

*Interconnected Findings*



**Collaborative Culture Increases Collective Efficacy**

A collaborative culture was identified in the Phase 1 interview with the highest-performing teacher team. Teachers on that team mentioned being non-judgemental when it came to trying new things and sharing ideas. Those teachers talked about how they could not live without their common planning time and were "constantly" communicating with one another about strategies that worked well for specific students. This study sought to replicate this constant communication around student learning by facilitating a process within the first session of professional development designed to ensure teacher teams began with a common goal. Relying on the improvement science framework described by Bryk et al. (2017) where teacher teams embrace a common problem and inquire about solutions while embracing the wisdom of others. Teacher teams engaged in team reflection around PLC critical issues defined by DuFour et al. (2016) by force ranking themselves as a 1, 2, 4, or 5, with 1 being "not yet started" and 5 signifying "complete". Three was intentionally left out to ensure they were committing to the force ranking process. After completing the 10-question survey that covered statements such as "We have aligned the essential learnings with state and district standards" and "we have agreed on how to best sequence the content of the course and have established pacing guides to help students achieve the essential learnings", the teacher teams talked through items that should be considered as "first steps" and evaluated what critical issues they felt would have the biggest impact on student learning. A facilitated discussion on the strategies that could be used to achieve this goal, how they would evaluate progress, and how they would navigate potential roadblocks was also conducted as a way to build a collaborative culture.

Defined as “an ongoing process in which educators work collaboratively in recurring cycles of collective inquiry and action research to achieve better results for students,” DuFour et al. (2013) set a high standard for what collaborative culture should look like. In addition to DuFour et al.’s (2013) definition, the interview data were reviewed in accordance with Fullan and Quinn’s (2016) definition that collaborative culture is where a non-judgmental culture of growth exists and where it is okay to make mistakes in mind as well.

After the completion of Phase 2 professional development, Phase 3 interview data reflected that three of seven teacher teams met this high standard. The teams that articulated this theme highlighted a strong and transparent relationship among all team members. There was no fear of failure among team members and a complete openness to discuss ideas and ask for help when needed. These three teams talked about how they could not survive without their common planning time—a time that was not a required team meeting time—but was utilized consistently as such by the three teams identified here. This researcher agrees with Fullan and Quin (2016) that collaborative culture is a critically important step in ensuring teacher teams recognize success.

The teacher team that had the second-highest growth rate did not exhibit collaborative culture. In fact, that teacher team had not met for weeks and described their collaborative culture as “meeting when they’re told to” and “throwing out things we’ve made and others can use if they want to.” However, this teacher team also identified that their team approach “just kept getting better”, referencing how the team members work together to meet the needs of students. The main reason this team was not marked as displaying a collaborative culture was their lack of interdependency. They were

accustomed to working independently on lesson plans—they even noted it never occurred to them to work together—and while they actively adjust their own instruction, they do not talk about it as a team. However, it should be noted that the individual capacity of this team was extensive, as evidenced by their ability to articulate how they were designing instruction to meet the specific needs of students within their own classroom. One teacher on this team talked about how she engaged the special education teacher and a paraprofessional so they could divide and rotate within three small groups each day. She talked about how these additional resources allowed her more one-on-one time to engage with students while ensuring each adult could gather important information about what skills students were mastering. She had changed her approach to include scaffolding of important information, more direct facilitation of learning, and adjusting materials to build missing skills. She was doing all of this within her classroom and without engaging her teammates in the planning or implementation perceivably because she considered this to be her responsibility. This particular teacher was able to take this approach due to having extensive content knowledge and many years of instructional strategies to pull from.

Wilcox and Zuckerman (2019) highlight developing will as the commitment to engage in district-wide and school-wide improvement efforts. So while this teacher team was not strong in a collaborative culture, they were still working toward the district-wide goal of increasing student growth. They did not meet their identified goal of getting 80% of students above the norm, but they recorded the second-highest amount of growth in the school. This team displayed the “will”, but perhaps equally interesting is that they also displayed significant capacity defined by Tichnor-Wenger et al. (2017) as “the

knowledge, skills, organizational routines, resources, and personnel available to support implementation" (p.8). In fact, this teacher team of veterans articulated significant content knowledge by talking in-depth about the specific skills they were teaching. Additionally, they highlighted how they made adjustments based on student data and how students demonstrated their skills before they moved on. So while this teacher team was not discussing these strategies openly among their teacher team, the adjustments that were expected to be collaborated upon were taking place in silos rather than as a whole teacher team. In this case, the teachers on this particular team described having strong individual efficacy without having strong collaborative culture.

It is because of the vast difference between Teacher Team 1 and Teacher Team 2 that the primary takeaway with regard to building a collaborative culture is that a collaborative culture is not sufficient on its own to increase student growth rates. Additionally, as evidenced by the interview data, a team can see success with regard to student growth rates without having a strong collaborative culture. It is true that the highest-performing teacher team also had a strong collaborative culture, but it is equally true that the second-highest-performing teacher team did not. It is because of this lack of consistency that we must fully consider the other factors regarding the whole picture and consider that individual teacher efficacy is not sufficient on its own to maximize student growth rates but that a collaborative culture can increase collective efficacy and maximize student growth rates.

**Using SMART Goals Unifies Teacher Teams**

The use of SMART goals was completely new for this district. Although setting goals is not new, learning what a SMART goal is and how to utilize it was new. DuFour

et al. (2010) emphasized the importance of goal setting as the single most powerful tool in a collaborative team's toolkit due to their ability to make a vision come to life. A goal that is Specific, Measurable, Attainable, Relevant, and Time-bound is more likely to be achieved and make a positive impact on student learning than a traditional goal that may lack focus and direction (DuFour et al., 2010). The teacher teams were supported in developing their SMART goals by all six administrators. A notable success of Phase 2 professional development is that six of seven teacher teams identified their SMART goal as the primary focus of their work during professional learning community (PLC) time each week. The reported consistency of using SMART goals indicates that teacher teams in this study found SMART goals to be valuable in keeping their team focused and organized during collaborative team time.

Unfortunately, only two teams met their goals, which focused on teacher tasks such as identifying essential learning standards and creating assessments that were fully dependent on adult action. These two teams ranked third and sixth in terms of student growth, making it difficult to conclusively determine that they increased student growth by meeting their SMART goals. Essentially, these two teams had goals that were not impacted by outside factors in the same way a student-performance goal might be, but the determination of whether achieving their goal influenced student growth remains inconclusive.

Miller's (2010) and Leithwood's (2010) research emphasizes that we must consider factors beyond student demographics though it is tempting to blame familial factors. It is critical to use this lens when evaluating why more teams did not meet their SMART goal by reflecting on whether we as a district have addressed the cultural

proficiency of our educators, provided equitable resources within the school system, and developed teacher efficacy to ensure equal access to high-quality instruction and presumably increased growth rates.

When I asked the teams about the roadblocks they had experienced, two teams blamed the students and three indicated that "the core was too difficult" for their students, all of which were in upper elementary. For a district with a free and reduced lunch population of 45% and virtually no language barriers, this attitude indicates a lack of both the will and capacity of teachers. These roadblocks indicate that the needs of these students have not been met at some point and that the teachers do not feel confident individually or collectively to meet these needs. It is possible that these students did not have access to quality instruction earlier in their elementary experience. It is also possible that their current teachers are lacking individual or collective efficacy in instructional knowledge, content knowledge, and the foundational skills needed to effectively realize growth within their classrooms. This attitude can be addressed; according to Tichnor-Wenger et al. (2017), this can be done by building the content knowledge, skills, and organizational routines of these teachers. These processes are essential in addressing this mindset, as Bryk et al. (2015) advised; without capacity, the teachers and students cannot make progress. Additionally, this district must consider if these roadblocks are occurring in the primary grades and simply not observed until later, or if there is a shift that occurs in upper elementary that causes this to be actualized for the first time. The answer to this will provide much-needed insight into where additional learning for the adults in the building should be focused.

Reflecting on the number of teams who met their SMART goals and what that context was, leads me to believe that there is much work to be done in this district with regard to ensuring equitable access to high-quality instructional materials. Teacher teams watering down grade-level content in hopes of securing an "A" in the classroom will eventually discover a lack of proficiency in the skills they need to be successful, as evidenced here. This is a slippery slope, as low expectations lead to low levels of growth, which further breeds the attitudes described here. A continual focus on building teacher capacity and actually utilizing the high-quality instructional materials provided by the district are important next steps.

**Adjusting Instruction Ensures Challenging Learning Opportunities**

The literature supports the importance of matching students with challenging and specific instruction and learning opportunities at their individual levels within a homogeneous setting as a powerful practice (Tomlinson & Imbeau, 2010; Tomlinson, 2018). In order to help teachers do this, they were taught how to dive deeper with MAP growth data by analyzing the Achievement Status & Growth Summary with Quadrant Chart that indicates which students in the classroom show the most growth and which students need to have instruction adjusted due to low levels of growth. Teachers were also provided with an opportunity to learn about the Learning Continuum and how MAP growth actually outlines small groups of students in the classroom who are ready to learn the same skill. I had anticipated the question, "when do we have time for this", so the remainder of the professional development session focused on integrating these questions and conversations into two protected times that are already built into teacher schedules: weekly team meeting time and collaborative team time every Wednesday.

Within these meeting structures, professional development was specific in ensuring each team had a specified meeting day and time each week in addition to collaborative team time on Wednesdays. A team meeting agenda was required of all teams that provided question prompts designed to help each team have deeper conversations about student learning. Questions such as “looking at our previous lesson plans, how can we adjust to ensure we’re meeting student needs?”, “Are there students that would benefit from shifting rooms?”, and “what strategies are being used in your colleagues' classrooms that they feel are effective? Can these be utilized in other areas?” were added as weekly conversation topics. The notes were then shared with SPED Teachers, MTSS Coordinators, the building principal, SPED Director, and Director of Teaching and Learning. Utilizing Google Docs’ “comments” feature, each of these colleagues that were not in the team meeting was able to help address many challenges throughout the year as they reviewed the team meeting notes.

Finally, teacher teams worked through the process of writing tiered objectives that intentionally utilize verbs that align with the Depth of Knowledge (DOK) outlined in the essential standard (Bloom, 1971). Teacher teams engaged in conversations around what DOK their objective should be and how they could adjust that objective by increasing the DOK by one category for those beyond-benchmark and how they could scaffold for students who needed assistance in meeting the grade-level DOK. Teachers utilized Bloom’s Taxonomy Chart to navigate through writing tiered objectives for an upcoming lesson. While this skill was taught, no product was required to demonstrate they were utilizing this process in their lesson planning nor were they required to report within their lesson plans the DOK they were working at for each objective.

After Phase 3, teacher interview data indicated that four of seven teams described adjusting instruction in ways that went beyond changing the pace of instruction, by going slower or faster or by changing the DOK of their objective to be less than what is expected for the grade-level skill, commonly known as "watering down the curriculum". The four teams that were able to articulate this process discussed specific strategies they use to make grade-level content more accessible, how they use small groups to support students with more direct adult assistance, and how their paras supported their instructional process by working one-on-one with students on specific skills they were ready to develop.

According to Tomlinson and Imbeau (2010), teachers must design instruction appropriate to students' stages of development, learning styles, strengths, and needs. Yet the data here show that there are three classrooms in the elementary where this is not happening. My interpretation of this information is that we still have some teachers that have not yet developed their individual capacity to truly differentiate instruction. While some teams can rely on their collaborative culture to develop their collective efficacy, those teams that lack this struggle to actualize high student growth rates.

**Additional Themes Supported in Literature**

As I reflected on the three professional development opportunities and the outcome of such, it was the lack of consistency among the top three teacher teams that led me to look at additional themes within the interviews. The usage of performance-based assessments, having data-focused conversations, shifting students once they have mastered the skill, and intentionally teaching within one's strengths are the additional themes identified.

**Table 5**

*Extended Themes Beyond Professional Development*

| | Extended Themes | | | | |
|---|---|---|---|---|---|
| Growth | Performance-Based Assessments | Data Focused Conversations | Truly Flexible Groups (opportunity to shift outside benchmarking) | Strengths-Based Teaching (team reflected on teacher strengths) | Met SMART Goal |
| 1st | X | X | X | | |
| 2nd | X | X | | | |
| 3rd | X | X | | X | X |
| 4th | | | | | |
| 5th | | X | X | X | |
| 6th | X | | | X | X |
| 7th | | | | | |

***Performance-based Assessments and Data-focused Conversations***

These two themes were the only consistent ones among all three of the teacher teams with the highest student growth rates. Teachers described using assessments that were focused on students demonstrating a specific skill. For example, rather than taking an assessment about the summary of a story, the teachers would prompt discussion with a question such as "what details in the story indicate the author's purpose was to entertain us?" This type of informal assessment required teacher teams to know their content, as well as interest on their part to ensure that students learn rather than simply follow a day-to-day progression outlined by a large-scale publisher. These team members also shared in their interview that they would record this information, often on a clipboard, with student names and skills listed, and would then discuss it with their team later. There was no reference to a student letter grade but rather if a student had mastered the skill or not. This made the data-based conversations easier for these teams, as they had specific skills they could talk about.

Blecker and Boakes (2010) argued that performance-based assessment strategies are one of the most important things to focus on when developing teacher training.

Further, DuFour et al. (2013) outlined that by employing performance-based assessment strategies, teacher teams can use this information to guide interventions, inform next steps, assess team member strengths and weaknesses, and develop better strategies for meeting the needs of students. Essentially, without performance-based assessments, data-focused conversations are difficult. In this case, performance-based assessments facilitate data-focused conversations.

Only one team's themes matched the positive outlier closely, and that team happened to have the fifth-highest growth rate among all seven teams. The one thing that was missing from this team's themes was the usage of performance-based assessments. This piece of information is worth exploring in that performance-based assessments may be a critical component in maximizing student growth rates. Based on the information shared by Blecker and Boakes (2010) and DuFour et al. (2013), it is possible that performance-based assessments are indeed the key to the process of differentiating instruction, as the assessments ensure the collection of critical information that can further be explored to ensure instruction can be customized; without the assessment data, teachers are simply shooting arrows in the dark.

### *Shifting Students Once they have Mastered the Skill*

The practice of tracking, or grouping students and leaving them in those same groups, is well documented in research to increase inequities for students. Oakes found repeatedly (2005, 1990, 1986) that tracking decreases equitable access to high-quality teachers and materials while increasing the achievement and opportunity gap for those students at the highest risk. Grouping, however, can yield increased growth rates, particularly for students who are at the benchmark level and above (Lou et al., 1996;

Murphy et al., 2017; Seleh et al., 2005). Flexible grouping provides a hybrid approach to instruction that allows teachers to realize the benefits of homogeneous groups without encountering the negatives.

While the ability to shift students frequently based on mastery of a skill was a primary conversation topic throughout the professional development sessions, only two teams clearly articulated during the interview that they shifted students as they mastered a skill and not based on an arbitrary date (after a benchmarking period or at the end of a quarter or semester). These two teams ranked first and fifth with regard to growth rates. This was disappointing considering nearly all of the teams claimed to be using flexible grouping, but when they were pressed for specifics, they would identify shifting students after benchmarking as being "flexible", when in reality, the groups were pretty stagnant between those periods. While this is an improvement from tracking that leaves a student in the same group for the entire year or longer, I hoped that this practice would be more common. It is my opinion that the reason this is so difficult for many teacher teams is that they are not collaborating on what skills they are teaching to ensure they maintain a constant pace so that students could easily slip in or out of a particular section without missing significant amounts of learning.

The ability to effectively adjust instruction is an issue of individual efficacy that can be mitigated through collaborative culture and collective efficacy but not in isolation. When groups become stagnant, the concern for inequity increases. As fluidity decreases, the potential that students, particularly those approaching benchmark, experience low-quality instruction and learning environment increases (Oakes, 1985; Gamoran & Hallinan, 1995). This research, combined with my interpretation of the interview data,

leads me to believe that this lack of flexibility negatively impacted student growth rates in this study. This particular situation outlines how a teacher who lacks the capacity to adjust instruction and does not have access to a collaborative culture can struggle to realize the kind of growth that is expected according to national norms.

### *Intentionally Teaching Within One's Strengths*

This theme came from the question "how do you decide who teaches what?" Four of seven teams identified "burnout" as their main reason for determining who took the low group and, similarly, who needed a "break" with the beyond-benchmark group. Only three of seven teams considered their individual strengths before determining who would teach each class, and only one of those teams was in the top three with regard to student growth. Outside the scope of this research project, the elementary principal had spent time with every grade level the previous year, encouraging them to evaluate their strengths and teach within their strengths. Thankfully, teacher teams with non-tenured teachers could identify the importance of having a brand-new teacher teach the benchmark group (i.e., fewer materials to design), but aside from that, it felt like the passing of a baton—"I had them last year."

This is an unfortunate revelation shared partially because these teachers are comfortable with me, but also is a reflection of the fact that we as a district need to develop a better process to meet the needs of students who require additional support. These data indicate a need for this district to develop individual efficacy with regard to meeting the needs of students approaching the benchmark. If these teachers felt that they could successfully meet the needs of the students in their classroom, they would not be as likely to feel "burned out." Similarly, by building these skills of the entire team, they can

increase their collective efficacy and share these practices through collaboration. As a district, we must further research the specific skills we can develop in our teachers that will ensure they feel confident and competent in meeting the needs of our most at-risk learners.

# CHAPTER 6

# CONCLUSION

The findings of this study suggest there is no one thing that enables all grade-level teams to observe the same kind of growth that the positive outlier team demonstrated. This is not surprising considering that the literature has well-documented the complexities of both teaching and learning. Dating back to the Coleman Report in 1966 when familial factors were considered to have the most influence on student outcomes, to more contemporary research that aims to understand the complex, dynamic transactions between nature and nurture, as well as the variations of these transactions across time, place, and individuals (Osher et al., 2020), teaching and learning have been considered more complex than any single study can dissect.

The interview data allowed us to gain a better understanding of what was learned. Only two common themes were identified from the interviews with the three teams with the highest, consistent student growth rate: 1) a primary focus on utilizing performance-based assessments and 2) collecting student data to adjust instruction based on that data. After Phase 1 data analysis, I expected to find that collaborative culture was the most impactful theme. However, the interview data revealed that even the team that had not met for two months produced strong student growth rates. This team followed a process to evaluate what students knew after each lesson so that even if the students were not flexing frequently, the teacher team could adjust how they designed and taught their lessons within their own classroom the next day.

These data lead me to believe that the ability to develop performance-based assessments and use those to guide instruction is the single most important skill to

increase student growth. Certainly, this study was not designed to rule one factor as more important than another, but there is a strong indication here that an individual teacher's ability to design high-quality instruction that aligns with the needs of students (individual efficacy) and their ability to collaborate about this process (collaborative culture) must exist simultaneously to allow teachers the opportunity to flexibly group students and maximize student growth.

**Implications for Future Improvement Efforts**

The most important part of any research project is not merely presenting what was learned, but what can be done with the information to achieve certain goals. Timperley and Philips (2003) outlined a three-pronged non-sequential approach that outlined the importance of confronting current teacher beliefs about student learning, focusing on building teacher efficacy and learning and developing new teacher skills. It is within this same context that this district will continue to operate moving forward in the Act phase of the PDSA cycle, acknowledging that once we have worked through this fourth phase, we will begin again with Planning, embracing the true nature of a cycle.

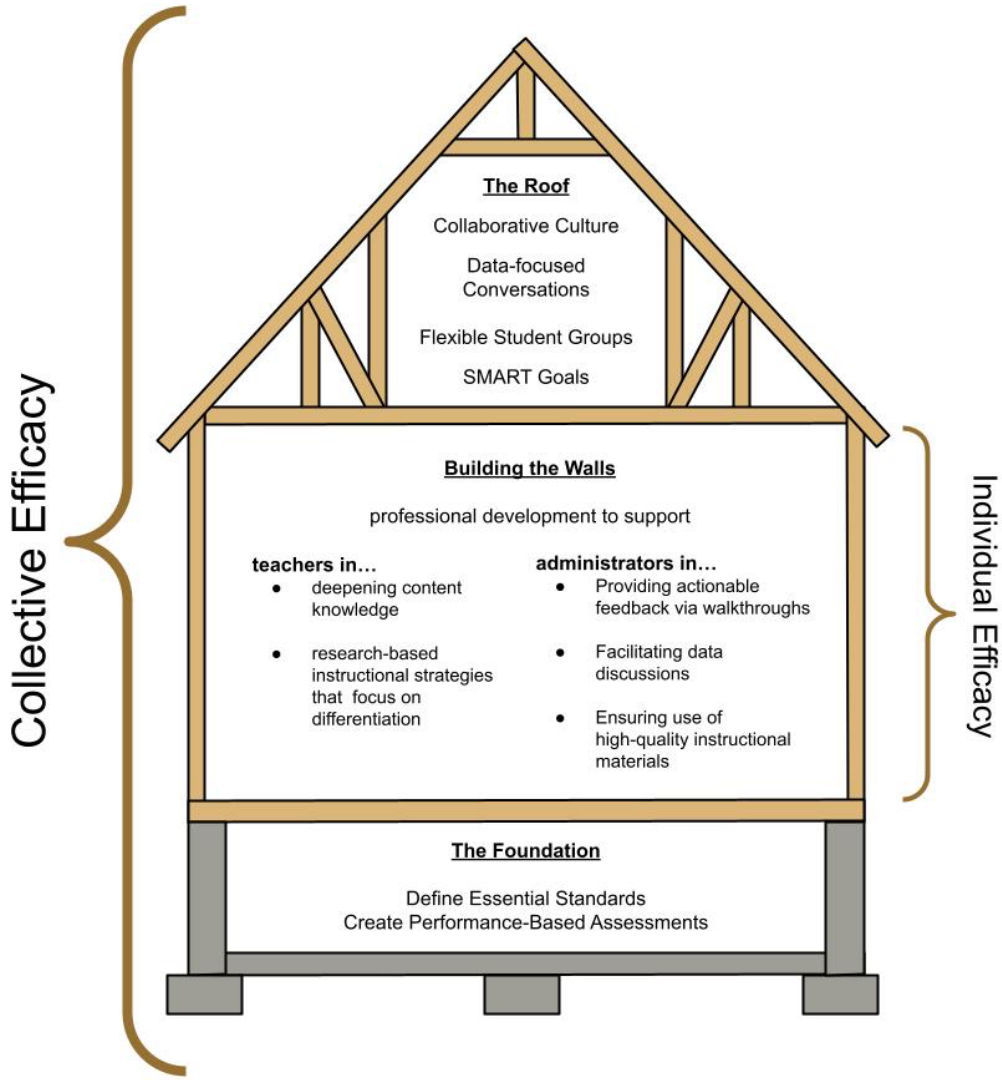
Based on this study's findings and conclusions, three overarching implications for our district and educational leaders can be outlined: 1) the importance of developing teacher capacity, 2) the need to build performance-based assessments, and 3) the importance of supporting and developing a collaborative culture among teacher teams. I intend to clarify here that I did not seek to construct new knowledge. Instead, my post-positivist lens was intentional in working to make sense of knowledge that already existed. So, while these conclusions may seem generic, these implications tell me that in this district, with these teachers, this is what we need to spend our time and efforts on as

we work to develop the capacity of our teachers. Until every teacher in this district can successfully master these skills, our hopes of maximizing student growth for every student will not be fully realized.

Figure 9 helps to visualize how each of these components fits together. One can bring to mind the process of building a house. To ensure strength, the foundation must be laid first, then the walls, and finally the roof. First, essential standards must be defined and performance-based assessments must be created. Next, individual teacher efficacy must be enhanced through professional development to deepen content knowledge and cultivate research-based instructional strategies that focus on differentiation. Finally, a collaborative culture supports an environment where effective flexible grouping is possible through data-focused discussions, inquiry-based teacher learning, and the achievement of SMART goals. It is only when each of these components exists in a collective effort that the structure can stand the test of time and maximize the growth rate of every student.

**Figure 9**

*Building the "House"*




***Foundation: The Need to Build Performance-based Assessments***

In addition to defining essential standards (i.e. teacher clarity), building formative performance-based assessments is an important step in ensuring teachers have data-focused conversations and properly adjust their instruction to meet the needs of students.

As the only theme observed in all three teacher teams with the highest student growth rates, ensuring teacher teams utilize formative performance-based assessments is the foundation for ensuring teachers have actionable information to adjust their instruction (Blecker & Boakes, 2010; Marzano, 2017; Tomlinson & Moon, 2013). DuFour et al. (2013) acknowledged that performance-based assessments allow teacher teams to guide interventions, inform steps, assess team members' strengths and weaknesses, and develop better strategies for meeting the needs of students. The teacher teams that utilized performance-based assessments also had these conversations outlined by DuFour et al. (2013). Without performance-based assessments, these other components would not be possible. Thus, developing performance-based assessments should be a primary consideration for this district to move forward.

The teacher teams that saw the most student growth identified performance-based assessments as "what they talk about" when they meet as a team. The fact that students are demonstrating what they are learning increases the likelihood that their grade will reflect what they know, but more importantly, that the teacher will know what they know and be able to use that information to further inform their instruction.

Educational leaders seeking to maximize student growth, particularly through flexible grouping, should take note that this particular strategy is the gateway to better conversations among staff. Although this step is listed second, it is critically important in order to take learning to the "next step." As I have transitioned to a new role, the work of our new Director of Teaching and Learning will need to focus on ensuring all grade levels develop and use performance-based assessments for essential skills. Based on the interview data collected, three of four teams are not using them at all; more importantly,

only two of those teams, the two highest performing teams, are using them well enough that they are actually informing their instruction accordingly. Additional professional development will need to focus on these next steps.

### *The Walls: The Importance of Developing Teacher Capacity*

Once the foundation has been built, a district can begin to focus on the "walls". These findings along with decades of research indicate that individual and collective teacher efficacy is more impactful than any other education input, strategy, or reform (Donohoo et al., 2018; Hattie, 2009; Wenglinsky, 2000). However, based on what we have learned here, we cannot truly focus on individual efficacy until the foundation has been laid. The findings of this study reiterate the importance of developing teacher capacity, as was evidenced by the team that showed the second-highest amount of student growth by using strong instructional strategies despite a lack of collaborative culture. Additional evidence in this case indicates that the teacher team with the lowest student growth rates displayed a lack of teacher capacity for instructional strategies when they shared that they were often "spinning their wheels" and relied on "Teachers Pay Teachers" for additional materials because the Tier 1 materials were "too hard" for their students. Their first response was that their lack of progress was the "student's fault" and that the benchmark materials were "just too hard", which bespeaks a lack of both individual and collective efficacy. Moving forward, it is critical we develop both the capacity of these teachers, as well as encourage the development of their will, in order to maximize student growth.

Further, supporting teacher teams in developing more hands-on activities and strategies to improve how they teach higher-order thinking skills is an important next step

(Wenglinsky, 2000). Additionally, ensuring access to high-quality instructional materials is critically important, as evidenced by Erberber et al. (2015) and Kane (2016), who indicated that the quality of instructional materials is as important as the quality of the teacher.

For instructional leaders in this situation, this requires intentional engagement. As I have transitioned to a new role within our district, I will be working with our building principals and the new Director of Teaching and Learning to ensure we have systems in place that promote classroom walkthroughs that focus on feedback, intentional engagement with teachers about student data and what to do with it, as well as a consistent review of instructional materials being utilized. This is the most critical step in the Act phase of our process, particularly with regard to those teams that have not yet observed student growth rates at or above the norm.

### *The Roof: The Importance Of Supporting And Developing A Collaborative Culture Among Teacher Teams*

Developing collaborative culture was not something I had sought to find initially. The positive outlier teacher team demonstrated a strong interdependency among one another that was difficult to replicate in professional development. Hoping to create a tangible experience (professional development) for an intangible quality (trust), the professional development in this area was successful in improving the collaborative culture in three teams; however, it cannot be recognized as the sole factor responsible for such improvement. Listed here as an important implication, developing a collaborative culture is about trusting one's teammates, knowing they're all on the same page, working toward the same goal, and moving in the same direction. Although this was not a

consistent finding among the top-performing teacher teams, it warrants further investigation for the sole purpose that the positive outlier team identified this as the major reason they are able to make a difference in their students' lives.

Educational leaders should make note of opportunities to build trust among teacher teams, provide opportunities for structured conversation using protocols, and consider the use of a facilitator in more challenging situations in order to develop collaborative culture. Fullan and Quinn (2016) assured that a shared depth of understanding about the nature of the work helps build capacity from the inside out, connecting the work of developing a collaborative culture with the importance of developing teacher capacity, which brings us full circle. Not a single implication operates on its own; instead, the three become interdependent in such a way that all are required to maximize student growth.

As such, it is the work of this district to further pursue ways in which teacher capacity can be developed, performance-based assessments can be designed and utilized, and collaborative culture can be further developed to ensure the interconnected skills and strategies learned about in this case study can be utilized for the benefit of students.

**Limitations and Areas for Future Research**

This project was designed to learn how a single positive case has been utilizing flexible grouping to address the diverse academic needs of their students and how those teachers work together as a grade-level team to impact student growth, in order to spread those practices to other grade-level teams within the same school. The narrow scope of this research project was both a strength and a limitation. Since I studied only one school building, I was able to develop an intimate setting where the data could be analyzed

closely while the experience was very fresh. The purpose of this case study was not to imply an understanding of other cases but to understand this one case exceptionally well (Schwandt & Gates, 2018).

However, the study was potentially limited by events outside of my control, such as the COVID-19 pandemic, which resulted in disrupted teaching and learning when students were sent home for the last 25% of the school year in 2019–2020. Continued disruptions in the form of rapidly changing quarantine guidelines, mask mandates, social distancing requirements, and the balancing of in-person, remote, and hybrid learning environments, as well as inequitable access to each of these environments, all contributed to a rapidly increasing skill gap of content knowledge between low- and high- performing students and contributed to a rapid increase in teacher burnout (Kuhfeld et al., 2020; Pressley, 2021). The potential impact may include growth rates that were impacted by a student's ability to access high-quality instruction, teachers lacking the will to design highly-engaging content due to the increased workload that accompanied the COVID-19 pandemic, and a classroom environment that had increased social-emotional needs that potentially created a learning environment that was not as conducive for learning as it would have been pre-pandemic.

In our own district, we had put a lot of time and effort into developing essential learnings prior to beginning this case study. These essential learnings reflected what we wanted students to know and be able to do once they completed each class. The only common thread among the top three performing teacher teams was their use of performance-based assessments and having discussions about that data. As a result of this study, we now recognize the importance of designing performance-based assessments

that connect back to these essential learnings as a critical next step in this process. This research also reiterates the importance of providing consistent collaboration time for teacher teams to have these data-focused conversations about student mastery of skills and adjusting instruction based on their current level of mastery.

Additionally, the data indicate that we must continue to focus on building individual teacher efficacy. In this study, we saw that those with strong content knowledge and research-based instructional strategies were able to differentiate their instruction regardless of whether they were teaching the approaching-, beyond-, or benchmark-level students. We can support this work as a district by promoting classroom walkthroughs that focus on providing specific feedback to teachers about their use of performance-based assessments and differentiated instructional strategies within the classroom.

Furthermore, we must recognize when a teacher needs to deepen their own content knowledge in order to feel more confident in adjusting instruction to meet the needs of students. Without this, true differentiation cannot exist beyond what is provided within mass-published instructional materials or, worse, what is quickly located on a shared resources website such as “Teachers Pay Teachers”, which is why a consistent review by the administration of the instructional materials being utilized is also important.

The most important lesson learned in this case study is that there is no single component that is sufficient to maximize student growth. We must first begin by laying the foundation by defining essential standards and creating performance-based assessments. We may then progress to building block-level components by developing

individual teacher efficacy through professional development that deepens content knowledge and cultivates research-based instructional strategies that focus on differentiation. Only after these pieces are in place can we maximize the impact of a collaborative culture where effective flexible grouping is possible through data-focused discussions, inquiry-based teacher learning, and the achievement of SMART goals. It is only when each of these components exists in a collective effort that we will be able to maximize the growth rate of every student.

## REFERENCES

Alexander, K., & Morgan, S. L. (2016). The Coleman Report at fifty: Its legacy and implications for future research on equality of opportunity. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *2*(5), 1–16.

Alexander, K. L., Cook, M., & McDill, E. L. (1978). Curriculum tracking and educational stratification: Some further evidence. *American Sociological Review*, *43*, 47–66.

Andrabi, T., Daniels, B., & Das, J. (2021). Human capital accumulation and disasters: Evidence from the Pakistan earthquake of 2005. *Journal of Human Resources, 58*(1). https://doi.org/10.3368/jhr.59.2.0520-10887R1

Ansalone, G. (2010). Tracking: Educational differentiation or defective strategy. *Educational Research Quarterly, 34*(2), 3–17.

Argys, L. M., Rees, D. I., & Brewer, D. J. (1996). Detracking America's schools: Equity at zero cost? *Journal of Policy Analysis and Management, 15*(4), 623–645.

Atteberry, A. C., & McEachin, A. J. (2020). Not where you start, but how much you grow: An addendum to the Coleman Report. *Educational Researcher*, *49*(9), 678–685.

Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one? *Child Development, 59*(1), 87–96. http://dx.doi.org/10.2307/1130391

Bahari, S. F. (2010). Qualitative versus quantitative research strategies: contrasting epistemological and ontological assumptions. *Sains Humanika*, *52*(1). https://doi.org/10.11113/sh.v52n1.256

Bailey, C., & Bridges, D. (2016). *Mixed ability grouping: A philosophical perspective*. Routledge.

Bandura, A. (1997). The anatomy of stages of change. *American Journal Of Health Promotion: AJHP*, *12*(1), 8–10.

Berliner, D. C. (2005). The near impossibility of testing for teacher quality. *Journal of Teacher Education*, *56*(3), 205–213.

Bernhardt. (2013). *Data analysis for continuous school improvement*. Routledge.

Biafora, F., & Ansalone, G. (2008). Perceptions and attitudes of school principals toward school tracking: Structural considerations of personal beliefs. *Education*, *128*(4), 588–602.

Bielinski, J., Brown, R., Wagner, K. (2020). *COVID slide: Research on learning loss & recommendations to close the gap.* Illuminate Education.

Black, P., & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *92*(1), 81–90.

Bloom, B. S. (1971). Mastery learning. In J. H. Block (Ed.), *Mastery learning: Theory and practice* (pp. 47–63). Holt, Rinehart and Winston.

Bloom, B. S. (1984). The search for methods of group instruction as effective as one-to-one tutoring. *Educational Leadership*, *41*(8), 4–17.

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American Educational Research Journal*, 42(2), 231–268.

Boser, U., Chingos, M., & Straus, C. (2015). The hidden value of curriculum reform. *Center for American Progress*. https://cdn.americanprogress.org/wp-content/uploads/2015/10/06111518/CurriculumMatters-report.pdf

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives, Handbook I: Cognitive domain.* Longmans.

Briggs, M. (2020). Comparing academically homogeneous and heterogeneous groups in an active learning physics class. *Journal of College Science Teaching*, *49*(6), 76–82.

Brookover, W. B., & Lezotte, L. W. (1979). *Changes in school characteristics coincident with changes in student achievement. Occasional paper no. 17.* Michigan State University.

Brulles, D., Saunders, R., & Cohn, S. J. (2010). Improving performance for gifted students in a cluster grouping model. *Journal for the Education of the Gifted*, 34(2), 327–350.

Bryk, A. S. (2020). *Improvement in Action: Advancing Quality in America's Schools*. Harvard Education Press.

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better*. Harvard Education Press.

Burris, C. C. (2014). *On the same track*. Beacon Press.

Burris, C. C., & Garrity, D. T. (2008). *Detracking for excellence and equity*. ASCD.

Carnegie Foundation (2020). *The six core principles of improvement*. Retrieved from https://www.carnegiefoundation.org/our-ideas/six-core-principles-improvement/

Carpenter, W. A. (2000). Ten years of silver bullets: Dissenting thoughts on education reform. *The Phi Delta Kappan*, *81*(5), 383–389.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2012). Great teaching. *Education Next*, *12*(3), 58–68.

Chingos, M. M., & Whitehurst, G. J. (2012). Choosing blindly: Instructional materials, teacher effectiveness, and the common core. *Brookings Institution*. https://www.brookings.edu/wp-content/uploads/2016/06/0410_curriculum_chingos_whitehurst.pdf

Coburn, C. E. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. *Sociology of Education*, 77(3), 211–244.

Coburn, C. E., & Woulfin, S. L. (2012). Reading coaches and the relationship between policy and practice. *Reading Research Quarterly*, 47(1), 5–30.

Coleman, J. (1968). The concept of equality of educational opportunity. *Harvard Educational Review*, 38(1), 7–22. https://doi.org/10.17763/haer.38.1.m3770776577415m2

Coleman, M. B. (2016). Commentary: Sorting it out. *Journal of Advanced Academics*, 27(2), 117–123.

Collins, K. M. (2013). *Ability profiling and school failure: One child's struggle to be seen as competent*. Routledge.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications.

Darling-Hammond, L. (2016). *Unequal opportunity: Race and education*. Retrieved October 28, 2021, from https://www.brookings.edu/articles/unequal-opportunity-race-and-education/

Datnow, A., & Park, V. (2018). Opening or closing doors for students? Equity and data use in schools. *Journal of Educational Change, 19*, 131–152. https://doi.org/10.1007/s10833-018-9323-6

David, R., Teddlie, C., & Reynolds, D. (2000). *The international handbook of school effectiveness research*. Psychology Press.

Denzin, N. K., & Lincoln, Y. S. (2018). *Handbook of qualitative research*. Sage.

Ding, C. S. & Davison, M. L. (2005). A longitudinal study of math achievement gains for initially low achieving students. *Educational Psychology, 30*(2005), 81–95.

Domina, T., Pharris-Ciurej, N., Penner, A. M., Penner, E. K., Brummet, Q., Porter, S. R., & Sanabria, T. (2018). Is free and reduced-price lunch a valid measure of educational disadvantage? *Educational Researcher*, *47*(9), 539–555.

Donohoo, J., Hattie, J., & Eells, R. (2018). The power of collective efficacy. *Educational Leadership*, *75*(6), 40–44.

Downey, D. B., & Condron, D. J. (2016). Fifty years since the Coleman Report: Rethinking the relationship between schools and inequality. *Sociology of Education*, *89*(3), 207–220.

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review, 101*(5), 1739–1774.

DuFour, R. P., & Schwartz, W. (1990). Addressing the tracking controversy by promoting educational opportunity. *NASSP Bulletin*, *74*(530), 88–94. https://doi.org/10.1177/019263659007453017

DuFour, R., & DuFour, R. (2013). *Learning by doing: A handbook for professional learning communities at work TM*. Solution Tree Press.

Edmonds, R. (1979). *A discussion of the literature and issues related to effective schooling*. ERIC.

Edmonds, R. R., & Frederiksen, J. R. (1979). *Search for effective schools: The identification and analysis of city schools that are instructionally effective for poor children*. Educational Leadership.

Eells, R. J. (2011). *Meta-analysis of the relationship between collective teacher efficacy and student achievement* [Doctoral dissertation]. Loyola University.

Elmore, R. F., & Burney, D. (1997). *School variation and systemic instructional improvement in community school district# 2, New York City. High performance learning communities project.* Pittsburgh University.

Elwood, S. A., & Martin, D. G. (2000). "Placing" interviews: location and scales of power in qualitative research. *The Professional Geographer*, *52*(4), 649–657.

Erberber, E., Stephens, M., Mamedova, S., Ferguson, S., & Kroeger, T. (2015). *Socioeconomically disadvantaged students who are academically successful: Examining academic resilience cross-nationally. Policy Brief No. 5.* International Association for the Evaluation of Educational Achievement.

Flores, A. (2007). Examining disparities in mathematics education: Achievement gap or opportunity gap? *High School Journal*, *91*(1), 29–42. https://doi-org.unk.idm.oclc.org/10.1353/hsj.2007.0022

Foote, M. Q., & Bartell, T. G. (2011). Pathways to equity in mathematics education: How life experiences impact researcher positionality. *Educational Studies in Mathematics*, *78*(1), 45–68.

Fridland, V. (2020). *Why do masks make it so hard to understand each other?* Nevada Today.

Fullan, M., & Quinn, J. (2015). *Coherence: The right drivers in action for schools, districts, and systems*. Corwin Press.

Fullan, M., & Quinn, J. (2016). Coherence making. *School Administrator*, *73*(6), 30–34.

Gallagher, J., & Harradine, C. C. (1997). Gifted students in the classroom. *Roeper Review*, *19*(3), 132. https://doi-org.libproxy.unl.edu/10.1080/02783199709553808

Gallop. (2017). *State of America's schools: The path to winning again in education.* Retrieved from file:///Users/allisonjonas/Desktop/Gallup%20Report%20--%20State%20Of%20Americas%20Schools.pdf

Gamoran, A., Hallinan, M. (1995). Tracking students for instruction. In M. T. Hallinan (Ed. ), *Restructuring schools* (pp. 113–131). Springer.

Gardner, H. (1998). A multiplicity of intelligences. *Scientific American*, *9*(4), 19–23.

Gillborn, D., & Youdell, D. (1999). *Rationing education: Policy, practice, reform, and equity.* Open University Press.

Gothenburg Public Schools (2021). *Strategic plan*. Retrieved Jan 22, 2022, from https://docs.google.com/document/d/1irIMBVY5jFF7LrigxdiPzLxCqKO2MmOwDHj89RDns3I/preview

Gorman, J. (2011). *The association between grades pre k-12 student achievement and differentiated instructional strategies in the Anytown Township School District*

*explored through units of study* [unpublished doctoral dissertation]. Rowan University.

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher*, *39*(1), 59–68.

Guba, E. G., & Lincoln, Y. S. (2005). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.). *The Sage handbook of qualitative research* (pp. 193–215). Sage Publications..

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–117). Sage.

Grix, J. (2019). *The foundations of research*. Macmillan International Higher Education.

Grusky, D., & Hill, J. (2018). *Inequality in the 21st century: A reader*. Routledge.

Guskey, T. R. (2010). Lessons of mastery learning. *Educational Leadership*, *68*(2), 52.

Guskey, T. R. (2009). Mastery learning. In T. L. Good (Ed.), *21st century education: A reference handbook* (Vol. 1, pp.194–202). Sage.

Guskey, T. R. (2005). Formative classroom assessment and Benjamin S. Bloom: Theory, research, and implications. *Online Submission*.

Hahn, J. (2019). *Nebraska approves NWEA assessments MAP growth reading and MAP reading fluency to meet early reading goals of Nebraska Reading Improvement Act.* NWEA. Retrieved November 20, 2021, from https://www.prweb.com/releases/nebraska_approves_nwea_assessments_map_gr

owth_reading_and_map_reading_fluency_to_meet_early_reading_goals_of_nebraska_reading_improvement_act/prweb16157295.htm

Hanushek, E. A. (2016). What matters for student achievement. *Education Next*, *16*(2), 18–26.

Hart, S., Dixon, A., Drummond, M. J., & McIntyre, D. (2004). *Learning without limits*. McGraw-Hill Education (UK).

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* Routledge.

Hertert, L., & Teague, J. (2003). *Narrowing the achievement gap: A review of research, policies, and issues.* EdSource, Inc.

Hinton, M. (2020). *Why teaching kindergarten online is so very, very hard*. Edutopia. Retrieved November 20, 2021, from https://www.edutopia.org/article/why-teaching-kindergarten-online-so-very-very-hard

Holmes, A. G. D. (2020). Researcher positionality—A consideration of its influence and place in qualitative research—A new researcher guide. *Shanlax International Journal of Education*, *8*(4), 1–10.

Honig, M. I. (2009). What works in defining "what works" in educational improvement: Lessons from education policy implementation research, directions for future research. *Handbook of education policy research* (pp. 333–347). Routledge.

Hornby, G., & Witte, C. (2014). Ability grouping in New Zealand high schools: Are practices evidence based? *Preventing School Failure, 58*(2), 90–95.

Huang, H., & Sebastian, J. (2015). The role of schools in bridging within-school achievement gaps based on socioeconomic status: A cross-national comparative study. *Compare*, *45*(4), 501–525. https://doi.org/10.1080/03057925.2014.905103

Hussar, B., Zhang, J., Hein, S., Wang, K., Roberts, A., Cui, J., Smith, M., Bullock Mann, F., Barmer, A., & Dilig, R. (2020). The Condition of Education 2020. NCES 2020-144. *National Center for Education Statistics*. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2020144

Jackson, K., & Makarin, A. (2018). Can online off-the-shelf lessons improve student outcomes? Evidence from a field experiment. *American Economic Journal: Economic Policy*, *10*(3), 226–254. https://www.nber.org/papers/w22398

Jencks, C. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. Basic Books.

Jenkins, L. (2012). *Speaking out: Stop the pendulum*. National Association of Elementary School Principals. Retrieved from https://www.naesp.org/principal-novdec-2012-stem-issue/speaking-out-stop-pendulum

Jimenez-Castellanos, O. (2012). Revisiting the Coleman Report: Deficit ideologies and federal compensatory funding in low-income Latino school communities. *Association of Mexican American Educators Journal*, *6*(2).

Johnson, D. W., Johnson, R. T., & Stanne, M. B. (2000). *Cooperative learning methods: A meta-analysis.*University of Minessota.

Kaffenberger, M. (2021). Modelling the long-run learning impact of the Covid-19 learning shock: Actions to (more than) mitigate loss. *International Journal of Educational Development*, *81*, 102326.

Kalogrides, D., & Loeb, S. (2013). Different teachers, different peers: The magnitude of student sorting within schools. *Educational Researcher, 42*(6), 304–316. https://doi-org.unk.idm.oclc.org/10.3102/0013189X13495087

Kane, T. J. (2016, March 3). Never judge a book by its cover: Use student achievement instead. *Education Next*. https://www.brookings.edu/research/never-judge-a-book-by-its-cover-use-student-achievement-instead/

Kane, T. J., Owens, A. M., Marinell, W. H., Thal, D. R., & Staiger, D. O. (2016). *Teaching higher: Educators' perspectives on common core implementation*. https://cepr.harvard.edu/files/cepr/files/teaching-higher-report.pdf

Karen, D. (2005). No child left behind? Sociology ignored! *Sociology of Education*, *78*(2), 165–169.

Kim, Y. (2012). Implementing ability grouping in EFL contexts: Perceptions of teachers and students. *Language Teaching Research*, *16*(3), 289–315. doi:10.1177/1362168812436894

Kozol, J. (2012). *Savage inequalities: Children in America's schools*. Crown.

Knapp, M. S., Copland, M. A., & Swinnerton, J. A. (2007). Understanding the promise and dynamics of data-informed leadership. *Yearbook of the National Society for the Study of Education, 106*(1), 74–104.

Kuhfeld, M. & Tarasawa, B. (2020). *The COVID-19 slide: What summer learning loss can tell us about the potential impact of school closures on student academic achievement.* NWEA.

Kuhfeld, M., Tarasawa, B., Johnson, A., Ruzek, E., & Lewis, K. (2020). *Learning during COVID-19: Initial findings on students' reading and math achievement and growth.* NWEA.

Kulik, C. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research*, *60*(2), 265–299.

Kulik, C. C., & Kulik, J. A. (1982). Effects of ability grouping on secondary school students: A metanalysis of evaluation findings. *American Educational Research Journal, 19*(3), 415–428.

Kulik, J. (1992). *An analysis of the research on ability grouping : historical and contemporary perspectives : Executive summary*. National Research Center on the Gifted and Talented.

Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in US schools. *Educational Researcher*, *35*(7), 3–12.

Langley, G. J., Moen, R. D., Nolan, K. M., Nolan, T. W., Norman, C. L., & Provost, L. P. (2009). *The improvement guide: A practical approach to enhancing organizational performance*. John Wiley & Sons.

Leithwood, K., Patten, S., & Jantzi, D. (2010). Testing a conception of how school leadership influences student learning. *Educational Administration Quarterly*, *46*(5), 671–706.

LeMahieu, P. G., Bryk, A. S., Grunow, A., & Gomez, L. M. (2017). Working to improve: Seven approaches to improvement science in education. *Quality Assurance in Education, 25*(1), 2–4. https://doi.org/10.1108/QAE-12-2016-0086

Lincoln, Y. S., Lynham, S. A., & Guba, E. G. (2011). Paradigmatic controversies, contradictions, and emerging confluences, revisited. *The Sage handbook of qualitative research*, *4th ed.* (pp. 97–128). Sage.

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, *66*(4), 423–458.

Loveless, T. (1999). The tracking wars: State reform meets school policy. Brookings Institution Press. 88

Loveless, T. (2013). The 2013 Brown Center Report on American education: The resurgence of ability grouping and persistence of tracking. Brookings Institute.

Loveless, T. (2009). *Tracking and detracking: High achievers in Massachusetts middle schools.* Thomas B. Fordham Institute. Retrieved from http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED507543

Loveless, T. (1998). Making sense of the tracking and ability grouping debate. *Fordham Report*, *8*(2), 4–30.

Lucas, S.R., & Gamoran, A. (2002). Tracking and the achievement gap. In J. E. Chubb & T. Loveless (Eds.), *Bridging the achievement gap* (pp. 171–198). Brookings Institution Press.

Mangan, K. (2021). Could the summer slide become a covid crash? *Chronicle of Higher Education*, *67*(23), 4.

Marsh, D., & Furlong, P. (2002). A skin not a sweater: Ontology and epistemology in political science. *Theory and Methods in Political Science*, *2*, 17–41.

Marzano, R. J. (2017). *The new art and science of teaching (revised and expanded edition).* Solution Tree Press.

Matthews, M. S., Ritchotte, J. A., & McBee, M. T. (2013). Effects of schoolwide cluster grouping and within-class ability grouping on elementary school students' academic achievement growth. *High Ability Studies, 24*(2), 81–97.

Mattos, M., Nielsen, M.O. (2017, July). *The fifteen day challenge: Professional Learning Communities at Work*. Solution Tree. https://www.solutiontree.com/events/plc-at-work.html.

McKown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology*, *46*(3), 235–261.

McLaughlin, M. W., & Talbert, J. E. (1993). *Contexts that matter for teaching and learning: Strategic opportunities for meeting the nation's education goals*. Center for Research on the Context of Secondary School Teaching, Stanford University

Mellard, D.F., & Johnson, E. (2008). *RTI: A practitioner's guide to implementing Response to Intervention.* Corwin.

Merriam, S. B. (1998). *Qualitative research and case study applications in education.* Jossey-Bass Publishers.

Merriam, S. B. (1991). Introduction to qualitative research. *Qualitative Research in Practice: Examples for Discussion and Analysis*, *1*(1), 1–17.

Merriam, S. B., Caffarella, R. S., & Baumgartner, L. M. (1999). *Learning in adulthood: A comprehensive guide*. John Wiley & Sons/Jossey-Bass.

Merriam, S. B. & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). John Wiley and Sons.

Mickelson, R. A., Bottia, M. C., & Lambert, R. (2013). Effects of school racial composition on K-12 mathematics outcomes: A metaregression analysis. *Review of Educational Research*, 83(1), 121–158. Doi:10.3102/0034654312475322

Milner, H. R. (2010). *Start where you are, but don't stay there: Understanding diversity, opportunity gaps, and teaching in today's classrooms*. Harvard Education Press.

Murphy, J. (2010). *The educator's handbook for understanding and closing achievement gaps*. Corwin Press.

Mourshed, M., Chijioke, C., & Barber, M. (2010). *How the world's most improved school systems keep getting better.* McKinsey & Company. Retrieved 3/1/17 from http://mckinseyonsociety.com/downloads/reports/Education/ How-the-Worlds-Most-Improved-School-Systems-Keep-Getting-Better_Download-version_Final. Pdf.

Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art—teacher effectiveness and professional learning. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, *25*(2), 231–256

Murphy, P. K., Greene, J. A., Firetto, C. M., Li, M., Lobczowski, N. G., Duke, R. F., Wei, L. & Croninger, R. M. (2017). Exploring the influence of homogeneous versus heterogeneous grouping on students' text-based discussions and comprehension. *Contemporary Educational Psychology*, *51*, 336–355.

National Association of Educational Progress. (2015). *NAEP - 2015 mathematics & reading assessments.* Retrieved from https://www.nationsreportcard.gov/reading_math_2015/#?grade=4

National Center for Children in Poverty. (2019). *Child poverty*. Retrieved from http://www.nccp.org/topics/childpoverty.html

National Center for Education Statistics. (2018). *Average class size in public schools.* Retrieved from https://nces.ed.gov/surveys/ntps/tables/ntps1718_fltable06_t1s.asp

National Center for Education Statistics. (2020). *Children's Internet access at home*. Retrieved from https://nces.ed.gov/programs/coe/indicator_cch.asp#info

National Center for Education Statistics. (2020). *Revenues and expenditures for public elementary and secondary education: FY19*. Retrieved from https://nces.ed.gov/pubs2021/2021302.pdf

Nebraska Children. (2020). *Gothenburg - Dawson county community well-being needs assessment* [Brochure]. Schmeekle Research.

Nesmith, B. M. (2018). *Deciding on classroom composition: Factors related to principals' grouping practices*. Georgia Southern University.

Neuman, S. B. (2016). The danger of data-driven instruction. *Educational Leadership*, 74(3), 24–29.

No Child Left Behind Act, 107th U.S.C. § 2412 (2001). https://www.congress.gov/bill/107th-congress/house-bill/1

Northwest Evaluation Association (2020). *2020 Norms Research Study*. Retrieved from: https://teach.mapnwea.org/impl/normsResearchStudy.pdf

Oakes, J. (2005). *Keeping track: How schools structure inequality*. Yale University Press.

Oakes, J. (1990). *Multiplying inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and science*. RAND.

Oakes, J. (1986). Keeping track, Part 1: The policy and practice of curriculum inequality. *Phi Delta Kappan*, 68, 12–17.

Oakes, J., & Guiton, G. (1995). Matchmaking: The dynamics of high school tracking decisions. *American Educational Research Journal*, *32*(1), 3–33.

Opfer, V. D., Kaufman, J. H., & Thompson, L. E. (2016). Implementation of K–12 state standards for mathematics and English language arts and literacy. *Santa Monica, CA: RAND*. https://www.rand.org/pubs/research_reports/RR1529-1.html

Ornston, R., Spencer, L., Barnard, M., & Snape, D. (2014). The foundations of qualitative research. *Qualitative research practice: A guide for social science students and researchers,* 2nd ed. Sage.

Osher, D., Cantor, P., Berg, J., Steyer, L., & Rose, T. (2020). Drivers of human development: How relationships and context shape learning and development. *Applied Developmental Science*, *24*(1), 6–36.

Pace, K. (2014). Start where you are, but don't stay there. *Michigan State University Extension*. Retrieved December 23, 2021, from https://www.canr.msu.edu/news/start_where_you_are_but_dont_stay_there

Park, S., Hironaca, S., Carver, P., & Nordstrum, L. (2014). *Continuous improvement in education*. Retrieved August 18, 2020, from https://www.carnegiefoundation.org/wp-content/uploads/2014/09/carnegie-foundation_continuous-improvement_2013.05.pdf

Perez, K. (2019). *Differentiated reading instruction: Multiple pathways to literacy success.* Solution Tree Blog. Retrieved March 11, 2023, from https://www.solutiontree.com/blog/differentiated-reading-instruction/

Piaget, J. (1932). *The language and thought of the child* (2nd ed). Routledge & Kegan Paul.

Pressley, T. (2021). Factors contributing to teacher burnout during COVID-19. *Educational Researcher*, *50*(5), 325–327.

Radencich, M. C., & McKay, L. J. (1995). *Flexible grouping for literacy in the elementary grades*. Pearson.

Ravitch, D. (1973). *Inequality, by Christopher Jencks et al.* Commentary. Retrieved December 11, 2021, from https://www.commentary.org/articles/diane-ravitch-2/inequality-by-christopher-jencks-et-al/

Reardon, S. (2018). The widening academic achievement gap between the rich and the poor. *Inequality in the 21st century*, 1st ed. (pp. 177–189). Routledge. https://doi.org/10.4324/9780429499821-33

Reardon, S. F. (2016). School segregation and racial academic achievement gaps. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *2*(5), 34–57.

Reardon, S. F. (2011). The widening socioeconomic status achievement gap: New evidence and possible explanations. *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 91–116). Taylor & Francis.

Reynolds, A. D., Crea, T. M., Medina, J., Degnan, E., & McRoy, R. (2015). A mixed-methods case study of parent involvement in an urban high school serving minority students. *Urban Education*, *50*(6), 750–775.

Riley, T. (2016). The importance of learning with like-minded peers through flexible grouping in inclusive educational settings. *International Journal of Learner*

*Diversity & Identities*, *23*(4), 33–47. https://doi.org/10.18848/2327-0128/cgp/v23i04/33-47

Roberts. (2016). The "more capable peer": Approaches to collaborative learning in a mixed-ability classroom. *Changing English, 23*(1), 42–51. https://doi.org/10.1080/1358684X.2015.1133765

Rollins, S. P. (2014). *Acceleration: Jump-starting students who are behind*. Retrieved March 11, 2023, from http://www.ascd.org/publications/books/114026/chapters/Acceleration@_JumpStarting_Students_Who_Are_Behind.aspx

Rothstein, R. (2004). *Class and schools*. Teachers College, Columbia University.

Rowe, W. (2014). Positionality. In D. Coghlan & M. Brydon-Miller (Eds.), *The Sage encyclopedia of action research*. Sage.

Rubie-Davies, C. M. (2010). Teacher expectations and perceptions of student attributes: Is there a relationship? *British Journal of Educational Psychology*, 80(1), 121–135

Rytivaara, A. (n.d.). Flexible grouping as a means for classroom management in a heterogeneous classroom. *European Educational Research Journal*, 10(1), 118–128. https://doi-org.unk.idm.oclc.org/10.2304/eerj.2011.10.1.118

Samuels, S. A. (2020). *Fighting for fairness amid a pandemic.* Education Week. Retrieved from https://www.edweek.org/leadership/fighting-for-fairness-amid-a-pandemic/2020/09

Savin-Baden, Maggi and Claire Howell Major. (2013). *Qualitative research: The essential guide to theory and practice.* Routledge.

Scammacca, N., Fall, A.-M., Capin, P., Roberts, G., & Swanson, E. (2020). Examining factors affecting reading and math growth and achievement gaps in Grades 1–5: A cohort-sequential longitudinal approach. *Journal of Educational Psychology*, *112*(4), 718–734. https://doi-org.unk.idm.oclc.org/10.1037/edu0000400

Schwandt, T. A. & Gates, E. F. (2018). Case study methodology. In N. K. Denzin & Y.S. Lincoln (Eds.), *The Sage Handbook of Qualitative Research*, 5th ed. (pp. 341–358). Sage.

Scotland, J. (2012). Exploring the philosophical underpinnings of research: Relating ontology and epistemology to the methodology and methods of the scientific, interpretive, and critical research paradigms. *English Language Teaching*, *5*(9), 9–16.

Saleh, M., Lazonder, A. W., & De Jong, T. (2005). Effects of within-class ability grouping on social interaction, achievement, and motivation. *Instructional Science*, *33*(2), 105–119.

Shafer, S. (2020). *Overcoming COVID-19 learning loss.* Education Week. Retrieved March 10, 2021, from https://www-edweek-org.unk.idm.oclc.org/leadership/overcoming-covid-19-learning-loss/2020/08

Shannon, G. S., & Bylsma, P. (2002). *Addressing the achievement gap: A challenge for Washington state educators*. Office of Superintendent of Public Instruction

Sikes, P. (2004). Methodology, procedures and ethical concerns. *Doing Educational Research*, (pp. 15–33). Sage.

Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best evidence synthesis. *Review of Educational Research*, 57(3), 293–336.

Slavin, R. E. (1990). Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. *Review of Educational Research*, 60(3), 471–99.

Slavin, R. E. (1991). Synthesis of research of cooperative learning. *Educational Leadership*, *48*(5), 71–82.

Slavin, R. E. (2011). Instruction based on cooperative learning. *Handbook of research on learning and instruction*, 4th ed. (pp. 388–404). Taylor & Francis Group.

Snyder, T. S., de Brey, C., & Dillow, S. A. (2019). *Digest of Education Statistics, 2017*. U.S. Department of Education, National Center for Education Statistics.

Sparks, S.D. (2021). *Is the bottom falling out for readers who struggle most?* Education Week. Retrieved October 1, 2021: https://www.edweek.org/teaching-learning/is-the-bottom-falling-out-for-readers-who-struggle-the-most/2021/06

Thompson, T. (2021). *What does COVID-19 learning loss actually mean?* EdWeek. Retrieved November 20, 2021, from https://www.edweek.org/leadership/opinion-what-does-covid-learning-loss-actually-mean/2021/02

Tichnor-Wagner, A., Wachen, J., Cannata, M., & Cohen-Vogel, L. (2017). Continuous improvement in the public school context: Understanding how educators respond to plan–do–study–act cycles. *Journal of Educational Change*, *18*(4), 465–494.

Timperley, H. (2008). Teacher professional learning and development. *Educational practices serioes-18.* International Academy of Education & International Bureau of Education.

Timperley, H. S., & Phillips, G. (2003). Changing and sustaining teachers' expectations through professional development in literacy. *Teaching and Teacher Education*, *19*(6), 627–641.

TNTP (2018). *The opportunity myth*. The New Teacher Project. Retrieved from: https://opportunitymyth.tntp.org/

Tomlinson, C. A. (2018). Complex instruction: A model for reaching up—and out. *Gifted Child Today*, *41*(1), 7. https://doi.org/10.1177/1076217517735355

Tomlinson, C. A. (2015). Teaching for excellence in academically diverse classrooms. *Society, 52*(3), 203. https://doi.org/10.1007/s12115-015-9888-0

Tomlinson, C. A. (1999). *The differentiated classroom: Responding to the needs of all learners.* ASCD.

Tomlinson, C. A., & Imbeau, M. B. (2010). *Leading and managing a differentiated classroom*. ASCD.

Tomlinson, C. A., & Moon, T. R. (2013). *Assessment and student success in a differentiated classroom*. ASCD.

Torraco, R. J. (2016). Writing integrative literature reviews: Using the past and present to explore the future. *Human Resource Development Review*, *15*(4), 404–428.

Valli, L., Cooper, D., & Frankes, L. (1997). Chapter 5: Professional development schools and equity: A critical analysis of rhetoric and research. *Review of Research in Education*, *22*(1), 251–304. https://doi.org/10.3102/0091732X022001251

Vila-Henninger, L., Dupuy, C., Van Ingelgom, V., Caprioli, M., Teuber, F., Pennetreau, D., Bussi, M., & Le Gall, C. (2022). Abductive coding: Theory building and

qualitative (re) analysis. *Sociological Methods & Research*. https://doi.org/10.1177/00491241211067508.

Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard University Press.

Wallace, F., Blase, K., Fixsen, D., & Naoom, S. (2008). *Implementing the findings of research: bridging the gap between knowledge and practic*e. Editorial Projects In Education.

Wang, Haertel, G. D., & Walberg, H. J. (1990). What influences learning? a content analysis of review literature. *The Journal of Educational Research (Washington, D.C.), 84*(1), 30–43. https://doi.org/10.1080/00220671.1990.10885988

Waters, T. J., & Marzano, R. J. (2006). School district leadership that works: the effect of superintendent leadership on student achievement. A working paper. *https://eric.ed.gov/?id=ED494270*

Wenglinsky, H. (2000). *How teaching matters: Bringing the classroom back into discussions of teacher quality*. Retrieved from: http://files.eric.ed.gov/fulltext/ED447128.pdf

Wilcox, K. C., & Zuckerman, S. J. (2019). Building will and capacity for improvement in a rural research-practice partnership. *The Rural Educator, 40*(1), 73–90. http://dx.doi.org/10.35608/ruraled.v40i1.534

Wilcox, K. C., Lawson, H. A., & Angelis, J. I. (2017). *Innovation in odds-beating schools: Exemplars for getting better at getting better*. Rowman & Littlefield.

Woulfin, S. (2015). Highway to reform: The coupling of district reading policy and instructional practice. *Journal of Educational Change*, *16*(4), 535–557. https://doi-org.unk.idm.oclc.org/10.1007/s10833-015-9261-5

Yin, R. K. (2018). *Case study research and applications*. Sage.

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Sage.

Yin, R. K. (2002). *Applications of case study research* (pp. 22–28). Sage Publications.

Zuckerman, S. J. (2016). *Organizing for collective impact in a cradle-to-career network*. State University of New York at Albany.

Zuckerman, S. J., Wilcox, K. C., Durand, F. T., Lawson, H. A., & Schiller, K. S. (2018). Drivers for change: A study of distributed leadership and performance adaptation during policy innovation implementation. *Leadership and Policy in Schools*, *17*(4), 618–646.