Spring 3-18-2019

Stratification, Tracking and Course-Taking Patterns: An Examination of the Impact of Mathematics Course Placement on Achievement in a Regional High School District

Charles B. Sampson
charles.sampson@student.shu.edu

Stratification, Tracking, and Course-Taking Patterns:

An Examination of the Impact of Mathematics Course Placement

on Achievement in a Regional High School District

Charles B. Sampson

Seton Hall University

Dissertation Committee

Christopher Tienken, Ed.D., Mentor

Daniel Gutmore, Ph.D.

Nicole Hazel, Ed.D.

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Education

Department of Education
Seton Hall University
2019

SETON HALL UNIVERSITY
COLLEGE OF EDUCATION AND HUMAN SERVICES
OFFICE OF GRADUATE STUDIES

APPROVAL FOR SUCCESSFUL DEFENSE

Charles B. Sampson has successfully defended and made the required modifications to the text of the doctoral dissertation for the Ed.D. during this Spring Semester 2019.

DISSERTATIONCOMMITTEE
(please sign and date beside your name)

Mentor:
Dr. Christopher Tienken _____ 3-18-19

Committee Member:
Dr. Daniel Gutmore _____ 3-18-19

Committee Member:
Dr. Nicole Hazel _____ 3-18-19

The mentor and any other committee members who wish to review revisions will sign and date this document only when revisions have been completed. Please return this form to the Office of Graduate Studies, where it will be placed in the candidate's file and submit a copy with your final dissertation to be bound as page number two.

Acknowledgments

I am forever indebted to my mentor, Dr. Christopher Tienken, who persevered through long years with me to finalize this study. Dr. Tienken's guidance, unwavering support, and educational acumen were invaluable. His passion for public education and commitment to the students of New Jersey have enriched the professional careers of countless educators and I am proud to have crossed paths with him on this journey. Dr. Tienken went out of his way to join me as my mentor and this dissertation did not come to life until I began to work with him. Thank you, Chris!

Thank you to Dr. Daniel Gutmore who served as my reader. I benefited greatly from Dr. Gutmore's expertise both in this process and in my time in the classroom at Seton Hall. I am also deeply indebted to Dr. Nicole Hazel who always made time to assist me in this process and to provide feedback and a helping hand along the way.

This study is not possible without my involvement in the ongoing work at the Freehold Regional High School District. I am proud of our shared commitment to children and of the opportunity to work alongside so many terrific educators dedicated to equity, access, and most importantly, support for all students to thrive in the modern age.

A special note of appreciation to my wife Janet, and my four children, Maggie, Brian, Katie, and Cora. Thank you for allowing me to complete this process and for finding time on the weekends for me to finish this work after long weeks and many nights spent leading a school system. I have been blessed by my family.

Dedication

To my grandfather, Leonard Manney. My best friend. Thank you for showing me the gift of time and for always believing in me. You taught me every meaningful life lesson I ever learned. I miss you so deeply every day.

Abstract

This study examined the impact of mathematics track placement on the academic achievement of students as measured by student performance on the 2014 New Jersey High School Proficiency Assessment (HSPA) in mathematics, and math course attainment in high school as measured by the level of math class completed by the end of grade 12. A substantive empirical base (Goodlad & Oakes, 1998; Slavin, 1995; Oakes, 2005; Darling-Hammond, 2010; Gamoran 2009) exists that outlines the inequalities often associated within tracked school systems. Despite widespread criticism, tracking remains dominant as a strategy to group students in American high schools (Loveless 2013). Regional high schools accept students from multiple sending districts and must by nature make placement decisions based upon, in part, middle school performance. As the student data utilized in this study was culled from sending districts that all use traditional tracking strategies, the measurement of student performance within the larger regional high school may identify both effective and potentially problematic student grouping practices. The current debate in the field of education regarding the passage of standardized assessments as a requirement for graduation demands that the impact of course-taking patterns and tracking decisions on performance on such assessments be more fully investigated.

This study indicated that while ninth grade course placement matters as it relates to standardized assessment scores, explaining 27.8 % of the student performance and 17.9% of grade twelve course attainment, other factors may positively alter an individual student academic trajectory regardless of track placement. The evidence indicates that school districts that take specific action to mitigate the impacts of tracking may provide students with opportunities and support that lead to increased academic outcomes. This study contributes to the growing research that indicates the use of standardized assessments as a sole measure of academic

achievement is deeply flawed if not taking into consideration other factors such as course-taking patterns, academic opportunities for advancement, and tracking decisions at the local level.

Keywords: tracking, ability-grouping, deleveling, standardized assessments, achievement gap, neo-tracking, growth mindset

Table of Contents

List Of Tables

Figure

Chapter 1:

Introduction

Background

Historically, public school administrators in the United States have used a variety of structures to group students for instruction. In seeking methods toward efficiency, grouping processes used by school administrators have been influenced by results from standardized assessments of student achievement and intelligence (Archbald, 2012).

In some cases, student grouping practices are based upon the belief that homogeneous grouping, having students of like ability and aptitude assigned to the same learning space, results in a more meaningful and effective educational experience for students. Such grouping practices have been favored by educators because historically, they have been thought to provide a more efficient means to meet individual student needs within the classroom (Turney, 1931). The large-scale utilization of homogenous grouping in American public schools can sometimes lead to a grouping condition called academic tracking (Persell, 1977). Jennie Oakes, in her seminal work, *Keeping Track,* characterized academic tracking as, 'the process whereby students are divided into categories so that they can be assigned in groups to various kinds of classes. Sometimes students are classified as fast, average, or slow learners and placed into fast, average or slow classes on the basis of their scores on achievement or ability tests. Often teachers' estimates of what students have already learned or their potential for learning more determine how students are identified and placed. Sometimes students are classified according to what seems most appropriate for their future lives. Sometimes, but rarely in any genuine sense, students themselves choose to be in 'vocational,' 'general,' or 'academic' programs.' Tracking, however, differs from ability grouping in scope, duration, and rigidity as students are placed

within course sequences that tend to have a permanency as a result of these sorting practices (Oakes, 2005). Homogenous grouping for the purpose of course placement can also result in tracking.

Thus, there are several ways through which students may be placed into academic tracks. Academic tracking originated as a means to sort students into instructional groups based on perceived intellectual capacity and is further defined more broadly as the grouping of students for instructional purposes based on an assessment of academic ability (Rubin, 2006). Tracking, whether through programmatic or course-specific means, is often a highly static process as students placed within particular academic tracks tend to remain within those same tracks throughout their public schooling. As the modern comprehensive high schools emerged in the early twentieth century, tracking students for classroom placement by perceived ability emerged as the dominant means to sort students for instructional purposes (Oakes, 1985, 2005). Students were placed in specific academic course sequences within the comprehensive high schools that were believed to correlate in some manner to their future endeavors within the modern industrial society. The relative permanency of track placement within the modern comprehensive high schools have historically minimized student opportunities within school systems and often profoundly impact student exposure to more rigorous course loads, decrease access to more capable teachers and blunt overall academic achievement over time (Kalogrides & Loeb, 2009).

The practice of tracking has been the dominant means of placing students for instruction and has remained the most prominent method to sort students into homogenous groupings in modern high schools (Boaler, 2007). The practice of tracking has survived legal scrutiny and an abundance of research that indicates tracking disproportionality impacts poor and minority students. The process of employing tracking as a means of grouping students for instruction was

challenged legally under the Equal Protection Clause of the Fourteenth Amendment as being discriminatory. In *Hobson v Hansen* (1967) the practice of tracking as it applied to the Washington D.C. public schools was found to violate the Equal Protection Clause of the Fourteenth Amendment as it promoted intentional segregation along racial lines as the various recommendation processes utilized to sort students, including standardized assessments (in particular IQ tests), teacher recommendations, and general beliefs and expectations of student capacity, were found to work against minority and poor students (*Harvard Law Review*, April 1989). Though this case resulted in a highly localized decision regarding the reverberations of utilizing tracking as a means to place students for instruction, a more generalized decision has not been rendered in a larger context by the courts. Still, the decision drew parallels between racial and class discrimination and school-based tracking practices that ultimately stagnated long-term student opportunities and expectations for student achievement. Subsequent cases have deferred the process to use tracking to educators at the local level, who have frequently maintained various forms of such practices to group students for instruction.

The comprehensive high school developed as a potential solution to meet the needs of the rapidly expanding and simultaneously diversifying student body of the early twentieth century. As schools faced the sharp influx of recent immigrants coupled with the rapid expansion of the urban working class, educators employed strategies to sort students into course pathways that were believed to be most appropriate to fill the needs of the rapidly modernizing economy (Oakes 2005). The publication of the *Cardinal Principles of Secondary Education* (1918), ushered in a novel attempt to address the needs of a rapidly expanding student population enlarged with hundreds of thousands of recent immigrants following the turn of the century. The comprehensive high school was seen as a potential solution to the varied interests and abilities of

a changing student populace with wide ranging and increasingly divergent backgrounds through curriculum specialization for anticipated roles in the industrial society and national unification through a common schooling experience (Fletcher 2012). As the numbers of public school-aged students continued to increase dramatically following the dawn of the new century, public school leaders grappled with how to most effectively and efficiently group this ever expanding multitude of students for academic instruction (Boaler, 2005). The typical course sequence still relied heavily upon the recommendations of *The Committee of Ten* (1892) as questions of curriculum content, course sequencing, and student grouping remained.

The egalitarian purposes of the comprehensive high school, initially predicated upon providing more curricular options for students to choose specialized paths through high school while simultaneously offering a unifying core curricula based on the study of democracy and social forces, quickly became subsumed by the financial needs of school officials to promote efficiency in student grouping practices. Increasingly, the initial attempt to provide differentiated curricular experiences for students of varying backgrounds devolved into sorting practices based on precepts of financial efficiency. As a result, students were increasingly sorted into highly static instructional tracks based upon academic subject matter where they experienced vastly divergent educational experiences (Oakes, 2005; Archbald, 2012; Ascher, 1994). The sorting practices, often based as much on subjective as objective criteria such as beliefs of diminished academic capacity for students from lower socioeconomic and minority backgrounds, fixed views of human intelligence, and an over reliance upon standardized assessments as determining factors for course placement, came to be the dominant method through which students were grouped for instruction in American public schools (Gamoran and Berends 1987; Mulkey, et al., 2005; Boaler, 2000; Ansalone, 2010). As schools transitioned

from the one schoolhouse model to the modern comprehensive high school, the separation of students for instruction resulted in what came to be known as tracking, where students were placed in academic courses based on perceived ability (Oakes, 1987, 2005).

Although the American secondary school system promulgated democratizing ideals as a means of educating the masses, the spurious tenets of Social Darwinism heavily impacted student grouping decisions. Students believed to possess the intellectual capacity for post-secondary education were separated from students believed to best serve as workers for the modern industrial economy (Van Houtte & Stevens, 2009). This separation was further supported by the rise of intelligence tests as a means to provide some form of external measure to support the separation of students by "ability" (Oakes, 2005). The belief that students differed sharply in their capacity to thrive within the modernizing economy and thus needed different academic experiences has continued to profoundly shape modern educational grouping practices. The separation of students into ability groups that followed prescribed pathways within the industrial economy would lay the basis for the development of a rigid tracking system in American public high schools. This system of tracking has persisted for the better part of a century while more specific and transitional practices such as ability grouping and age-based grade assignment have remained the primary means of organizing students in American public schools (Rubin, 2008).

School systems that employ tracking assign students to classrooms or programs that are based partially on prior academic achievement (Kalogrides & Loeb, 2013). For the vast majority of the twentieth century, tracking was the administrative sorting and grouping of students into programmatic paths within the comprehensive high school based on some form of assessment of academic ability, i.e., standardized assessment, intelligence test, or teacher perception. In the

modern American high school this administrative practice, in many instances, has given way to a more nuanced method of tracking where students are grouped within particular subject areas rather than across entire academic programs (Rubin, 2006; Mulkey 2005). However, the shift from a rigid over-arching system of tracking in American public schools has resulted in a highly complex process of student placement that varies dramatically across districts. This process is highly locally contextualized as schools within the same districts differ dramatically in their sorting practices. Students are often subject to varied course-taking requirements that ultimately group students for instruction based upon a mixture of external and internal indicators determined largely at the local level (Kelly & Carbonero, 2012).

Ability Grouping

Ability grouping is the practice of organizing students into small groups for instruction. This within-class grouping is inherently different than a tracking system where students are placed into separate classes based on achievement (Loveless, 2009). This sorting or ability grouping may take the form of between-class grouping, mixed grouping, or within-class grouping (Mosteller et al., 1996). These processes are typically more fluid in comparison to the system of tracking which tends to be more deterministic for students within specific courses in the modern American high school (Lucas, 1999). As Deever (1995) notes, regardless if the term applied is tracking or ability grouping, the separation or designation of student placement by ability is so interchangeable in its impact to student grouping practices that they are each merely "two different ends to the same horse."

The process through which students have been grouped into tracks has changed over time. For most of the twentieth century traditional comprehensive high schools utilized a tiered structure that included three tracks: academic, general, and vocational. These tracks were largely

programmatic distinctions where students were sorted into specific course pathways based upon their perceived academic ability (Wheelock, 1992). The espoused goal of such practices was to provide curriculum differentiation to organize the curriculum in order to facilitate the learning of a diverse group of students with wide-ranging backgrounds while simultaneously managing their behavior (Hallinan & Kubitschek, 1999). In the 1970s, the rigid tracking system in American schools began to evolve into a more flexible and fluid system of ability grouping that was more discipline specific (Bernhardt 2014; Mickelson & Everett, 2008). More recently, the primary tracking system in American high schools has continued to evolve; it is more subject specific and separates students amongst "Advanced Placement," "honors" and "general" or "basic" courses (Hallinan, 1994). The modern incarnations of ability grouping are often determined through student choice; students are offered divergent curricula experiences with specific subject area differentiation. This more fluid and highly subject specific form of tracking is known as "neo-tracking" (Mickelson & Everett, 2008). Although programmatic tracks are less common, the process of separating students by external determinants within the public school system remains deeply entrenched in our modern schools (Oakes, 2005).

Sorting and Segregating

As early as the 1920s, the methods through which students were being grouped for instruction were being questioned for their effectiveness, particularly the practice of sorting students by ability into different courses and subsequently different academic tracks (Forman, 1975). Students were separated based on varying criteria including age, teacher assigned grades, perceived intelligence, scores on achievement tests, and general beliefs about ability (Miller & Otto, 1930). As these tracking processes expanded to assist in efficiently grouping the rapidly increasing student masses, the subjectivity of selection criteria came under renewed scrutiny

from researchers. The gap between student performance and course placement highlighted the often subjective nature of these processes as several studies noted wide-ranging discrepancies between student performances on achievement tests, teacher assigned grades and overall track placement (Miller & Otto, 1930). As early as 1930, several studies challenged the effectiveness of tracking and the sorting of students into various programs and courses based on perceived ability without also considering the need to alter instructional practices (Keliher, 1931). The different academic experience for students placed within the lowest tracks has been well chronicled in the literature as students placed within the lowest tracks were found to experience a dramatically truncated curriculum compared to students placed in higher tracks (Gamoran 2009, 1991).

Regardless of such challenges throughout the 1930s, tracking, both through ability grouping and programmatic distinction, remained the primary means utilized by school administrators to sort students for instructional purposes. With the advent of the Cold War following the close of World War II, the use of between-class ability grouping, where students were tracked into core courses by perceived ability, increased as a newfound emphasis on math and science swept across American schools following the launch of *Sputnik.* The process of grouping students for academic courses continued throughout the next several decades and was not critically examined in a widespread manner until the 1960s (Havighurst & Neugarten, 1967). During this era, criticism specifically identified tracking as a means of enforcing de facto segregation and reinforcing existing social inequalities (Esposito, 1973). Throughout the 1970s and 1980s, the form of tracking shifted as schools began to refrain from using the term tracking and student grouping practices came to focus on individual courses rather than entire programs (Lucas & Gamoran, 1991). Increasingly, such grouping practices came under intense criticism

as the way the use of tracking was determined to reinforce social inequalities, sort students into different groups based on status and prestige, and have negative psychological implications for those students placed in lower tracks (Oakes, 1985; Gamoran, 1987; Hallinan 1994). With the seminal publication of Jeanie Oakes *Keeping Track: How Schools Structure Inequality* in 1985, the fundamental student grouping mechanism utilized within the comprehensive American high school and larger public education system came under increasing criticism for its disproportionately negative impact to students from poor and minority backgrounds.

There is strong evidence tracking has had pernicious impact to students placed in lower tracks. Often, the curriculum is less challenging. Students are exposed to modified versions of the curriculum content, engage in fewer higher-order thinking activities, and experience redundant teaching practices focused on repetitive drills as opposed to subject area mastery (Oakes, 1985; Gamoran, 1992; Ansalone, 2010). Students placed within lower tracks are often identified in middle and sometimes elementary grades and struggle to move beyond these designations throughout the duration of their public schooling experience (Mulkey, 2005). Minority and disadvantaged students are also over-represented within lower tracks where such systems exist (Oakes, 1985; Ansalone, 2010; Rubin, 2006). The use of tracking and its deleterious impact on poor and minority students has continued to increase the segregation within public schools throughout the United States (Braddock & Dawkins 1993, Oakes 2005, Darling-Hammond 2010).

Statement of the Problem

The education reform movement predicated upon the use of results from standardized assessments as a core determinant of student and teacher proficiency continues to advance in New Jersey with the implementation of the *Common Core State Standards* and the

accompanying high stakes testing umbrella of the *Partnership for Assessment of Readiness for College and Careers* (PARCC). To gauge college and career readiness, all public school high students in New Jersey will be expected to pass a battery of end-of-year course assessments in specific content areas, including mathematics. The use of results from standardized assessments as both a course requirement and primary means of measuring student achievement in secondary mathematics may have profound implications for how students are grouped or tracked into specific academic courses. As school administrators are pushed by accountability measures to have students demonstrate proficiency on standardized assessments as the primary means of demonstrating student growth, performance on standardized assessments becomes a de facto guiding principle in determining student course mastery.

This allegiance to standardized assessments as the means to assesses student achievement was fully manifest during the *NCLB* era where high stakes testing resulted in the use of tracking as a rationale for the hierarchical sorting of students (Ascher, 1994; Burris, 2004; Rubin 2008). Students are often placed within lower academic tracks based upon their performance on these standardized assessments. Although potential changes to statewide accountability systems loom under the *Every Child Succeeds Act* (ESSA), all students in New Jersey are expected to meet the demands of the *Common Core* as evidenced by performance on end-of-course PARCC assessments. Grouping students into different academic tracks with differing academic expectations and outcomes may profoundly impact overall student achievement on these standardized assessments. Students in lower tracked courses are exposed to fragmented or slowed curriculum pacing, lowered expectations, stigmatizing labels, and in many instances, less experienced teachers, the difficulties in achieving proficiency on external assessments in the same manner as students in faster moving courses that may be exposed to more cognitively

demanding tasks and upon whom higher expectations and status labels are placed within the American schoolhouse, we run the risk of widening rather than closing existing achievement gaps in the near future (Gamoran, 2009). As these assessments will be used in increasing numbers in New Jersey to determine graduation eligibility, the persistence of tracking in our schools may lead to increasing numbers of students who are unable to meet graduation requirements.

Rubin (2006) explored the transformation of the purpose of tracking in American public schools from a potential means of providing individualized educational experiences to a modern structural reinforcement of the status quo. While a substantive empirical base (Goodlad & Oakes, 1998; Slavin, 1995; Oakes, 2005; Darling-Hammond, 2010; Gamoran 2009) exists that outlines the inequalities often associated within tracked school systems, limited evidence exists regarding the correlation between divergent mathematics track placement in middle school and high school achievement within a regional high school setting that receives students from multiple sending districts that employ different tracking strategies to group students for instruction. There is a lack of research examining the influence of tracking in the middle years upon the academic outcomes of students merging into a regional high school setting. As regional high school systems must intake students from multiple sending districts, the determinations of how students are sorted in the middle years and those impacts upon performance in the regional setting will provide insight into what measures might benefit the grouping of students from such diverse backgrounds while enrolled at the regional high school. With 31 regional high schools in New Jersey, an examination into how tracking in sending middle school systems impacts student outcomes at the regional high school and how actions within the regional system may or may not mitigate tracking impacts will profoundly benefit school leaders of those 31 districts.

Specifically, such an examination will provide insight into grouping impacts from school systems that utilize different versions of the traditional tracking strategy described by Oakes (2005) that group students into broadly tracked categories that recognize low, middle, and high achievers in middle school mathematics courses. Regional high schools accept students from multiple sending districts and must by nature make placement decisions based upon, in part, middle school performance. As the student data utilized in this study emanates from sending districts that all use such traditional tracking strategies, the measurement of student performance within the larger regional high school may illuminate both effective and potentially problematic student grouping practices. As students from all sending districts populate each of the six regional high schools in the study, the examination of the impact of tracking upon their performance allows for a comparison between similar students that attend separate middle schools but are exposed to the distinct sorting practices prior to entering the regional high school. Examining the academic progression through math attainment and overall performance on a state assessment may prove invaluable in unlocking the complex impact of tracking on student academic outcomes.

Purpose of the Study

My purpose for this non-experimental, correlational, explanatory quantitative study is to explain the influence of middle school mathematics track placement on the academic achievement of students in a regional high school district as measured by student performance on the NJ High School Proficiency Assessment (HSPA) in mathematics, and math course attainment in high school as measured by the level of math class completed by the end of Grade 12. The results from this study will help to explain the relationship between middle school course assignments and academic achievement within one regional high school context that will

assist in informing the course placement strategies of the 31 regional high school systems in New Jersey and other public education systems.

Research Questions

The overarching research question for this study is: How does mathematics course level placement in the middle school influence student academic achievement in a regional high school setting?

1. How is performance on the New Jersey High School Proficiency Assessment influenced by ninth grade mathematics course placement when controlling for student demographic factors?
2. How is mathematics course attainment by Grade 12 influenced by ninth grade mathematics placement when controlling for student demographics?

Null Hypotheses

The null hypotheses developed for this study were based upon the stated purpose of the study and the two research questions presented for research and analysis. The following hypotheses are presented for this study:

Null Hypothesis 1: No statistically significant relationship exists between ninth grade mathematics placement and student performance on the NJ HSPA when controlling for student demographics.

Null Hypothesis 2: No statistically significant relationship exists between twelfth grade mathematics placement and student mathematics placement in Grade 8 when controlling for student demographics

Design and Methodology

This correlational, explanatory design with quantitative methods utilized data from the graduating class of 2015 from a large suburban high school district in central New Jersey. This type of design was appropriate since I examined a number of variables and the relationship to a major complex variable and to what degree this relationship exists (Gay, Mills, & Airasian, 2012). This design will allow the researcher to predict the influence of the variables on the major complex variable.

The sample for this study consisted of students from the graduating class of 2015 of six suburban high schools in central New Jersey comprising a regional high school setting. The study utilized data from four sending districts into the regional high school system and excluded data from three sending districts as such data was not appropriately identifiable from the eighth grade year of the class of 2015. The six schools in this study are public schools, contain Grades 9–12 and report all student demographic information to the New Jersey Department of Education through their data warehousing system known as NJ SMART. The sample consisted of 1,232 students comprising four of the seven sending district data for the regional high school. Data from three of the sending districts was excluded because of the inability to confirm all accuracy of student placement and achievement in the middle grades prior to transition to the regional high school. While the study controlled for free and reduced lunch status, special education classification, and English Language Learner status, there are some socioeconomic and demographic differences between the four sending districts used for the purposes of this study. The 1,232 students entered into the regional high school setting of six high schools from seven different sending K–8 districts. These districts all used some form of tracking in mathematics

and had varied systems to indicate traditional tracking practices that constituted advanced, college preparatory, or below grade level designations.

Independent Variables

Each of the four sending districts to the regional high school used for this study utilizes a form of tracking that divides students into advanced, college preparatory, or below grade level course designations. These students then progress to the regional high school system. These categories of math track placement will be examined as predictor variables for student outcomes including HSPA performance and overall math attainment as the highest math course taken by Grade 12.

Dependent Variables

Until 2015, HSPA performance was used as a measure of school performance for New Jersey high schools. Students were required to receive a passing score on both the language arts and the mathematics section of the HSPA in order to graduate. Math attainment is measured as the highest level math course completed by the end of Grade 12. All factors are indicators of a student's academic achievement while in high school.

Conceptual Framework

Oakes (1985, 2005) explored the negative impact of utilizing tracking systems for student grouping in schools for students placed within lower tracks. These outcomes included a narrowing of the curriculum, increased likelihood to remain in lower tracks for the duration of schooling, and a widening of achievement and opportunity gaps over time. Oakes and Guiton (1995) explored the role of selective flexibility in tracking decisions where personal perceptions of student capacity and acceptance of tracking as a cultural norm often impacted teacher and administrator decision making when placing students in particular tracks. Watanabe (2006)

extended the work of Oakes to demonstrate that although the curriculum looked remarkably familiar in the same courses of different tracks, the classroom instruction and curriculum exposure varied significantly and substantially for students placed in lower tracks. Mulkey et al. (2005) also extended the Oakes' research to demonstrate that students in lower tracks become locked into their track placements as early as elementary school and experience a more rigid curriculum as they progress through high school. Gamoran (2009) demonstrated the curricular differences and more limited educational experiences of students placed within lower tracks. Kelly (2007), Bernhardt (2014), Ansalone (2009), and Boaler (2000) each examined the subjectivity of tracking practices and how that subjectivity negatively impacted student performance, teacher expectations for student success, and student expectations for their own success. Tracking was found to broaden student differences and result in a fundamentally different experience for students relegated to lower tracks throughout their schooling. Tracking has been found to be a highly localized construct that varies greatly, even from school to school within the same district (Archbald, 2012; Esposito, 1973; Oakes 2005).

The regional high school district selected for the study is located in the Northeastern section of the United States. The regional high school district is comprised of six high schools that receive students from eight surrounding communities. Each high school has one or more magnet programs that receive students from all eight sending communities, thus it is possible for students from all eight communities to be represented to varying degrees within each individual high school. Typically, however, the vast majority of students attend their locally designated high school. This study examines the course-taking patterns and outcomes on the mathematics portion of the New Jersey High School Proficiency Assessment (NJHSPA). The study consists

of the four-year mathematics course-taking patterns for 1,232 students enrolled in the regional high school as ninth graders in the fall of 2011.

Beginning in the 2011–2012 school year, the regional high school district undertook several measures to begin to dismantle the rigid tracking that had been a component of the mathematics course progression. While these processes took several years to implement, they are important to note as they might have positively impacted the student experiences represented in the study. Teacher recommendations were reviewed to identify students in non-advanced levels who would be strong candidates for success in more challenging math courses based on their grades. Standardized assessment data including "Advanced Placement Ready" data provided by PSAT scores was used to identify students in lower tracked mathematics courses who might be misplaced. Requirements for entrance into advanced and honors courses were relaxed to provide more mobility in courses. The regional system also began to offer a wide range of original credit core mathematics courses over the summer months that enabled students to progress through the continuum of mathematics courses beyond the traditional school day. Although the regional system in the study did not fully embark on a multi-year detracking initiative in mathematics until 2015, a similar movement was underway in English and social studies from 2011–2014 and the changes to leveling may have also impacted the mindsets for school administrators and other faculty when placement decisions were under review.

Significance of Study

The study will extend the existing body of research on tracking and will have a significant impact for policy makers as well as school and district leaders in several ways. First, no study exists examining mathematics tracking in the middle school and the corresponding impact upon academic achievement in a regional high school setting. The overwhelming

majority of middle schools employ a form of tracking in mathematics. Examining the impact of divergent sending districts tracking practices and the corresponding impact to academic achievement in a regional high school setting may provide guidance on how to effectively group students (or not) in middle school mathematics courses for academic success at the high school level. This study has the potential to explore the longitudinal achievement of students within the same high school systems that were tracked in differing ways in their middle school experience. The ability to gather longitudinal data from different middle school systems that feed into the same regional high school system offers the possibility of further synthesizing the convoluted research base that has struggled to differentiate achievement outcomes as a result of track placement or other factors such as teacher experience and curriculum pacing. Particularly in the era of high stakes accountability assessments and increasingly complex academic expectations through the *Common Core State Standards*, determining the most effective means to group students for instructional purposes will prove invaluable. The research on homogeneous and heterogeneous grouping clearly points to potentially problematic practices for many students.

However, these outcomes are hotly debated in the educational community and despite widespread criticism, tracking remains dominant as a strategy to group students in American high schools (Loveless 2013). This study will offer potential guideposts on how to mitigate these negative outcomes within an academic environment that has expectations measured by standardized assessments for all students regardless of course enrollment. Unwrapping how the complexities of mathematics course placement decisions unfold for students across a regional system will provide insight toward ensuring a consistent academic experience for students held to the same standards but placed in wildly different courses ostensibly engaging in the same curriculum.

Finally, the study will enrich the decision-making processes regarding mathematics placement throughout the 31 regional high school systems that currently exist within New Jersey. Many of these systems receive students from districts that employ varied student placement practices. Engaging in a meaningful examination of how students from divergent sending districts fair when placed within one large regional high school district will have generalizability for other similarly situated districts in New Jersey and the nation. In particular, the wide-ranging demographics and comprehensive nature of the six high schools included in the study, as well as the significant sample size, offer intriguing opportunities to inform the work of school district leaders across the spectrum all public school districts. The impact of tracking decisions from diverse school districts that send to one large regional high school district comprising six high schools, where each high school receives students from all sending districts, will provide a unique opportunity to examine how tracking practices and policies may impact the performance of students within the same peer groups as it relates to specific outcome variables, including state standardized assessment performance and overall math attainment. How tracking decisions in the middle school may affect students from different middle schools who attend the same regional high school provides an exceptional opportunity to examine how grouping decisions impact students from the same academic peer groups over time. The potential to examine within a school and across school differences for students who demonstrate the same academic aptitude on standardized assessments or have earned the same grades in similarly titled courses prior to high school but are placed within different academic tracks will provide a more nuanced way to examine the impact of tracking decisions than what presently exists within the literature. The impact of the placement decisions upon the long-term academic achievement of a large sample size of 1232 students could provide a deep understanding of student grouping procedures. The

research implications are rich for articulation practices as well as for a detailed examination of the longitudinal impacts of tracking decisions upon students who may demonstrate equitable academic achievement but have been placed within different academic tracks.

Limitations

As a correlational study, where a mutual relationship between two variables may be examined, it must be noted that correlation does not equal causation. As the variables examined are interacting with one another in a complex manner, the resulting outcomes examine relationships that may provide directional for further examination and for school responses, however, these findings are specific to the experience of the regional school system included in the study.

Delimitations

All data pertain to the class of 2015 from a large comprehensive regional high school district. The data is unique to the context of this school system. Data from other types of K–12 public systems, private, charter, or other schools are not represented. Additionally, the HSPA is no longer used in New Jersey high schools as PARCC subject area examinations are now used for state assessment purposes. As a high school system the work achieved by students within that system does not follow the typical K–12 progression; rather, the student achievement data represents a combination of the cumulative experiences of students within two separate school systems. The school district represents a cross section of demographics, however, is largely suburban in context. The wealth of the district, overall, is higher than the state norm. While schools within the system receive students from all eight sending municipalities, overall they represent the core student demographics of the towns in which they are located. As a result the differences in wealth across the divergent high schools may be significant. Any student data

from free and/or reduced lunch background, with a special education classification, and from English Language Learner designation has been excluded for the purposes of this study.

Definition of Terms

Ability grouping: flexible tracking system accomplished by academic ability grouping that is subject specific (Bernhardt 2014).

Common Core State Standards: Set of academic standards that include college and career readiness and academic expectations along a K–12 continuum that have been adopted by 43 states as of 2014.

Curriculum differentiation: A framework or philosophy for effective teaching that involves providing different students with different avenues to learning (often in the same classroom) in terms of acquiring content; processing, constructing, or making sense of ideas; and developing teaching materials and assessment measures so that all students within a classroom can learn effectively, regardless of differences in ability (Tomlinson, 2001).

Curriculum polarization: belief that students placed in lower tracks have decreased access to the richness of the full curriculum (Boaler, 2007).

Heterogeneous grouping: grouping students of divergent academic backgrounds in one classroom.

Homogeneous grouping: grouping of students of similar abilities in one classroom.

Neo-tracking: modern use of ability grouping rather than traditional programmatic tracking used predominantly in schools throughout the 1970s (Mickelson and Everett 2008).

New Jersey High School Proficiency Assessment: New Jersey state assessment administered during the junior year that measures achievement in reading, writing, and mathematics content. First-time eleventh graders are eligible to take the NJHSPA during the

month of March of their junior year. The HSPA was replaced by PARCC End of Year Assessments beginning in the 2014–15 school year (NJDOE, 2014a).

Partnership for Assessment of Readiness for College and Careers: Assessment consortium comprised of a number of states to develop a series of K–12 assessments mathematics and English Language Arts/Literacy, some of which are known as "End of Year" (EOY) assessments.

Tracking: the sorting and grouping of students for instruction, based on an assessment of academic ability (Rubin, 2006). Traditionally tracking has been associated with various class or program distinctions in schools such as advanced, honors, regular, and remedial (Oakes, 2005). Most recently tracking occurs in a much more course-specific manner within schools rather than the large-scale programmatic differences from the turn of the century. Tracking occurs as students are most commonly grouped for instruction based on some predetermined conception of intelligence or ability, e.g., intelligence as measured by a standardized assessment, teacher recommendation, past performance in school, or some other identified criteria. Such grouping has remained the predominant form of student grouping in American public schools to the present date (Darling-Hammond, 2010: Watanabe, 2004; Hallinan 1994).

Chapter II:

Review Of Literature

The purpose of this non-experimental, correlational, quantitative study was to explain the influence of middle school tracking practices and placement processes on the academic achievement of students in a regional high school setting. The author's emphasis was to examine how tracking practices in the middle school setting may impact student achievement within a regional high school setting. The literature has been reviewed for the purposes of tracking, and includes arguments both supporting and arguing against the use of tracking as a means of sorting students for instruction within the American public school system. The first section of the literature review assesses the origins of tracking and the uses and purposes of tracking systems in American public schools beginning with the turn of the twentieth century. Literature considered focuses on various political, social, and economic factors that cultivated a system of tracking in American schools while further examining the immediate and long-term outcomes related to tracking systems. The development of tracking systems across school systems is examined from both a contextual and a practical standpoint.

The second section of the literature review focuses upon the varying forms of tracking that have been, and remain, prevalent within the American public school system. The shifting forms that tracking has undertaken in American school systems over the course of the past century include broad programmatic tracks that sorted the college bound from the workforce bound to a modern form of tracking that centers upon class ability groups within specific subject areas.

The subsequent section of the literature review focuses upon studies examining student outcomes and attitudes within tracked school environments and the identification of academic

attributes related to tracking. This section critically examines the debate regarding the uses of tracking and encompasses recent literature from the past 25 years identifying positive and negative outcomes associated with tracking. This section also includes the overview of detracked environments for students and the differences in student outcomes along a variety of metrics when examining heterogeneous and homogenous grouping practices. The literature review further examines relevant research that provides a lens toward identifying the implications for students within each environment. The review of literature is comprised of the following areas: historical antecedents of tracking, studies examining outcomes and impacts related to tracking structures in schools, studies examining the sociological, economic, and political implications of tracking, and outcomes related to academic achievement for students placed within tracked environments.

Purpose of the Review

The purpose for the review was to identify studies that attempted to determine the significance of tracking practices on student outcomes within a regional school context. School specific variables of tracking procedures within the mathematics courses were examined and weighed against the performance of student outcomes as it relates to performance on standardized assessment and mathematics course progression across the student high school experience. The send/receive relationship of the high school regional and the distinct school systems that feed into the larger high school system offers an opportunity to critically examine the impact over time of tracking decisions that sort students who demonstrate comparable academic achievement into divergent academic pathways. For students with comparable achievement on external measurements, the impact of tracking as it relates to course placement may have profound implications for subsequent student sorting practices.

Literature Search Procedures

In order to thoroughly attend to the topics included in this study, searches were conducted to identify rich, relevant literature on each variable. The literature reviewed for this study came from a variety of texts, government reports, and academic articles obtained from EBSCOhost, ERIC, JSTOR, Sage, the United States Department of Education (USDOE) website, and the New Jersey Department of Education (NJDOE) website. General internet-based searches were also conducted utilizing Google Scholar.

Keywords used in the search for literature included tracking, ability grouping, academic achievement, state standards, differentiation, comprehensive high school, graduation rate, dropout rate, academic achievement, socioeconomic status, special education students, limited English proficient students, Black and Hispanic students.

Inclusion and Exclusion Criteria for Literature Review

Studies that met the following criteria were included in this review:

1. Involved public schools in the United States and to a minor extent, the United Kingdom.
2. Included a sample that consisted of Grades K–12 in a variety of combinations.
3. Used experimental, quasi-experimental, correlational, and meta-analysis designs.
4. Used peer-reviewed literature.
5. Dissertations from accredited universities but not online universities like Walden, Berry State, Capella, or Nova Southeastern.
6. Used quantitative and qualitative methodology.
7. Were published within the last 30 years unless considered seminal work that provided the beginning of later developments.

8. Included government reports.
9. Referenced federal and state legislation as background and contextual information.
10. Provided descriptive information that added clarity to the topic.

Tracking and the Competing Interest of American Public Schools

The American public education system has historically been answerable to two competing goals: the creation of a shared experience and personalized learning that meets the needs of unique learners. Schools are charged with instilling the democratic principles of the nation and providing individualized educational experiences that most appropriately prepare students for the divergent roles they may take within that larger American system (Gamoran, 2009). This dilemma, how to provide a diversified curriculum that reflects the universality of the democratic ideal, is reflected in the century-long discourse and debate on how students should be most effectively grouped for instruction. As our nation expanded from its agricultural background because of westward expansion and increasing urbanization, the shape and focus of the modern American public school system emerged. Buoyed by the educational reform movements spearheaded by Horace Mann and grounded in the Jeffersonian ideal of free public education that encapsulated the masses, school leaders turned their attention toward how to most effectively and efficiently group this increasingly wide range of students for instructional practice (Tanner and Tanner, 2007).

As the free and compulsory system of American education emerged following the turn of the twentieth century, the establishment of a system created a significant challenge for public school leaders to find methods and structures to effectively instruct students in the same courses who possessed widely differing backgrounds, knowledge, interests and capacity (Boaler, 2007). With the rapid urbanization of American cities fueled by immigration beginning in the late

nineteenth century, schools were viewed as the natural tool to democratize the new American masses. These new students needed to be prepared for productive lives within the modern industrial American society and school leaders looked to student grouping practices that would prepare students for the modern economy. The grouping practices that emerged from this quest for preparation for the industrial economy formed the basis for the large-scale sorting that placed students within academic tracks that approximated their perceived role within the industrial United States. Developing these tracking methods to effectively meet the disparate needs of the growing and divergent student groups fueled the early deployment and utilization of specific academic tracks. These tracks provided specific pathways through the American school system and were based upon broad programmatic differences—vocational, general education, and college preparatory (Archbald, 2012). These three broad pathways served the modern economy by creating a small class of students prepared for university life, whereas the majority of the remaining students were prepared for the workforce needed to sustain the industrial economy. Thus the use of academic tracking, "the process whereby students are divided into categories so that they can be assigned in groups to various kinds of classes" came to dominate the schooling practices throughout the early twentieth century (Oakes, 2005). Proponents of efficiency argued that the use of tracking served to tailor specific instruction to students of varying backgrounds, needs, and experiences (Turney, 1931; Kirkland, 1971; Ansalone, 2010). This structure was based on the premise that students should be separated by perceived intelligence or learning ability as the most effective means of delivering instruction (Ansalone, 2010). The use of tracking from a pedagogical standpoint assumed that students differ in abilities, prior knowledge, and in living environments conducive to learning (Lucas and Berends, 2002). The underlying rationale for the use of tracking was the assumption that it is easier to teach a group of

homogenously grouped students who are possessed of similar abilities and capacities. Furthermore, students must be appropriately prepared for their divergent futures within the American capitalistic system, which necessitates a fundamentally different academic experience. Thus tracking comports to future positions in society and the economy as students historically have been placed in courses or programs based upon expectations of their capacity to contribute to, and flourish within, the American workplace (Van Houtte and Stevens, 2009).

Many have posited that the use of a tracked system, where the most able students are sorted for instructional purposes, benefits high-achieving students (Argys, Rees, & Brewer, 1996, Loveless, 1999). These supporters also argue that the elimination of tracking stagnates the achievement of high-performing students within schools, pointing to the disparity between the lowest and highest quartile growth on the Scholastic Aptitude Test (SAT) against the growing use of heterogeneous grouping strategies (Loveless, 2009). Others have argued that the very premise of heterogeneous grouping flies in the face of student needs. As students arrive to school with divergent backgrounds, experiences, intellect, and expectations, to expect equal educational outcomes is simply not realistic nor within the capacity of the schoolhouse to produce (Benbow and Stanley, 1996). Other defenders of tracked systems believe a certain type of tracking must occur to ensure equitable educational outcomes. Nevi (1987) states, "Unless everyone is going to be taught everything simultaneously, tracking is necessary. To ignore different experiences, aptitudes, and capacities does a disservice to students." Once the curriculum is differentiated you cannot expect the same outcomes or use the same metrics to gauge student success. Yet this is exactly what has occurred during the *NCLB* era and now with required state assessments such as those developed by the multi-state consortia from PARCC and Smarter Balanced. Students in courses with the same title, with wide variance in the quality of

instruction, curriculum pace, teacher expectations, and student backgrounds, are all expected to achieve at the same levels on external standardized assessments. Students who are identified as potentially struggling on these assessments are grouped for remediation and targeted with highly specific and prescriptive instruction to raise their scores on said standardized assessments even though their fundamental experience with the course content is often strikingly different from higher achieving peers. The result is a problematic cycle where academic and experiential gaps are widened, particularly for students placed within lower academic tracks through an attempt to close performance gaps on standardized assessments. Student growth is thus often ignored in a flailing attempt to have students reach a specific cut score on a standardized assessment for accountability purposes (Tienken & Orlich, 2013).

In tracked systems, students are most commonly grouped for course placement for instruction based on some predetermined conception of intelligence or ability, e.g., intelligence as measured by a standardized assessment, teacher recommendation, past performance in school, or some other identified criteria. Such grouping has remained the predominant form of student grouping in American public schools to the present date (Darling-Hammond, 2010; Watanabe, 2004; Hallinan 1994). The process of sorting students in such a manner has historically been highly deterministic, as the tracks through which students are placed tend to be stagnant. Students more often remain in the tracks they are assigned to, sometimes as early as elementary school, for the duration of their schooling (Wheelock, 1992). As Oakes (2005) notes, tracking has both changed and endured with the primary shift in focus being that the responsibility of choice for a track has shifted from the school to the students. Regardless, tracking has remained the primary means to group students for instruction in American public schools (Slavin, 1993, Oakes, 2005).

Various attempts have been undertaken to respond to the competing interests of this dynamic free and compulsory education system. *The Committee of Ten* (1892) was formed with the express purpose of codifying what learning was of most use to the increasing numbers of students being educated as a result of the wave of immigration in the late nineteenth century. Debates over curriculum, course sequencing, and the purpose of secondary education were the focus of the Committee's work. These debates continue today. Ultimately, the *Committee of Ten*, chaired by Harvard University President, Charles Eliot, forged a pathway that was primarily geared toward students who were college bound. Importantly, the Committee saw this college preparatory curriculum as valuable for all students regardless of socioeconomic status, racial composite, or family background. The Committee laid the foundation for subsequent steps toward an education system that prepared all students for meaningful contributions toward work and life through access to a rich curriculum. Despite these steps by the *Committee of Ten*, the structure that emerged within high schools made it difficult for non-college bound students to aspire toward rich academic learning as the pathways quickly became highly stratified through tracking, as school officials sought to efficiently deliver the college curriculum to students who were perceived to represent different academic capabilities than their workforce bound peers (Tanner and Tanner, 2007).

While tracking has remained the dominant means to sort students for instruction, the rigid tracking systems that dominated much of the twentieth century have given way to more flexible tracking systems that are presently more course specific, predicated (ostensibly) on student choice. As Oakes (2005) notes, tracking has both endured and transformed over time. Now, the responsibility of choice for a track has shifted from the school to the students. Still, the majority of American schools use some sort of determinant regarding student perceived ability and

intelligence to sort students into clear groupings for instructional purposes (Darling-Hammond, 2010; Watanabe, 2006; Hallinan, 1994). The premise that students differ in academic abilities and therefore instruction must be provided at differing levels is so manifest in the American psyche that rankings such as those published in *Newsweek* sort and stratify schools based upon the number or percentage of students placed within the highest track courses—Advanced Placement or International Baccalaureate (Matthews, 2013).

The debate between efficiency and democracy was further addressed with the publication of the *Cardinal Principles* in 1918. The focus of the *Cardinal Principles* was an attempt to outline a curriculum that included "something for everyone" and was firmly aligned to the Dewian notion of student-centered learning, where the interests and experiences of the students determine the curriculum. The comprehensive high school was seen as a tool that could both serve and further the dual interests of unification and specialization, where all students would receive a democratizing academic experience within a comprehensive high school rich in academic opportunities. In direct contrast to European elitism, the children of America "would grow, learn and play, together" (Noddings, 2015). The *Cardinal Principles* were conceived as a means of allowing students choice in the selection of their coursework and areas of study. The *Cardinal Principles* tried to create curriculum with something for everyone and indicated that the comprehensive high school would fulfill the need for unification and specialization (Oakes, 2005).

However, the system within the American comprehensive high school designed to provide students an egalitarian opportunity in rich academic coursework devolved over time as school administrators, in seeking convenient means to group students from wildly divergent backgrounds, turned to the practice of tracking where students were placed in academic courses

based on perceived ability (Oakes, 1987). The lofty goals espoused within the *Cardinal Principles* were quickly subjugated to the need for efficiency as students were separated into tracks based upon their perceived academic potential and status. Over time these practices focused increasingly on separating students into instructional groups based on academic subject matter (Oakes, 2005; Archbald, 2012; Ascher, 1994). The concept of tracking in the late 19th and early 20th century focused largely on the belief that intelligence was static. Students were possessed of an intelligence that was fixed and could be measured. Therefore, students should be placed within academic, vocational, or college preparatory tracks in direct correspondence to their intelligence, which was measurable. The intelligence tests popular during the era of World War I were significant determining factors during this time and were the primary criteria for student grouping decisions into the 1930s (Mulkey et al., 2005, Oakes, 1992).

Multiple researchers have attempted to measure the value and impact of tracking. As early as 1931, Turney discussed the advantages and disadvantages of placing students in designated courses by perceived ability. He noted several positive attributes of tracking students, including the benefit of adapting instruction to more effectively meet student needs, tailored instruction, and the opportunity for teachers to work more closely with smaller groups of struggling learners. Conversely, he noted that disadvantages included a stigmatizing effect to slower learners, learning losses as a result of the removal of more academically advanced students from lower tracks, and the inability of teachers to effectively differentiate and meet the diverse student needs within their classrooms (Turney, 1931). These arguments, so clearly set forth in the 1930s, continue to reverberate in the modern era. Such arguments have comprised the bulk of the discussion on how to most effectively group students for instruction in American public schools. Both proponents and critics of tracking have attempted to understand the full

impact of instructional grouping practices for almost a century. Arguments in support of tracking center on differentiation: Students within tracked environments are taught at an academic level and pace that is commiserate with their specific academic aptitude and capability. Advanced students benefit from such environments as they progress through the curriculum at a more rapid pace and work within an atmosphere of equally able and similarly placed college-bound peers (Loveless, 1999). These early studies tended to focus explicitly on student outcomes from tracked environments and on the procedures for sorting students (Slavin 1990). The underlying premises of tracking, a fundamental belief in divergent student capacity, was not significantly challenged within the literature (Turney, 1931).

A clear case for a more specific and individualized curriculum without tracking was evident with the publication of *The Eight Year Study* (1942). *The Eight Year Study* proved that when given the freedom from traditional constraints on the core curriculum, and where an emphasis was placed on conceptual thinking, real world applicability, and individualized programs, students could perform admirably at the secondary and post-secondary level. This study had the potential to dramatically alter the prescriptive and highly stratified curricular approach and corresponding sorting practices that had dominated public education since the turn of the century. Instead, the outcomes were lost in the haze of WWII and soon sacrificed upon the altar of the Cold War; in an effort to accelerate instruction for the science and mathematics fields, our school systems turned back to highly mechanistic grouping practices designed to sort and stratify students based upon perceived abilities (Tanner and Tanner, 2007).

Rise of Tracking in Public Education

The advent of the Cold War ushered in an era in America that saw a sharp justification for the use of tracking in its public schools. Following the launch of *Sputnik* by the Soviet Union

in 1957, tremendous pressure was placed upon school officials to ensure there was ample opportunity for advancement of "gifted" students, particularly in the mathematics and sciences. The work of James Conant and Jane Bruner proved significant in ushering in a new era of explicit tracking. The publication of James Conant's report, *The American High School Today*, firmly re-established tracking in academic subjects as an effective means to educate all and particularly as a means to provide expanded and enriched opportunities for the very highest achievers (Conant 1959). Although Conant envisioned the democratizing principles of a comprehensive high school for students from all walks of life that furthered democratic and egalitarian principles, he recommended the separation of students within academic tracks to provide the education most suitable to all manner of learners. Conant argued forcibly that the comprehensive American public high school must correspond to the needs of all students; however, he clearly identified divergent tracks as a means of most effectively educating those students thought to be intellectually gifted. This argument would be utilized across American schools to develop clear sorting practices that afforded a basic academic education for all while ensuring a separation of those students deemed academically more capable than their peers. Conant's recommendations, though well intended, were a product of their time and ignored underlying societal forces, expectations, and beliefs that resulted in biased sorting practices in schools for poor and minority students. Conant's work reaffirmed the commitments laid out in the *Cardinal Principles*. However, the perception of school leaders continued to distinguish between those students who were deemed capable of receiving the most challenging academic coursework and experiences and those students more suited for other experiences within the comprehensive high school. *The American High School Today* became the bible by which modern high schools were measured against for decades and firmly re-established tracking as the

dominant method through which students were sorted into courses of study in American schools (Wragge, 1994)

In 1960, Jerome Bruner published *The Process of Education.* Bruner strongly supported the use of a more highly academic curriculum amongst the core disciplines, particularly science in the wake of *Sputnik*. Bruner's theories related to student intellectual development, notably the ability of students to understand and relate to complex content within core disciplines in a highly intellectual manner. Bruner criticized the American public school system for its inability to meet the needs of its most gifted learners and argued that the institution, particularly in the sciences, may be limiting the development of many students. Bruner's theories on learning stood, at the time, in stark contrast to the Dewian notion of student experiential learning and instead recommended heavy content by discipline to fill each course requirement so that students most fully developed intellectually. The Bruner Report ultimately fueled the calls to more aggressively sort students within tracks, particularly in the mathematics and sciences. The culling of the students deemed the most intellectually gifted according to varying local criteria came to be standard process in most American high schools. The expansion of more advanced course opportunities for those students deemed gifted rapidly emerged. At face value, such developments seemed in the best interest of pursuing more rigorous experiences to identify and nurture exceptional students to keep pace during the height of the Cold War. However, Conant and Bruner ultimately ushered into public education a more comprehensive fixed view of intelligence that echoed student placement decisions based on intelligence tests and other external measures of a student's intellectual capacity for placement. The call for a rigid separation of students within academic tracks remained largely uninformed and ignored flawed

processes for selection, individual backgrounds impact, and issues of wealth and prejudice that so often marred such placement decisions.

These seminal works served as a basis to further stratify students in public schools and helped to fuel a growing separation in course opportunities between poor and minority students. These arguments furthered the course and program segregation of schools. These processes would begin to be challenged during the Civil Rights Movement, particularly in the literature during the latter part of the 1960s.

Tracking as a Lens to Societal Inequalities

During the 1960s additional research related to tracking practices began to draw links between race, cultural norms, economic standing, and student grouping practices in schools. Studies began to scrutinize how underlying external and subjective factors related to race, class, and individual perceptions impacted tracking decisions within schools. The pioneering work of Cicourel and Kitsuse in the Educational Decision Makers (1963) found that track placements were not simply derived from students' curriculum choices but were influenced also by counselors' personality assessments. Thus, the link between societal perceptions and larger cultural norms impacting student track placements was established. Researchers also began to explore the impact of subjective factors on track placement. As the unrest of the Civil Rights Movement permeated the larger American culture, the processes that resulted in segregated academic experiences for students came under closer scrutiny, particularly the factors that shaped adult and student perceptions related to student placement and how those factors worked against large groups of students. The belief that students were placed into tracks by objective measures of ability alone was challenged, and ultimately, the emerging understanding of the complexity of tracking practices became more pronounced. Still, the research base had little if

any impact to changes in the grouping practices for instructional purposes employed by school administrators.

As the 1960s waned, the stark differences in academic performance and opportunities within low income and minority schools as well as the sorting practices for poor and minority children came under increasing fire. The report from the US Commission on Civil Rights in 1967 admonished public officials and noted academic track placements would likely affect student outcomes related to educational attainment. The report also explained that the heavily segregated school systems across the nation resulted in vast differences in teacher capacity. Poorer schools contained greater numbers of teachers lacking a major in the subject area they instructed, while teacher attitudes reflected a clear and pressing desire to work in more affluent areas as social perceptions related to class impacted teacher views of student motivation and capacity. Perhaps most strikingly, the Commission's report set the stage for more expansive research that closely examined the academic experience of lower tracked students, including more specific examinations utilizing survey data on how tracking impacted student and teacher expectations and beliefs regarding the use of such processes (Persell, 1976). This research linked student perception and racial and socio-economic isolation as critical components negatively impacting the long-term outcomes related to student achievement for lower tracked students. Heather (1969) found that teachers in lower tracks explicitly stated they used drill and basic facts while utilizing conceptual thinking with higher tracked students. These determinations were made on the basis of individual teacher perceptions of the academic and intellectual capacity of the students they were responsible for instructing.

Throughout the late 1960s and the decades following, there was increased research of teacher-student relationships as it related to student track placement. Teachers were found to

generally possess better relationships with students in higher tracks than in more remedial tracks, while students expressed better relationships with higher track peers than lower track peers (Hargreaves, 1967; Heynes, 1974). Hargreaves' work confirmed the findings of the US Commission on Civil Rights which identified that novice teachers were more likely to be assigned to lower track courses rather than experienced teachers. The National Education Association (NEA, 1968) found that teachers valued working with lower tracked classes less than with higher tracked classes and were comprised more often of the least experienced teachers. An experiential gap further eroded the learning opportunities for lower tracked students while reinforcing teacher perceptions related to race and class as minority and poor students typically comprised the majority of low track courses. In Persell (1976) Kariger found that when examining three junior high schools that there would have been greater student diversity if scores on standardized tests alone served as the determinant to course track placement (Data reported in Persell, 1976: 43). Clearly other factors aside from standardized test performance resulted in student placement to lower and higher tracked courses. The resulting impact to instruction and teaching and learning has been a constant theme in the literature related to tracking over the past forty years (Gamoran, 2009).

During the 1970s, research related to potentially pernicious impacts of tracking practices on student subgroups emerged in earnest. Breaking from the stagnant view espoused by Conant's work, an important shift in the literature occurred as researchers began to argue that the negative consequences of relegating students to lower tracks might have profound long-term impacts that far outweigh the perceived benefits for the highest achieving students. Findley and Bryan (1970) collected and analyzed research regarding tracking practices from 1920s—1970s and found that IQ scores and other standardized assessments played a significant role in course

placement practices within schools. They found that negative consequences impacted far more students within tracked environments. Findley and Bryan argued that ability grouping produced at best mixed results for higher groups (some studies show yes) and clearly negatively impacted lower and general groups, which were comprised of disproportionate poor and minority students. Thus the groundwork for the discussion around the difficulty of discerning tracking impacts was becoming more prominent as the nation emerged from the Civil Rights Movement. Though the larger educational community was not yet embracing the ramifications of tracking and its impact on the vast majority of students, Findley and Bryan pointed out that the research suggested strongly that assignment to lower courses and groups generally carries a stigma that proves more debilitating than relatively poor achievement in heterogeneous groups. The impact of such stigma and the indication that only the very highest achieving groups may benefit from tracking practices at the expense of the vast majority of students ran counter to the work espoused by Conant and Bruner. The literature of the 1970s challenged the existing structures so wildly popular throughout American schools. This important shift helped to pave the way for subsequent research throughout the 1970s that built the basis for the pioneering work of Jeannie Oakes in the 1980s.

In his synthesis of existing research related to tracking, Esposito (1973) began to draw the linkage between the complexity of localized tracking decisions and the larger social, economic, and political forces that often shaped how school districts grouped students for instruction. Esposito noted that the critical examination should center upon which specific student grouping plan results in better conditions for teaching and learning. Such a question, Esposito argued, was almost impossible to answer for a host of reasons. These included difficulties of comparing grouping practices, divergent durations of student groupings, and other

school and district specific factors that contributed to the final decisions surrounding how and when students were grouped for instruction. Esposito argued forcibly that convenience and efficiency should not trump student needs and experiences when determining how to group students for instruction. He concluded, "If one of the principle objectives of the American education system is to provide each child with an equal educational opportunity to maximize and develop his potential so that he may benefit himself and thereby contribute more effectively to the larger society, then the present status and predicted trends with respect to homogenous grouping suggests that this cardinal objective will *not* be realized" (p 177).

Esposito's work provided a framework for subsequent researchers to explore tracking with the understanding that it is difficult to draw specific long-range conclusions when examining the research. Calling for a more nuanced examination into the social and emotional aspects associated with tracking, Esposito lamented the disjointed research on tracking that focused almost universally on specific and limited academic outcomes. Esposito, in reviewing the work of Findley and Bryan and other research on tracking, found that homogeneous grouping itself tends to reflect social, political, and economic differences of already marginalized groups. Esposito hypothesized that the instructional practices may be wildly divergent across tracks and that it is critical to subsequently examine how grouping patterns impact classroom environment and quality of instruction, divergent teacher expectations of students across tracks, and differences amongst the types of student-teacher interactions these groupings bring to the fore. Esposito concluded that tracking practices sanction substandard education for a large component of children. This would be a significant focus for subsequent research regarding tracking.

Academic Outcomes and Societal Inequalities

Researchers began to examine the potential post-secondary outcomes related to tracking in earnest throughout the 1970s. In re-examining the data from the Coleman Report (1966) that covered 48 urban high schools and over 30,000 students, Heynes (1974) found that the curriculum placement for students was largely explained by aptitude on standardized tests; specifically, how well students performed on verbal components of standardized assessments had a significant impact on curriculum placement decisions in secondary schools. Heynes found that student stratification occurred largely through administrative decisions. Additionally, students were often counseled into lower tracked courses in subsequent years of being stratified based upon performance upon standardized assessments. The work was important in describing the formation of an academic hierarchy based largely on "within school" variables predicated upon decisions and perceptions of the adults within the schools. These adult-based decisions on grouping practices often thwarted the meritocratic principles of the American comprehensive high school.

Throughout the decade of the 1980s, it had been well established in the literature that tracking resulted in varied benefits for the very highest achievers, while uniformly negatively impacting middle and low achievers (Froman, 1981). The 1970s had witnessed a marked increase in the examination of the academic and social impacts on students (Esposito 1973, Heynes 1974, Findley 1975). In examining the literature on tracking and grouping students by perceived ability, Froman concluded in 1981, "As a group, the qualitative discussions of ability grouping, its logic and documented effects are surprisingly severe." Discussions or reviews advocating its use with the majority of students are rare. Such literature is virtually nonexistent in the general educational journals published in the past ten years. Given this comment by

omission, it is surprising that ability grouping continues to be practiced on such a broad scale in the United States. Increasingly the alignment of social perceptions and school officials' decision-making processes on how to sort students for instruction emerged as a critical point of examination in the literature. Researchers began to examine the hypothesis that tracking decisions and practices might be based on preconceived notions of aptitude rather than intellectual capacity. Rosenbaum (1980) examined longitudinal data from the Educational Testing Service (ETS) involving 16,683 seniors from over one thousand high schools from the class of 1972 and found that students rarely recognized the potential debilitating effects their track placement might have on their college aspirations, and absent such knowledge, did not foresee how their course selection and placement stifled their own long term post-secondary plans. The National Longitudinal Study (NCES) used by Rosenbaum was critical to examining how tracking placements impacted students' college plans, as the data included both student perceptions of the track they were placed within as well as the specific school indicator of the track. This allowed Rosenbaum to compare perception to reality and draw conclusions of how tracking decisions and placement may impede collegiate plans. Rosenbaum found that late into their senior year students had unrealistic expectations related to their college plans based on their track placement.

Rosenbaum's examination also pointed out that tracking decisions mirrored social perceptions of student ability and capacity. Rosenbaum questioned whether aptitude was important because it indicated some sort of real ability or because perceived aptitude mirrored a particular social classification that staff in schools responded to, often resulting in the perpetuation of societal divisions. Tracking was ideally supposed to place students in the most appropriate academic course, however, this did not explain the sharp distinction between college

plans and track placement. Rosenbaum speculated that students may have accepted track placements blindly because they did not realize the full implications such placements would have on their ultimate college plans. Rather than repudiate perceptions of their ability, students failed to realize the long-term impact sorting decisions had upon their college aspirations. Rosenbaum found that the complex relationship between perceived aptitude and the impact of subjectively viewing student capacity through the lens of track assignment race and socio-economic status all seemed to indicate that track placement was so intertwined with social classification that true assessment of student capacity was continually clouded by the subjective and complexity of the tracking systems in schools. Rosenbaum's work paved the way for the seminal works of Oakes (1985) and Gamoran (1989), who began to examine how tracking resulted in fundamentally divergent academic experiences for students within the same schools.

The research related to tracking throughout the era prior to the publication of *A Nation at Risk* (1983) begged the question: is funding the advantaged or disadvantaged of greater public use? As the debate over tracking would begin to intensify in earnest in the 1980s, it was apparent that although tracking remained the primary method through which students were grouped within schools, particularly secondary schools, differential tracking in secondary schools introduced academic inequalities that did not previously exist, contributing to increased societal inequalities. Increasingly the arguments that began to form between those who favored tracking for efficiency purposes as well as for the purpose of isolating the perceived gifted from other students from those individuals who advocated for a heterogeneous mixing of students began to take on a more heated debate (Oakes, 2005; Loveless, 1999).

As the tracking debate intensified in the literature throughout the decade of the 1980s, two major questions of empirical research emerged: Does tracking work and what factors

contribute to decisions to place students within particular tracks? Oakes (1985, 1987) began to look at how tracking practices themselves affected schooling through the development of strata representative of societal inequalities. Examining tracking practices within schools against larger societal inequalities became a fixture of the research throughout the decade.

The compound impact of being placed in lower tracks, starting in the elementary grades, also served to diminish student post-secondary accomplishments. For many students placed in lower tracks, the slower pace of the curriculum often resulted in a sharply divergent introduction of critical concepts that ultimately increased the likelihood that students within these lower tracks would struggle to gain entrance into the courses needed for college acceptance (Ball 1981). Indeed**,** so many of the influences of senior high school are established long before high school. In particular, elementary grouping patterns may play a critical role in high school course attainment. In the 1980s researchers began to examine specific curriculum differences rather than outcomes of tracking practices. The compound impact of elementary grouping patterns may contribute to teacher perceptions and administrative decision making in placing students in lower tracks. As it is the academic record rather than social background that had the most significant impact on track placement, the contributing factor of multi-year differences in course material almost certainly impacts teacher perceptions and administrative decisions to place students within certain tracks. Jencks and Brown (1975) had concluded that track placement was not as relevant as it relates to student achievement as it merely reflects existing differences in test performance. However, this simplistic argument ignores how those test differences are impacted by marginal teaching, slower curriculum pace, reduced access to high-achieving peers with specific goals for college and preconceived notions of ability from teachers and administrators. The work of Alexander, Cook, and McDill and Rosenbaum paved the way for more expansive

examination of long-term impacts of tracking as well as the instructional variance experienced by students placed in lower tracks that rapidly expanded during the decade of the 1980s leading to a sharper and more focused battle over grouping practices in schools (Froman, 1981; Oakes, 1985; Gamoran 1989).

The work of Jennie Oakes began in earnest in the 1980s and has continued to the present. This work has examined the issue of tracking through an equity lens focusing upon the interconnectedness of potential constitutional, societal, and academic concerns exasperated by the continued use of tracking as a means of sorting students in American schools. Oakes (1981) challenged the constitutionality of a system where resources are deployed in a manner that further intensifies the segregation of schools. As tracking ultimately mirrored societal discrepancies in wealth and status along racial and socio-economic lines, Oakes challenged the use of tracking as a specific barrier to equal opportunity for an appropriate education. Students in the highest tracked classes were predominantly non-minority and non-poor students and had access to the knowledge most valued by society in a fundamentally different manner than lower tracked students who were disproportionality non-white and poor. Therefore, Oakes argued that the system was unconstitutional as it naturally inhibited student development along racial and socio-economic lines. Such systems, Oakes argued, determine the type and quality of an education received and further the marginalization of already marginalized groups. These systems also operated under a belief that aptitude would not likely to be altered by different experiences. This static or fixed mindset was a fundamental cornerstone of the premise of tracking. Oakes (1981) noted that studies of tracking continually find a high correlation between race, SES, and status and track level. Ultimately, Oakes' work would dramatically expand the

previous research on tracking that was largely focused upon academic barometers and begin to build specific linkages to larger societal ramifications of tracking.

With the publication of *Keeping Track* in 1985, Oakes ushered in a new era of challenging the most traditional of all grouping practices in American public schools. Oakes expanded upon the earlier examinations of tracking that focused on academic impacts and often ignored the status and prestige issues associated with a system predicated upon dividing resources along student groupings. Expanding upon work that had suggested tracking and the resulting differences in instruction that followed such stratification denied student equal educational opportunity (Fenstermacher, 1983), Oakes identified three barriers to changing tracking practices; they included political, normative, and technical challenges. Oakes examined 25 secondary schools from across the nation that had participated in Goodlad's *A Study of Schooling (*Oakes, 1985, 2005). Oakes updated her work in 2005 in order to demonstrate the continued use of tracking practices in public schools across the United States. The schools examined represented a cross section of district types: poor, wealthy, suburban, urban, and rural and had wide-ranging student populations. Oakes surmised that many of the findings had strong generalizability to the larger American education system. The data was based on a six-week period in 1977 where researchers spent considerable time conducting interviews of principals, students, teachers, and counselors and examining curricular documents from these schools. Classes were also observed regularly as a component to the data collected. Ultimately, over 1000 classes were used to cull specific data related to the tracking practices of these schools. The examination highlighted the ubiquitous use of tracking, as all 25 schools contained a process to track, though the degree to which those practices were public and understood within the school community varied tremendously.

The unintended consequences of grouping practices were also noted by the researchers as students often became tracked unintentionally in courses initially not designed to be grouped in such manners as a result of the tracking practices for other specific courses. These unintended consequences served to further stratify the students within the schools examined; often to a degree not fully understood nor intended by school officials. Such unintended consequences are a critical consideration when looking at how schools engage in tracking. These findings suggest that within each highly localized context, any student grouping process may have significantly broader consequences than what may be readily apparent to school officials and school community members, including parents and students. Detailed analysis of how students move through course sequences as a result of tracking practices would need to be undertaken at each school that employs such practices to fully understand the impact that tracking may have on the students within a given school.

Oakes (1985) also found that teachers and counselors shared the primary decision-making power in determining where a student was placed in a given track and that the overwhelming majority of movement within the 25 schools was downward in particular courses. Such trends may indicate that the natural tendency for school officials and teachers is to believe a student is not prepared to succeed in a given course. As teacher expectations play a critical role in student success and achievement in school, this may profoundly impact student opportunity. Oakes' work also established the bold premise that any tracking ultimately represents an unfair educational experience for students as students placed within divergent tracks had fundamentally different academic experiences within the same schools. Oakes examined a core component of 297 English and math classes and focused upon the instruction within these classrooms to support this assertion. What the research team found was that students in higher tracked classes

had a disproportionate exposure to what was deemed "high status" knowledge, knowledge that was a proven commodity within the larger competitive world of post-secondary schooling. In fact, Oakes found that students in lower tracked classes examined significantly less course content than high tracked classes and were overwhelmingly composed of minority and poor students. This light shined upon what actually occurred within the classrooms of similar courses that had been grouped high, middle, and low within specific schools provided the impetus for the tracking debate that escalated in earnest in the late 1980s and exploded in the 1990s. The argument that the school itself actually promoted differences that negatively impacted a large proportion of the student body was novel. In fact, the overwhelming exposure to "low status" knowledge and lack of exposure to "high status" knowledge effectively locked students into tracks that were perpetuated year after year of their schooling experience.

Oakes' research expanded upon the work of Bowles and Gintis (1976) who had argued the structural elements of the school socialization process served to build upon existing capabilities and preparedness for different work environments within students from differing social backgrounds and that these differences were expanded upon based on the tracking in their respective school systems. Oakes suggested that students from the lower rungs of society from an economic standpoint were being prepared to remain in those lower rungs as a result of tracking practices employed by most public schools. The opportunities that such school systems failed to provide—indeed, in many instances might be argued, blocked—as a result of tracking practices employed at the school level perpetuated the existing social order and provided for sharp differences in the academic experience and preparation of students that cut between racial and economic lines. Oakes' seminal work served as an impetus to examine the specific classroom practices in tracked environments.

The examination of the impacts within schools of tracking decisions was further broadened following the publication of *Keeping Track*. Gamoran and Berends (1987) questioned how the root causes of achievement gaps could truly be identified if research methods did not effectively disengage prior conditions to conditions exacerbated by track placement. They criticized earlier examinations of tracking for lack of specificity and ignorance of contextual based decisions at the local level, noting that survey research failed to adequately address the complexity of local tracking decisions as students often misidentified their placements while educators often did not possess a full understanding of their own tracking practices. Gamoran and Berends centered in on the work of Oakes as a prime example as all of the 25 schools examined for *Keeping Track* had processes for grouping students not fully manifest from the outside looking inward. Individual building grouping practices vary so much it is difficult to distinguish track placement from amongst differing schools.

In examining the research to date, Gamoran and Berends noted that while some studies seemed to indicate tracking had negative impact to students' academic aspirations and achievement, in reality not enough of the prior conditions were controlled for while the school system policies and practices around tracking were not always fully examined. They found it difficult to discern between tracking practices and prior track conditions that were difficult to control for, particularly within school differences regarding student experiences. It was unclear if tracking had impacted student goal attainment or if prior conditions and aspirations had more significant impacts. Gamoran and Berends, like Oakes, were clear; the critical issue was in examining the instructional quality of the academic experience for within school differences, which most studies had failed to examine because of the simplistic nature of much of the previous survey-based research on tracking. After examining ten data sets in 16 different studies

concerning tracking, they found that very little guidance could be found in any ethnographic research on determining when academic differences are substantial and when academic differences are trivial and lacking such specificity, found it difficult to draw clear conclusions regarding the growth of academic gaps and how those gaps might be expanded because of school-related decisions on grouping. However, Gamoran and Berends noted that clearly the survey and ethnographic studies did find distinctions in academic achievement and attitudes toward school based upon track placement. Students had been found to more likely view higher track teachers as more patient and satisfied with their work (Vanfossen and Jones, 1987). Still, Gamoran and Berends called for more substantive quantitative examinations of longitudinal data related to track placements within a local context. They also noted the need for more substantive qualitative work on the impact of track placement on student self-perceptions and goal development. The effects of specific track placement could only be examined over time where meaning could be derived of the various grouping and stratification processes utilized in schools. Oakes (1987) found that tracking decisions often represented an eclectic mix of choices within the local context that were comprised of political, social, and cultural factors relevant to larger societal issues. Noland and Taylor (1986) conducted a meta-analysis of approximately 50 studies related to tracking practices from the 1960s through the early 1980s and found that grouping students according to perceived ability had no discernable positive impact to academic achievement but had an adverse impact upon student self-concept. Oakes argued that by the late 1980s even though substantial empirical evidence existed on the negative impacts of tracking, and court decisions supporting the abandonment of such practices had occurred, the use of tracking remained a nearly universal practice in secondary schools. Citing the work of Oakes as

a critical first step, Gamoran and Berends established the framework for more substantive examinations of tracking as the debate over tracking took the national stage during the 1990s.

Following the seminal work of Oakes (1985) various large scale meta-analyses attempted to illuminate the impacts of tracking on student outcomes. These works, conducted by Kulik (1992), Slavin (1990, 1993), and Mosteller et al. (1995) examined specific studies that attempted to discern if tracking provided benefits for specific educational groupings as well as if tracking should continue to be supported as a means to group students for instruction.

Slavin (1990) conducted a meta-analysis of 29 studies related to ability grouping in secondary schools and found that between class ability grouping and ability grouping by subject had no appreciable impact to student achievement across the studies. Slavin noted that ability grouped class assignment by subject was the predominant means of tracking explored in the literature. Slavin suggested that high-achieving students may have gains in academic achievement through their course and content acceleration in school, not necessarily as a result of tracking. This differential course experience and lack of exposure to high quality curriculum, engaging content, and teacher expertise had been substantiated in a number of previous studies (Persell, 1977; Rosenbaum, 1980; Oakes 1985, Gamoran 1989). Slavin, in conducting a meta-analysis of the tracking literature, concluded that tracking unduly caused institutional academic harm to the majority of students who were not designated as the very highest achieving students early in the tracking process. Slavin also refuted the work of Borg (1965) in showing ability grouping had a positive impact for high achievers as his exploration of Borg's work showed such differences were only true in some of the studies Borg examined. Slavin argued in his meta-analysis that one cannot control for the large differences between students and that many of the previous work in the literature were flawed as a result as they lacked proper controls. Slavin also

hypothesized that the subjective nature associated with student tracking placements could not be controlled for and as a result made it difficult to cite specific differences as a result of course placements. Slavin's subsequent work (1993) resulted in a call to figure out how to most effectively improve instruction in heterogeneous groups.

Braddock and Slavin (1992) found that when examining student performance of lower achieving students placed in lower tracks within tracked environments that those students generally performed significantly lower than similar ability low achievers in non-tracked schools. They surmised that this might occur as a result of the level of rigor expected within the non-college preparatory courses the majority of low-achieving students would be placed within a system that practices tracking. Slavin and Braddock argued that the connection to lower self-esteem for lower track students related to their schooling was a far more pressing issue than the purported impact to grouping students in mixed ability classrooms. Slavin (1993) argued that the existing research regarding the presence of low achievers in mixed ability classrooms did not appear to slow down high achievers; rather, such groupings actually held benefits for a far larger number of students as wider swaths of students would be exposed to the same curricular and instructional pace shared by students traditionally placed within highest tracks. It was noted that when examining truly gifted students, such students generally only comprised 3–5% of the overall student population and no there was no evidence that such opportunities for enrichment would not benefit all students. Braddock and Slavin (1992) noted that the social costs of tracking far outweighed the benefits of utilizing mixed ability classrooms as such practices further stratified schools based on societal differences and as a result limited student exposure, opportunities to make friends outside of their tracked environments, and separated students along racial and class lines.

Mosteller referred to the practice of tracking as skill grouping and noted the temporary nature of student groupings within classrooms. Four major types of grouping were found to be evident within the literature and common throughout American schools. These included: age grouping, mixed ability grouping, xyz grouping (within grade level stratification by high, medium, and low), and within class groupings based on student skill (Mosteller, 1995). Mosteller further expanded the understanding of how tracking did (or did not) impact overall student achievement by including in the meta-analysis only those studies that had a control and treatment group and a design that included a randomized field trial. Mosteller noted that the difficulty in examining the literature regarding tracking was the sheer lack of size, scope, duration, and the wide-ranging variability in the outcomes indicated within prior studies. When placing such specific parameters around their examination of the literature related to tracking, it was demonstrated that using tracking resulted in an overall 0 net effect in relation to student groupings and academic performance. Ironically, while slight positive gains may be seen for very high ability students grouped together for instruction, the overall lack of curriculum differentiation may have stymied their overall academic growth. It was evident within the meta-analysis that tracking did not benefit the large majority of average and lower performing student groups. Furthering the work of Slavin, the researchers speculated that it was a dangerous path for schools to so blindly follow such grouping processes given the lack of empirical data to support maintaining such distinctions. They noted the success of the Joplin Plan (Lovell, 1960) where multi grade level groupings proved effective for reading instruction in elementary grades. The complexity of drawing any meaningful conclusions from the literature was manifest within these meta-analyses as the studies pointed to the need for long-term longitudinal data that may be impossible to fully examine given the local complexity of tracking decisions, local practices

around tracking, and other long-term impacts that may be replicated throughout a student's schooling as a result of reinforced societal expectations and considerations being replayed within the schoolhouse to the detriment of students from lower socio-economic and minority backgrounds.

The expansion of the use of longitudinal data related to the impacts of tracking did occur at the end of the 1990s and included an increase in the use of large, nationally representative longitudinal survey data for the purpose of examining the impact of tracking. As longitudinal data became available, examinations of the long-term impacts of tracking became more prominent. Hallinan and Kubitschek (1999) conducted analysis that utilized two data sets that included the *National Educational Longitudinal Study* (National Center for Education Statistics, 1994) and the Ability Grouping Project which had occurred throughout several high schools in the Midwest. The NELS contained data related to approximately 11,000 students who answered survey questions as eighth graders in 1988 and as tenth graders in 1990. A second cohort included data from eighth graders in 1989 who answered follow-up survey questions in 1991. The Ability Grouping Project contained data related to 4,000 students across seven high schools who had track level and grades recorded four times a year beginning in their eighth grade year in 1987. The study examined the impact of track placement on tenth grade achievement with previous ability grouping as a predictor as well as scores on standardized tests and grades earned and socioeconomic status, gender, and race. The study revealed a strong and cumulative effect on student achievement from tracking and suggested that track placement has a direct impact on the degree to which a student learns course content. Two high quality data sets found that regardless of differences in samples and definitions of track, effects were the same; track placement inhibited student learning in lower tracks and removing the effect of previous track

placement demonstrated a gain or loss solely based on track placement over a two-year period. The results also revealed a strong cumulative effect on tracking as it related to student achievement as the track level students were placed in during eighth grade was shown to have a statistically significant impact on achievement in Grade 10. Braddock and Dawkins (1993) also examined NELS data from the *National Educational Longitudinal Study* of 1988. Despite some variations among groups when controlling for grades, achievement, aspirations, and socio-economic status, a student's placement in eighth grade served as a strong indicator of his or her specific tenth grade placement. Both studies found that movement within tracks was unlikely to occur on an upward trajectory as students overwhelmingly remained locked into their tracks from eighth grade throughout high school. In short, academic destinies of student experiences were largely concluded before students stepped foot in a high school. As minority and poor students were disproportionately placed within lower tracks in eighth grade, Braddock and Dawkins concluded that if schools were going to help produce the workforce needed for success in the modern era, tracking must change, as it suppresses the overall academic experiences of minority and poor students prior to the onset of their high school academic experience

The implications for the cumulative impact of student grouping decisions over the span of a K–12 continuum are thus frightening as students become locked into instructional groupings that are routinely and regularly exposed to more simplified instruction that are often a manifestation of a watered-down version of the established and expected curriculum (Roirdan, 1997). Slavin (1995) explored the difficulty with most of the existing studies on tracking that were primarily comprised of very different comparisons that included comparisons between no track/track environments as well as high track/low track environments. Slavin surmised that these simple comparisons often failed to control for drastic pre-test differences. Brewer, Rees,

and Argys (1996) argued that detracked environments would work against the high-achieving students and that high-achieving students benefited from tracked environments when examining overall academic performance. Slavin refuted this assertion arguing that while ability grouping may have some positive impacts for high achievers the research to date had done nothing to prove that high-achieving students would not perform as well in untracked schools. Slavin strongly criticized the practice of tracking for the overall negative impact to the majority of children within tracked schools. "Lacking any evidence that ability grouping is beneficial for students in general, it is important to ask how we can justify this practice. Ability grouping by its nature works against democratic and egalitarian norms, often creates racial or ethnic divisions, risks making terrible and long-lasting mistakes, and condemns many children to low-quality instruction and low-quality futures" (p. 2).

Development of the Tracking Wars

In defending the use of tracking, the literature tends to focus on the impact of heterogeneous grouping upon high-achieving students, while positing that practices aside from the grouping strategies used by schools result in lower student achievement outcomes for students placed within lower tracks (Argys, Rhees, & Brewer, 1996; Loveless 1999). Guill, Ludtke, & Koller (2016) examined students over a four-year period and found that students in higher tracks demonstrated higher gains over time on cognitively demanding tasks than peers not placed within higher academic tracks. Thus they concluded that tracking may also allow for the growth not only of subject specific competencies over time but also the improvement of general cognitive abilities. Betts and Shkolnick (2000) argued that as a result of poor comparisons between students in tracked and non-tracked classes was too convoluted to draw any specific policy decisions. They argued that as a result the failure to compare apples to apples has allowed

for an overstatement of the impact of tracking on inequality in test score performance. The inability to discern how students were specifically grouped at the classroom level by individual teachers throughout the literature prohibited the proper examination of the impact of tracking on student academic performance as it related to expanding existing achievement gaps. Kulik and Kulik found that students that are accelerated gained almost a full year over their non-accelerated peers (2004) and that grouping high-achieving students together may lead to increases in academic performance for those high achievers (1992).

Hidalgo-Hidalgo (2013) argued that the grouping of students into lower ability low track classrooms may have negative outcomes not because of the grouping strategies but because of the lack of teacher expertise, differential curriculum pacing, and general expectations for performance. In a comprehensive examination of teacher assignment in the Miami-Dade County Schools in Florida, Kalogrides and Loeb (2013) found that generally teacher leaders with more experience were assigned to higher track courses within schools. They hypothesized that given the larger academic gains made by students with more experiences teachers compared to novice teachers, lower tracked students may make smaller gains as a result of teacher assignment practices within schools rather than as a result of the utilization of tracking to sort student by ability.

The decade of the 1990s saw an increase in debate regarding the merits of tracking throughout the 1990s that has continued to the modern day. The 1990s saw tracking condemned by numerous national organizations with broad political strength including the National Governors Association and the National Education Association (Rubin, 2008). Those who sought to follow a pattern of detracking argued that tracking had more significant impact on the overall academic achievement of students than the school they attended (William and

Bartholemew, 2004). In her study of three comprehensive high schools in the Northeast, Rubin (2003) found that teachers spoke about children as levels, i.e., "honors" or "regular" students, and that local constructs of ability played a critical role in the daily teaching and learning in each school. Such descriptors of students are common throughout the hallways of many of the schools I have worked in and visited in my time in public education. Rubin additionally concluded that teachers and administrators had sharply developed preconceived notions of student ability as it related to their specific track placement. Students placed in honors or advanced courses were thought to be more academically capable than those students placed within regular or remedial courses. Thus student placement became a self-fulfilling prophecy. Students were viewed as having been placed within their respective course levels as a result of their academic capacity rather than as a cumulative impact of teacher and administrator perception related to capacity that may have contributed to the continual placement within a given course level over time.

Rubin argued that in order to move away from tracking, educational leaders must modify school-based structures significantly to provide for continual opportunity of students to move out of original course placements that might perpetuate placement decisions for high school as early as the elementary grades and provide access and support and resources to a wide range of students with substantial curricular revisions. In a groundbreaking study of racial sorting practices across three larger urban high schools, Kalogrides and Loeb (2009) found that within school sorting may minimize poor performing students' access to both high-quality teachers and high-performing peers. The study is vital in that it highlighted the vicious cycle that achievement differences in a given year on particular assessments or other narrow areas of focus may have a significant long-term impact on the academic experiences and exposure to a number

of students who find themselves perpetually placed in courses that move at a slower pace and contribute to watering down the entire educational experiences for many students. Over time students are continually sorted by performance on these arbitrary measures viewed as academically less capable by many faculty and staff and engaged in curriculum that often differs dramatically from the curriculum experiences by students perceived to be higher achieving and more capable so that as a result, over time student opportunities are diminished in a variety of ways including access to high-quality teachers, more rigorous courses, and disproportionate grouping with students of the same race and similar socioeconomic status.

Loveless (1999) posited that tracking may potentially have a negative impact on high achievers and low achievers alike. Tracking could be harmful to high achievers in that they would be bored and not achieve as readily in the course while harming the self-esteem of low achievers placed in courses where they struggled to keep pace with the academic concepts. Loveless again examined the issue of tracking in schools (2009) when he examined tracking within subject and within grade level at the middle school level in Massachusetts during the period 1991–2009. Loveless made the broad generalization that tracking had been significantly minimized in middle schools. However, the generalization ignored other ways through which tracking continues in less apparent school-based processes that may not be detected in simply examining a course title. Loveless defined tracking as the grouping of students in different classes based upon achievement. He subsequently updated his initial study with a specific examination of tracking on the achievement of high achievers (2009). He found that school size influenced tracking policies as larger schools tended to be more tracked than smaller schools, noting that the majority of detracking efforts follow socioeconomic lines as higher wealth schools were found to have been less likely to undertake detracking initiatives. Loveless argued

that middle schools with more tracks tended to have an overall increase in the number of advanced math students in mathematics even when controlling for socioeconomic status and that tracked schools tend to have higher number of students scoring at advanced levels. However, as stated in his own argument, Loveless ignored the complexity of the tracking process as a whole as detracked schools were more likely to represent lower socioeconomic areas regardless of specific wealth controls. His study failed to take into consideration several vital variables related to the overall instructional environment, including the presence of savvy parents who would request advanced levels of mathematics for their children, the long-term negative impacts of students placed in tracks beginning as early as elementary schools and the biases of adults working within the schools when making placement decisions (Oakes, 2005; Rubin, 2008; Slavin 1992). The impact of tracking was also examined in the Chicago public schools following the requirement for all ninth graders to enroll in English I and Algebra I. In exploring the academic impact of placing all students in Algebra I across the Chicago schools, Allensworth et al. (2009) concluded that the use of more rigorous content for all acted as a null gain over the approximate decade for which the change had been made. Authors noted that they did not examine the instruction, only the outcomes, and did not explore in depth what if any curricular changes occurred nor how, if at all, the course-taking changes followed any significant changes at the K–8 scope. In failing to do so, they ignored a critical component at the very heart of tracking: the differences in the quality of instruction, and often, the quality of instructors that often varied dramatically from high track to low track courses (Gamoran, 2009).

Argys, Rees, & Brewer (1996) found that ability grouping had small positive effect on achievement for high-performing students. They argued that detracking initiatives would ultimately create winners and losers and that while the academic performance of low-achieving

students would increase, high-achieving students would have a corresponding decrease in achievement. Their argument was put forth at the height of the public debate regarding the use of tracking and criticized the belief that schools could detrack at no cost to students. In their estimation, the potential loss in achievement for high-tracked students would offset the gains made by low-tracked students if the practice of schoolwide heterogeneous grouping was followed. Additionally, advocates of tracking argued that for the gifted, such separation allowed for advanced academic growth and the development of positive self-perception and high esteem (Rogers, 2007). Slavin found the same indication for a very small percentage of the gifted population in earlier studies (1987). In these cases, the numbers impacted were related specifically to the very highest achieving students in advanced classes. Certainly an important consideration, however, the overwhelming reviews in the literature found tracking had disproportionality negative impacts on the low and average-achieving students. There was one other notable exception to the positive impact of tracking and that involved the education of English Language Learners. In some small instances tracking may play a pivotal role in ensuring the adequate academic experience for select students groups, in particular English Language Learners (ELLs) that seem to benefit from being placed within such environments (Robinson 2008). While these findings point toward the possibility of particular experiences being improved through the use of tracking for extremely gifted and ELL students, the umbrella use of course-specific tracking that presently dominates American public schools simply does not suffice for the vast majority of students and exasperates racial and economic differences using the public school as the primary agent in such division.

The decade of the 1990s ushered in a more sophisticated utilization of longitudinal data that moved beyond student survey data and began to integrate a deeper examination of long-term

student outcomes related to being placed in specific academic tracks. The findings across the literature clearly identified negative impacts to students to the use of tracking. The ways through which adults grouped students for instruction in schools were often subjective and had a potentially significant negative impact to student outcomes. The use of tracking may perpetuate a vicious cycle where teacher expectations and student expectations enter a downward spiral (Page, 1992). The use of such practices where students were believed to possess a certain academic capacity and no more assumed a static or fixed view of student ability. As the longitudinal studies of the decade began to make clear, tracking mattered regarding how student capacity was or was not maximized as the majority of students placed in lower or non-advanced tracks had their academic growth curtailed in a manner wholly inconsistent with peers placed in accelerated or advanced tracks (Slavin and Braddock, 1992). The impact of such decisions were compounded over time making the examination of high school tracking practices more complex as many of the root causes of student differences in academic achievement were exasperated over time while achievement gaps widened between students placed year over year in different academic tracks. As the tracking decisions of school leaders often mirrored societal differences as assigned by race and class, the impact to poor and minority students in such processes was significant (Oakes and Guiton 1995).

As the shift in raising academic expectations through accountability measures for all occurred throughout the 1990s and through the turn of the century with the passage of *No Child Left Behind* (NCLB 2001) the move toward mixed ability classrooms was often overshadowed by a return to tracking as standardized test accountability measures came to be the dominant manner through which school success was measured. As pressure mounted to demonstrate student growth along very limited and highly specific language arts and mathematics

standardized assessment scores, curriculum polarization fragmented the academic experience throughout American public schools. This polarization resulted in significant learning outcomes for students as a result of advanced academic tracks often moving too rapidly through content while lower academic tracks often made considerably less progress *with the same curriculum*. In examining almost one thousand students in six schools over a three-year period (Boaler et. al., 2000) were able to utilize a longitudinal study that examined the performance of students in mixed ability classes against the performance of students in more specifically tracked environments. Using research led observations, assessments, and interviews, Boaler found that impact to students' beliefs about mathematics was severe. Many students reported that the curriculum moved too fast or too slow and as a result they became lost or disengaged in the curriculum. Lower tracks were also comprised of disproportionality poor and working class populations. The study also revealed a general belief amongst teachers that students placed within explicitly tracked classes were assumed to be possessed of identical abilities and as a result, limited differentiation of instruction was implemented when compared to the classroom strategies employed in mixed ability classes. Placing students into tracked classrooms may create a set of expectations of teachers that overrides their awareness of individual capabilities. As a result, curriculum polarizations may lead to a majority of students achieving below their potential as their instructors perceive their abilities as being comparable in all areas of their academics and rather than assessing individual strengths, weaknesses, and interests, proceed with classroom instruction that shortchanges the individual academic experience.

The debate over the student impact of tracking continued in earnest throughout the 1990s and into the new century (Loveless, 1999). As schools began to re-examine the use of their student grouping processes, the difficulties in providing a common experience for all students

while also effectively differentiating the curriculum for all became more fully manifest in how schools approached their grouping strategies. This tension had been explored in the literature in depth in the previous decades (Gamoran, 1989: Oakes 1995). The modern use of school and district rankings such as those in the *US News and World Report* and *Newsweek*, where schools were categorized based on the number of their students enrolled in Advanced Placement (AP) courses became prevalent in the post *NCLB* world and created political pressure on schools to move increased numbers of students into more rigorous courses, further exasperating the curriculum experience for those students placed in lower tracked courses. Tracking, though vastly different from the programmatic distinctions of the past remains steadily ingrained in the practices of American schools (Moller and Stearns, 2012).

Tracking may work against both low-achieving and high-achieving students. Boaler, William, and Brown (2000) examined almost one thousand Grade 8 students in six schools as they transitioned from homogeneous to heterogeneous math course placements in four of the six schools in the United Kingdom (UK). Using classroom observations, interviews, and survey data, they found that students in the lower tracks entering the ninth grade had their learning opportunities sharply restricted as they conformed to the generally lower expectations of their performance. The lower expectations associated with their coursework caused a loss of opportunity as they progressed through high school in lower tracks from grade level to grade level thereby limiting their overall performance. Conversely, the researchers found that almost a third of the high track students performed poorly due to the rapid pace of the high track curriculum and amount of content coverage and pressure to succeed. The curricular pace produced disaffection in lower tracks and student anxiety in higher tracks resulting in negative implications for all groups examined. Although the sample size was relatively small, the

pressures and expectations associated with the higher tracks have clear and demonstrable connections to advanced coursework in the United States including AP and International Baccalaureate (IB). Students in higher tracks recognized that even in mixed ability classrooms the range of student aptitudes and interests were sharply varied. The expectations of each class were exaggerated by the adults; too rapid a curricular pace and too much content in higher tracks, too little engagement and lower expectations in the lower tracked courses. The adult expectation for students based on their track assignment had a pernicious impact on the students as a whole.

Mickelson (2005) utilized survey data of 24 middle schools from the Charlotte-Mecklenburg school system (CMS). The survey data of approximately 1,800 Grade 8 students in 1997 was further enhanced with school level controls for prior academic achievement, socio-economic status, and used school officials to corroborate the specific tracks students were placed within. Students placed in schools with higher percentages of minority students were found to be less likely to be placed in a college preparatory middle school track, thus inhibiting the opportunities to be placed in higher level high school courses. The subjective nature of track placement clearly was impacted by the student race in terms of track placement. The study corroborated the work of Downey and Pribesh (2004) who found that white teachers and administrators were more likely to judge African American students in a negative manner. In fact, Mickelson further posited that allowing tracking in such a manner worked against a compelling state interest in promoting the most expansive forms of academic achievement as tracking in the CMS school district clearly had racial overtones that precluded large numbers of black students from achieving at the same level of their white counterparts. As our schools become increasingly segregated by race, particularly in New Jersey, this study has profound

implications for how tracking serves as a segregationist proxy to limit the opportunities of minority students. With increased standardized testing resulting in the further segregation of students by test scores, the limitations and loss of opportunities will be the result and indeed has resulted in a vicious cycle of underperformance by minority groups. Certainly these practices contribute to increasing achievement gaps from as early as elementary school and are fully manifest in high school post-secondary outcomes.

Chambers (2009) argued that tracking leads to the development of a receivement gap that helps to precipitate an achievement gap because tracking inherently focuses on inputs. Students in lower academic tracks are exposed to slower paced, watered down curricula, less experienced teachers, and are saddled with lower expectations that have a profound compound impact over time in schools. This receivement gap helps to perpetuate the achievement gap while also perpetuating societal differences based on race and socio-economic status. As school leaders and teachers have deeply-rooted beliefs about what a successful student should look and act like, the subjective nature of tracking becomes manifest in course placement decisions over time, forever altering the academic program of students based on preconceived societal differences. Thus, stigma and separation becomes a normal and routine component to the schooling process. As this separation becomes apparent, students placed in lower tracks naturally begin to feel less intelligent. These overt reminders of being separated based on "ability" lead to the further belief that they are deemed less capable in the eyes of the school system. The receivement gap becomes a self-fulfilling prophecy that further divides students based on race and class and leads to the widening of performance on standardized tests which are ultimately utilized to further that divide throughout the course of a child's schooling.

The types of tracking students experience seem to differ by socioeconomic status. In examining divergent schools Attewell and Domina (2008) found that high SES schools tend to offer more advanced and honors courses. Lucas (1999) found that social status seems to play a critical role in student course taking even after controlling for individual aspirations and achievement. In examining NELS data for the class of 1992, Lucas found that the mitigating factor in course selection was not race but socioeconomic status. Poor whites took less demanding courses than blacks of the some wealth. Statistically significant disparities in outcomes that cannot be explained by prior academic performance are most directly related to socioeconomic status. Lower SES students were over-represented in lower tracks regardless of similarities in prior academic achievement. So although there were increased opportunities for honors and advanced classes in high SES schools, the poor students within such schools were still disproportionately represented in the lower academic tracks when examining course-taking data.

Archbald & Farley-Ripple (2012) further examined whether probabilities in enrollment in advanced level mathematics courses vary by demographic factors when controlling for prior achievement. Their work built on the work of Mickelson (2001) who as discussed had found black students were significantly underrepresented in in advanced levels of English while being over represented in lower level English classes throughout the Charlotte-Mecklenburg schools. Archbald and Farley-Ripple found that Mickelson had few controls for prior academic achievement. The complexity of transcript data and difficulty in obtaining that data has clouded similar studies.

In examining NELS data from 1988–1992 and controlling for standardized test scores in Grade 8 Schiller and Hunt (2011) found there was no statistically significant difference in course

taking by ethnicity. However, this ignored eight years of schooling prior to the standardized assessment in Grade 8 where course placement may have been stagnant and the difficulties with curriculum and pacing and content exposure may be fully manifest but not captured within the scope of the study. Archbald and Fraley-Ripple examined 707 Grade 8 students within one school district to examine their course-taking patterns in mathematics over a three-year period. The study controlled for prior mathematics courses taken, prior standardized test scores, and prior grades earned in mathematics courses. While their findings were consistent with previous research that had shown no statistically significant difference in course-taking patterns by race or ethnicity (Schiller & Hunt, 2011, Stone, 1998) the researchers noted the significant gaps that existed between student performance by demographic by Grade 8 and the resulting impact the Grade 8 math placement played in determining the high school attainment in mathematics was striking and troubling. In examining any high school data, it is important to include middle school measures in the data as these have powerful implications for high school course selection and opportunities. Fletcher (2012) also found that the course-taking decisions students make by ninth grade have profound impact on their long-term career and that such choices unbeknownst to students often have a profound impact on career aspirations.

Regardless of academic impact there is clearly little movement between tracks in many instances where tracking is employed. Using the concept of an inter-track mobility coded system using specific data sets to understand the likelihood of a student moving up or down a track, Archbald and Keleher (2008) found that the gulf between research and complex student placement practices is profound and that schools must more effectively use student demographic, performance, and transcript data to monitor student movement in tracks over time and make more specific policy decisions. Without a highly detailed analysis in making decisions about

how tracking is working within a given system is likely to result in an incomplete understanding of the impact to students. Kelly (2008) also found that students were highly likely to remain in the same tracks they were placed within in Grade 9 through their entire high school experience, regardless of academic performance. This has been a common occurrence in most school systems I have entered in my career and without significant adjustments to identifying individual students for movement along the academic continuum, it sadly remains the reality for most students.

The use of specific longitudinal data enriched the understanding of the impact of tracking practices upon long-term student achievement and success in post-secondary endeavors as well as the examination of workforce income statistics. Moller and Stearns (2012) examined the National Educational Longitudinal Study (NELS) data from 1988–2000 and found that when controlling for socio-economic status, prior achievement, and discipline infractions students who studied in more rigorous academic tracks beginning in the eighth grade and through high school clearly benefitted in their annual income as adults. Their examination relied upon specific transcript data and followed a large cohort of students from high school through college and entrance into the workforce and pointed to the disparity in earning capacity where academic tracks negatively impacted the earning potential of students placed in less demanding courses as a result of tracking practices within their respective schools.

Student Motivation

The impact of tracking practices on student self-esteem and self-perception as well as in the variance of instructional quality between tracks became a focus of examination throughout the 1970s. Tracking has been found to have a significant impact on student self-perception and motivation and to strongly shape teacher perceptions of student ability and motivation. The

assignment to level specific classes, whether as a pre-determinant to ability grouped levels or as a larger assignment into a general academic track, has profound implications for student self-perceptions, self-concept, and views on their own academic capabilities (Darling-Hammond, 2010). Students assigned to lower tracks or less demanding courses may prefer being assigned more low-level work including the assignment of worksheets over other lessons in order to not be exposed for what they do not know or may struggle with (Metz, 1978). In fact, the use of tracking may exasperate student self-concept in a manner that reinforces existing academic differences and expands rather than closes achievement gaps. Tracking has been found to increase student self-concept in higher tracks while lowering student self-concept in lower tracks; thus student achievement becomes a self-fulfilling prophecy as students who are already identified as not performing to a standard are labeled and stigmatized in a way that may further erode their academic confidence and self-esteem, resulting in lower overall student achievement (Ansalone, 2009).

Two longitudinal studies published in the late 1970s and 1980 provided more insight into the influence to track placement on long-term student outcomes. Alexander, Cook, & McDill (1978) expanded upon the severe limitations of most previous work regarding tracking in their examination of longitudinal data from 2400 students across ten high schools in the late 1960s. The research focused upon selection processes within schools for student track placement and controlled for a number of pre-curriculum variables including expectations, peer plans, and plans to attend college. The resulting conclusions found that whether one enrolled in a college or non-college track often varied amongst students even when controlling for ninth grade achievement and academic ability and that a number found to be of broad consequence across a number of ranges. High status students were seen to have benefitted from a series of modest advantages

over time that propelled them toward the college-bound track including the peer effects of their curricular placements, stronger parental support, and the expectations of their friends in ninth grade. Critically, the study found that 60% of the variance of track placement within schools was unexplained, speaking to the subjectivity of student placement processes within schools. Most powerfully, simply being placed in a college preparatory track increased the likelihood of students applying to college by over 20%. The authors concluded, "Even with pre-enrollment controls, however, the importance of curriculum placement for junior year and senior year outcomes is marked. In particular, tracking consistently affects educational goals, achievements, and goal-oriented behaviors in the twelfth grade, and is often the most important factor of those included in our model" (p. 62). The irony within schools that such subjectivity produces is that students of equal ability and motivation often ended up in divergent academic tracks, sharply altering their pathway toward college attainment and enrollment.

The impact of tracking on self-esteem and student and staff expectations for students emerged prominently in the literature from the 1970s. Kelly (1975) found that track placement had a direct relationship to student self-esteem with different groups of students being cognizant of their respective places and corresponding expectations for academic success based upon their track placement. Lower track placement was found to result in lower expectations for learning by peers, teachers, and counselors (Persell, 1977). In Persell's synthesis of tracking literature, a number of studies were found to suggest that tracking has a profound stigmatizing effect on students. Metz (1978) suggested that tracking introduced inequalities that did not previously exist as the main purpose of lower level tracks was to maintain order and discipline rather than to support learning. By basing the grouping on the need to maintain order over learners who were viewed as generally more disruptive or unruly, lower tracks effectively provided fundamentally

different academic experiences for the students assigned to such tracks. Kerchhoff (1976) found that the roles of societal forces were clearly evident in tracking, as curriculum tracks were representative of societal divisions by race and wealth.

Long-term assignment to lower tracked classes had a negative impact to teachers in some instances, impacting their competencies (Finley, 1984). Finley utilized a qualitative study in a large suburban high school over the course of a year to determine if and how teachers impacted the specific tracking processes within a school. Though the research was conducted in the mid-1980s, much of the findings are commonplace in comprehensive high schools today. Finley detailed a curious battle between teachers with influence over parental course placement decisions and influence over guidance counselors. Finley also identified specific teacher practices including threatening high workloads with difficult writing and conceptual understanding as a means through which teachers sorted and stratified students to further reproduce societal inequalities. While Finley's examination was school specific, the influence that teachers had on structuring the tracking system within their school was profound. Finley argued that her study could indeed be replicated at many other schools who serve a diverse student body and utilize some form of tracking. Finley speculated that the characteristics of the suburban schools she examined would lend themselves to expecting similar practices in a majority of suburban high schools across the nation: "A heterogeneous student population, ability grouping or tracking in academic subjects, teachers who migrate from inner city to suburban schools to find middle class clientele, teachers who favor ability grouping although they do not want to teach low-level groups, an organizational structure of isolated classrooms, which separates teachers and contributes to their preoccupation with individualistic rewards and

participation in the national selection of selecting and preparing elite students through AP and accelerated programs" (p. 242).

The equity issues raised by Finley's work would be examined in earnest throughout the literature. Student perceptions of their teachers also differed according to track placement as students were found to perceive higher track teachers as being more patient and more effective with classroom management (Vanfossen, Jones and Spade 1987). Boaler (1997) found that teachers in higher tracked classes did tend to alter their instructional practices to a higher degree than lower tracked instructors, further reinforcing status labels that perpetuated the belief that the academic experience was being dumbed down for students in lower tracks. Teachers of lower tracks were found to stick to more rigid instructional practices as a result of their expectations of student capacity to engage in the curriculum.

When examining track placement effect on friendship choices and controlling for school size utilizing data from a subsample of the High School Beyond study that involved over 50,000 sophomores from over 1000 schools across the country, and controlling for socio-economic status and gender and prior achievement as designated by grades earned, Hallinan (1998) found a link between organizational characteristics of tracking and student friendship choices, suggesting that school-based decisions on tracking practices may impact the social development of students as it increases the likelihood of interaction with students within a given track. Tracking also emphasized existing similarities between students. As schools grew larger the likelihood that students interacted with other students outside of their academic tracks decreased. The research suggests that through their tracking practices schools impact social and emotional development of students though formation of peer relationships. This is an important impact in determining how students are placed or grouped in their academic courses over the course of their high school

years and suggests that schools can be more proactive in engaging in practices that foster increased student interaction across tracks through detracking initiatives in various courses. It also suggests that schools should pay great attention to scheduling specifics on an individual student level in order to foster increased social interactions. As students stratified in tracks in testing era that divides students by test scores may also restrict social and emotional development by limiting opportunities to engage and befriend other students outside of tracks. Additionally, students were more likely to identify friends in higher tracks than lower tracks suggesting a stigma to academic hierarchy in schools.

Hallinan (1994) extended the work of Oakes (1985) in examining the variance in the quantity and quality of instruction across differing track levels and the effects of track level on student self-esteem and motivation and found that much of the inequity in student performance could be explained by the impact track placements had on student motivation, self-esteem, and the quality of instruction received. Hallinan's work followed two cohorts of students from public and private schools from Grade 7 and included all public schools in two large Midwestern cities. The study relied upon multiple measures of student prior ability, including math and English scores on standardized tests, and contained approximately 4600 students. Hallinan found that the majority of student math tracks moved down through high school and that ninth grade track placement had a statistically significant impact on students' growth in achievement. This growth, while stronger in mathematics, was notable in English as well. Schools differed in the degree to which tracking processes mirrored existing inequalities by race and SES and the tracking impacts varied from school to school with clear gains for some honors and advanced level courses in Grade 9 and in other schools less clear gains made for similar placements. Some schools were also more successful in promoting student achievement of students placed in lower

tracks. Hallinan noted that to truly understand the complexity of tracking and the impact to student achievement, school level tracking practices must be examined in isolation from district level practices as the complexity of the decision-making processes and the subjective nature of track placements varied widely from one school to another, even for those schools within the same larger school district. Hallinan argued the work moving forward should be focused on specific ways schools ameliorate or exaggerate existing achievement gaps with the tracking and programmatic structures utilized

Marsh argued in what became known as the Big Fish Little Pond Theory (BFLP, 1994) that tracking had negative effects on high-performing students. Marsh found that students in high-achieving classes felt more threatened and insecure regarding their school status because of comparisons to high-performing peers and may actually harm academic self-concept of highest performing students. Marsh used propensity analysis to compare students in tracked and untracked schools in an attempt to mitigate pre-existing differences that might be evident in comparing high- and low-tracked students within the same schools. Students within the higher tracks in tracked environments were found to be less sure of themselves because they are constantly compared to the very highest achieving students within their respective schools. High-performing students in untracked environments were found to have higher academic self-concept. The social-psychological impacts of tracking may have a negative impact on the highest performing students. Marsh and Yueng (1998) found that student grades were more sensitive to student motivation and teacher feedback than standardized assessments. If student perceptions are clearly impacted by track placement and their subsequent exposure to teachers who may perceive them as lacking the capacity to perform well within an academic setting increases when placed within a lower track the long-term, consequences may be dire as a

perpetual cycle of underperformance and under exposure to an appropriate curriculum naturally compounds over time.

Hallinan was explicit in the findings that the effects of track level placement on student self-esteem and motivation, coupled with the wide-ranging variation in the type and quality of instruction afforded to higher level tracks as opposed to lower level tracks, profoundly impacted student achievement (Hallinan 1994). The issue of motivation and what students bring to classes prior to being placed in a track must be considered when examining these effects. In fact, most research fails to adequately account for these variations in instruction after students are placed within particular academic tracks (Gamoran, 1992; 2009). The quality of instruction and the educational climate of the classroom cannot be taken out of context in determining student self-esteem and perceptions of academic worth when examining what impacts tracking may have on a student's academic achievement and outcomes from a school system.

The psychological reminders and impact to students placed within lower tracks are profound as students are faced with multiple reminders that they are not as academically capable as many of their peers; these include separation, feeling dumb, moving at slower pace, being viewed as less capable by the educational professionals around them. These impacts result in what might be viewed as an opportunity gap that contributes to a widening gap on test performance which, ironically, is often the major determinant in placing students in lower tracks in the first place (Chambers, 2009). These psychological impacts were found to have profound academic consequences on overall student achievement. Van Houtte and Stevens (2009) found that students in lower tracks experience a loss of status and lose ability to benefit from positive peer effects. In examining 11, 872 valid student questionnaires from 85 schools with equivalent 9th and 11th grades of the majority of students placed in an academic rather than vocational track

they found students in lower tracks in schools with divergent tracks may actually exacerbate student negative feelings toward schooling and a feeling of status depravation as students in lower tracks had more negative self-concept and that those views became more manifest as such tracks become more visible to students.

Kelly and Carbanaro (2012) found that the use of tracking produces powerful status labels. Students alter behavior to conform to the expectations of their track placement, teachers change expectations and use labels identified with track placements, and student expectations for college attendance is generally higher than those of teachers in lower tracks. They examined NELS data for the graduating class of 1992 and focused on students placed in divergent course specific tracks to see if teacher expectations for students differed depending upon track placement. The research offered unique methodological advantage by focusing on discrepant course-taking students. Low-track students as early as second grade may have very low views of their academic capacity (Stipek and MacIver). Interestingly, students respond differently to varying classroom contexts. Focused on N of 5,852 students in discrepant tracks that had reports from two teachers, who also identified student tracks for courses they instructed. While 84% of the students expected to attend college only 60% of the teachers believed the students would attend college. For both teacher and student, track level plays a critical role in expectations or beliefs that the student will attend college. Teachers of the same students in lower tracks had statistically significant lower expectations from teachers of the same students in higher tracks. Teacher expectations must be examined and included in any data set of a school trying to optimize student achievement and opportunity.

Hallam and Ireson (2007) examined almost 8,000 students across 455 secondary schools in the United Kingdom and found that large percentages of students in lower tracks did not find

the work appropriate and wanted to move into higher and advanced tracks. This was true of students in the middle tracks as well. Students felt not able to progress at a pace that was conducive to learning and points to potential benefits of increasing mixed ability classrooms for all students. Students in mixed ability classes were overwhelmingly more inclined to remain in class than students in mathematics that were more rigidly tracked.

Kelly and Price (2011) examined the course guides of 128 public high schools in North Carolina for the 2007–2009 school years. By coding guides related to the number of tracks for particular subjects as well as course-taking corequisites and prerequisites, Kelly and Price found that both skill and subjective requirements (such as teacher recommendations) and the use of test and pre-requisite scores had increased in the majority of the schools examined in the subsequent decade. As low income and minority students enter high school with overall lower achievement levels, these practices would seem to exacerbate the educational gap in high school while limiting the opportunities for low income and minority students to move through more rigorous academic course loads in high school. They found that more highly selective tracking systems are in place in more affluent schools with fewer poor and minority students. Thussystem of trcking was highly nuanced and hyper-localized, making it more difficult to fully gauge potential impacts.

Yonezawa and Jones (2006) examined student perceptions of tracking across 12 high schools and conducted 75 group meetings with over 500 students. Students expressed a recognition that tracking practices were also shaped by political and social factors and demonstrated a clear understanding that external factors beyond academic record, such as perceptions around race and class and belief in student capacities, impacted student placement decisions. While the students condemned these practices and viewed such practices as unfair to

certain students, few understood how the tracking practices worked within their own schools. Students also expressed concern over what they referred to as irrational practices that required students to take multiple AP or Honors courses or none at all. Such course groupings worked to exclude larger numbers of students from the higher academic tracks and also ignored subject area expertise that might be demonstrated by individual students. Students were found to be in strong agreement that test scores were not valid measures of student ability and to effectively detrack, teachers needed to believe in all students. Student opinions clearly expressed a sentiment that higher track teachers had higher student expectations and felt that tracking as a whole left many students behind in their academic development.

The myth of student choice in tracking decisions perpetuated by Loveless (1999) who argued that low-track students choose lower tracks to engage in easier and less demanding coursework and others was debunked in several studies. Yonezawa, Wells, and Serna (2002) examined the detracking efforts of six racially mixed high schools from various regions in the country as well as four similarly situated middle schools through interviews, surveys, and classroom observations over a two-year period and found several factors that inhibited detracking efforts and maintained the existing tracking structures. These included institutional barriers such as access to information, limited information and awareness by low-track students, particularly minorities of the opportunities for higher tracked classes, hidden prerequisites, and selective flexibility in meeting course requests of minority and non-minority students that seemed to fall along socio-economic lines. Yonezawa, Wells, and Serna (2002) further challenged Loveless' argument (1999) as overly simplistic that low-track students choose low-track classes because they think the courses will be easier as they found that open access manifested itself in very different ways from school to school and classroom to classroom within

the same district. These differences included the manner through which information was put forth to school community members, including students; the process for moving students into higher tracked courses also had significant distinctions across divergent schools. The schools were marked by what they deemed "selective flexibility" where non-minority, wealthier students ran into fewer barriers in opting into higher tracked courses. Ironically, it was found that higher track students met most resistance when trying to drop to lower track courses. There were clearly hidden barriers to access even in schools that claimed to have open enrollment processes for all students. Additionally, different screening processes also served as detriments to students enrolling in more challenging courses. The argument that lower tracked students simply chose the easiest path in course selection was clearly not the finding once schools hidden barriers were evident in open-enrollment grades or courses in previous courses. Through screenings as well often scared students away from taking more challenging courses.

Most tracking occurs in modern day as between class groupings where students in the same course are sorted into classes of high, medium, and below average based upon perceived ability (Ansalone 2010). Ansalone posits that the arguments for tracking have remained relatively constant for almost a century and are predicated upon efficiency through the maximized use of resources to meet the needs of diverse learners through appropriate separation for instruction. Additionally, the pro tracking argument states that students have higher self-concept if grouped with like ability students. The difficulty with the veracity of these arguments are the subjective nature of tracking decisions. What teachers know about students more often comes from the specific track assignment as teachers' perceptions are shaped by the groupings students are placed. Boaler (1997) found that teachers alter teaching practices in higher tracks because they expect students to pick up the subject area content more rapidly based on their track

assignment. Ansalone and Biafora (2004) found that even though teachers understood that tracking may have a negative impact to students, they still believed they needed to sort students into tracks as a managerial tool and as a means of targeting instruction for specific students. Thus, an understanding of a different curriculum in delivery, experience, and expectation becomes apparent for the sake of efficiency. The constant pressure of educating the individual within a system that must educate all students has clearly resulted in many students being inappropriately limited in their overall academic experience.

Historically, teachers and administrators have played a significant role in the development and implementation of a school's policies and practices surrounding the use of tracking. Various criteria for student placement decisions have historically included teacher impressions and expectations of student ability and intellectual capacity (Useem 1992, Gamoran and Berends 1987). The importance of teacher impressions and perceptions of student aptitude in shaping school policy and practice cannot be understated.

The role of the individual teacher in tracking decisions evolved as the need for an efficient method to educate rapidly increasing numbers of poor and non-English speaking students bubbled over throughout the late 19th and early 20th century. Tracking was viewed as a means to differentiate the curriculum in a manner that prepared divergent student groups for their perceived roles in the workplace. The vision of Horace Mann, in providing the same academy experiences for all students in public schools, was crowded out by the social efficiency movement and what it wrought. From this cauldron of conflict emerged the modern comprehensive high school, where these competing interests play out in the modern day (Gamoran, 2009). Educators still struggle with balancing efficiency and meeting the needs of all

learners. In an era of high stakes testing such as PARCC, these balances may be tipped in ways that work against students sorted into lower level or test prep classes.

Teacher recommendations have been a primary means of determining student placement for courses and course levels, since school officials first sought efficiency in grouping practices as the numbers of poor and immigrant children began to rapidly expand following the turn of the century (Oakes, 2005). Teacher recommendations remain an important component to the overall student placement process in most schools in the modern day (Yonezawa, Wells, & Serna 2002). As such, the inherent biases that have been uncovered in the literature need to be given careful consideration. In fact, placement decisions are often beyond the control of formal school policies and are most significantly impacted by the perceptions of the staff and faculty within the school (Archbald, 2012). This "selective flexibility," determining which students will or will not be placed in each particular course or course level, is often clouded by external factors aside from academic achievement including race, class, and ethnicity (Oakes and Guiton, 1995; Kelly, 2007). With the increasing demands of meeting the expectations set forth through the *Common Core State Standards a*s well as high stakes assessments increasing in frequency and duration, headlined by those states utilizing high stakes standardized assessments from PARCC and SMARTER Balanced Consortia respectively, as students are subjugated to lower tracks as a result of educator perceptions, the consequences could be drastic as the very students placed within courses that move through the curriculum at a slower rate and are comprised by less experienced teachers (Kelly, 2004; Kalogrides and Loeb, 2013) will be expected to perform to the same levels as their peers with more experienced teachers and wider ranging curricular experiences (Gamoran, 2009). This two-pronged disservice may have the impact of further obscuring opportunities for students of color and from poor backgrounds from partaking of the

same opportunities provided to other students. Teachers are under duress to ensure that students are moving through the curriculum at an appropriate pace and are being prepared to perform well in regards to the more demanding assessments expectations. As teachers of higher tracked classes tend to move through the curriculum at a more rapid pace, the opportunity and achievement gaps for students in lower tracks grows disproportionately wider over time (Chambers, 2009; Oakes, 2005). Higher tracked courses are exposed to a more rapid pace through the curriculum, more challenging content and a wider range of challenging instructional strategies (Hallinan, 1994; Gamoran, 2009; Watanabe, 2006).

Some teachers reported that they felt unprepared to appropriately address the educational needs of students without a sorting mechanism within the school, whether that be open enrollment of untracked environments where students of perceived ability groups populate the same courses (Welner and Burris, 2005). Even in untracked environments, teachers often seek ways to re-track students to more effectively deliver instruction to students perceived to have differing means of assessing and managing that instruction (Rubin and Noguera, 2004). Students who move onto more demanding course work may still be subjugated to lower teacher expectations as a result of the search for efficiency. Teachers who were themselves placed within tracked school systems are also more likely to be opposed to practices that mitigate tracking (Welner and Burris, 2005). In many ways the cycle of tracking may be inadvertently perpetuated by those who benefited in schools from such structures. Teachers with more experience and seniority are also more likely to be assigned to higher level courses leaving lower tracked courses with a disproportionate number of new and less experienced instructors (Kelly, 2004; Kalogrides and Loeb, 2013).

The literature in identifying that traditionally more test preparation work is conducted in lower tracked classes is profound. Teachers in higher tracked classes tend to focus less on test preparation (Watanabe, 2006). Teachers, when surveyed, identified a higher degree of test preparation in lower tracked classes. With increasing demands to move all students through a more rigorous curriculum, the research suggests that teachers of lower tracks will spend more time focusing on test preparation in order to meet expanding accountability demands. The converse may ring true within this age of accountability. As the research suggests that teachers of higher tracks engage in more meaningful learning by avoiding as heavy a reliance on test preparation as teachers of lower tracked courses, this divide in instructional quality will only continue to grow. Additionally, students in lower tracked classes tend to be assigned less work outside of class (Page, 1989) which may exacerbate student differences even further. At the least the disparity in expectations as conveyed through work expectations established for students in higher and lower tracked courses magnifies the lower expectations set forth for students placed in low-track classes. These differential learning opportunities have a profound and cumulative impact over time (Rubin, 2008).

Administrators also believe in utilizing tracking as a means of efficiency in student grouping practices. Following their work in (2004) where an extensive survey of elementary teachers indicated that over 70% of elementary teachers believed in some form of tracking as the most efficient means to group students for instruction, Ansalone and Biafora (2008) conducted a survey of 216 Long Island principals regarding the uses of tracking in their schools. The principals found that tracking was overwhelmingly used and that principal perceptions of tracking as an extension of magnifying societal inequalities differed based on the socio-economic level (SES) of the school. Higher SES school leaders were more concerned with inequalities

associated with tracking yet continued to use tracking as a sorting tool. While the general views espoused regarding the practice of tracking tended to be negative, the principals tended to still utilize such practices as they believed it to be the most efficient means to group students, regardless of their perceptions of the equity in dividing students between high- and low-tracked classes.

Weaknesses in the Existing Literature

Examining the role of tracking in student achievement is a difficult process as a result of the divergent ways that tracking is identified in the literature, as well as the complex relationship between other factors such as student motivation, family background, socioeconomic backgrounds, and parental support that may serve to further complicate the impact of student placement decisions at the schoolhouse level. Additionally, the forms and type of tracking employed by school leaders has changed significantly over time and is often highly localized in its form and process. The dramatic shift in tracking practices from broad programmatic distinctions to course specific ability grouping has resulted in difficulty in examining the impact of the practice in a consistent manner over time (Archbald, 2012). The rigidity inherent in tracking systems and design so evident in most of the 20th century has now largely fallen away to more politically acceptable forms predicated upon school choice and differentiated instruction (Loveless, 2009). Most formal school and district-based policies regarding tracking have disappeared, making it even more difficult to discern individual grouping practices at the school level and their potential impact upon student achievement for comparative purposes (Hallinan, 2004). The use of some form of tracking has defined the American public school system for so long that often the practice is so deeply embedded in the local culture that little or no attention may be paid to its impact on student outcomes (Watanabe, 2006). The decision of what and how

tracking systems will be employed remain a local decision, heavily influenced by administrators, teachers, and counselors within individual school systems. These realities make it critical to examine how students end up in certain courses and examining the impact of the sorting practice used by disparate middle schools that send to the same regional high school district provide a rich opportunity to extend gaps in the current literature.

Two major methodological challenges that faced researchers examining tracking practices had been the reliance upon self-reported data in the 1970s and 80s as well as the difficulty in separating the impact of course placement and grouping practices from pre-existing student differences that were already evident amongst students (Lucas, 1999; Gamoran, 2009). As researchers grappled with more modern and shifting forms of tracking they began to look at more specific student transcript data rather than rely upon student surveys and self-reporting as a source of examining track differences. Researchers in the 1980s and 1990s also began to explore in earnest the differences in classroom instruction that plagued different level courses. Clearly, numerous researchers had established sharp distinctions between the teaching and learning in higher level or advanced and honors courses and lower level and remedial courses in schools. More affluent areas also offered a wider range of honors and advanced courses for students to explore and challenge themselves (Spade, Columba, and Vanfosson 1997). By the close of the century, it had been established thoroughly in the literature that tracking and the practices associated with student grouping in schools was overall a highly subjective process that varied from district to district often in dramatic fashion (Useem, 1992; Oakes, 1995; Riehl, 2001). The sorting practices used by school officials also contributed to the widening of student gaps and began with de facto tracks established in the middle and sometimes elementary grades (Archbald, 2012).

Tracking in its modern form is often a stealth variable, making the research and exploration of such practices difficult, as there are highly divergent ways a student may end up placed in a particular course or academic pathway. School districts may employ automatic placement determinations that rely on specific factors related to prior achievement such as assessment scores or grades or rely upon more covert methods such as teacher recommendations and student choice in placing students in classes (Rubin, 2008). However, whether an overt or covert process, the research clearly indicates that movement between tracks is highly limited and that students tend to remain within whatever track they may be placed regardless of criteria (Oakes, 2005). How students end up placed in particular courses result from school-based practices that most often reflect a broad and eclectic mix of social and political factors (Oakes and Guiton, 1995).

Hallinan (2004) made the case that limited longitudinal data that can clearly compare tracked and non-tracked environments existed even in states such as California and Massachusetts that had adopted no tracking policies at the state level. The complexity of tracking, Hallinan argued, clearly allowed for individual school systems to employ a number of ways to continue to group students that were detrimental to those placed within lower level courses. The emergence of grit and persistence as a means of promoting student academic self-worth (Duckworth 2016) can be simply mitigated by policies, and more importantly, practices at the school level that continue to reflect a view of static student ability. Classes bearing the same name from one school to the next may have very different curriculum content. These hidden differences further complicate the issue of examining the impact of tracking as the curriculum may appear the same on paper but may be implemented with extreme variation based on student course assignment. Such subtle and nuanced differences are difficult to discern (Gamoran, 1993;

Watanabe, 2006). The disconnect between school grouping practices that promote tracking and federal and state accountability systems that measure student proficiency based on how students perform on the same standardized assessment further complicates the process of examining tracking practices in earnest. Students exposed to less curriculum, inferior instruction, lower expectations, and who are stigmatized from an early age within a K–12 system will not perform to the same level on standardized assessments, nor should they be expected to when examining the totality of their educational experience. Students are marginalized repeatedly within the K–12 experience based on their standardized assessment performance, which further complicates how the impacts of tracking may be determined. Tracking, in many ways, still remains a "black box" described by Oakes in 1987.

The difficulties of student self-reporting had also plagued earlier examinations of tracking (Rosenbaum 1980). Lucas and Gamoran (1993) re-examined the work of Lucas and Mare (1989) and concluded that tracked environments resulted in smaller achievement gaps than non-tracked environments with the use of specific course-based indicators utilizing *High School and Beyond* data of transcript data for 12,198 students in high school for the class of 1982. The transcript data was broken down into five categories related to post-secondary alignment. Gamoran found that black students were more likely to over report track position. Once school level variables were added and utilized student assessment performance in Grade 10 and Grade 12, race appeared as a non-factor but SES had persistent effects on track assignment. However, because blacks have on average lower achievement and higher percentages of students falling within a low SES category, tracking has a persistently negative impact on black achievement. Gamoran found the same phenomena holds true for Hispanic students. Self-reported data regarding track placement was useful in that there were clear links between student aspirations

and track placement, but the overall academic impacts were clouded unless multiple variables were controlled for in the examination. Any changes to tracking caused much more ambiguity as to where students actually were placed and were highly fluid and contextual to local districts.

Oakes and Guiton (1995) examined three large high schools in a west coast urban area in proximity to one another. All three had elaborate processes for student placement and faculty saw the process for course enrollment as merit based. However, when examining student achievement and track levels, there were a disproportionate number of students with the same achievement from poor and minority backgrounds as that of higher tracked students placed within lower level courses. They found that this had occurred as a result of more advantaged parents partaking of existing opportunities to use a waiver policy for admittance into higher level courses. This was a significant shift in the literature as the specific processes that favored one group of students over another in the course selection and enrollment process were being decontextualized to illuminate the subjectivity of local tracking processes. The structures in place and the corresponding communication to parents and guardians tended to favor the most advantaged students in the course enrollment process. The divergent ways students ended up in particular courses even when examining three distinct schools requires further examination as the researchers noted that tracking decisions do not necessarily follow a rational order but are often a combination of structural and cultural factors that come together in a subjective process to shape tracking decisions in schools. The impact of such processes favoring particular parent groups cannot be overstated, as those parents who are educated and are aware of the possibilities of challenging school-based sorting practices to change the placement in low-track courses for their children are rewarded in a disproportional manner to the students who have parents not aware of

such opportunities. These processes tend to occur along social and economic lines to mirror existing societal inequalities. Thus school-based practices extend preexisting social differences.

Betts and Shkolnick (1998) argued that the problem with most studies on tracking could be found in the difficulty of trying to place statistical controls over differences that cannot be controlled for and that most studies compared high-tracked and low-track classes which too often contain students with significant IQ differences and pre-test capabilities to begin with. The examinations of tracking were too often comparing apples and oranges and were not a fair examination of the overall impact of tracking on the student experience. In simplest terms, the pretest differentials in the study by Brewer, Argys, and Rees (1996) are too significant to adequately place controls over. In examining studies of track/no track studies where pretest differentials were similar it is important to note there were no significant gains for high ability students. This work included the examination of 15 randomized or matched experiments, none had effect size larger than .25 in favor of ability grouping for high achievers. As the benefits for high achievers cannot be substantiated, the argument that tracking provides the maximum benefit for high-achieving students is not fully supported within the literature. The corresponding overwhelming evidence in the literature over the past several decades on the negative impact to low and "average" tracked courses on overall student achievement serves as a stark contrast. This reality calls into question the need to continue to utilize grouping practices for efficiency at the expense of providing all students opportunities to maximize their academic experience across the K–12 spectrum.

The complexity of tracking and in determining how students are sorted into particular tracks remains a barrier to effectively making broad assumptions about the role of academic achievement in tracking decisions on a school or district-by-district basis. Race and economic

status have been identified as contributing factors to track placement (Wheelock, 1992; Oakes, 1985; Yonezawa, Wells and Serna, 2002). Moving forward from this paradigm, Archbald, Glutting, and Qian (2009) found that the research is inconclusive concerning how much track placement is determined by academic achievement and how much such placements are determined by other variables such as race or socioeconomic status, as high schools typically offer courses with title modifiers such as Advanced Placement (AP), honors, etc., and that these titles imply the principle of meritocratic selection which may mask subjective school-based processes underlying student placement decisions. Many researchers are highly skeptical of the existence of any form or meritocracy in student sorting decisions. (Wheelock, 1992; Oakes, 2005 and Yonezawa, Wells and Sema, 2002) all corroborate the impact of race and other external factors on tracking decisions in schools. A significant shortcoming of most tracking studies is that many rely on self-reported data where a full understanding of track placement is not always apparent or an oversimplification of more complex systems to sort students result in a limited understanding of the various factors at play in determining how a student is ultimately placed within a particular course. Also, many complex tracking systems are simplified for the purpose of conducting the research and lead to an incomplete understanding of the complexity of the particular school system being examined. These layers of subjective processes are not able to be effectively sorted through to gain a full understanding of the cacophony of practices that result in students being sorted for high- or low-track courses and mask much of the bias that may be inherent in such practices from school to school.

Mulkey et al. (2005) explored the multidimensional nature of tracking that made it difficult to understand from fragmented components of the literature, using large scale data from the *National Educational Longitudinal Study* from 1988. This data set allowed for the

comparison of students in tracked schools against students in untracked schools. Tracking in middle school begins to sequester students into smaller more homogeneous groups that limits the flexibility students are provided at the elementary level to demonstrate a range of ability. The curricular experience for students then becomes even less flexible in tracked systems as students enter high school. Students in tracked middle school environments are then exposed to a much more restrictive high school curriculum. Although students in tracked and untracked systems may enroll in the same courses and engage in the same curriculum, over time, tracking mediates the instructional pathways for students through the pace and content of subject matter presented in class. Classes bearing the same name from one school to the next may have very different curriculum content, further complicating the issue of examining the impact of tracking. Though the *Common Core National Standards* may limit this academic range explored by grade level, the sheer volume of courses used to sort students in tracked systems, particularly in the mathematics at the high school level, may still make overcoming this sharp variability of student academic experience impossible. Curriculum may look similar but instructional variation may be significant within schools (Watanabe, 2006). Although Watanabe only examined a small group of instructors, his findings of instructional variance within one school are disturbing when placed within the context of what students are learning from one course to the next (or not) as a result of curriculum divergence across academic tracks. However, this is important to demonstrate a divergence between accountability systems and tracking systems. This more modern form of tracking with hybrid programmatic distinction and subject specific tracking has become known as neotracking (Mickelson and Everett, 2008). As state and federal policies have increased academic expectations across the nation, preexisting tracking practices work against increasing student academic achievement as students in lower tracks are exposed to larger

qualities of test preparation and a general lower quality of instruction (Ansalone, 2010; Gamoran 2009).

Archbald, Glutting, & Quan (2009) argued that the shortcomings of prior examinations of tracking included a reliance upon student self-reporting and the collapsing of multiple course levels into college preparatory and non-college preparatory courses for comparison sake as well as the subjectivity of the appropriate measurements of prior grades. While they only examined data from one high school of approximately 1,100 students, they explored three factors that led to ninth grade course placement for students. These included end-of-course grades in Grade 8, state assessment scores, and race. They found no statically significant role of race in placement decisions but did find that prior grades were the most critical determinant for subsequent course placement. However, the authors noted other factors such as teacher expectations, curriculum experience, and perpetual assignment to lower academic tracks are ignored in such an overly simplistic determination of track impact on overall academic achievement. If students were performing poorly in Grade 8, the cumulative possibilities as to how and why those students were performing poorly are not captured in such an analysis. The long-term impacts of the subjective student placement decisions that ultimately resulted in poor academic performance was simply not able to be captured in such examinations.

The complexity of tracking and how schools utilize ability grouping makes examining true tracking impacts a convoluted endeavor as it must be examined school by school (Archbald and Keleher, 2008). To thoroughly examine the impacts of tracking, schools must monitor detailed patterns of course placement decisions over time, examine the progressions of students moving upward or downward through inter track mobility statistics, identify the specific academic effects of each track placement, and course-taking patterns. In short, in many ways,

the gap between research and specific practices at the school level are profound (Loveless, 1999; Hallinan, 2004). Without such robust and nuanced data systems, schools are ignoring the impacts of their own tracking decisions. Often, as a result of *NCLB*, standardized assessment scores have been allowed to serve as the singular metric to continue traditionally subjective sorting processes that limit opportunities for students, whether through leveling, teacher recommendation, or school practice that ignored the complexity of an individual student journey through a K–12 institution. Schools need to embrace a systems approach to identify individual student needs, strengths and aspirations within the larger system, and to utilize data in a more unconventional and district specific manner to truly unlock the potential of each student. This would broaden the scope of our understanding of opportunities and achievement on a school-by-school basis while simultaneously expanding ongoing work grounded in the concepts of equity and excellence. Our modern educational systems where standardized assessments are used as sorting mechanisms in the guise of accountability measures is the perfect system for tracking practices to unleash the most nefarious outcomes possible. School leaders need to develop the sophistication to embed the proactive use of programmatic, instructional, course-taking patterns and curricular experiences to ensure every student receives the education deserved, and needed, to fuel our democracy.

Chapter III:

Methodology

This quantitative research study was conducted in order to explain the influence of mathematics course placement on student achievement in relation to achievement on standardized assessments and math attainment in a regional high school setting. Presently, there is limited quantitative research that explains student math achievement in a regional high school setting. This study is a one district case study, examining the mathematics outcomes for students enrolled in Grade 8 in 2011 who graduated from high school in 2015. The study measured the influence of ninth grade mathematics course placement on math attainment as measured by the number of years a student enrolled in a mathematics course, the highest math course attained, and performance on the state standardized assessment while controlling for specific demographic variables including free and reduced lunch status and special education classification. I utilized a sorting mechanism to code course placements according to pipeline levels generated on a 0–9 scale, building upon previous examinations of mathematics course progression and outcomes (Ready, Hatch, Warner & Chu, 2013, Table 1). This study adds to the extensive literature on school tracking providing policy makers and school officials with specific data and evidence to most appropriately design courses of study and course-taking patterns that promote the long-term academic success of all students.

Table 1

Pipeline Levels Assigned to Mathematics Courses

Mathematics Pipeline Levels	
Level	
0	No Corresponding Math Courses
1	Pre-Algebra, Basic Math, Consumer Math
2	Algebra 1
3	Geometry
4	Algebra II
5	Advanced Algebra and/or Geometry (Trigonometry, Discrete Math etc.)
6	Pre-Calculus
7	Calculus
8	AP Calculus, Statistics, or Computer Science
9	Advanced topics (multivariate, etc.)

Research Design

The research design was a correlational, explanatory longitudinal design with quantitative methods. The study utilized data from the graduating class of 2015 from a large suburban regional high school district in central New Jersey. This type of design was appropriate since I examined how a number of variables were related to a major complex variable and to what degree this relationship exists between two or more quantifiable variables (Gay, Mills, & Airasian, 2012). This design will allow the researcher to predict the influence of the variables on the major complex variable. While high correlation between variables does not imply that one variable causes another, high correlation allows increased accuracy in predicting outcomes (Gay et al., 2012).

A correlational, explanatory, longitudinal design was used to explain the relationship that exists between the course placements for mathematics in ninth grade on student math attainment in the regional high school setting. The longitudinal design allows for the explanation of

relationships over time. The correlational design was used to determine the relationship between mathematics course placement in the ninth grade and overall academic performance in the regional high school setting. Correlational studies attempt to understand patterns of relationships amongst variables. Correlational research is useful in determining whether and to what degree a relationship exists between two or more quantifiable variables (Johnson, 2001). For the purpose of this study, the determination between completion of a higher level mathematics course was represented by the completion of pre-calculus during high school. Additional higher level mathematics courses included calculus, statistics, computer science, and multi-variable calculus.

School achievement data for this study included NJASK 8 assessments for the year 2011, and Grade 11 HSPA assessments for mathematics for 2014. The New Jersey Department of Education officials designate schools into eight distinct district factor groups that correspond to census data related to wealth. These designations include A, B, CD, DE, FG, GH, I, and J as codes (DFG) for categorizing each New Jersey school district. The four sending districts utilized for this study represent DFGs of B, GH, FG, and I respectively.

Regression analysis is a form of predictive modeling that includes examining relationships between dependent and independent variables to predict future outcomes (Price, 2016). Simultaneous regression is useful when predicting a dependent variable from two or more independent variables (Leech, Barrett, & Morgan 2011). The researcher utilized a simultaneous regression model to determine the relationship between performance on the New Jersey High School Proficiency Assessment (HSPA) [dependent variable] and mathematics pipeline levels for Grade 9, Grade 11 as [independent variables]. With multiple independent variables, multiple regression was an effective method to examine the relationship between mathematics pipeline levels and performance on the HSPA (Figure 1).

Hierarchical regression models simultaneously investigate relationships within and between hierarchical levels of grouped data, thereby making it more efficient at accounting for variance among variables at different levels other than existing analysis (Woltman, Feldstain, Mackey, & Rocchi 2013). A hierarchical regression model was used to determine the influence of each independent variable on the dependent variable. The researcher was able to distinguish between the variables with both significance and strength of the correlation by performing a hierarchical regression. Data normality was met by running a test for skewness on the HSPA data. The data were within the normal range of +/- 1.000. The researcher then calculated frequencies for dichotomous variables including gender, Grade 9 mathematics pipeline level, Grade 12 pipeline level, and ran frequency tables (Tables 2–4). A Pearson correlation matrix was developed to identify any significant relationships between the independent and dependent variables. Each pipeline level and gender was examined for statistical significance on HSPA performance. Each pair was examined to determine a positive, negative, or no correlation.

Research Questions

1. How is performance on the New Jersey High School Proficiency Assessment influenced by ninth grade mathematics course placement when controlling for student demographic factors?
2. How is mathematics course attainment by Grade 12 influenced by ninth grade mathematics placement when controlling for student demographics?

Null Hypothesis

Null Hypothesis 1: No statistically significant relationship exists between ninth grade mathematics placement and student performance on the NJ HSPA when controlling for student demographics.

Null Hypothesis 2: No statistically significant relationship exists between twelfth grade mathematics placement and student mathematics placement in Grade 8 when controlling for student demographics

Sample Population/Data Source

The sample for this study consisted of students from the graduating class of 2015 attending six suburban regional high schools in central New Jersey. The six high schools in this study are public schools, contain Grades 9–12, and report all student demographic information to the New Jersey Department of Education through NJSMART, the New Jersey Department of Education Data System (NJDOE). The sample consisted of 1,223 students. The regional high school system receives students from seven separate and independent public K–8 school systems. Data from three of the sending districts was excluded because of the inability to confirm all accuracy of student placement and achievement in the middle grades prior to transition to the regional high school. While the study controlled for free and reduced lunch status, English Language Learner status as well as special education classification, there are some demographic and socioeconomic differences between the four sending districts used for the purposes of this study as the districts represent divergent DFGs, and include B, GH, FG, and I districts. Students who were not continuously enrolled from 2011–2015 in the regional high school district or those who did not have a valid math placement from Grade 8 in 2011 and Grade 12 in 2015 were excluded from this study.

Data Collection

The data for this study were retrieved from the regional high school district data base utilized for student placement for incoming ninth graders in 2011. Additionally, student performance data representing course progression, earned grades, and standardized assessment

scores were identified over a four-year period from school year 2011–2012 to school year 2014–2015. The data was migrated to a single excel spreadsheet. Specific school performance data was retrieved from each of the six high schools comprising the regional high school district. Students with Individual Educational Plans (IEPs), English Language Learners, and students from free and reduced lunch backgrounds were excluded from the data collection. Students who did not attend the schools continuously for four years were also excluded. Sending school data that did not report portions of their student performance in Grades 7 and 8 and who could not verify the accuracy of their data system for Grade 8 placement in 2011 were excluded from the study. A master spreadsheet was then developed that included performance on the Grade 7 New Jersey Assessment of Skills and Knowledge (NJASK), Grade 8 NJASK, raw score on the mathematics section of the High School Proficiency Assessment (HSPA), course assignments in mathematics for all courses, Grades 7–12, grades earned in each course, and college acceptance (Table 5). The researcher then assigned a pipeline level of mathematics to gauge math attainment over four years for each math course 9–12 in order to designate between advanced and non-advanced course attainment (Ready, et. al., 2013). The pipeline level was used to distinguish between advanced mathematics study and non-advanced mathematics study in relation to the highest level course taken by a student. For the purpose of this study, advanced mathematics attainment was designated by completion of any of the following courses in high school: pre-calculus, calculus, statistics, or computer science and multivariate calculus. Longitudinal data are presently not available for the current state PARCC assessment and thus those data were not used in this study. Mathematics courses were coded in the following manner: higher tracked 8th grade mathematics courses were assigned a 1. At grade level or lower mathematics courses were assigned a 0. Higher tracked 11th grade mathematics courses were

assigned a 1. At grade or lower mathematics courses were assigned a 0. Students assigned to intermediate Algebra 1 were coded 1; all other ninth grade algebra students were coded 0. Students who took four years of mathematics were coded 1. Those students who took fewer than 4 years of mathematics courses were coded 0. Gap years, or years where students did not take a mathematics course were coded 0. Pipeline levels were assigned to all students in Grade 9, Grade 11, and Grade 12.

The clean and formatted data were in correct form to be imported into IBM's SPSS statistical software.

Table 2

Grade 9 Course Levels 2011-2012

9th Grade Pipeline Levels		
Level	Frequency	Percentage
2	669	54.3
3	495	40.2
4	68	5.5

Table 3

Grade 12 Course Levels 2014-2015

12th Grade Pipeline Levels		
Level	Frequency	Percentage
0	129	10.5
1	0	0
2	0	0
3	4	.3
4	19	1.5
5	431	35
6	243	19.7
7	101	8.2
8	295	23.9
9	10	.8

Table 4

Gender Breakdown of Class of 2015

Mathematics Gender Count 2011–2015		
Gender	Number	Percentage
Male	622	50.5
Female	610	49.5

Table 5

Data retrieved from Regional High School Data Base and Utilized in Excel Spreadsheet

Student State Identification
Resident District
Attending High School
Gender
NJASK 7 Math Score
NJASK 8 Math Score
HSPA Math Score
9th Grade Math Course Code
9th Grade Math Grade Earned
10th Grade Math Course
10th Grade Math Grade Earned
11th Grade Math Course
11th Grade Math Grade Earned
12th Grade Math Course
12th Grade Math Grade Earned

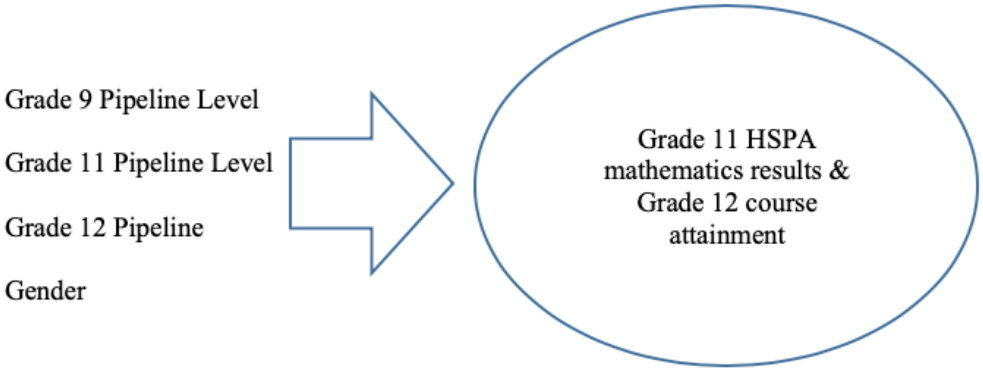

Figure 1: Simultaneous regression framework

Data Analysis

The research methodology used within this study examined the data to ensure it met the assumptions of regression. According to Morgan et al. (2013), "the relationship between each of the predictor variables and the dependent variables are linear, the errors are normally distributed, and the variance of the residuals (difference between actual and predicted scores) is constant" (p 164). The data met the assumption for regression. The researcher conducted a simultaneous multiple regression model and correlation coefficient matrix that included all independent variables for course placement, Grade 9 mathematics pipeline level, Grade 12 mathematics pipeline level, and gender. This matrix allowed the researcher to recognize variables that are potentially statistically significant, insignificant, and allowed for the identification of variables that had potential for multi-collinearity. The simultaneous regression model allowed me to identify potentially statistically significant variables. Variables were all under 5 in terms of VIF statistic. Statistical significance of the models were calculated on being able to identify a $p <$ .05. The coefficients matrix was analyzed to determine statistical significance of the predictor variables. After, simultaneous regression variables were ranked in order from highest to lowest beta value. Hierarchical models allowed for the identification of the influence of specific predicator variables on the dependent variable. The initial analysis involved determining if the dependent variable met the test of normality. A test of skewness was run on the dependent variable. A general rule is that if skewness is less than -1 or greater than 1 the distribution is skewed. The skewness was within normal limits. Descriptive statistics were also run including means and standard deviations. Additionally, I calculated the means for dichotomous variables. The dichotomous variables included gender, Grade 9 mathematics pipeline level, and Grade 12 mathematics pipeline level. Frequency tables were run for the dichotomous variables (Table 3).

Chapter IV:

Analysis Of Data

My purpose for this correlational, explanatory, longitudinal case study with quantitative methods was to explain the influence of Grade 9 mathematics course placement on student outcomes in relation to achievement on standardized assessments, and ultimate math course level attainment in a regional high school setting. This study focused on independent school level variables for the class of 2015, including ninth grade math course placement, eventual eleventh grade math course placement, and gender to predict the variance in the dependent variable which was performance on the 2014 New Jersey High School Proficiency Assessment (HSPA). By focusing on course-taking patterns and their impact to overall student achievement this study extended the work of Oakes, 2005, Gammoran, 1992, and Ansalone, 2010, that described the negative outcomes associated with the use of tracking in American public schools. This study utilized simultaneous and hierarchical regression models to analyze the impact the independent variables on the dependent variables.

Research Questions

1. How is performance on the New Jersey High School Proficiency Assessment influenced by ninth grade mathematics course placement when controlling for student demographic factors?
2. How is mathematics course attainment by Grade 12 influenced by ninth grade mathematics placement when controlling for student demographics?

Procedure

The analysis conducted was to determine the normalcy of the continuous dependent variable of HSPA data to ensure the data met the assumptions of normality. Table 6 shows the

descriptive statistics for the dependent variable. To determine data normalcy, a test of descriptive statistics and test of skewness on the HSPA test data was run. 1,231 valid cases were identified. The mean was 241.85, the standard error was .598, the minimum was 160, and the maximum was 300. The skewness was within normal limits at -.833 with a standard error of .07. The HSPA data met the skewness requirements.

Table 6

Descriptive Statistics

Case Processing Summary

	Cases					
	Valid		Missing		Tota	
	N	Percent	N	Percent	N	Percent
HSPA	1231	99.9%	1	0.1%	1232	100.0%

Descriptives

			Statistic	Std. Error
HSPA Math	Mean		241.85	.598
	95% Confidence Interval for Mean	Lower Bound	240.68	
		Upper Bound	243.03	
	5% Trimmed Mean		242.80	
	Median		246.00	
	Variance		440.793	
	Std. Deviation		20.995	
	Minimum		160	
	Maximum		300	
	Range		140	
	Interquartile Range		29	
	Skewness		-.633	.070

Next, I calculated the frequencies for dichotomous variables. The dichotomous variables included gender, Grade 9 pipeline level, and Grade 12 pipeline level. A frequency table was run through the SSPS software. The sample included general education students who were not on

free and reduced lunch, not classified, did not receive English language assistance and who had spent all four high school years as a student within the regional high school district. The researcher then assigned a pipeline level of mathematics to gauge math attainment over four years for each math course, 9–12 in order to designate between advanced and non-advanced course attainment (Ready, et. al., 2013). The pipeline level was used to distinguish between advanced mathematics study and non-advanced mathematics study in relation to the highest level course taken by a student. For the purpose of this study, advanced mathematics attainment was designated by completion of any of the following courses in high school: pre-calculus, calculus, statistics, or computer science and multivariate calculus. Longitudinal data are presently not available for the current state PARCC assessment and thus those data were not used in this study. Mathematics courses were coded in the following manner: higher tracked eighth grade mathematics courses were assigned a 1. At grade level or lower mathematics courses were assigned a 0. Higher tracked 11^{th} grade mathematics courses were assigned a 1. At grade or lower mathematics courses were assigned a 0. Students assigned to Intermediate Algebra 1 were coded 1, all other ninth grade algebra students were coded 0. Students who took four years of mathematics were coded 1. Students who took less than four years of mathematics courses were coded 0. Gap years, or years where students did not take a mathematics course were coded 0. Pipeline levels were assigned to all students in Grades 9, 11, and 12.

The clean and formatted data were in correct form to be imported into IBM's SPSS statistical software. For gender there were 622 females or 50.5% of the sample. There were 610 males or 49.5% of the sample (Table 7).

Table 7

Frequency Table for Gender

Statistics

			12th Pipeline	9th Pipeline
N	Valid	1232	1232	1232
	Missing	0	0	0

Gender1 Male

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	622	50.5	50.5	50.5
	1.00	610	49.5	49.5	100.0
	Total	1232	100.0	100.0	

Grade 9 pipeline levels included three potential categories: pipeline level 2 (Algebra 1), pipeline level 3 (Geometry), or pipeline level 4 (Algebra 2). A frequency table was run. 669 students were in pipeline level 2 taking Algebra 1 and representing 54.3% of the total student population. 495 students were in pipeline level 3, taking Geometry and representing 40.2% of the total student population. 68 students were in pipeline level 4, taking Algebra 2 as ninth graders, representing 5.5% of the total population (Table 8). A ninth grader determined to be at grade level would enter the regional district enrolled in some form of algebra. The frequency table indicated that a total of 45.7% of the Grade 9 population entered the regional district in advanced mathematics courses, geometry and above, as entering ninth graders.

Table 8

Grade 9 Pipeline Levels

9th Pipeline Level

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	669	54.3	54.3	54.3
	3	495	40.2	40.2	94.5
	4	68	5.5	5.5	100.0
	Total	1232	100.0	100.0	

A frequency table was then run to identify the mathematics course distribution and corresponding pipeline levels for the same students in Grade 12 (Table 9).

Table 9

Grade 12 Pipeline Levels

12th Pipeline Level

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	129	10.5	10.5	10.5
	3	4	.3	.3	10.8
	4	19	1.5	1.5	12.3
	5	431	35.0	35.0	47.3
	6	243	19.7	19.7	67.0
	7	101	8.2	8.2	75.2
	8	295	23.9	23.9	99.2
	9	10	.8	.8	100.0
	Total	1232	100.0	100.0	

Grade 12 pipeline levels included 129 students who did not take a mathematics course. This represented 10.5% of the sample. Four students took some form of pre-algebra or basic math representing .3% of the total population. 19 students took Algebra 1 representing 1.5% of

the total population. 431 students took advanced algebra, trigonometry, or discrete mathematics representing 35% of the sample. 243 students took pre-calculus, representing 19.7% of the total population. 101 students took calculus, representing 8.2% of the total population. 295 students took Advanced Placement calculus, Advanced Placement statistics, or computer science representing 23.9% of the population. Ten students took multivariate calculus, representing .8% of the population. In examining the pipeline levels the number of students from Grade 9 to Grade 12 who were enrolled in courses considered advanced as designated by the pipeline levels rose from 45.7% of the total population to 52.6% of the total population over the four-year duration captured in the study.

Pearson Correlation Matrix

Following completion of the descriptive statistics, a Pearson Correlation Matrix was run. Pearson correlations examine relationships between and among independent and dependent variables. Correlation coefficient values range from -1.00 to 1.00 and can be positive or negative. A positive value implies a positive relationship between variables while a negative value implies a negative relationship between variables. A moderate positive statistically significant correlation between Grade 9 pipeline level and Grade 12 pipeline level .422 was found. There was a weak, non-significant correlation with gender and pipeline levels (Table 10).

Table 10

Correlation Table for Grade 9 & 12 Pipeline Levels and Gender

Correlations

		12th Pipeline Level	9th Pipeline Level	Gender1Male
Pearson Correlation	12th Pipeline Level	1.000	.422	.045
	9th Pipeline Level	.422	1.000	.031
	Gender1Male	.045	.031	1.000
Sig. (1-tailed)	12th Pipeline Level	.	.000	.056
	9th Pipeline Level	.000	.	.136
	Gender1Male	.056	.136	.
N	12th Pipeline Level	1232	1232	1232
	9th Pipeline Level	1232	1232	1232
	Gender1Male	1232	1232	1232

Descriptive statistics were run on the HSPA math scores for the student population (Table 11). The mean HSPA math score was 241.85 with a standard deviation of 20.995.

Table 11

Descriptive Statistics HSPA Math Scores

Descriptive Statistics

	Mean	Std. Deviation	N
HSPA Math	241.85	20.995	1231
9th Pipeline Level	2.51	.599	1231
Gender1Male	.4947	.50018	1231

Following completion of the descriptive statistics, a Pearson Correlation Matrix was run. Pearson correlations examine relationships between and among independent and dependent variables. Correlation coefficient values range from -1.00 to 1.00 and can be positive or

negative. A positive value implies a positive relationship between variables while a negative value implies a negative relationship between variables. A weak positive statistically significant correlation between gender and HSPA performance of .161 was found. A moderate statistically significant correlation between Grade 9 pipeline level and HSPA performance of .532 was found. A moderate, statistically significant correlation between Grade 11 pipeline level and HSPA performance of .589 was found (Table 12).

Table 12

Correlation Table Gender, Grade 9 & 11 Pipeline Levels & HSPA Performance

Correlations

		HSPA Math	Gender1Male	9th Pipeline Level
Pearson Correlation	HSPA Math	1.000	.161	.532
	Gender1Male	.161	1.000	.029
	9th Pipeline Level	.532	.029	1.000
	11th Pipeline Level	.589	.017	.789
Sig. (1-tailed)	HSPA Math	.	.000	.000
	Gender1Male	.000	.	.152
	9th Pipeline Level	.000	.152	.
	11th Pipeline Level	.000	.275	.000
N	HSPA Math	1231	1231	1231
	Gender1Male	1231	1231	1231
	9th Pipeline Level	1231	1231	1231
	11th Pipeline Level	1231	1231	1231

Results and Interpretation of Hierarchical Regression Model for HSPA Mathematics Performance

In order to determine the relationship between the independent and dependent variables, the hierarchical regression models determine the influence between Grade 9 pipeline level, Grade 11 pipeline level, gender, and performance on the HPSA Math Assessment. Three models

were created. Table 13 shows how the independent variables were placed into the hierarchical regression model. The models were assessed at the .05 level of significance where $p < .05$ (Gay, Mills & Airasian 2012).

Table 13

Variables Entered for HSPA Math Performance

Variables Entered/Removed [a]

	Variables Model Entered	Variables Removed	Method
1	Gender1Male[b]	.	Enter
2	9th Pipeline Level[b]	.	Enter
3	11th Pipeline Level[b]	.	Enter

a. Dependent Variable: HSPA Math
b. All requested variables entered.

For Model 1 the predictor percentage of gender reports and *R* Square of .026 and explained 2.6% of the variance in the dependent variable. In Model 2, the predictor percentage of Grade 9 pipeline level reports an *R* Square of .304 and explains 30.3% of the variance in the dependent variable. Model 2 contains an R Square change of .278, indicating that Grade 9 pipeline level accounts for 27.8% of the combination of predictors and explains of the variance in the dependent variable. The R Square change from Model 2 to Model 3 is .077, which shows that an additional 7.7% of the variance is now added with Grade 11 pipeline level. Model 3 explains 38.1% of the variance in performance on the mathematics section of the HSPA (Table 14). The R Square change was statistically significant $F(1, 1{,}229) = 251.923$, $p = .000$. Therefore, Model 3 explains the greatest amount of variance in the dependent variable.

Table 14

Model Summary Independent Variables Variance on Dependent Variable

Model Summary [d]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics		
					R Square Change	F Change	df1
1	.161[a]	.026	.025	20.728	.026	32.898	1
2	.551[b]	.304	.303	17.530	.278	490.414	1
3	.617[c]	.381	.380	16.536	.077	152.972	1

The Interpretation of Two-way ANOVA Table for Hierarchical Model for 2015 NJ HSPA mathematics

This two-way ANOVA estimates the impact of three main effects on the dependent variable in the model of best fit. The ANOVA demonstrates that the selected Model, Model 3, is statistically significant at the .000 level F=251.923, *df*= 3, 1,227 (Table 15).

Table 15

Two-way ANOVA for 2015 HSPA Mathematics

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	14134.642	1	14134.642	32.898	.000[b]
	Residual	528041.154	1229	429.651		
	Total	542175.795	1230			
2	Regression	164830.919	2	82415.459	268.206	.000[c]
	Residual	377344.876	1228	307.284		
	Total	542175.795	1230			
3	Regression	206660.256	3	68886.752	251.923	.000[d]
	Residual	335515.540	1227	273.444		
	Total	542175.795	1230			

a. Dependent Variable: HSPA Math

b. Predictors: (Constant), Gender1Male

c. Predictors: (Constant), Gender1Male, 9th Pipeline Level

d. Predictors: (Constant), Gender1Male, 9th Pipeline Level, 11th Pipeline Level

Coefficient for 2015 HSPA Mathematics

The coefficient table within the hierarchical model demonstrates how each predictor influences the dependent variable. In Model 1 the predictor percentage of gender reports a beta =.16. It is statistically significant at the .000 level, t=5.736. The beta is positive, which means that as the population of males increases, scores on the 2015 HSPA mathematics section increase. In Model 2, the predictor percentage of males population decreases in power from .161 to .146. It is significant at the .000 level, t=6.129. The predictor added in Model 2, Grade 9 pipeline level, reports a beta of = .527. It is significant at the .000 level, t=22.145. The positive beta indicates that as the Grade 9 pipeline level increases, performance on the 2015 HSPA mathematics increases. In Model 3, the predictor percentage of gender increases slightly in power with a beta = .149. It is significant at the .000 level, t= 6.620. The predictor of Grade 9 pipeline level loses power to a beta = .171. It is significant at the .000 level, t= 4.678. The predictor added in Model 3, Grade 11 pipeline level, reports a beta= .452. It is significant at the .000 level, t= 12.368.

Table 16

Standardized Coefficient Betas and Tolerance for Hierarchical regression Model for 2015 NJ HSPA Mathematics

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	238.498	.831		286.961	.000
	Gender1Male	6.777	1.182	.161	5.736	.000
2	(Constant)	192.415	2.196		87.603	.000
	Gender1Male	6.128	1.000	.146	6.129	.000
	9th Pipeline	18.481	.835	.527	22.145	.000
3	(Constant)	194.736	2.080		93.602	.000
	Gender1Male	6.244	.943	.149	6.620	.000
	9th Pipeline	5.990	1.281	.171	4.678	.000
	11th Pipeline	6.136	.496	.452	12.368	.000

Coefficients[a]

Model		Collinearity Statistics	
		Tolerance	VIF
1	(Constant)		
	Gender1Male	1.000	1.000
2	(Constant)		
	Gender1Male	.999	1.001
	9th Pipeline Level	.999	1.001
3	(Constant)		
	Gender1Male	.999	1.001
	9th Pipeline Level	.378	2.648
	11th Pipeline Level	.589	.333

a. Dependent Variable: HSPA Math

The VIF for all predictors in all models fell below the threshold of 5 which indicate the models do not have a multicollinearity problem.

Grade 12 Pipeline Level

The first simultaneous regression analysis of the independent variables gender and Grade 9 pipeline level will determine the relationship to the dependent variable Grade 12 pipeline level. The variables entered, model summary, and ANOVA tables for the first simultaneous regression are reflected in Tables 6, 7, and 8 respectively. The model was assessed at the .05 level of significance where $p < .05$ (Gay, Mills & Airasian 2012). The means and standard deviations were calculated for the dependent and independent variables used in the regressions (Table 17). The mean percentage of Grade 12 pipeline level was 5.57 with a standard deviation of 2.267. The mean percentage of Grade 9 pipeline level was 2.51 with a standard deviation of .600. The mean percentage of gender/male was .4951 with a standard deviation of .50018.

Table 17

Grade 12 Pipeline Level Descriptive Statistics

Descriptive Statistics

	Mean	Std. Deviation	N
12th Pipeline	5.57	2.267	12
9th Pipeline Level	2.51	.600	12
Gender1Male	.4951	.50018	12

Pearson Correlation Grade 12 Pipeline Level

After the descriptive statistics, I ran a Pearson Correlation Matrix. I found a moderate positive statistically significant correlation of .422 between Grade 9 pipeline level and Grade 12 pipeline level. There was a weak, non-significant correlation of .031 with gender and Grade 12 pipeline level.

Table 18

Correlations

Correlations

		12th Pipeline Level	9th Pipeline Level	Gender1Male
Pearson Correlation	12th Pipeline Level	1.000	.422	.045
	9th Pipeline Level	.422	1.000	.031
	Gender1Male	.045	.031	1.000
Sig. (1-tailed)	12th Pipeline Level	.	.000	.056
	9th Pipeline Level	.000	.	.136
	Gender1Male	.056	.136	.
N	12th Pipeline Level	1232	1232	1232
	9th Pipeline Level	1232	1232	1232
	Gender1Male	1232	1232	1232

Simultaneous Regression Model Grade 12 Pipeline Level

The simultaneous regression model determines the influence of the independent variable on the dependent variable. The simultaneous regression model was statistically significant at the .000 level. The R Square value was .179 with an F Change value of 133.803. The model explained 17.9% of the variance in the dependent model.

Table 19

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Gender1Male, 9th Pipeline Level[b]	.	Enter

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics		
					R Square Change	F Change	df1
1	.423[a]	.179	.177	2.056	.179	133.803	2

This two-way AOVA estimates the impact of two independent variables on the dependent variable in the model of best fit. The ANOVA demonstrates that Model 1 is statistically significant at the .000 level F=133.803, *df*= 2, 1,229.

Table 20

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1131.237	2	565.618	133.803	.000[b]
	Residual	5195.305	1229	4.227		
	Total	6326.542	1231			

a. Dependent Variable: 12th Pipeline Level

b. Predictors: (Constant), Gender1Male, 9th Pipeline Level

The coefficients table indicates that Grade 9 pipeline level contributes most to Grade 12 pipeline level with a standardized beta of .421. Grade 9 pipeline level was statistically significant at the .000 level.

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.505	.257		5.854	.000
	9th Pipeline Level	1.588	.098	.421	16.264	.000
	Gender1Male	.146	.117	.032	1.247	.213

The VIF for all predictors in all models fell below the threshold of 5 which indicate the models do not have a multicollinearity problem.

Model		Tolerance	VIF
1	(Constant)		
	9th Pipeline	.999	1.001
	Gender1Male	.999	1.001

a. Dependent Variable: 12th Pipeline Level

Research Questions

Research Question 1: How is performance on the New Jersey High School Proficiency Assessment influenced by ninth grade mathematics course placement when controlling for student demographic factors?

Answer: Ninth grade mathematics course placement explained explains 27.8 % of the student performance on the NJHSPA Mathematics section. Grade 11 pipeline level explains

7.7% of the student performance on the NJHSPA Mathematics section. Gender explains 2.6% of the student performance on the NJHSPA Mathematics section.

Research Question 2: How is mathematics course attainment by Grade 12 influenced by ninth grade mathematics placement when controlling for student demographics?

Answer: Grade 9 mathematics course placement accounts for 17.9% of Grade 12 course placement when controlling for student demographics.

Null Hypothesis 1: No statistically significant relationship exists between ninth grade mathematics placement and student performance on the NJ HSPA when controlling for student demographics.

The null hypothesis was rejected. A statistically significant relationship exists between Grade 9 mathematics course placement and performance on the NJHSPA Mathematics section.

Null Hypothesis 2: No statistically significant relationship exists between twelfth grade mathematics placement and student mathematics placement in Grade 9 when controlling for student demographics.

The null hypothesis was rejected. A statistically significant relationship exist between Grade 9 mathematics course placement and Grade 12 mathematics course attainment.

Chapter V:

Conclusions And Recommendations

The purpose of this study was to explain the influence of Grade 9 mathematics course placement on student outcomes in relation to achievement on standardized assessments, and ultimate math course level attainment in a regional high school setting. This study focused on independent school level variables for the class of 2015, including ninth grade math course placement, eventual eleventh grade math course placement, and gender to predict the variance in the dependent variable which was performance on the 2014 New Jersey High School Proficiency Assessment (HSPA). By focusing on course-taking patterns and their impact to overall student achievement this study extended the work of Oakes, 2005; Gammoran, 1992; and Ansalone, 2010 that described the negative outcomes associated with the use of tracking in American public schools.

Using a correlational, explanatory, longitudinal case study with quantitative methods this study focused on two overarching research questions:

1. How is performance on the New Jersey High School Proficiency Assessment influenced by ninth grade mathematics course placement when controlling for student demographic factors?
2. How is mathematics course attainment by Grade 12 influenced by ninth grade mathematics placement when controlling for student demographics?

The results of this study support the findings of previous research by Oakes (2005), Boaler (2007), Ansalone (2010), and Kelly and Carbanero (2012) that found tracking to be the dominant means to sort students for instruction and that such course placements were highly static within secondary schools. However, this study differs in scope as it followed a core group of students

through their four-year mathematics course placement within a regional school system that was in the early stages of dismantling traditional tracking structures. As a result, this study is a more nuanced examination of how specific school systems may mitigate tracking impacts over time. The results of this study extend the literature and research in this area to provide predictive power to the importance of course placement decisions on larger measures of student outcomes on standardized assessments and through course attainment.

The results of this study illuminate the negative impacts of the entrenched manner through which students are placed within courses at the secondary level. Without direct interventions to provide for more flexibility in course taking for all students, school systems are negatively impacting student opportunities for growth and potentially impacting performance on gatekeeping standardized assessments for entrance into college and university. This study confirms previous research that students often placed in courses in a highly stratified manner for ease of placement through teacher recommendation and administrative process may have far reaching negative impacts to future opportunities.

The results of this study indicate that school leaders should be actively investigating the practices and policies within their own schools and school systems around placement decisions and simultaneously work toward developing direct interventions to mitigate traditional course placement strategies throughout the system to ensure that all students are provided the same opportunities and support for long-term academic success.

Traditionally, the academic trajectory of individual students within schools has been relatively fixed (Kelley & Carbonero, 2012). Additionally, many of the practices used to group and sort students often have more far-reaching consequences for future student outcomes than known by school officials making those placement decisions (Oakes, 2005). However, this

research sheds light on a number of practices and actions that if followed by school leaders, can ensure more equitable outcomes for all students by specifically altering the academic trajectories of individual students within the larger school structures. Rather than following perfunctory pathways strengthened by the status quo, student scheduling should be approached from a much broader and more holistic approach that takes into consideration multiple sources of information to make individual student recommendations. These scheduling decisions need to be supported by additional school-based opportunities and supports that both close academic gaps and provide year over year opportunities to take more challenging course loads. The placement of a ninth grader is important for overall academic success, but a secondary school can indeed alter each individual student journey by embracing practices that mitigate tracking decisions while pressing policymakers to provide increased flexibility in course-taking requirements and expectations for all students.

Recommendations for Policy

Placement decisions have a direct impact on standardized assessment outcomes. If we insist on having high-stakes graduation tests, then we are doing a disservice to a large group of our students not placed in more challenging courses over time, test validity arguments aside. The traditional course placement stratification that occurs in schools works against a larger number of students—particularly those repeatedly placed in less challenging courses. If we are to continue to use these high-stakes tools then policymakers need to adopt requirements for school districts to earnestly examine their student sorting practices. As sorting practices are highly subjective, highly localized, and have a profound impact on overall student achievement, school systems need to be required to engage in more earnest examination of how sorting practices and the use of tracking is or is not occurring throughout their respective systems. In

New Jersey, the heated legislative and philosophical debate regarding the effectiveness of the use of PARCC standardized assessments as graduation requirements has unfolded for the larger part of the past decade. Legislation that is singularly focused upon passage of a high-stakes assessment as a primary graduation requirement cannot be supported. A far broader scope of attention to how course-taking patterns and tracking within systems impacts students within all school systems need to be undertaken so that each school district understands the nuances of the individual student journey within their respective system. If PARCC or other standardized assessments are going to be used as hurdles for graduation, then school district course-taking patterns need to be more deeply scrutinized and district leaders need to develop increased awareness of how their tracking decisions impact the students they serve.

Tracking practices start long before high school so the very act of mandating a certain cut score on an external assessment for graduation places students in lower-tracked courses at a distinct disadvantage. A disadvantage that widens year over year. In light of the compelling impact that subjective course placement decisions have on student performance, policymakers must refrain from mandating passage on any singular assessment as a graduation requirement. Given the expansion of multiple mathematics pathways to success in higher education and the resulting increases to student proficiency (Mangan, 2019), policymakers should allow schools to develop their own mathematics pathways and measures of success rather than mandating compliance on individual exit or course specific assessments. Schools must be allowed to develop their own unique metrics of student success that correlate to the individual student experience. The top-down categories of standardized test scores and school performance reports in New Jersey fail to capture the unique work that may be underway for students. Policymakers must allow districts to generate their own yardsticks of student success and move away from the

rigidity of course specific assessments as the determinant of student success. At the secondary level, policymakers can support the work of improving student outcomes by allowing more customized student experiences in course taking rather than rigid and outdated seat time requirements connected to individual courses of study that are often isolated from one another and redundant in nature. Giving school leaders increased flexibility in developing mathematics pathways for individual students would allow for schools to recognize the individual talents and capacities of each student at the school level. Additionally, policymakers can provide financial support in conjunction with this flexibility to provide more original credit, summer academies, and after-school opportunities to mitigate the impacts of poor student placement decisions. In turn, this research indicates that the flexibility in providing these opportunities and supports to students can indeed improve academic outcomes for larger numbers of students and alter the impact of tracking.

This research indicates that clear steps can be taken to mitigate the long-term impacts of tracking decisions within a school system. Policymakers should look to progressive models like the district in this study to fully understand the differences between systems that have actively dismantled tracking systems and those that have continued to utilize outdated means of sorting students for instruction. Clearly there are steps such as opening access to more challenging course work, providing additional school day and summer supports, and de-leveling courses so that all students have a rich curricular experience that impacts student mathematics attainment. Unless all of these facets are realized for all, mandating a graduation exit examination requirement is a deeply flawed and misguide policy pursuit.

Recommendations for Practice

In order to adequately prepare students for the modern age traditional comprehensive high schools must serve students in meeting both career and college aspirations and this is only possible with a strong academic foundation for all (Oakes & Saunders, 2008). The stratification of students grouped by perceived ability for instruction is deeply ingrained into the fabric of school systems. The processes to group students are often based on highly subjective criteria with a particular over-reliance upon standardized assessment performance as determining factors for course placement (Boaler, 2000). School administrators have incredible leverage in determining the processes schools will use for student placement. Much of the long-term tracking that occurs is a by-product of school-based practices that often are based on subjective factors and have long-term negative consequences for students (Kalogrides & Loeb, 2009). Developing the capacity of school administrators to understand how specific tracking practices are working within their building or district is critical to systemically improve the processes for student grouping. Currently course selection at the secondary level is a largely perfunctory practice where students move from one course to the next in the same progression they may have been following since elementary school. If administrators lack understanding of how this process fundamentally alters the schooling experience for some students compared to peers who expressed a capacity to excel academically at an earlier age, the subjective and pernicious cycle that is the reality for many students in largely tracked environments will remain. The work of Archbald and Keleher (2008) indicated that schools have narrow understandings of how their own tracked systems worked. To mitigate these entrenched practices, school leaders need to engage in more comprehensive work around the academic journey of individual students within the larger school system. There are a number of specific practices school leaders can take to

improve the status quo. Research indicates that schools can take measured steps to mitigate the subjectivity in course placement decisions while limiting the impact of external factors that weigh so predominantly in determining the course opportunities for students. This study has validated those findings while outlining additional measures that might prove useful in ensuring more robust academic experiences in the mathematics for all students.

Analyze teacher recommendation trends over time by course and course level. Often school districts perpetuate the status quo by relying on recommendation processes that simply matriculate a student from one mathematics course to the next within the same pre-defined "level" year over year. A school leader with the capacity to examine recommendation trends and then corroborate those trends with a more nuanced understanding of the entirety of a student transcript using other sources of student information may allow for increased flexibility for a larger numbers of students to enroll in more challenging courses regardless of highly subjective recommendation procedures that tend to maintain the status quo for students.

Education leaders should focus more attention on course-taking progressions and patterns within their systems. If a growth mindset is embraced, it is logical that students will develop different capacities at different times throughout their academic experience. In order to rebuff tracking decisions that might have occurred years prior to Grade 9, school leaders at the secondary level must be willing to engage in a more sophisticated analysis of individual student journeys within their mathematics courses. Understanding where opportunities are to provide students more challenging coursework as well as understanding which course-taking patterns may lead to lower outcomes, both in course attainment and on standardized assessments, school leaders can begin to rebuild their own systems from the inside out to ensure more positive mathematic outcomes for larger numbers of students within their systems. The district in this

study developed its own unique metric to understand student journey's within their system. The metric known as the "Opportunity Index" is a visual depiction that identified key access points, including AP coursework, honors programs, and magnet programs. District leaders then created a formula that compared the group's percentage in the district overall to its percentage in these programs and was boiled down to a single number illustrating the over- or under-representation of a disaggregated group of students. This graphic representation of access highlighted participation in AP coursework or specialized programs helped leaders understand course progressions for all students. The resulting metric helped to bring to light critical areas of inequity that needed to be addressed. One of these areas was mathematics course-taking progression, which helped to reinforce the decision to detrack as students in lower tracks were found to take fewer years of mathematics overall and were less likely to reach advanced levels of mathematics. Regardless if policymakers provide increased flexibility to districts to demonstrate their student achievement and effectiveness as a system, school leaders have a responsibility to build their own metrics of student achievement and to educate their public and larger school community about the impacts of those metrics. By doing so, schools can provide greater equity, opportunity, and support to all students.

Additionally, school leaders need to be aware of the impacts to students who are required to progress through lower-level courses at the high school level. Students who do not complete Algebra 2 may be at a distinct disadvantage when applying to more selective universities. Students who do not take challenging courses such as statistics or computer science after completing Algebra 2 or other mathematics courses may struggle more during the college application process. Students in lower level courses are often less likely to take four years of mathematics, and their overall representation to colleges and universities may suffer negative

repercussions during the application process. The course-taking processes utilized by schools have real consequences on how students arrive on the shore of the college application process and must be more fully understood, and where necessary, altered by education leaders mitigating hyper-localized impacts of tracking practices. Finally, course names may serve as an impediment during the college application process. As schools create courses to meet state math requirements, careful consideration needs to be given to how these courses appear to college admissions officers. Often, lower level courses are readily identified by course title and may work against students during the college application process.

This work cannot be done without an accompanying examination of the guidance practices when it comes to scheduling students. The district in this study focused specifically on guidance counselors as a key entry point to identify and encourage students to move beyond their ninth grade course level. Counselors were brought together from across the district and engaged in deep analysis collectively related to student potential for more rigorous course loads. Principals then worked with counselors in many instances to develop heat maps that tracked student progress in more challenging courses. Lists of "students to watch" were developed to monitor student movement upward within and between tracks, while ensuring that students received additional support once they moved into more challenging courses. Original credit offerings were dramatically increased over the summer to provide students opportunities to complete state-mandated courses over the summer months in order to take on courses of greater interest and challenge during the school year. Providing students enrolled in more challenging courses as well as students placed in lower mathematics tracks real-time support during the school day can also alleviate the negative impacts of tracking and positively alter outcomes for students. Developing accompanying workshop courses that run on special schedules, and

providing peer tutoring, faculty office hours, and more extensive and individualized counseling for students related to their course-taking journey can be accomplished readily and without financial burden.

Another recommendation for practice includes educating educators about the negative impacts of students being placed in "lower" tracks in an effort to engage in the long-term work of changing mindsets around student capacity. In the post *NCLB* world, educators have been conditioned to equate student performance and capacity with standardized assessment scores. Recent research has clearly debunked such reliance (Tienken 2015). Still, a generation of educators and stakeholders have been conditioned to utilize such measures to not only sort and rank students but to utilize school rankings around such narrow concepts. The research is clear that teacher mindsets around student ability follow the track the student has been placed within. This means teachers have higher expectations for students in higher-tracked classes than students placed in lower-tracked classes (Boaler 2000). When coupled with lower expectations and the clear research that placement in lower-tracked courses also impacts performance on the same standardized assessments that so much public faith has been placed in, there is extensive work that must occur throughout our school systems to provide evidence and data that support moving away from such practices. There are clearly models, including this research, that can provide definitive data to support the claims of moving away from rigidly tracked mathematics pathways. While work around "mindsets" is both long and difficult, a logical starting point is for school leaders to create a sense of urgency for such work while providing clear rationales for alternative mathematics pathways that might better serve all students. All of the above work can be done in other secondary schools and should be approached as potential further research.

Recommendations for Future Research

This research examined the impact of ninth grade course placement on overall mathematics performance on the 2014 NJHSPA and overall mathematics course attainment in a regional high school setting. Further research in this area should be conducted. Recommendations are as follows:

- Conduct a similar study using school level variables specific to tracking and course placement not researched in this study to determine the impact they have on student achievement and student mathematics outcomes. As the regional high school district used in this research was in the formative stages of addressing inequities in tracking decisions, a follow-up study examining student progression through the regional system now that these practices have been put in pace may provide more nuanced information related to how high schools can support the mathematic experience for all students.
- Replicate this study and control for student demographics and compare the same independent and dependent variables within multiple regional systems to see if there are certain variables that explain the most variance and make the best predictions.
- Conduct a study that examines different independent variables to explain variance within the dependent variables to see if certain school specific variables are significant.
- Conduct a study that utilizes qualitative methods in order to explain the impact of school-based tracking decisions on individual students.
- Conduct a similar study to this research that uses the same independent variables to see which variables explain the most significance in the 2014 HSPA ELA.
- Conduct a qualitative study to see how teacher mindset might have impacted student placement decisions and expectations within the same regional high school setting.

- Conduct a study that examines how the independent variables in this study impacted college acceptances for the class of 2015.
- Conduct a study that maintains all independent variables including English Language Learner status, special education classification, and free/reduced lunch status to see how these variables impact mathematics achievement and attainment relevant to course placement.

Overall Summary

This study examined the impact of ninth grade mathematics course placement on performance on the 2014 NJHSPA mathematics and Grade 12 mathematics course attainment. Ninth grade mathematics course placement explains 27.8 % of the student performance on the NJHSPA mathematics section. Grade 11 pipeline level explains 7.7% of the student performance on the NJHSPA mathematics section. Gender explains 2.6% of the student performance on the NJHSPA mathematics section. Overall, Grade 9 mathematics course placement accounts for 17.9% of Grade 12 course placement when controlling for student demographics. This study indicated while ninth grade course placement matters as it relates to standardized assessment scores and Grade 12 course attainment, other factors may positively alter an individual student academic trajectory. The evidence indicates that school districts that take specific action to mitigate the impacts of tracking may provide students with opportunities and support that lead to greater academic outcomes. The impact of specific actions to mitigate tracking is an area of great concern that needs to be examined in further research.

Conclusion

This research demonstrated that decisions related to course opportunities and student sorting practices at the secondary district and school level to support student growth have critical

impacts to long-term student outcomes and can mitigate the negative impacts of placement decisions made years earlier. I believe ninth grade course placement would have a greater impact on overall academic achievement and attainment in the mathematics. While my research validated prior research by Boaler (2000) Oakes (2005) and Rubin (2008) that indicated tracking processes utilized in secondary schools have potentially harmful long-term impacts to student outcomes, I was heartened to find how school-based interventions and practices designed to specifically mitigate tracking impacts could offset many of the profoundly negative outcomes found in prior research. After conducting the research, policies such as those being debated in New Jersey dictating as a graduation requirement meeting a specific cut score on a standardized assessment, are woefully misguided and harm students as they fail to adequately capture the educational experience of a child. These policies may more readily reflect local student grouping decisions rather than provide any accurate measure related to gauging overall student achievement. School leaders need greater latitude to demonstrate student success, measure student outcomes, and create new mathematics pathways for all students to succeed. With legislative and school community support this is possible as the ways schools respond to student course-taking patterns and actively intervene where appropriate can provide more equitable outcomes for all.

References

Ainsworth, J. W. & Roscigno, V. J. (2005). *Stratification, school-work linkages and vocational education* [article]

Akos, P., Lambie, G. W., Milsom, A., & Gilbert, K. (2007). Early adolescents' aspirations and academic tracking: An exploratory investigation. *Professional School Counseling,* 11(1), 57–64.

Alvarez, D. & Mehan, H. (2006). Whole-school detracking: A strategy for equity and excellence. *Theory into Practice,* 45(1), 82–89. doi:10.1207/s15430421tip4501_11

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

Ansalone, G. & Biafora, F. (2004). Elementary school teachers' perceptions and attitudes to the educational structure of tracking. *Education,* 125(2), 249–258.

Archbald, D. & Farley-Ripple, E. (2012). Predictors of placement in lower level versus higher level high school mathematics. *The High School Journal,* 96(1), 33–51.

Archbald, D. & Keleher, J. (2008). Measuring conditions and consequences of tracking in the high school curriculum. *American Secondary Education,* 36(2), 26–42.

Ascher, C. (1994). Successful detracking in middle and senior high schools. *Emergency Librarian,* 21(5), 30.

Bernhardt, P. E. (2014). How do I get in? criteria shaping the high school course recommendation process. *Current Issues in Education,* 17(1), 1.

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

Björk, L. G. & Gurley, D. K. (2005). Superintendent as educational statesman and political strategist. In L. G. Bjork, & T. J. Kowalski (Eds.), pp. 163–185. Thousand Oaks, CA US: Corwin Press.

Boaler, J. (2006). *How a detracked mathematics approach promoted respect, responsibility, and high achievement.* Lawrence Erlbaum Associates. doi:10.1207/s15430421tip4501_6

Boaler, J. (2006). How a detracked mathematics approach promoted respect, responsibility, and high achievement. *Theory into Practice,* 45(1), 40–46. doi:10.1207/s15430421tip4501_6

Booher-Jennings, J. (2006). Rationing education in an era of accountability. *The Phi Delta Kappan,* (10), 756. doi:10.2307/20442152

Braddock, Jomills Henry, II, & Center for Research on Effective Schooling for Disadvantaged Students. (1990). *Tracking: Implications for student race-ethnic subgroups. report no. 1*.

Braddock, Jomills Henry, II, & McPartland, J. M. (1990). Alternatives to tracking. *Educational Leadership,* 47(7), 76–79.

Burris, Carol Corbett Welnerc, Kevin G. Bezoza, & Jennifer Weiser. (2010). Universal access to quality education: Research and recommendations for the elimination of curricular stratification. *Poverty & Race,* 19(1), 3.

Burris, C. C., & Garrity, D. T. (2008). *Detracking for excellence and equity.* Alexandria, Va.: Association for Supervision and Curriculum Development, c2008.

Burris, C. C., & Garrity, D. T. (2009). Equity and excellence. *American School Board Journal,* 196(1), 28–31.

Burris, C. C., Heubert, J. P., & Levin, H. M. (2006). *Accelerating mathematics achievement using heterogeneous grouping.* American Educational Research Association.

Burris, C. C. & Welner, K. G. (2005). Closing the achievement gap by detracking. *Phi Delta Kappan,* 86(8), 594–598.

Carbonaro, W. J. & Gamoran, A. (2002). *The production of achievement inequality in high school English.* American Educational Research Association.

Catsambis, S., Mulkey, L. M., Buttaro, A., Steelman, L. C., & Koch, P. R. (2012). Examining gender differences in ability group placement at the onset of schooling: The role of skills, behaviors, and teacher evaluations. *Journal of Educational Research,* 105(1), 8–20. doi:10.1080/00220671.2010.514779

Comer, James P. (2010). Sage Publications.

Crosby, B. & Bryson, J. (2005). A leadership framework for cross-sector collaboration. *Public Management Review,* 7(2), 177–201. doi:10.1080/14719030500090519

Cullen, J. P. (2010). A comparison of the academic performance of college bound high school students in regional vs. community high schools in Connecticut. *Current Issues in Education,* 13(2)

D.I. Rees, D.J Brewer, & L. M. Argys. How should we measure the effect of ability grouping on student performance? *Economics of Education Review,* (19) 17–20. doi:10.1016/S0272-7757(98)00050-8

DiMartino, J. & Miles, S. (2005). Reaching real equity in schools. *Education Digest,* 70(5), 9–13.

Dreeben, R., & Gamoran, A. (1986). *Race, instruction, and learning.* American Sociological Association.

Elrnore, R. F. (1987). Instruments and strategy in public policy. *Policy Studies Review,* 7(1), 174–186.

Entwisle, D. R. & Alexander, K. L. (1993). *Entry into school: The beginning school transition and educational stratification in the United States.* Annual Reviews, Inc.

Esposito, D. (1973). *Homogeneous and heterogeneous ability grouping: Principal findings and implications for evaluating and designing more effective educational environments.* American Educational Research Association.

Findley, W. G., Bryan, M. M., & Georgia University, Athens Center for Educational Improvement. (1971). *Ability grouping: 1970—status, impact, and alternatives.*

Finley, M. K. (1984). *Teachers and tracking in a comprehensive high school.* American Sociological Association.

Fletcher, E. C., Jr. (2012). Demographics, tracking, and expectations in adolescence as determinants of employment status in adulthood: A study of school-to-work transitions. *Career and Technical Education Research,* 37(2), 103–119.

Fowler, F. C. (2000). *Policy studies for educational leaders: An introduction.* Upper Saddle River, N.J.: Prentice-Hall, c2000.

Franklin, C. C. (2012). *The Wisdom and Teachings Of Stephen R. Covey*. Business Wire. (English).

Froman, R. D. (1981). *Ability grouping: Why do we persist and should we.*

Fullan, M. (2001). *The new meaning of educational change. Third edition*

Gamoran, A. (1992). Is ability grouping equitable? *Educational Leadership,* 50(2), 11.

Gamoran, A. (1992). *The variable effects of high school tracking.* American Sociological Association.

Gamoran, A. (1993). Is ability grouping equitable? *Education Digest,* 58(7), 44.

Gamoran, A. & Mare, R. D. (1989). *Secondary school tracking and educational inequality: Compensation, reinforcement, or neutrality?* University of Chicago Press.

Gamoran, A. & Wisconsin Center for, E. R. (2009). *Tracking and inequality: New directions for research and practice.* WCER working paper no. 2009–6. Wisconsin Center for Education Research.

Goodlad, J. I. & Oakes, J. (1988). We must offer equal access to knowledge. *Educational Leadership,* 45(5), 16.

Greenleaf, R. K., Beazley, H., Beggs, J., & Spears, L. C. (2003). *The servant-leader within: A transformative path.* New York: Paulist Press, c2003.

Hallinan, M. T. (1994). School differences in tracking effects on achievement. *Social Forces,* 72(3), 799–820.

Hallinan, M. T. (1994). Tracking: From theory to practice. *Sociology of Education,* (2), 79. doi:10.2307/2112697

Hallinan, M. T. (1996). *Track mobility in secondary school.* University of North Carolina Press.

Havighurst, R. J. & Neugarten, B. L. (1967). *Society and education.* Boston: Allyn and Bacon, 1967; 3rd ed.

Kalogrides, D. & Loeb, S. (2013). Different teachers, different peers: The magnitude of student sorting within schools. *Educational Researcher, 42*(6), 304–316.

Kalogrides, D., Loeb, S., & Béteille, T. (2013). Systematic sorting: Teacher characteristics and class assignments. *Sociology of Education,* 86(2), 103–123. doi:10.1177/0038040712456555

Kao, G. & Thompson, J. S. (2003). *Racial and ethnic stratification in educational achievement and attainment.* Annual Reviews.

Keliher, A. V. (1940). *Pernicious I. Q.* Inor Publishing Co.

Kelly, S. (2007). The contours of tracking in North Carolina. *High School Journal,* 90(4), 15–31.

Kelly, S. & Carbonaro, W. (2012). Curriculum tracking and teacher expectations: Evidence from discrepant course-taking models. *Social Psychology of Education: An International Journal,* 15(3), 271–294.

Kelly, S. & Price, H. (2011). The correlates of tracking policy: Opportunity hoarding, status competition, or a technical-functional explanation? *American Educational Research Journal,* 48(3), 560–585. doi:10.3102/0002831210395927

Laprade, K. (2011). Removing instructional barriers: One track at a time. *Education,* 131(4), 740–752.

LaPrade, K. (2011). Removing instructional barriers: One track at a time. *Education,* (4), 740.

Lee, B. (2014). The influence of school tracking systems on educational expectations: A comparative study of Austria and Italy. *Comparative Education,* 50(2), 206–228. doi:10.1080/03050068.2013.807644

LeTendre, G. K., Hofer, B. K., & Shimizu, H. (2003). *What is tracking? Cultural expectations in the United States, Germany, and Japan.* American Educational Research Association.

Levin, B. & Fullan, M. (2008). Learning about system renewal. *Educational Management Administration & Leadership,* 36(2), 289–303.

Levin, D. E. (2015). *Putting law to work to improve public health: A national convening.* American Society of Law, Medicine & Ethics.

Levin, H. M. & McEwan, P. J. (2000). *Cost-effectiveness analysis: Methods and applications.* Thousand Oaks, Calif.: Sage Publications, c2001; 2nd ed.

Levin, P. (2011). *Excellent dissertations! [electronic resource].Peter Levin* Maidenhead: McGraw-Hill Open University Press, 2011; 2nd ed.

Lewis, T. & Cheng, S. (2006). *Tracking, expectations, and the transformation of vocational education.* The University of Chicago Press. doi:10.1086/506494

Lotan, R. (2006). Teaching teachers to build equitable classrooms. *Theory into Practice,* 45(1), 32–39. doi:10.1207/s15430421tip4501_5

Loveless, T. & Fordham, T. B. (1998). *The tracking and ability grouping debate.* Volume 2, number 8.

Loveless, T., & Fordham, T. B. (2009). *Tracking and detracking: High achievers in Massachusetts middle schools*. Thomas B. Fordham Institute.

Lucas, S. R. (1999). *Tracking inequality: Stratification and mobility in American high schools.* Sociology of Education Series.

Marzano, R. J., Waters, T., & McNulty, B. A. (2005). *School leadership that works: From research to results*. Alexandria, Va: Association for Supervision and Curriculum Development.

Mathews, J. (2013). *Ability grouping is back and likely to stay despite qualms.* Washington Post News Service with Bloomberg News.

Mau, W. & Bikos, L. H. (2000). Educational and vocational aspirations of minority and female students: A longitudinal study. *Journal of Counseling & Development,* 78(2), 186–194.

Mehan, H., National Center for Research on Cultural Diversity and Second Language Learning, & Others, A. (1994). *Tracking untracking: The consequences of placing low-track students in high track classes. research report:* 10.

Miller, W. S. & Otto, H. J. (1930). *Analysis of experimental studies in homogeneous grouping.* Public School Publishing Company.

Moller, S. & Stearns, E. (2012). Tracking success: High school curricula and labor market outcomes by race and gender. *Urban Education,* 47(6), 1025–1054.

Mulkey, L., Catsambis, S., Steelman, L., & Crain, R. (2005). The long-term effects of ability grouping in mathematics: A national investigation. *Social Psychology of Education,* 8(2), 137.

Oakes, J. (1987). Tracking in secondary schools: A contextual perspective. *Educational Psychologist,* 22(2), 129.

Oakes, J. (1992). Can tracking research inform practice? technical, normative, and political considerations. *Educational Researcher,* 21(4), 12–21.

Oakes, J. (1994). One more thought. *Sociology of Education,* 67(2), 91–91.

Oakes, J. (2000). Grouping and tracking. In A. E. Kazdin (Ed.), pp. 16–20. Washington, DC, US; New York, NY, US: American Psychological Association; Oxford University Press.

Oakes, J. (2008). Keeping track: Structuring equality and inequality in an era of accountability. *Teachers College Record,* 110(3), 700–712.

Oakes, J. & Lipton, M. (1990). Tracking and ability grouping: A structural barrier to access and achievement. In J. I. Goodlad, & P. Keating (Eds.), (pp. 187–204). New York, NY, US: College Entrance Examination Board.

Oakes, J., Saunders, M., & Harvard University, Graduate School of Education. (2008). *Beyond tracking: Multiple pathways to college, career, and civic participation.* Harvard Education Press.

Oakes, J. & Wells, A. S. (1998). Detracking for high student achievement. *Educational Leadership,* 55(6), 38–41.

Pool, H. & Page, J. A. (1995). *Beyond tracking: Finding success in inclusive schools* Bloomington, Ind.: Phi Delta Kappa Educational Foundation, ©1995.

Ravitch, D. (1983). *The troubled crusade: American education,* 1945–1980. New York: Basic Books, c1983.

Rosenbaum, J. E. (1980). Track misperceptions and frustrated college plans: An analysis of the effects of tracks and track perceptions in the national longitudinal study. *Sociology of Education,* 53(2), 74–88.

Rosenbaum, J. E. (1999). If tracking is bad, is detracking better? *American Educator,* 23(4), 24–29, 47.

Rubin, B. C. (2003). Unpacking detracking: When progressive pedagogy meets students' social worlds. *American Educational Research Journal,* 40(2), 539–573. doi:10.3102/00028312040002539

Rubin, B. C. (2006). *Theory into Practice,* 45(1), 1–3. doi:10.1207/s15430421tip4501_1

Rubin, B. C. (2006). Tracking and detracking: Debates, evidence, and best practices for a heterogeneous world. *Theory into Practice,* 45(1), 4–14. doi:10.1207/s15430421tip4501_2

Rubin, B. C. & Noguera, P. A. (2004). Tracking detracking: Sorting through the dilemmas and possibilities of detracking in practice. *Equity and Excellence in Education,* 37(1), 92–101.

Schiller, K. S. (1999). *Effects of feeder patterns on students' transition to high school.* American Sociological Association.

Sergiovanni, T. J. (2007). *Rethinking leadership: A collection of articles.* Thousand Oaks, Calif.: Corwin Press, 2007; 2nd ed.

Slavin, R. E. & Braddock, J. H. I., II. (1993). Ability grouping: On the wrong track. *College Board Review,* (168), 11–17.

Slavin, R. E., Lake, C., & Groff, C. (2009). *Effective programs in middle and high school mathematics: A best-evidence synthesis.* SAGE Publications.

Slavin, R. E., Madden, N. A., & Center for Research on the Education of Students Placed At Risk. (1995). *Effects of success for all on the achievement of English language learners.*

Smith, J. S., Akos, P., Lim, S., & Wiley, S. (2008). Student and stakeholder perceptions of the transition to high school. *The High School Journal,* (3), 32.

Smith-Maddox, R. & Wheelock, A. (1995). Untracking and students' futures. *Phi Delta Kappan,* 77(3), 222.

Southworth, S. & Mickelson, R. A. (2007). The interactive effects of race, gender and school composition on college track placement. *Social Forces,* (2), 497.

Spade, J. Z., Columba, L., & Vanfossen, B. E. (1997). *Tracking in mathematics and science: Courses and course-selection procedures.* American Sociological Association.

Spillane, J. P. & Healey, K. (2010). Conceptualizing school leadership and management from a distributed perspective: An exploration of some study operations and measures. *Elementary School Journal,* 111(2), 253–281.

Tienken, C. H. (2011). Common core state standards: An example of data-less decision making. *AASA Journal of Scholarship & Practice,* 7(4), 3–18.

Tomlinson, C. (2005). This issue: Differentiated instruction. *Theory into Practice,* 44(3), 183–184. doi:10.1207/s15430421tip4403_1

Tyson, K. (2011). *Integration interrupted: Tracking, black students, and acting white after "brown."* Oxford University Press.

Van Houtte, M. (2006). Tracking and teacher satisfaction: Role of study culture and trust. *The Journal of Educational Research,* 99(4), 247–254. doi:10.3200/JOER.99.4., 247–256.

Van Houtte, M., Demanet, J., & Stevens, P. A. J. (2013). Curriculum tracking and teacher evaluations of individual students: Selection, adjustment or labeling? *Social Psychology of Education: An International Journal,* 16(3), 329–352.

Van Houtte, M. & Stevens, P. A. J. (2009). *Study involvement of academic and vocational students: Does between-school tracking sharpen the difference?* SAGE Publications.

Wells, A. S., Hirshberg, D., Lipton, M., & Oakes, J. (1995). *Bounding the case within its context: A constructivist approach to studying detracking reform.* American Educational Research Association.

Wells, A. S. & Oakes, J. (1996). Potential pitfalls of systemic reform: Early lessons from research on detracking. *Sociology of Education,* 135. doi:10.2307/3108461

Welner, K. G. (2001). Tracking in an era of standards: Low-expectation classes meet high-expectation laws. *Hastings Constitutional Law Quarterly,* 28, 699.

Welner, K. G. & Oakes, J. (1996). (Li)ability grouping: The new susceptibility of school tracking systems to legal challenges. *Harvard Educational Review,* 66(3), 451.

Welner, K., & Burris, C. C. (2006). Alternative approaches to the politics of detracking. *Theory into Practice,* 45(1), 90–99. doi:10.1207/s15430421tip4501_12

Werblow, J., Urick, A., & Duesbery, L. (2013). On the wrong track: How tracking is associated with dropping out of high school. *Equity & Excellence in Education,* 46(2), 270–284.

Wheelock, A. (1992). The case for untracking. *Educational Leadership,* 50(2), 6.

Wheelock, A. (1993). More students learn more with untracking. *Educational Leadership,* 51(2), 82.

Wheelock, A. & American Association of, S. A. (1994). *Alternatives to tracking and ability grouping.*

Yonezawa, S. & Jones, M. (2006). Students perspectives on tracking and detracking. *Theory into Practice,* 45(1), 15–23. doi:10.1207/s15430421tip4501_3

Yonezawa, S. & Jones, M. (2009). Student voices: Generating reform from the inside out. *Theory into Practice,* 48(3), 205–212. doi:10.1080/00405840902997386