

**The Perception of the Final Exam as Preparation for College-Level Assessments and
Its Reliability as a Measure of Student Understanding of Course Content**

A Doctoral Capstone Project

Submitted to the School of Graduate Studies and Research

Department of Secondary Education and Administrative Leadership

In Partial Fulfillment of the
Requirements for the Degree of
Doctor of Education

Lori A. Pavlik

California University of Pennsylvania

California University of Pennsylvania
School of Graduate Studies and Research
Department of Education

We hereby approve the capstone of

Lori A. Pavlik

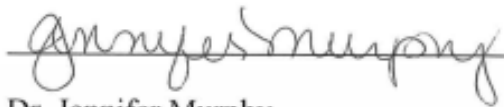
Candidate for the Degree of Doctor of Education

7/25/2021



Dr. Kevin Lordon
Doctoral Capstone Faculty Advisor
Doctoral Capstone Faculty Committee Chair

7/28/2021



Dr. Jennifer Murphy
Deputy Superintendent, Peters Township School District
Doctoral Capstone External Committee Member

Dedication

This work is dedicated to my husband, Daniel, and daughters, Katharine and Anna, who have given me the encouragement and confidence to persevere through this process. To my daughters, I hope you see that your resolve is stronger than the obstacles put in front of you. You can do anything you set your mind to.

Acknowledgements

This completion of this work would not have been possible without the assistance and participation of my colleagues at Peters Township School District, especially the teachers at the high school. Thank you to the teachers for their honesty and willingness to share their practices and feedback. Their commitment to providing their students with the best possible educational experiences continues to motivate me to do the same.

I am forever grateful for the support and encouragement of Dr. Jennifer Murphy and Dr. Kevin Lordon. Their patience and guidance gave me the confidence to complete this work.

A special thank you Mitch Wentzel for his brilliance with statistical coding and help with the grading data analysis.

A final note of appreciation to the students of Peters Township and the alumni who participated in this study for letting me a part of their journey and continually showing me the potential in this world.

Table of Contents

Dedication	iii
Acknowledgements	iv
List of Tables	viii
List of Figures	ix
Abstract	xi
CHAPTER I. Introduction to Research	1
Background	2
Purpose of this Research	4
Research Questions	4
Expected Outcomes of the Action Research	5
Fiscal Implications	5
Summary	6
CHAPTER II. Review of Literature	7
Final Exam Beliefs and Practices	7
The Evolution of Testing	11
<i>The Rise of Modern-Day Methods</i>	11
<i>Education Reform through Standardized Assessments</i>	14
<i>Standardized Assessments as Graduation Requirements</i>	15
The Shift to College Readiness	19
Measures of College Readiness	21
<i>Entrance Exams and GPA</i>	21
<i>Advanced Placement Programs</i>	22

<i>High School Exit Exams</i>	23
Testing Considerations	24
<i>The Testing Effect</i>	24
<i>Final Exams and Retention</i>	28
<i>The Psychology of Testing</i>	32
<i>Final Exam Weighting</i>	38
Summary	41
CHAPTER III. Methodology	44
Purpose	46
Setting and Participants	49
Research Plan	53
Research Design, Methods & Data Collection	56
Validity	59
Summary	62
CHAPTER IV. Data Analysis and Results	64
Data Analysis	65
Results	66
<i>Teacher Survey Results</i>	66
<i>Alumni Survey Results</i>	84
<i>Grading Data Analysis</i>	94
Discussion	105
Summary	110
CHAPTER V. Conclusions and Recommendations	113

Conclusions	115
<i>Application of the Results</i>	122
<i>Fiscal Implications</i>	125
Limitations	125
Recommendations for Future Research	126
Summary	128
References	130
APPENDIX A: Survey Disclosure Letter to Teachers	136
APPENDIX B: Teacher Survey with Consent to Participate	138
APPENDIX C: Survey Disclosure Letter to Alumni	143
APPENDIX D: Alumni Survey with Consent to Participate	145
APPENDIX E: IRB Approval	148

List of Tables

Table 1. PTHS Final Grade Calculation	96
Table 2. Quality Point Conversion Table	97
Table 3. Quality Point Impact on the Final Grade	98

List of Figures

Figure 1. Content Areas of Teacher Participants	67
Figure 2. Teacher Experience of Participants	67
Figure 3. Teacher Survey Question 1 Results	68
Figure 4. Teacher Survey Question 2 Results	69
Figure 5. Teacher Survey Question 3 Results	70
Figure 6. Teacher Survey Question 4 Results	71
Figure 7. Teacher Survey Question 5 Results	71
Figure 8. Teacher Survey Question 6 Results	72
Figure 9. Teacher Survey Question 7 Results	73
Figure 10. Teacher Survey Question 8 Results	73
Figure 11. Teacher Survey Question 9 Results	74
Figure 12. Teacher Survey Question 10 Results	75
Figure 13. Teacher Survey Question 11 Results	75
Figure 14. Teacher Survey Question 12 Results	76
Figure 15. Teacher Survey Question 13 Results	76
Figure 16. Teacher Survey Question 14 Results	77
Figure 17. Education Level of Alumni Participants	85
Figure 18. Category of College Study of Alumni Participants	86
Figure 19. Alumni Survey Question 1 Results	86
Figure 20. Alumni Survey Question 2 Results	87
Figure 21. Alumni Survey Question 3 Results	88
Figure 22. Alumni Survey Question 4 Results	88

Figure 23. Alumni Survey Question 5 Results	89
Figure 24. Final Exam Impact per Course	99
Figure 25. Final Exam Grade vs. Year-End Grade Scatterplot.	99
Figure 26. Total English and Math Scatterplot	100
Figure 27. Final Exam Impact on Final Course Grade	101
Figure 28. Precalculus Academic Scatterplot	102
Figure 29. Precalculus Honors Scatterplot	102
Figure 30. English 10 Academic Scatterplot	103
Figure 31. English 10 Honors Scatterplot	103
Figure 32. Percentage of Final Exam Grades Aligned with Quarter Grades	105

Abstract

With the goal of increasing both student accountability for learning and college readiness skills, many high schools have adopted the collegiate practice of the using final exams. Peters Township High School is one such school whose practice of using final exams has come under scrutiny. Very little research exists on the efficacy of this practice at the secondary level. The purpose of this research was to conduct a case study at Peters Township High School to determine if final exams were a valued practice that provided students with the skills and practice needed to be successful in college and if they were an accurate measure of student learning. The framework for this study was the four research questions focused on discovering teacher perceptions about the purpose of the final exam and its reliability as a measure of student learning, analyzing the historical impact of final exam grades on student grades, and comparing the high school and college experiences. A mixed-methods approach was taken using qualitative and quantitative data from teacher and alumni surveys and analyzed alongside quantitative data on historical final exam grades and final grades. The data revealed that high school and collegiate practices were not aligned and that final exam grades were not an accurate representation of student learning. However, the research showed that when weighted properly and incorporated as part of a comprehensive assessment plan, the use of final exams can improve student retention and capacity to study.

CHAPTER 1

Introduction to Research

A major emphasis in secondary education is student preparation for post-secondary training and learning. To accomplish this, schools have worked to improve both the rigor of their courses and student accountability for their own learning. One strategy that is commonly used in the high school setting is a comprehensive final exam like those used at the collegiate level. Grades on the final exam are factored into the students' final course grades. This is meant to hold students accountable for their learning and provide them with the practice of preparing for a cumulative assessment. Peters Township High School uses final exams at the end of a semester course and at the end of a full year course for this purpose.

In addition to the inclusion of the final exam into high school courses, other end of course assessments may be required of students, depending on the enrolled course. For example, the Department of Education in Pennsylvania requires end of course Keystone assessments to test general knowledge in Algebra, Biology, and Literature. Students taking Advanced Placement (AP) courses take a standardized end of course College Board exam in their course of study to earn college credit. School-based final exams, Keystone Exams, and Advanced Placement Exams occur within weeks of each other during the last month of school. As a result, students have multiple exposures to cumulative, comprehensive, and standardized tests based on course content throughout their high school career. While the original purpose of adding a final exam to high school courses was to provide students this exposure, questions are now being raised about the need for and usefulness of the final exam in a high school course. Additional concerns

include loss of instructional time to high stakes testing and testing overkill. If the final exam is not measuring student learning as intended, then the practice may not be worth the time or stress it costs.

This research study will explore the calibration of the final exam grade to the final course grade, the comparability of the final exam experience to the collegiate evaluation process, and the perception of both teachers and former students towards the final exam.

Background

Peters Township High School is part of a high performing suburban school district in Western Pennsylvania with over 90 percent of graduating students continuing to college. As such, the high school has a rigorous course of study with numerous college-preparatory courses. For the last few decades, the high school has also used a unique grading system that includes a comprehensive final exam. Final course grades are determined by a calculation that uses letter grades and quality points rather than percentages. Each quarter a student's percentage is converted to a letter grade. At the end of the year, letter grades for each quarter are assigned a quality point. The final exam is a stand-alone grade that is also assigned a letter grade and quality point. For semester courses, the final exam counts for 1/5 of the course grade. For a full year course, the final exam counts for 1/9 of the course grade. Grades and grading practices have been the focus of discussion among students, faculty, and parents. It is often a source of confusion as it does not directly align with a typical percentage-based grading scale.

In 2014, the Peters Township School Board created a subcommittee to investigate the final exam and final grades process and make a recommendation to the school board. The committee included school board members, parents, administrators, teachers, and

students. The committee's focus was on the mathematical weight of the final exam. Teachers complained that the final exam held no weight for most students, therefore students did not take it seriously or give their best effort. Teachers felt that they were not preparing students to study for comprehensive exams given in college and they were dedicating valuable class time to prepare students for a test that did not matter. Parents who volunteered for the committee expressed the same concern and questioned whether the final exam prepared students for college level testing. Additional concerns were raised that the system was unfair and caused undue stress. Parents expressed fear, however, that a change to a more heavily weighted final would negatively impact student grades, particularly students who had high achievement throughout the year. After a year of looking at data and research, the committee developed a recommendation to have two semester finals rather than one and to increase the weight of the exam slightly. Though the recommendation did not move forward, debate continues about the use of the final exam.

As the high school principal, I serve as instructional leader whose responsibility it is to ensure programming meets or exceeds state standards and district goals with a heavy emphasis on college and career readiness. I have worked with teachers and students to understand this calculation, but each year I field questions from both parents and students about the value and fairness of this system. While most of the debate has focused on the weight of the final exam, little discussion has occurred regarding the efficacy of the final exam. The prior committee analyzed the impact on student grades in various scenarios of weighted finals, but they did not look at the purpose of the final exam or the perceptions

of staff and former students. As a school focused on college readiness, this needs to be investigated before considering any changes.

Purpose of this Research

While anecdotal discussions call for change, data needs to be collected and analyzed to determine if change is warranted and if so, what that change should be. There are three significant ideas that need to be examined. The first is to determine if the final exam is a true measure of student learning. The second is to evaluate the fairness of the impact of the final exam grade to each students' final course grade. The third is to assess the perceptions and attitudes of both teachers and students towards the final exam.

Considering the time dedicated to developing and preparing for a final exam by teachers and students, is it critical to investigate this practice. Many teachers draw on their own college experiences to prepare students, but current college trends may be different. The goal of this action research project is to analyze the reliability of the current final exam process and to determine if it aligns with current practices at the post-secondary level.

Research Questions

Question 1: What are teachers' perceptions of the purpose of the final exam?

Question 2: What are teachers' perceptions of the reliability of the final exam as a measure of student learning?

Question 3: What is the impact of final exam grades on students' grades?

Question 4: How does the current final exam structure compare to current assessment trends at the collegiate level?

Expected Outcomes of the Action Research

With a focus on college-readiness, high schools need to better understand the value of using a comprehensive final exam to both improve student learning and prepare students for college. It is important that the perceptions of stakeholders be continually reviewed. Surveying teachers and former students will provide information that can be used to improve current practices. Reviewing final exam and final grades data from prior years will identify any potential discrepancies between student final exam performance and their quarter grades. Through the analysis of the current final exam process at Peters Township High School, I expect to affirm its use or to make recommendations for improvement or change.

Fiscal Implications

Any costs associated with both this research study and the use of final exams are indirect. Google Forms is a free platform that is used for the survey questions to teachers and alumni. PowerSchool and Excel are district programs already in use that will be used to compile the data from previous years. RStudio, a free statistical coding software in the language R, will be used to assist with data analysis. Time will be the biggest investment for this research study. Time for participants to complete the survey questions and time for the researcher to compile and analyze the data.

There is a broader fiscal implication for students who are attending college. Providing students with a strong college-ready foundation could eliminate the cost associated with taking remediation coursework in college, the cost of which is incurred by the student.

Summary

This study seeks to create an accurate picture of how final exams are used in Peters Township High School to both measure student learning and prepare students for college. This study will reveal how teachers use data from final exams and how they value the practice of giving final exams. It will also provide insight from college students on the final exam process and their current college experience. Furthermore, the data analysis will determine how calibrated a final exam grade is to overall student performance. The study will provide valuable guidance that will be utilized by teachers and administrators to make decisions on how final exams are used at Peters Township High School.

CHAPTER II

Review of Literature

A major emphasis in secondary education is student preparation for post-secondary training and learning. But do high schools adequately prepare students for college? This topic has come to the forefront of educational policy as the phrase “College and Career Ready” is driving school reform. Standardized tests like the Scholastic Aptitude Test (SAT) or the American College Test (ACT) have claimed for years to measure this “college readiness” in students. Training programs and the military use their own standardized tests to determine the readiness of applicants. High schools have adopted the practice of using cumulative assessments like semester or final exams to provide students with a more rigorous high school experience modeled after the collegiate level. Grades on these final exams are often factored into the students’ final course grade for both semester and full year courses. But in a modern-day world of state-required testing at all levels, questions are being raised about the value of these final exams. Some of these include: Do these tests accurately measure student learning? Are they a fair reflection of student work throughout the course? Do they match the experience of students in college? A review of current literature seeks to investigate these questions and shed light on the practice of cumulative end of course testing.

Final Exams Beliefs and Practices

In their focus to better prepare students for college, more and more high schools have adopted a final course examination experience for students. This process varies widely from school to school with some final grades heavily impacted by the final exam score and others not impacted at all. Furthermore, professors at the collegiate level

continue to debate not only the merits of a final exam but also best practices for incorporating them. While the college level final exam practice has been the subject of multiple studies, very few studies look at the use of final exams at the secondary school level. Easier to find are editorials about final exams and anecdotal beliefs, but specific research on the pros and cons of administering final exams at the high school level is in need of study. Research completed prior to the implementation of No Child Left Behind (NCLB) does look at one high school's experience with final exams. In this study, researchers set out to determine the effectiveness of the school's approach to final exams as preparation for future college exams.

The study was conducted by Indiana State University in partnership with Terre Haute South Vigo High School (Hall & Ballard, 2000). Terre Haute South Vigo had moved to a trimester schedule and partnered with Indiana State to review their grading policies and use of major assessments as another area of possible reform. As typical of many high schools, the faculty had previously debated the use of final exams. Questions about test validity, the length and format of the exam, whether students could opt-out if they performed well during the year, and whether the exams truly prepared students for college were being asked. Terre Haute South Vigo had nearly 1,900 students from diverse socio-economic backgrounds that consistently scored above state and national averages on the SAT (Hall & Ballard, 2000). Additionally, nearly 90 percent of graduates indicated they would continue their education beyond high school.

To conduct their study, researchers used surveys to faculty, administrators and alumni who graduated in 1997 and went on to college. They strategically wanted to survey alumni who had experienced college finals multiple times (Hall & Ballard, 2000).

The alumni who responded represented college students who attended both in-state and out-of-state schools, both large and small schools, and a variety of declared majors. The study looked at both quantitative conclusions and qualitative insights. Some results were consistent with what the researchers expected, while others were surprising.

Both faculty and graduates indicated that “assessing knowledge and increased retention were the primary purposes” of final exams (Hall & Ballard, 2000, p. 49). There was agreement on the practice of preparation for final exams through either teacher-led reviews or teacher-provided study guides. This also seemed to carry over into college practices. Both groups agreed that a student who performed exceptionally through the school year should be exempt from taking the final exams. This idea was not supported by administrators who took the survey. Faculty did not agree on whether finals should be mandatory, but most faculty did say they would still give one regardless. Finally, a slight majority of respondents felt that performance on final exams did not reflect the abilities of the teacher (Hall & Ballard, 2000).

Alumni perspectives and beliefs also varied on some experiences. While most alumni indicated that the finals in their high school classes and college classes were the same type and both were generally comprehensive, they disagreed on how well those exams prepared them. Only one-third stated that their high school final exams prepared them for college. Ironically, the majority of faculty felt that the final exams were college preparatory (Hall & Ballard, 2000). This shows a potential disconnect between high school and college practices.

Other unexpected findings revealed further divisions between teachers and alumni perspectives. Teachers felt that finals helped students get organized, while graduates did

not. Discrepancies among the graduates' responses appeared to be related to their individual high school experience. For example, one graduate said, "I took AP level classes in high school in which the teachers tried to prepare us for college" while another said, "High school finals were a joke compared to college finals...even compared to college tests" (Hall & Ballard, 2000, p. 49). In addition, students who were fine arts majors felt that high school finals prepared them well, while those who were science majors said that finals in college were much more intense. This suggests that high schools should consider a differentiated approach to the use of final exams.

The final findings from this study focused on the role of the teacher as perceived by all three groups and the quality of the actual assessment. Starting first with the assessment, alumni referenced "good tests versus bad tests" (Hall & Ballard, 2000, p. 49). One graduate stated, "A good final tests cumulative knowledge of concepts or key areas that were focused on in the course. Unfortunately, bad finals are too common in high school" (Hall & Ballard, 2000, p. 49). This thought process may align with how graduates responded to questions about their preparation for college. While the majority indicated that final exam outcomes were not a reflection of the teacher, many also commented that final exams should be used to help teachers evaluate their teaching.

This study revealed inconsistencies among teacher and student experiences and perceptions which impacted the perceived effectiveness of the high school final exam in preparing students for college. It also hinted at the difference in the college assessment experience depending on field of study being pursued by the alumni respondent. This study begins to provide context to how the current practice of testing seems to function, but a next step is to look at how this came to be.

The Evolution of Testing

The Rise of Modern-Day Methods

To understand the patterns that currently exist in American education and the perceptions about them, it's important to start by tracing its origins and the evolution of cumulative testing. From 1635, when the first American private high school was founded, to 1821 which marks the arrival of the first American public high school, the major focus of high schools was to prepare young men for college at Harvard or service work for the church or government (U.S. Department of Education, 2003). Education through the elementary level was considered sufficient for the general population while secondary education was reserved for the elite. High school education was considered rigorous and designed to prepare young men for further education. Prior to 1840, formal assessment consisted of oral examinations. A change occurred after 1840, when education began to shift in response to the need for more skilled workers during the Industrial Revolution. More students progressed to high school, but they were not all college or service bound. In response to this growth, educational testing shifted from oral exams to written tests. This was due to the ease of efficient administration for large numbers of students (National Education Association [NEA], 2020) All of this led high schools to transition towards vocational training rather than only college readiness. In fact, students and families in urban areas viewed high school as a “shortcut to the new skilled jobs in the burgeoning factories and agricultural enterprises” (U.S. Department of Education, 2003).

With the rapid growth in high school enrollment and a demand for more skilled workers, education became more of a societal concern. As public discourse focused on the quality of education, it was apparent that schools lacked consistent expectations and

standards. As a result, the end of the twentieth century witnessed the first external standardized written exams developed and used to assess student progress in core areas. These were often mandated and used to shape administrative and policy decisions (NEA, 2020). The first real academic standards for high school education came about in 1892, when the National Council of Education was formed. Known as the Committee of Ten, a group of Ivy League professors mapped out a liberal arts education that would prepare students for college and for life (U.S. Department of Education, 2003). This guide gave high schools a common roadmap to prepare students for future education.

New testing instruments were also at the forefront of American education at the turn of the century. From measuring mental ability to assessing student preparedness for college, testing had become a key component to education (NEA, 2020). In 1890, the president of Harvard, Charles William Eliot, proposed that colleges and professional schools utilize a common entrance exam rather than each giving separate exams (NEA, 2020). This led to the establishment of the College Entrance Examination Board (CEEB), which administered the first standardized examinations in 1901 on nine subjects (NEA, 2020). High schools began to calibrate around a common experience of studying core subjects and using written assessments to show student achievement, but a flood of new immigrants in the early twentieth century took education in another direction.

Educating a new wave of immigrants presented schools with a different challenge. Learning a common language and assimilating to American culture began to supersede a liberal arts education. High schools during this movement became the precursor to the modern American high school. In 1918 the Commission on the Reorganization of Secondary Education issued *The Cardinal Principles of Secondary Education*. Primary to

the purposes of high schools were “health, citizenship, and worthy home-membership and, only secondarily, command of fundamental processes” (U.S. Department of Education, 2003, p. 2). The socialization of students to American culture was at the forefront of high school life. Education became more general and less academically rigorous. Achievement testing grew and statewide testing programs become more common (NEA, 2020). The National Education Association endorsed the use of standardized testing in the hopes of bringing consistency to schools (NEA, 2020).

By 1916, The College Board, formed by the CEEB in 1899, offered comprehensive examinations in six subjects and by 1926 the first Scholastic Aptitude Tests (SATs) were nationally adopted. (NEA, 2020) The military began using intelligence tests developed with the American Psychological Association in collaboration with Stanford University professor Lewis Terman. A student of Terman, Arthur Otis, introduced the multiple-choice test format which became widely used (NEA, 2020). This marked a turning point for standardized testing as it became a mainstreamed practice in education, though not opposition.

Testing controversy erupted around 1930 when the University of Iowa initiated the first major statewide testing program for high school students. Critics worried that the rapid spread of the multiple-choice testing format encouraged memorization and guessing rather than authentic learning. The efficiency and perceived objectivity of this type of testing, however, superseded any concerns and it became widely adopted (NEA, 2020).

Technology made it possible to rapidly process a large number of tests which led to many states also adopting The Iowa Assessments. Testing of this nature created a

standard tracking system used by primary and secondary schools to monitor student learning, shifting it from individual oral assessments to mass standardized testing.

By the late 1960s, global competition led to a pendulum swing back to academic rigor in high school education. The launch of Sputnik in 1957 and the economic rise of Germany and Japan became a cause for concern in American education (U.S. Department of Education, 2003). Concerns were raised that American students had fallen behind competing countries in science and math. At the same time, the Civil Rights Movement's demand for equal access to education led President Lyndon B. Johnson to sign the first sweeping legislation by the federal government focused on K-12 education. The Elementary and Secondary Education Act (ESEA) passed in 1965 was significant in changing the future landscape of American education. It generated decades of testing as a means by which student and school progress were monitored. It opened "the way for new and increased uses of norm-referenced tests to evaluate programs" (NEA, 2020). It also brought the federal government into the education policy conversation.

Federal policy continues to be a significant driver of educational practices. The events of the 1960s brought to light the need to improve educational access for all students and to improve academic rigor for students in math and science. Since then, educational policy has focused on how to accomplish and measure this.

Education Reform through Standardized Assessments

The growth of secondary education for the masses led to political pressure to increase rigor in schools. One of the most impactful reports came in 1983 when "A Nation at Risk" was released from the National Commission on Excellence in Education (U.S. Department of Education, 1983). Similar to the Committee of Ten, the Commission

renewed the demand for a more rigorous academic curriculum for all students. This marked a sharp shift in educational philosophy in America which previously used testing to identify students' aptitude to determine post-secondary opportunities. Among the recommendations from "A Nation at Risk" was for states to administer "standardized tests of achievement...at major transition points from one level of schooling to another and particularly from high school to college or work" (1983). This report is often cited as the most impactful call for education reform in history. Since that time, the phrase *education reform* has become synonymous with education itself and a common campaign promise of both federal and state leaders. Though education policies are frequent, progress through the years has been slow and inconsistent.

Standardized Assessments as Graduation Requirements

Each state sets its own education policies and graduation requirements. With different needs, both social and economic, states have not been consistent in their requirements. The federal government uses funding to incentivize states to adopt more rigorous and consistent education standards. Even so, how states interpret academic rigor and college readiness continues to vary.

A wide-range of norm-referenced tests are used to evaluate students and programs. In 2001, President George W. Bush reauthorized ESEA with No Child Left Behind (NCLB) which sought to "improve education for disadvantaged students and increase the academic achievement of all students" (Hunt Institute, 2016). This law, once again, increased the federal government's role in public education. Most notable in this law was the requirement of states to establish goals for "adequate yearly progress" (AYP) which "established minimum levels of improvement as measured by standardized tests

chosen by the state” (Hunt Institute, 2016). Assessment testing in reading and math was put into place for grades 3 through 8 and at least once in high school. Further, each state was required to establish core content standards, review and disaggregate their data, and provide this information to the public. While this gave states common goals and expectations, critics argued that it forced educators to teach for high stakes testing and it placed too heavy an emphasis on math and reading at the expense of non-tested subjects (Hunt Institute, 2016).

In December of 2015, another reauthorization of ESEA was passed and replaced NCLB. The Every Student Succeeds Act (ESSA) signed by President Obama was a response to concerns over the cost of implementation of NCLB on states and school districts, emphasis on test-based instruction resulting in narrowed curriculum with a heavy emphasis on math and reading, and little support for schools making improvements but identified as failing (Hunt Institute, 2016). This version gave states a bit more control over accountability measures and lessened the reliance on standardized assessments. It maintained mandated testing in grades 3 to 8 and once in high school, but it allowed states to diversify academic indicators to include SAT or ACT tests and college-readiness evidence (Hunt Institute, 2016). While this still held true to many of NCLBs expectations, particularly standardized testing, it provided states with more opportunities to use other measures.

During the NCLB years, a trend that was previously motivated by public concerns over students’ post-secondary preparedness, gained significant momentum and that was to adopt State High School Exit Examination (HSEE) policies (Warren & Kulick, 2007). In 1976, only two states had required students to pass a standard high school exit

examination to obtain a high school diploma. By 2003, that number grew to 26 and by 2016 two more states had adopted HSEEs. Warren and Kulick looked at this trend and found that HSEE policies were tied to the desire of states to improve their economic standings and to apply standards equally to diverse student populations (2007).

The Journal of Educational Research published a study on the implementation of HSEEs in various states between 1990 and 2013. By applying economic literature and theory to educational accountability policies, researchers identified three “rather strict assumptions” that must be met for the full benefits of HSEE to be realized. These include “perfect content, perfect measurement, and perfectly aligned incentives” (Caves & Balestra, 2016, p. 187). The authors identified perfect content as that which every student should learn to be ready to graduate. They did not consider simply being literate and numerate, but also having a broader development of citizenship and morality. Testing these ideals would be difficult with just an HSEE, so they suggested additional measures like teachers’ grades be equally important. Schools would otherwise focus their efforts on student performance on the HSEE at the expense of other classroom learning.

With the understanding that perfection was impossible, the researchers analyzed whether HSEEs effect on graduation rates and student achievement helped, hurt or had no impact (Caves & Balestra, 2016). Data from the Center on Education Policy, state data, and the NCES, specifically the National Assessment of Educational Progress (NAEP) were used for their analysis. They discovered that the use of HSEEs created a short-term decrease in graduation rates (right after enactment) followed by a long-term increase. There was no statistically significant change to student achievement in states that introduced HSEEs, so therefore the results were inconclusive. This research focused

solely on review of limited data and the authors acknowledged that other factors could have contributed to improvements in both graduation and achievement. One was expected since the intent of an HSEE was to improve curriculum and the behavior of schools, teachers, and students. The second was to encourage schools to spend more time on tested material. This could also be viewed as a limitation if the tests did not perfectly align with what learning was needed for students to be successful post-graduation. The long-term improvement in graduation rate also did not consider changes that may have been made to the HSEEs along the way. Adjustments that lowered the difficulty level or the required passing score may have occurred as a result of social and political pressure over the difficulty of the tests or the fairness to certain groups such as English Language Learners (Caves & Balestra, 2016).

Since the 2007 and 2016 studies, much has changed in the testing landscape in American schools. With the passing of ESSA, states have been able to differentiate their requirements, and many have postponed or abandoned HSEEs. By 2017, only 13 states still had a required HSEE for graduation. To better align with the college readiness goals, many states have simply adopted the SAT or ACT exams. Twenty-four now require students to take one of the college entrance exams, up from only seven states in 2005 (Gewertz, 2017). States have also shifted away from the Common Core Standards-Based assessments, Partnership for Assessment in Readiness for College and Careers (PARCC) and Smarter Balanced Test Consortium, in favor of state-created assessments. Only one-third of states still use one of these assessments while 32 have created their own (Gewertz, 2017). Twenty-five states require end of course assessments in specific curricular areas, usually reading, writing, mathematics and science during high school.

Nine of those states even require the score on the assessment to be factored into the students end of course grade in ranges from 20 to 30 percent (National Center for Educational Statistics [NCES], 2017).

Popularized during the Twentieth Century, standardized testing has been the most widely accepted measure of student achievement and college-readiness used today to the point that states have adopted them as graduation requirements. These tests can broaden or limit opportunities for students. Knowing this, high schools have worked to prepare students for these tests and those they will take in college usually by creating their own models of cumulative tests in the form of a final exam. Criticisms of the stress of cumulative finals has plagued campuses of both high schools and colleges. Further research has sought to determine whether there is a true benefit to cumulative testing as a learning measure or if the stress of testing acts as a deterrent to real learning. With a shift to college readiness for students graduating high school, the practice of cumulative testing has remained.

The Shift to College Readiness

Concurrent to the growth of standardized testing was the expectation that students graduating from high school will continue to some type of post-secondary education, usually a technical, two-year or four-year institution. The pressure on secondary schools to produce a more educated and better trained workforce to compete in an ever evolving and demanding global economy is intense. However, there exists an important discrepancy between the number of students who begin high school with the intent to continue their education after high school and those who complete it. The U.S. Bureau of Labor Statistics reported that 66.2 percent of high school graduates were enrolled in a

post-secondary school, but of those who attended a four-year college, only 62 percent completed their degree within six years (2020). Upon deeper investigation, researchers found many students were unprepared for college. Those who seemingly did everything expected of them in high school, like taking full ownership of their work and learning, still found that they were struggling at college (Perkins-Gough, 2008). Many of these students had to enroll in remedial classes. Statistically students in college who take remedial courses are more likely to drop out (Perkins-Gough, 2008). This needs to be a consideration for high schools' college-readiness planning as the cost to students of leaving high school ill-prepared can be significant.

In a report by the non-partisan group Strong American Schools titled “Diploma to Nowhere,” 688 students who were enrolled in remedial college classes were surveyed (Perkins-Gough, 2008). Surprisingly, many of these students reported that they believed they were prepared for college when they left high school. Most claimed they had taken the most challenging courses in high school, got good grades, and had a grade point average over 3.0. As a result, they blamed their high schools for not preparing them. Nearly 40 percent of the students thought their high school did not prepare them with the knowledge and skills to be successful in college and 60 percent said that their high school classes were too easy. “Eighty percent said that they would have worked harder if their high school had set higher expectations” (Perkins-Gough, 2008, p. 89). The report called for high schools to provide more rigorous and engaging instruction and to build collaborations with postsecondary schools to align standards and create better understanding of college expectations for students (Perkins-Gough, 2008).

A first step for high schools is to understand what college readiness really means. One study that looked at the relationship between high school coursework and college course success concluded that the definition of college readiness is that students can complete first-year college-level courses without needing remediation (Woods et al., 2018). While skills like reasoning and critical thinking were also identified as important components of college-readiness, the authors noted a disconnection between the definition of college readiness between K-12 schools and higher education. For example, high school math teachers emphasized higher order content while college professors stressed a solid grasp of basic skills (Woods et al., 2018). As researchers in their respective fields, college professors are also continually updating their content so having students with a strong understanding of basic skills gave them a better foundation to build.

Measures of College Readiness

Entrance Exams and GPA

Additional studies have focused on the relationship of high school performance to college performance. Because high schools vary, sometimes drastically, in how they determine high school grade point average (HSGPA), colleges and universities have relied more heavily on college readiness tests like the SAT and ACT to compare applicants. Koretz and Langi evaluated the predictive value of tests scores and HSGPA on future college success both within and between high schools (2017). They wanted to determine for two students from different schools, if a 60-point difference in SAT scores aligned with a 60-point difference in HSGPA as a determinant of first year college GPA (FGPA). They also weaved into their study how an HSEE factored in, specifically the

New York Regents and the Kentucky KCCT scores. A two-level regression model was used, and multiple scenarios studied. Their findings were interesting. First, when they looked at each predictor by itself (HSGPA, SAT composite score, HSEE score) they found that each predicted FGPA more strongly between schools than within. However, when combining the HSGPA with either test, they found that the net predictive value was stronger within a high school than between high schools. Test scores, in contrast, were 3.5 to 5 times larger between schools than within (Koretz & Langi, 2017). This supports the use of college-readiness tests (SAT and ACT) by college admissions as a statistically moderating factor to evaluate high school graduates from schools with vastly different grading standards. While the researchers admit that other factors play a role in student performance between schools, this study supported the universally accepted beliefs about standardized testing.

Advanced Placement Programs

Another universally established expectation is that the College Board's Advanced Placement (AP) program which can earn a student college credit for receiving a score of three or higher on the accompanying exam is a strong predictor of student success at post-secondary institutions. This program has become not only an accepted pathway to college credits, but also recommended for any student to gain college level experience. With pressures to continue to post-secondary education, the past twenty years has witnessed a tremendous growth in the AP program enrollment with a 9.3 percent increase in the number of exams being administered (Sadler & Tai, 2007). Some states require or give incentives to high schools to offer AP programs to their students. These programs are enticing to students who want to begin college with advanced standing.

Critics of AP contend that since high school teachers do not have the research-focus of college professors, students may miss learning the most current findings and issues in their fields. Sadler and Tai's research focused on the validity of science AP courses as a predictor of college performance in the equivalent introductory college course and the value-added to students for taking AP sciences in high school (2007). They found that students who scored a three or higher on the AP exam earned college grades that were higher than the student average, but they did not find the predicted performance for students who scored a four or five in science that the College Board claims. They also found little to no advantage for students who scored a one or two on the AP science exam. Survey data revealed that AP students felt that their high school courses were a good preparation for college even though some students still felt they benefited from taking the equivalent college science course (Sadler & Tai, 2007). The study was limited because it only looked at students who repeated the equivalent college course and only focused on science, it did not look at students who took the college credit and advanced. In any case, it still provides a critical piece of evidence of the popularity of testing as a measure of student mastery of content and the limitations of high school teachers to mimic the college experience.

High School Exit Exams

Other possible measures of college readiness are standards-based high school exit exams. A study conducted in 2009 looked at this relationship in one state, Arizona. Researchers used the Arizona Instrument to Measure Standards (AIMS) that students took in tenth grade and compared it to University of Arizona freshman GPAs (D'Agostino & Bonner, 2009). They sought to understand the trend of the rising number

of incoming freshmen enrolling in remedial coursework. They noted a previous study that revealed that at least 42 percent of students in two-year institutions and 20 percent of students in four-year institutions took remedial courses. They hypothesized that this was caused by a disconnect between secondary and post-secondary expectations.

“Researchers have theorized that students receive erroneous messages through sources such as academic content standards, tests of those standards, and grading practices regarding their degree of preparedness for college” (D’Agostino Bonner, 2009, p. 26).

While many states have moved towards increasing rigor through standards-based testing, the question of whether these were enough to prepare students for college was answered affirmatively by this study. Looking at the three AIMS tests, the researchers found that those tested standards were higher than the University of Arizona’s expectations (D’Agostino & Bonner, 2009). The math standards were the most rigorous and exceeded the expectations. The writing standards either met or exceeded expectations. The reading standard, however, fell below expectations and was not a good indicator of college performance. Though this study was limited to one scenario, it was representative of a typical public college experience. It is evidence that the gap between college norms and grading practices is not entirely disconnected from high school expectations. While these similarities do exist and frequent testing is common in both K-12 education and beyond, the value of testing needs to be considered.

Testing Considerations

The Testing Effect

As students advance through middle and high school, they are faced with frequent quizzes and tests. Post-secondary schools, whether technical or collegiate,

continue this trend with more heavily weighted cumulative assessments. Educational research and various studies have proven that assessments can positively affect retention of learning (McDermott et al., 2014). These studies have identified a phenomenon known as the *testing effect* in which “taking a test on studied material promotes subsequent learning and retention of that material on a final test” (McDaniel et al., 2007, p. 495; see also McDermott et al., 2014). Experiments in both laboratory and classroom settings have shown the testing effect to be true. “These experiments reveal that low-stakes multiple-choice quizzes with immediate correct-answer feedback can indeed enhance student learning for core course content” (McDermott et al., 2014, p. 3). Two particular studies, one at the college level and one at the high school level, provide key findings that can be used by educators to improve pedagogical practices.

At the collegiate level, final exams are a staple of nearly every college course. There is consensus among research studies that final exams are necessary to encourage student ownership of their learning and improve long-term retention of material. Where studies diverge is how to best prepare students throughout the semester for the final exam to achieve the desired goal.

A study conducted in 2007 on a class called “Brain and Behavior” at The University of New Mexico investigated the testing effect on 35 students (McDaniel et al., 2007). In their study, researchers used the term *testing effect* to refer to the memory gains produced by intervening tests such as recall and recognition tests administered throughout the semester (McDaniel et al., 2007). Students were given weekly quizzes and two unit tests prior to a final exam. Because most previous research was done in a laboratory setting, this study wanted to apply the testing effect to an actual classroom

over a longer period. Weekly quizzes and unit exams included multiple choice, short answer and read only items. The read only items were content statements that did not require a response. The final exam consisted of multiple-choice questions from both unit exams. Half of the questions were worded the same as previous tests and half were worded differently. At first, researchers discovered that multiple-choice performance was better than short answer performance on initial tests. This supports the idea that recognition is a less demanding task than recalling information (McDaniel et al., 2007). However, short answer performance did appear to play an important role in future recall. On the unit exams and the final exam, students performed overwhelmingly better on questions that had previously been quizzed in both the short answer and multiple-choice format. Though there was a significant advantage for content previously tested with multiple-choice and short answer quizzes, there was no advantage for read only items. Interestingly, short-answer quizzes were more beneficial to students on later multiple-choice exams than were multiple-choice quizzes (McDaniel et al., 2007). The testing effect was present even when the testing stems were different. Overall, this study was consistent with previous research on memory and the testing effect and supports the idea of test enhanced learning at the collegiate level. It suggests that courses that are heavily fact-based may benefit the most from this approach.

How does the testing effect look when studied at the secondary level? A similar study published in 2013 used some of the same premises as the college study and applied them to both middle school and high school courses. Researchers asked the question, “could frequent low-stakes testing be used within normal classroom procedures to enhance retention of important classroom material” (McDermott et al., 2014, p. 3)? Just

like the previous study, short-answer and multiple-choice classroom quizzing were used in the context of a seventh-grade science classroom and high school history classroom using typical instructional content and practices. Multiple testing strategies were applied to discern if there were any differences. These included four experiments with different conditions. The first condition had students tested three times prior to a unit or end-of-course exam, some with multiple-choice format and some with short-answer format. The second type involved taking three short-answer quizzes compared to reviewing targeted material the same number of times. The third and fourth experiments looked at the impact of the number of quizzes and the wording of questions on future performance at the middle school and high school levels and between science and history (McDermott et al., 2014). Researchers identified as their most important and surprising finding that the format of the quizzes did not have to match the format of the unit exam to be beneficial nor did it matter if the previous quiz used multiple-choice or short-answer questions. “Even quick, easily administered multiple-choice quizzes aid student learning, as measured by unit exams (either in multiple-choice or short-answer format)” (McDermott et al., 2014, p. 15). This was a departure from the findings of the 2007 study at the University of New Mexico which found that short-answer quizzes were more beneficial for future memory retrieval (McDaniel et al., 2007).

End of semester exam results showed that the benefits were long lasting. Low-stakes quizzing had a greater benefit than selective re-exposure or restudying of key material. Just reviewing in class or giving students targeted readings to do this did not have the same impact. They concluded that two quizzes would be sufficient exposure and questions did not need to be identical between assessments. Finally, the same benefits

applied to both middle school science and high school history. These findings, both expected and unexpected, provide clear support for the use of low stakes testing in classrooms to increase student learning.

Both studies viewed the final exam or end-of-course exam as a required assessment of student learning with their focus being on strategies to ensure student learning prior to the final cumulative test. Both studies showed evidence to support the idea of using quizzes, regardless of format, to enhance student learning. It is also important to note that feedback after each quiz or test was given to students in both studies which indicated that student feedback after an assessment was a key factor in providing the benefit of low stakes testing to enhance student learning. While these studies show the benefit of low stakes testing to improve student learning and retention as evidenced by student performance on a final exam, the next step is to review how long-term retention is impacted by final exams.

Final Exams and Retention

Though testing over the past few decades has become synonymous with accountability and school improvement, it is not a new phenomenon. People often compare a final exam or cumulative end of course exam with standardized tests, but these are two different tests with different purposes. Testing in the classroom setting is meant to measure student comprehension of learned content. Cumulative tests like midterm and final exams are a vehicle by which educators encourage students to remember content learned over a span of time. Simply put, when students have an expectation that information will be tested again, they are more likely to retain it. As previously stated,

prior and intermittent quizzing on the same content increases student retention, but what role does student expectation for a future assessment have on their long-term retention?

A study conducted in 2007 asserted that having the expectation of a final test was more influential on student retention than even intermittent testing (Szpunar et al., 2007). This study combined both concepts of frequent testing and the expectation of a final test to determine the influence of the future testing expectation. Washington University undergraduates participated in this experiment with 40 students each randomly assigned to four groups with a different experimental testing condition. The first group was the “aware” group who were told that there would be a final test and they received initial testing on content along the way. The second group was the “untested control” group. They knew there was a final test, but they were not tested during learning. The third group was the “unaware” group who received all of the same conditions as the “aware” group with the exception that they were not told there would be a final exam. The fourth group was the “unaware-cue” group. Students were initially tested and were unaware of the final tests. Additionally, students were instructed to forget the content right after being tested on it (Szpunar et al., 2007).

As expected, researchers found that the final exam performance was enhanced by taking an initial test on the content. Students who had the expectation of the final test performed better than students who did not have that expectation. Students who were cued to forget their learning and unaware of the final exam scored the same as those who were only unaware of the final exam. Therefore, cuing students to forget had no impact. Finally, “the expectation of a final test kept studied words in a more accessible state reducing forgetting between initial and final testing and aiding reminiscence” (Szpunar,

et al., p. 1010). In essence, students who knew that they would need to recall information for a final test kept that information accessible in their memories between the initial and final test. This study reinforces the use of final exams to encourage students to retain information.

Rutgers University published a study in 2013 that looked at the testing effect in relation to long-term retention. How well do students retain the information that was tested on the final exam compared to information that was not tested on the final? Researchers conducted an in-depth look at memory and retention (Glass et al., 2013). They identified the practice of distributed questioning as a key component of the testing effect. Distributed practice involves testing students on content followed by intervals of exposure to other learning, content, and tasks before additional testing on the original content. Pedagogically, the authors assert there are three reasons for giving a final exam beyond it serving as a summative measure of student learning. They give students an “opportunity and an incentive to further study the course content and increase their knowledge” (Glass et al., 2013). Cited was a 2009 study that found that on average, students increased their correct response to repeated unit questions that appeared on the final by 11 percent. The second reason involves the known research that distributed study increases long-term retention after the course has ended. “Since long-term retention is virtually always the goal of academic instruction, the effect on long-term retention by itself justifies administering final exams” (Glass et al., 2013, p. 231) The third and final reason for giving a final exam is that regardless of prior studying, answering a question on a final exam increases long-term retention of the answer to that question. Studies have supported this assertion, finding “a positive effect on retention even when feedback as to

the correct answer was not given” (Glass et al., 2013, p. 231). This suggests that the content tested should be comprehensive to the expectation of what should be retained.

This experiment that tested students from a psychology lecture course four months after taking the final exam and completing the course reinforced the long-term retention effect that taking a final exam has. What was also interesting about this study was how students scored on items that appeared on the final exam and items that were not on the final but appeared on previous unit tests. They found that four months after taking the final, students scored much better on items that appeared on the final exam. In fact, the percent correct was seventy-nine on finals questions and only sixty-seven on previous unit test questions (Glass et al., 2013).

Another study recommended that instructors use cumulative finals to increase both short term and long-term retention based on experiments done with Introductory Psychology students. They were motivated by the little research that existed on whether giving a final exam improves student learning. “In lieu of solid research evidence, professors often make their decisions about the use of cumulative exams based on heuristics from the personal preferences or student opinions” (Khanna et al., 2013, p. 3). While most editorial comments show that students are not fond of comprehensive final exams, there is survey evidence that most students admit that they felt that they had a fuller understanding of courses when they had to review previous class material for a cumulative final (Petrowsky, 1999).

In this study, researchers conducted two experiments. The first asked the question: “Do students who have a cumulative final at the end of a semester score higher on a measure of class content knowledge than students who do not have a cumulative

semester final” (Khanna et al., 2013, p. 5)? The second experiment measured student retention up to eighteen months after course completion, comparing students who took a cumulative final in their course and students who did not. Psychology course students in both introductory and upper-level courses were the subject of these two experiments. They found that students who took a cumulative final in an introductory course outperformed those students who did not have a cumulative final. In both introductory and upper-level courses, all students who had taken a cumulative final performed better on assessments given soon after their final exam. The long-term advantage was not obvious for upper-level students, most likely because they had been repeatedly exposed to similar content throughout the multiple related courses they had taken. The authors acknowledge that other factors like motivation and interest were unmeasured but important contributors to student learning (Khanna et al., 2013). However, efforts made by educators to know and understand this paradigm and use it may improve student overall retention and comprehension of important course material. Being transparent with students about the benefits of the final exam and retaining information for future use is an important step in this process and one often overlooked by educators. Khanna et al. strongly recommended the use of cumulative finals to colleagues for the simple fact that, “Improved learning through cumulative exams not only benefits students, but also has the potential to benefit the profession, by producing graduates who retain more of the information they acquire during training” (2013, p. 21).

The Psychology of Testing

Much of the studies on cumulative assessments focus on the testing effect as it relates to retention, but another potential benefit of frequent assessments is its impact on

human behavior. A common complaint about final exams from both professors, teachers, and students is that students wait until the last minute to study and attempt to cram as much information as possible prior to the test. A study published in 2015 investigated how a cumulative assessment program that includes frequent testing might impact student procrastination and test performance. According to Kerdijk et al., “Cumulative assessment combines the principles of spaced testing, repetition of content and compensation among tests in order to stimulate students to study regularly and to benefit students’ test performance” (2015, p. 710). Seventy-eight undergraduate medical students at the University of Groningen in the Netherlands participated in this study. The researchers randomly assigned students to one of the two conditions for the experiment. In the first group, the cumulative assessment condition, students were assessed in weeks 4, 8 and 10 while students in the second group, the end-of-course assessment condition, were only assessed in week 10. Each week students self-reported the number of hours they committed to studying (Kerdijk et al., 2015). As expected, students in the first group spent 69 more hours on self-study throughout the course than students in the second group. The only exception was during the finals week when the second group spent seven more hours studying than the first group (Kerdijk et al., 2015). Unlike previous studies that show clear benefits to test performance for students who have cumulative tests, this study found no significant benefit for the first group on test performance. Considering this study was conducted on medical school students, this is not surprising. However, the study did note that students in the first group did noticeably better on the final exam test questions from the last part of the course. The researchers found that students in the first group had more time to focus on the most recent content since they had spent more time

previously studying earlier content (Kerdijk et al., 2015). Though this study focused on a specific demographic of college student, it provided a sound framework for establishing a comprehensive and clearly articulated cumulative assessment plan that encourages students to distribute their learning thus avoiding procrastination and potentially increasing retention.

The utility of a comprehensive and cumulative assessment plan following the testing effect model extends beyond student test performance and long-term retention. It can greatly increase a students' confidence along with their study habits. Faculty at the collegiate level consider prior knowledge, study habits, confidence, and motivation to be significant factors in a student's academic success. These ideas were explored by researchers at Montana State University who wanted to examine the effect of two different testing strategies on academic achievement and end-of-course assessments in an introduction to statistics course (Myers & Myers, 2007). During their initial research, they focused on social cognitive theory and self-efficacy, the idea that students are more motivated to learn if they believe they can get positive results. They wanted to see what type of assessment strategy would positively influence self-efficacy and therefore motivate students to study. They proceeded with the belief found in research that tests should be given "early and often so students receive continuous formative classroom assessment" (Myers & Myers, 2007, p. 229).

This study also used an introductory college course for its experiment. The courses were the same with instruction and content but had two different assessment models used. The first course assessed students with bi-weekly exams and a final exam while the second course assessed students with two midterm exams and a final exam. The

bi-weekly exams were 25 minutes each and the midterms were one-hour each so that the amount of time tested was identical between the two courses. Both groups took the same two-hour final exam. The results of this study were conclusive in that students in the bi-weekly exam group received a final grade that averaged fifteen percentage points higher than students in the midterm group. When comparing the same test questions between the bi-weekly and midterm tests, the students in the bi-weekly group scored better. Not only were scores better, but students were more confident and positive about the bi-weekly course. No students dropped the course as opposed to the midterm group which lost 11 percent of students who withdrew after the first midterm (Myers & Myers, 2007). All of the students in the bi-weekly test group rated the course “One of the Best” or “Better than Most,” while only 69 percent of students in the midterm group ranked the class the same. Asked if they would recommend the course to friends, 49 percent of students in the bi-weekly group said they would “Definitely” recommend it but only 14 percent in the midterm group felt the same (Myers & Myers, 2007). Finally, while both male and female students benefitted from the biweekly format, females did so at a greater rate. This revealed another unexpected finding, that gender differences exist in some courses and that educators need to be cognizant of it.

The testing effect was once again reinforced with study results illustrating that distributed assessment and evaluation practices reduced procrastination and cramming, increased student self-efficacy and confidence, and improved overall learning and achievement. Nearly a year later, another study wanted to go a step further and see if the type of assessment made a difference in student performance. Knowing that the Myers and Myers study only used a multiple-choice testing format, another researcher wanted to

compare close-ended and open-ended formats. In this 2007 study, researchers looked at student exam and achievement data from three psychology courses to compare the reliability and validity of close-ended and open-ended assessments (Bleske et al., 2007).

As previously stated, multiple choice assessments gained popularity in the early twentieth century when they came into use by the Army in order to conduct large-scale intelligence testing (NEA, 2020). This coincided with the advent of standardized testing and provided a more efficient and objective method of assessment. At the time, Army developers lauded the multiple-choice format for providing, "...a way to transform testees' answers from highly variable, often idiosyncratic, and always time consuming oral or written responses into easily marked choices among fixed alternatives, quickly scorable by clerical workers" (Office of Technology Assessment, 1992). Modern standardized tests rely heavily on the multiple-choice format, from college readiness tests like the SAT and ACT to classroom unit and final exams.

Early critics of the multiple-choice format worried that it encouraged students to memorize or guess (Office of Technology Assessment, 1992). Modern critics of multiple-choice exams argue that it serves the opposite effect on motivation because preparing for an exam leads to extrinsic motivation, rather than intrinsic motivation to learn (Myers & Myers, 2007). Extrinsic motivation can negatively impact student's natural curiosity. One research study looked at the relationship between motivation and procrastination and concluded that students who were extrinsically motivated had poorer academic performance and greater procrastination tendencies (Myers & Myers, 2007). There is some research that suggests that students who perform poorly on closed-ended assessments do better on open-ended assessments. This may put some students at a

disadvantage if they are in a course that solely relies on closed-ended assessments, which is the trend in higher education (Myers & Myers, 2007).

As the number of students attending both high schools and universities has grown, educators have become more reliant on multiple-choice exam formats to manage an ever-increasing teaching load. Furthermore, it is recognized that “closed-ended assessments demonstrate greater content validity and inter-rater reliability than do open-ended assessments” (Bleske et al., 2007, p. 90). In order to further study whether some students are at a disadvantage when only multiple-choice or similar closed-ended assessments are used, researchers looked at three groups of students at various levels in college, either freshman students in an introductory course or upper level and advanced psychology majors. After conducting correlational and performance discrepancy analysis on the exam and achievement data from all three groups of students, researchers concluded that the continued use of multiple-choice items to assess student learning is valid. Their findings were contrary to critical and anecdotal reports that there was a large discrepancy between performance on multiple-choice and short answer assessments. They took their study a step further and found a strong correlation between how students performed on multiple-choice assessments and other measures of general student aptitude, like high school ranking, ACT score, term GPA and cumulative GPA (Bleske et al., 2007).

Critics have also argued that testing increases anxiety and students with existing anxiety experienced greater extrinsic motivation (Myers & Myers, 2007). This is an important factor to be explored when considering the significance of instituting a cumulative testing program into any grading system. When faced with high-stakes tests that may impact a final grade or their future educational opportunities, some students can

underperform. Studies have shown that when a student feels anxiety to perform at a high level, their worry competes with their working memory. Working memory refers to the short-term memory that is used to access information immediately needed for a task. “If the ability of WM [working memory] to maintain focus is disrupted because of situation-related worries, performance can suffer” (Ramirez & Beilock, 2011, p. 211). Stress is a precursor to anxiety and when faced with an environment that is heavily test-laden, stress around testing will naturally increase. Everyone manages stress differently, so one researcher sought to better understand how students handle the stress and how it impacts their performance (2004).

The study focused on the relationship between test anxieties that students experienced during the final exam and to their expectations, performance, and level of preparation (Burns, 2004). Burns discovered that the expectation of performance on the final exam was the biggest factor that increased student anxiety. This was most notable for students who had already performed well on previous assessments in the course. There was no relationship found between student’s expectation of their final grade at the beginning of the course and no evidence of a negative relationship between test anxiety and actual performance on the final exam (Burns, 2004). The conclusion that test anxiety was not a universally limiting factor in student performance is important. Knowing and understanding what is causing test anxiety can help instructors find strategies to ease students’ anxiety during the time of the final exam.

Final Exam Weighting

Student anxiety can be greatly impacted by the weight of the final test to their final course grade. An important component of designing a course is creating a grading

scheme that fairly and accurately represents student performance. Instructors at both the secondary and post-secondary levels have incorporated multiple measures throughout a course that factor into a student's overall final grade. Oftentimes greater weight is placed on cumulative exams like unit tests, midterms and final exams. How that weight is determined is a highly debated and little researched topic. Franke wrote, "The lack of scholarship on assignment weighting in general, and on final exam weighting in particular, is mirrored by the lack of scholarly consensus on the relationship between grades and learning" (2018, p. 91). Franke asserted that while collegiate teaching has been redefined "to improve student learning and classroom participation through innovative uses of new technology, transparency in grading, and strategic use of learning objectives," grade weighting has largely been ignored (2018, p. 91). Though many in the educational community believe that grades are an accurate reflection of student achievement, they also understand that students can become so obsessed with their grades that it works against their learning. This is somewhat counter to the previous research that asserted students can learn simply through the act of being tested (Glass et al., 2013).

After intensive review of previous literature, Franke proposed suggestions on how to assign weight to an assessment based on the assessment's importance. "For although grades are supposed to reflect growth or proficiency, it's possible to assign weights to final exams and assignments that fail to reflect students' abilities" (Franke, 2018, p. 92). For example, if an exam is too lightly weighted that it will have no impact on a student's grade, there is no value for the student's effort. If it is weighted so heavily that it gives no opportunity for student grade improvement, students lose motivation. Franke identified a phenomenon of *reverse inflation* that can occur when an exam weighting is such that the

final exam grade will only cause a final grade to go down or stay the same. Since some instructors may not even realize the impact of their calculations, Franke focused attention on the mathematical opportunities of weighting systems that can be applied to many various teaching scenarios and subject matters. Final exams should have a weight that allows a “reasonable percentage of students to improve” (Franke, 2018, p. 93). This is the only fair way that grades can fairly reflect student performance. The only exceptions would be at the extreme ends: students who have already reached the highest possible grade or students who are failing. To make a final exam meaningful to reach the desired testing effect, the weight must allow for student growth. That means that weight should be intentionally determined, not that it should be the same from course to course. As Franke points out, “One instructor might think it appropriate for a student whose performance in class has been undistinguished at best to be able to reach the next grade level by acing the final, because a high score on the exam indicates increased mastery of course content; another might offer a low-stakes final of less than 10% because the deciding work of the course occurred during the semester and cannot be measured by a test” (2018, p. 93).

Franke looked at multiple weighting systems and evaluated their mathematical opportunities for students coupled with how they might impact student motivation and learning. He noted three problems that emerged from his analysis: when finals are worth too little, student’s scores make little to no difference in their final grades so students have little motivation to put forth their best effort and tests do not reflect student learning; weights can cause a ceiling effect for some students who do not have an equal opportunity to improve their grade; and a too heavily weighted final exam can cause high

stress levels in students and diminishing returns on their grades (Franke, 2018). Picking a final exam weight between 10 and 30 percent is most likely to allow students to show grade improvement. A weight of 15 percent, for example, avoids the motivation issue of a lower weight since most students must at least score a 50 percent to maintain their grade while a weight of 25 percent expands the number of students who can move their grade up to the next level. One creative solution offered by Franke is to “decouple final exams from course grades” (2018, p. 100). This provides the final exam as an opportunity for all students to improve their score by some amount if they score over a certain percent while also allowing students who have hit the ceiling grade to opt out of the exam. Considering an extra credit process would reward students for their content mastery prior to the final exam but may hamper long-term retention as evidenced by the testing effect. Furthermore, in a high school setting it may prevent some students from exposure to a collegiate-style final exam. Regardless of the weight chosen or the system used, instructors in both high school and college should be transparent with students about the effect of the final exam weight on student grades.

Summary

The use of standardized and cumulative testing as a practice in education are well established and unlikely to diminish in importance in the future. The literature review demonstrated that testing has always been synonymous with college preparation. How it has looked over time has shifted from rigorous oral exams of a limited population to lengthy standardized multiple-choice assessments of masses of college bound students. With the adoption of ESEA in 1965 and its subsequent renewals with the No Child Left Behind and Every Child Succeeds Acts, assessment testing is widely used as a graduation

requirement. Education reform relies on standardized testing as a means of monitoring the progress of students and schools. With such a strong emphasis on testing and college readiness, it is not surprising that some high schools have adopted the collegiate final exam model as a way of increasing rigor and preparing students for college.

Simply incorporating a final exam into a course, though, may not yield the desired results of increased learning and retention. There are several factors that educators and schools should consider. The first is the idea of the testing effect which occurs when there are intervening tests such as recall and recognition on content throughout the course. Having a comprehensive and clearly articulated testing plan that includes frequent testing checks teaches students to distribute their learning. It also can increase student confidence, motivation, and study habits which are key college-readiness traits identified by college professors (Myers & Myers, 2007). The next two considerations are grade weighting of final exams and the stress of high-stakes testing. Because test anxiety can negatively impact performance and learning, schools should consider only using final exams in content-rich courses and they should consider the number of tests students are taking during a given period. This should include state standardized tests, AP exams, and college entrance exams like the SAT and ACT. The final consideration should be the impact of the grade weight of a final exam on the students' overall course grade. Too little weight decreases student motivation to retain learning, appropriate weight allows students to improve their grade, and too heavy weight leads to high stress and diminishing returns. Finding the right balance to achieve the desired results may mean adjustments after a plan is put into place. A continual review of the final exam process should be embedded into the plan itself.

Finally, and perhaps most importantly, is clarity of the purpose and process of testing to students. To be successful at achieving the desired results of improved learning and development of college-readiness, teacher and student beliefs and perceptions about the process need to align with the purpose. When teachers are not clear about the benefits or purpose, they cannot clearly articulate the value of the final exam to students. There is significant research on the benefits of cumulative testing, but it is not consistently reflected in the research on teacher and student beliefs about testing. As the key to the success of any assessment plan, teachers need to understand the research and see evidence of it working in their classrooms.

CHAPTER III

Methodology

The role of secondary education as the most impactful preparation for societal success, both economically and ethically, cannot be underscored. From the mission of the first high schools to modern day, secondary coursework has laid the groundwork for students' success at the post-secondary level. Though there is a general agreement on the need to prepare students for college, the literature review revealed that what that means and how that looks varies widely from state to state and school to school. One thing that is consistent is the use of standardized testing to measure student aptitude and knowledge and the efforts of high schools to adopt collegiate practices like final exams. While extensive research supports using final exams in college to increase retention and student learning, little research exists on the high school model. Because final exams have an immediate impact on a student's course grade and grade point average and a potential long-term effect on retention and college preparedness, further study of the practice was warranted by this study, especially with the increasing demands of standardized testing on instructional time.

Research on using final exams revealed one misconception and one truth. The misconception was that final exams caused undue stress that negatively impacted student performance (Burns, 2004). It is well known that students do not like taking final exams, but the research suggested that any stress caused by final exams does not have a widespread negative impact. In fact, most students perceived that they learned more in courses that had a final cumulative exam (Khanna et al., 2013). The biggest factor to increased anxiety was not found to be the final exam, but a students' expectation of how

they should perform on the final exam (Burns, 2004). The truth was found in a phenomenon known as “the testing effect.” Studies demonstrated that there were long-lasting benefits to student learning when a clear cumulative testing plan that included a final exam was implemented (McDermott et al., 2014). Developing good study habits, increasing long-term retention, and building student confidence were among these (Kerdijk, et al., 2015; Myers & Myers, 2007; Glass et al., 2013). Content-heavy courses benefitted the most from the testing effect. Having low-stakes assessments on important content throughout the course helped students to distribute their learning (McDaniel, 2007).

Understanding the misconceptions and truths about final exams found in the literature review provided a baseline for this study of the current beliefs and practices at Peters Township High School. This research was designed to gather the perceptions from teachers and alumni on the purpose of the final exam, its reliability at measuring student learning, and its use as college preparation. Additional data that was relevant to understanding whether the perceptions revealed truths or misconceptions was the historical grading data, including the grades earned by students each quarter, their final exam grade, and their final course grade. This data provided a complete picture of how well the final exam grade reflected quarter grades and the final course grade, and whether the weight hurt or helped students improve their grade. Both quantitative and qualitative information on the experiences and perceptions of teachers and former students was collected. The outcome of this research will provide the groundwork for actionable next steps to keep, revise or eliminate the current final exam system based on its performance

as a measure of student understanding of course content and its alignment with the college-readiness indicators identified in the research.

This chapter outlines the methodology used for this research study. The organization of this chapter includes the following topics: Purpose, Setting and Participants, Research Plan, Research Design, Methods and Data Collection, Validity, and Summary.

Purpose

This study focused on the factors that need to be considered by the Peters Township School District regarding the current final exam process used at the high school. Research was conducted on three areas: identification of the prevailing final exam beliefs and practices of teachers, comparison of final exams in high school and college, and analysis of the final exam grades. This research study used a mixed-methods approach that included both qualitative and quantitative strands through surveys of teachers and alumni that included Likert-style and open-ended questions and a quantitative analysis of historical grades stored in the student information management system.

The beliefs and practices of the teachers who administer the final exams can influence how students perceive and prepare for the final exam. All teachers at the high school were invited to participate in this study. This included teachers of all disciplines and grade levels. In addition, former students were invited to participate in this study to provide feedback on their experiences with final exams at Peters Township High School and in college. The Terre Haute study from the Literature Review provided a comparison for this research since it also focused on teacher and alumni attitudes about final exams.

That study found that teachers and alumni agreed on the purpose of the final exam but did not always agree on its effectiveness at measuring student learning or preparing students for college. The variation in alumni responses was based on their course selections in high school and their courses of study in college (Hall & Ballard, 2000). This could be evidence of differentiation and trends in colleges based on disciplines. Several studies have also shown a disconnect between what high school teachers and college professors think students need to be ready for college (Woods et al., 2018). To effect meaningful assessment practices for students, understanding what teachers perceive about the exam purpose, the exam reliability, and the exam impact was essential to this study. This research sought to identify if the same pattern exists in Peters Township that existed in Terre Haute.

Comparing teacher perceptions to feedback from former students was essential to discover similarities and differences. The desire was to gather information from alumni who had spent several semesters at college and could provide information about their college experiences. Of particular importance was how they were being assessed in college, how that may differ from how they were assessed in high school, and whether they felt their high school final exam experience adequately prepared them for their college final exam experiences. Having a comprehensive understanding of the impact that the high school final exam experience had on students who were currently in college was more valuable than simply researching college assessment trends. Since this research study was focused on the practices used at Peters Township High School, it was necessary to track the experiences of students who graduated from Peters Township High School.

While studies on the effectiveness of cumulative and final exams at the college level provided a good starting point for this research, different factors exist at the high school level that also needed to be considered and further examined. For example, high school final exams may be based on a full year of coursework rather than a semester like in college. Students have choices of academic level in high school which contributes to the rigor of the course. Because high school courses meet daily, students have more opportunities to demonstrate learning and improve their course grades. Former students provided key insights into how the two systems compare and whether their high school courses provided them with adequate college-readiness opportunities.

Finally, grades matter to students. This is true for several reasons but most notably because students are building their transcripts to send to prospective colleges. With most Peters Township students planning to attend college, they seek to have a higher GPA to demonstrate their college readiness. Historically, colleges weighed this information alongside student performance on college entrance exams like the SAT or ACT. This gave them an overall picture of the potential future college freshman seeking admission. Recently, this process has shifted. It is important to note that many colleges have waived college entrance exams beginning in 2020 during the Covid-19 pandemic. The continuation of this trend is unknown, but this has placed even more emphasis on high school GPAs. Previous studies have shown that the high school GPA was a strong predictor of college success (Koretz & Langi, 2017). This supported the grade motivation of college-bound students. Szpunar (2007) and Franke (2018) in their respective research found that the expectation of future testing increased retention of learning and that the weight of the exam influenced student motivation. Peters Township High School is a

high-achieving school with most students planning to attend college or post-secondary training, some competing for admission into highly selective schools. The evaluation of the current final exam beliefs and practices was necessary to determine if the current system is helping, hurting, or even having an impact on student preparation for college or admission to highly selective schools.

Setting and Participants

Peters Township High School is part of a high-performing suburban school district that covers 19.73 square miles located in Washington County, Pennsylvania. Peters Township School District educates approximately 4,000 students each year in grades K – 12. In 2019, Peters Township School District was ranked first in the state for Keystone Exam Scores based on the percent proficient and advanced in all three tested subjects combined and third in the state for PSSA Exam scores based on percent proficient and advanced for grades 3 through 8 in all subjects combined. The district has three elementary schools, one middle school and one high school. Two of the five schools, Pleasant Valley Elementary and Peters Township Middle School, have been recognized as National Blue Ribbon Schools by the U.S. Department of Education. In 2020, Peters Township School District was ranked #4 among the 104 school districts in Southwestern Pennsylvania by the Pittsburgh Business Times (Lott, 2020). The U.S. News & World Report's Best High School Rankings (2021) placed Peters Township High School #18 out of 754 high schools in Pennsylvania and #5 out of 147 high schools in the Pittsburgh Metro Area and #727 out of nearly 15,000 high schools in the United States. These rankings included traditional, charter, and magnet schools.

For the purpose of this study, demographic information for Peters Township High School was based on data submitted to the Pennsylvania Department of Education (PDE) and reported via the Future Ready Index (2020). The Future Ready Index is a tool used by PDE to report information and progress measures for each school in Pennsylvania. Information reported on the Future Ready Index included both the 2019-2020 and 2018-2019 school years. The high school enrollment for the 2019–2020 school year was 1,371 students in grades 9 through 12. The following student statistics accurately represent a typical school enrollment by percent of students in each category at Peters Township High School:

- Male 50.3% and Female 49.7%
 - Gifted 7.7%
 - Special Education 12.1%
 - Economically Disadvantaged 8.4%
 - English Language Learner .1%
- Percent Enrollment by Race/Ethnicity
 - White 91.5%
 - Asian 3.2%
 - Hispanic 2.3%
 - American Indian/Alaskan Native .4%
 - Black .3%
 - 2 or more races 2.1%
- Student to Teacher Ratio 16:1

An important consideration in this study is the number of students graduating from Peters Township High School with the intent of pursuing their studies at the college level. Information on the Career Standards Benchmark relative to this include the following:

- Percent of students meeting the Career Standards Benchmark 100% (10.2% higher than the state average)
- Graduation Rate
 - Four-Year Cohort 98.7%
 - Five-Year Cohort 99.4%
- Percent of students in a rigorous course of study 82.4% (24.9% higher than the state average)
 - Percent AP Participation 77.9%
 - Percent College Course Enrollment 22.9%
 - Percent CTE Program of Study 2.1%
- Percent of graduates transitioning to post-secondary education or military 91.1%

With over 90 percent of students who attend Peters Township High School seeking to continue their education, the emphasis on instructional practices that will best prepare them for this transition is a high priority.

Instructional staff includes 86 full time teachers and five part time teachers. Additional support staff includes five school counselors, one full time and one part time gifted coordinator, a librarian, a school nurse, a social worker, and ten paraprofessionals. The administrative staff is made up of a principal, two assistant principals, a dean of college and career readiness, an athletic director, and an assistant athletic director.

The school district has invested time and resources into improving students' college and career readiness. In 2015, the school district participated in the College Readiness Program with the National Math and Science Initiative (NMSI). This two-year program provided intensive professional development for Advanced Placement and pre-Advanced Placement teachers in English, math, and science. It also provided workshops for students and incentives for students to participate in and receive a passing score on at least one AP exam. The NMSI partnership yielded incredible growth in the number of students who take AP classes and AP exams. The goal of the program was to ensure every student planning to attend college would take at least one AP course before graduating. In the AP Score Report for Equity & Excellence reported on October 16, 2020, Peters Township High School had 62.4% of graduates from the class of 2020 who took and scored a three or higher on at least one AP test during their high school career. As a comparison, the class of 2016 had only 43.4% of graduates who met this benchmark.

In 2017, the position of Dean of College and Career Readiness was created. This individual directly supervises the counseling office, works with post-secondary institutions, and develops programs for staff and students related to college-readiness. The school began administering the PSAT to all students in grades 9 through 11 at district expense to better prepare students for college entrance exams. The counseling office works with students and families to understand these scores including identified strengths and weakness. Students connect their PSAT scores to Khan Academy for individualized practice to improve their skills since most Peters Township students take the SAT exam.

There are 91 teachers at the high school, both full time and part time. All were invited to participate in this study and 45 consented to participate. The Class of 2019 was selected for alumni invitations to participate in the study for two reasons. The first was that with a least three semesters of college, they would have enough college experience to give meaningful feedback and that having only been out of high school for three semesters, they still have accurate memories of their experiences.

Ethical guidelines were followed to ensure that teachers and alumni were not harmed during the research process and that their identities would not be recorded. Participating in the surveys was voluntary. A survey disclosure (Appendix A) was sent with the electronic survey via email to all prospective teacher respondents describing the action research project and how survey data would be used. The survey was administered via Google Forms and included a consent to participate (Appendix B). The alumni survey was sent to 347 graduates of the class of 2019 of which 65 consented to participate. Alumni received a survey disclosure describing the purpose of the action research project and how the survey data would be used (Appendix C). The survey was administered via Google Forms and included a consent to participate. (Appendix D) In both teacher and alumni disclosures, participants were assured that their anonymity would be preserved.

Research Plan

It is evident in the review of literature that the beliefs of teachers and students as well as the practices of teachers and institutions significantly influenced the effectiveness of using a final exam as a cumulative assessment. For example, the Terra Haute study revealed a frequent disconnect between the purpose of a final exam and its effectiveness and the beliefs of students compared to teachers (Hall & Ballard, 2000). While students

and teachers agreed that the tests were meant to assess knowledge and increase retention, they disagreed on whether the tests accomplished that and prepared students for college (Hall & Ballard, 2000). Some teachers may believe that their practices are preparing students for college, but that might not be the case. College instructors may gage college-readiness by students' prior knowledge, study habits, confidence, and motivation (Myers & Myers, 2007). Some studies evaluated college-readiness by students' completion of their first year of college without needing to take a remediation course (Woods et al., 2018). The goal of this research is to determine where the previous research aligns with or differs from teachers' perceptions and alumni experiences at Peters Township.

The benefits of using cumulative final exams in high school courses are supported by research regardless of whether their purpose is driven by measuring student learning or preparing students for college. According to researchers at Rutgers University there are benefits to using final cumulative exams to measure student performance, incentivize students to distribute their study and revisit previously tested content and to increase long-term retention (Glass et al., 2013). In order for a school to successfully use final exams, there needs to be an understanding of the purpose, benefits, and limitations of using them. Pedagogical practices need to align to research and college trends.

This research study was planned to evaluate the current final exam process at Peters Township High School by collecting information from teachers and alumni and analyzing historical grades. Information was gathered from current teachers to assess their perceptions of purpose, reliability, and impact of their current final exam practices.

This provided data for the first and second research questions:

1. What are teachers' perceptions of the purpose of the final exam?

2. What are teachers' perceptions of the reliability of the final exam as a measure of student learning?

Also measured were the perceptions of former students currently enrolled in college regarding their preparation for college and practices in high school compared to their current college experience. This provided data for the research question:

3. How does the current final exam structure compare to current assessment trends at the collegiate level?

Information from teachers and alumni was used to identify common and differing beliefs. This was further differentiated by teachers' experience and content area and students' pursued course of study. This information provided a groundwork with which to evaluate the historical grading data for two previous school years in math and English courses. This data provided evidence to address the final research question:

4. What is the impact of final exam grades on students' grades?

English and math grades were selected as representative of all content courses. Student quarter grades provided a baseline to identify a pattern of grades to identify the whether the final exam grade was representative of the pattern.

The fiscal implications of this action research project are both minimal and indirect. There were no financial costs associated with conducting the surveys or collecting historical grades data as all electronic tools were free. In-direct costs include time to conduct the study for both the researcher and participants. Potential future costs may include professional development time and resources for teacher training should they be recommended as a result of this study. In-service time during the school year may be used to review the conclusions of this research study and develop action plans. This

time is already built into the school calendar. A revision of the current process may involve time to create new assessments or curriculum revisions.

Research Design, Methods & Data Collection

This action research study used a mixed methods approach with both qualitative and quantitative strands collected in a convergent parallel design. Both quantitative and qualitative survey data was gathered using a Likert scale and open-ended questions. Simultaneous to the surveys, quantitative data was collected from the PowerSchool student management database. The mixed methods approach was selected because one data source was not sufficient to examine the breadth and depth of the research questions.

Qualitative data was collected in the form of open-ended questions to both teachers and alumni. These questions asked teachers what they preferred as their assessment method and what recommendations they had for giving final exams. The alumni survey asked two open-ended questions to elicit feedback on their college assessment experience and identify skills they wish they had learned in high school. Quantitative data was collected through Likert scale survey questions and collection of historical grading data. Both surveys used a 5-point Likert scale from “Strongly Agree” to “Strongly Disagree.” The teacher survey asked eight questions specific to final exams at Peters Township High School and six general questions about final exams. Teachers also identified their content area and years of experience. Alumni were asked five questions about their high school and college assessment experiences. They also identified their current educational level and planned course of study. Two questions on the alumni survey paralleled two questions on the teacher survey.

The final piece of quantitative data analyzed was historical grading data from English and math courses at each grade level from the 17-18 and 18-19 school years. Data included each quarter grade, the final exam grade, and the final course grade. Since the grading system only used letter grades and their equivalent quality points rather than percentages, no percentages were used in this study. The exam weight was analyzed to determine its impact on student motivation and learning. It was compared against the three potential problems of too little weight, the ceiling effect which prevented any grade improvement, or too much weight (Franke, 2018). Both academic and honors level classes were used to identify potential trends based on course rigor.

Using the mixed methodology, trends, and patterns both within and between data sources were identified. The collection of qualitative data for teachers and alumni provided context to the responses and gave insights into what teachers believe about college-readiness, how they envision any future changes to the final exam process, and what students believe would have helped them better prepare for their college courses.

The research and data collection procedures included in this capstone research project were submitted to the Institutional Review Board on August 14, 2020 for review. Additional information was requested and provided. After clarifications were made to the role of the researcher, changes were made to ensure an external email was used to invite participants to this study. Final IRB approval was received on September 17, 2020 to conduct this study using a mixed-methods approach (Appendix E). During the months of November and December 2020 data from the 2017–2018 and 2018–2019 school years was pulled from the PowerSchool system and put into an Excel file. To provide a purposive sampling, grades were collected for two core classes at each grade level

representing two of the four subject areas of English and math at both the honors and academic levels. Data was pulled for the following courses from the 17–18 school year: English 9 Honors, Geometry Honors, English 11 Academic, Calculus Honors. Data was pulled for the following courses from the 18–19 school year: Precalculus Academic, Precalculus Honors, English 10 Academic, English 10 Honors, English 12 Honors. Calculus Honors and English 12 Honors were selected as they represented students from the academic track who are college bound and elect this honors level course their senior year. Excluded from this study were Advanced Placement courses whose students take the AP test in lieu of a final exam. All identifying student and teacher information was replaced with either an unidentifiable number (Student 15) or letter (Teacher A). Courses were divided by tabs and each student data strand included four quarter grades, a final exam grade, and a final course grade.

Two surveys were created using Google Forms (Appendices B & D). Per the IRB approval, each survey included a disclosure statement indicating the purpose of the study, the role of participants, how the data would be used, and that participation was voluntary and anonymous. The survey further stated that consent to participate was given through submission of the survey. Google Forms was selected to administer the surveys due to its ease of use for electronic surveys and the quick turnaround time. Surveys were delayed in being sent due to school closure and transitioning of the high school to a new facility in December and January. The teacher survey was sent on February 1, 2021 with a reminder sent on February 8, 2021. The Survey was closed on February 14, 2021. Of the 91 teachers invited, 45 or 59 % participated. While this return was lower than expected, it was understandable considering changes to teacher workload and schedules this year as a

result of the Covid-19 pandemic and a mid-year move to a new high school. On February 12, 2021 the alumni survey was sent to 347 former students. A reminder was sent on February 19, 2021 and the survey closed on February 26, 2021. Some emails were returned as undeliverable. Nineteen percent of alumni responded, which was greater than expected and met the recommended threshold of thirty participants (Mertler, 2019).

Survey data was converted into two Google Sheets for analysis as soon as surveys were closed. Survey responses were analyzed for similarities and differences in responses between and among groups. Pie charts were used to illustrate responses to each survey question. This allowed for easier cross-referencing between alumni and teachers. Grading data in Excel was converted from letter grades to quality points to better analyze the results, specific focus was put on the relationships between the final exam grade and quarter grades and between the final exam grade and the final course grade. Grading data was coded in the language R then run through the RStudio program, a statistical coding software. This produced scatterplots and correlation data.

Validity

This action research project was intended for use by the Peters Township School District to evaluate the current final exam process and identify research-based actionable next steps. While the data is singular to Peters Township, the outcome of this study was intended to be applicable to other educational entities who can use it as a comparison to their own systems and processes. To ensure the validity and reliability of this research study, strategies were employed to minimize potential bias, promote objectivity and recognize study limitations. This included using an accurate electronic data recording system. For the survey, the Google Suite of products was used through a password

protected, secure server. Google forms and Google sheets were used to record data. Excel was used to record historical grading data. To invite both teachers and students to participate in this study, a non-school district email was used. The California University of PA email system originated the invites and was used to gather information for this study. This was done to ensure separation of my role in the school district from my action research study. Emails of respondents were not recorded to guarantee anonymity of participants. The use of a digital survey allowed for consistent and accurate recording of responses. Lastly, all high school teachers and all graduates of the class of 2019 were invited to participate.

As a means of ensuring validity with the mixed-methodology approach to research, Mertler (2019) recommends using a set of guiding questions developed by Leedy and Ormrod (2013). These questions were used to benchmark the legitimacy and validity of this research. The questions are as follows:

Question 1: Are the samples used for the quantitative and qualitative components of the study sufficiently similar in order to justify comparisons between the two types of data?

Question 2: Are the quantitative and qualitative data equally relevant of the same or similar topics and research questions?

Question 3: Are the two types of data weighted equally in drawing conclusions? If not, what is the justification for prioritization of one type over the other?

Question 4: Are you able to use specific qualitative statements or artifacts in the study in order to support or illustrate some of the quantitative results?

Question 5: Can obvious discrepancies between the two types of data be resolved?

In response to questions 1, 2 and 3, both the teacher and alumni surveys were written to reflect the perceptions of the respondent to the final exam process at Peters Township High School. The Likert-scale and open-ended questions were similar in both surveys in format and question stems. The Likert-scale questions identified trends in perceptions of teachers and students based on their content focus and experience. They also identified patterns in beliefs about the reliability of the final exam and its impact on college-readiness. Open-ended questions for both teachers and alumni provided information about individual preferences and experiences. This established a basis for comparison. Finally, the grading data analysis was used concurrently to prove or disprove perceptions indicated in the surveys. All data points, quantitative and qualitative, were equally considered in this study. Since the purpose of this study was the perception and reliability of final exams, both the qualitative questions and conclusions and the quantitative analysis of data was of equal consideration to drawing conclusions.

For question 4, the qualitative statements support the results of the quantitative data from the Likert-scale questions and grading data. Participant comments also illustrated the trends found in responses based on content areas and experience. There are no obvious discrepancies between the quantitative data and qualitative data collected. Open-ended survey responses reflected the same perceptions and trends as the Likert-scale questions but provided further explanation and context. Furthermore, many of the opened-ended survey responses helped to explain the patterns of performance noted in the grading data.

The final means of ensuring validity was through the triangulation of data. According to Mertler (2019), “triangulation is the process of relating multiple sources of data to establish their trustworthiness or verify the consistency of the facts while trying to account for their inherent biases.” Multiple sets of data were included in this study to verify the accuracy of the data and add credibility to the findings. In addition to the triangulation of open-ended responses to Likert-style responses discussed earlier, historical grades were compared to student and teacher perceptions of the accuracy of final exam grades as a measure of student learning and an indicator of college-readiness. This process provided clarity and meaning to the study’s overall findings.

Summary

The goal of this action research was to use the findings to determine if the current final exam system at Peters Township High School was supported by the research on final exams and aligned with students’ college experiences. A successful final exam process should include a clear and consistent college-readiness plan that articulates a comprehensive assessment framework. This framework would include both low-stakes and high-stakes testing and provide opportunities for students to demonstrate their knowledge and improve their course grades. Presentation of these findings to stakeholders within the school system will allow for discussion of the current process and recommendations for changes.

The mixed-methods approach to data collection was used to provide adequate depth and breadth to the analysis. Quantitative data was collected through 17 survey questions, 14 of which were Likert-scale, to teachers and 7 survey questions, 5 of which were Likert-scale, to alumni. Qualitative data was collected through 2 open-ended

questions to teachers and 2 open-ended questions to alumni. Quantitative data was also used from historical grades of nine courses across the two core content areas of English and math covering all four grade levels. Two survey questions on the teacher and alumni surveys paralleled each other providing an opportunity for comparison of their perceptions of the same topics. Google products including forms and sheets along with Microsoft Excel were used for data collection and analysis.

The four research questions provided the framework for this study. Data on teachers' perceptions of the purpose and reliability of the final exam (research questions one and two) was used as a comparison to current assessment trends at the collegiate level as reported by alumni (research question four). Historical grades data was used to determine the impact of final exam grades on student grades (research question three). This data was also used to identify truths and misconceptions from teacher and alumni perceptions. Through this investigation, a more complete picture was created of the purpose, reliability, impact, and relevancy of the current final exam process.

Chapter 4 will outline the data analysis and results of this study by answering the four research questions and drawing conclusions based on the research included in the Literature Review. Suggestions for future planning related to the use of a final exam are provided.

CHAPTER IV

Data Analysis and Results

This study sought to determine the level of understanding of teachers and students of the usefulness and effectiveness of a final exam and to determine the actual impact of the current final exam on student grades for Peters Township High School. Embedded in this research process is a two-fold analysis of how final exams are used to both measure student learning and prepare students for college. A key factor in this analysis was the perceptions of both the educators creating and using the exams and the students taking them.

To accomplish this, all part-time and full-time teachers were invited to participate in a survey. The survey revealed how teachers use data from final exams and how they value the practice of giving final exams. Former Peters Township High School students were also invited to participate in this research through a survey. The former students (or alumni) provided insight into the final exam process as they experienced it in high school and shared their current college experience. In addition to the survey results, a data analysis was completed on grading data for two years of final exams in math and English courses. This data analysis was used to determine the calibration of the final exam grade to student overall performance as measured by quarter grades and end of year grades.

The study provides valuable information that can be utilized by teachers and administrators to make decisions on how final exams are used at Peters Township High School in the future. The survey results, grading analysis, and key findings are outlined in this chapter. This chapter is organized into the following sections: Data Analysis, Results, Discussion, and Summary.

Data Analysis

Quantitative and qualitative survey data were collected from both teachers and former students. This included demographic data for teachers including content areas, and years of experience. Alumni provided their current college level and course of study. That allowed survey responses to be disassembled and categorized to determine trends in common beliefs or themes. A narrative review of the qualitative data from the open-ended survey responses also looked for trends in perceptions. This was used to draw conclusions and identify any data that requires clarification or need for additional research.

Analysis of stored grades for two school years and nine courses provided quantitative data for analysis. Data was gathered from the PowerSchool student management system from the 2017-2018 and 2018-2019 school years for the following courses: English 9 Honors, English 10 Academic, English 10 Honors, English 11 Academic, English 12 Honors, Geometry Honors, Precalculus Academic, Precalculus Honors, and Calculus Honors. Courses were selected to cover all four grade levels as well as multiple academic levels. Advanced Placement courses were excluded from this study as the AP test replaces a final exam for most students.

Excel was used to organize grading data for each course. Grading data included quarter grades, final exam scores, and final grades for each student. No student names were used, rather each grading record was given a number. To assess the impact of final exam grades on final grades and to determine if the average of the quarter grades correlated to the final exam grade, letter grades were converted to the quality points used in the Peters Township High School grading system. This allowed for an accurate

analysis of the impact of the final exam under the current grading practice. Additionally, a column was added to average the quarter grades to allow for easy comparison with the final exam grade.

The R programming language, commonly used for statistical computing, was used to code the data and then it was run through the RStudio program to generate scatter plots, correlations, and comparisons. This revealed the relationship between the final exam grade and both quarter grades and final grades. It also provided evidence of similarities and differences between content areas and academic level.

The final step of the analysis process involves the triangulation of the quantitative and qualitative data points. Areas of corroboration between the three sources of data has been identified and noted.

Results

Teacher Survey Results

The teacher survey was sent to 91 teachers at the high school of which 45 completed the survey and consented to participate in the study. Survey questions were divided into two categories with additional identifier questions. The first set of questions was specific to Peters Township High School and the second set of questions was general to perceptions about final exams.

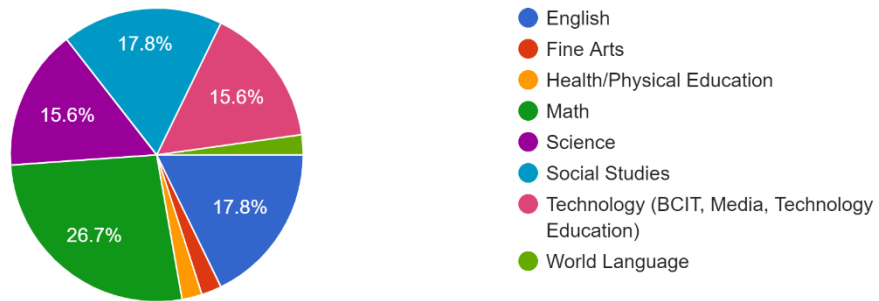
The demographic or identifier questions were meant to provide context to the perception questions. These included questions identifying the type of final exam used, the content area taught, and years of teaching experience. All departments were represented in the survey (Figure 1). The largest responding group was math with 12, followed by English and social studies with 8 each, science with 7, and electives making

up 10. Electives include teachers of fine arts, health/physical education, technology, and world language.

Figure 1

Content Areas of Participants

I teach in the following content area
45 responses

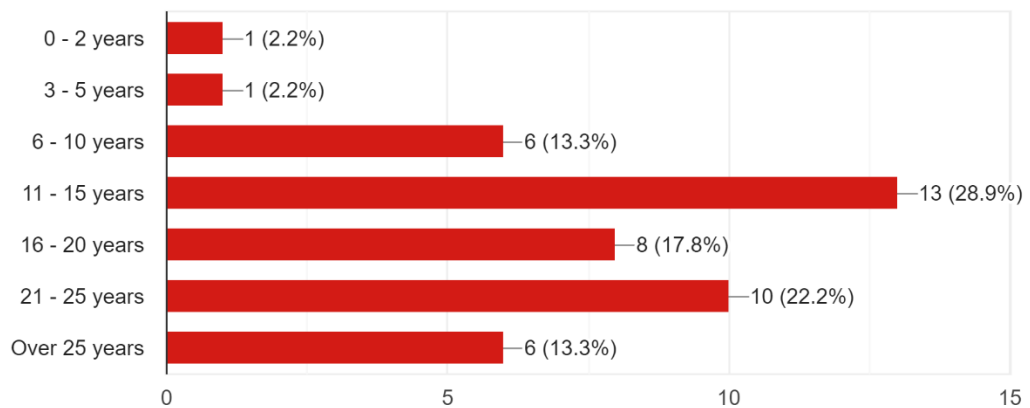


All levels of teaching experience were represented by respondents (Figure 2).

Teachers with between 11 and 25 years of teaching experience made up most respondents with 31, while only 2 teachers with between 0 and 5 years responded. This is representative of the experience level of the full teaching staff.

Figure 2

Teacher Experience of Participants



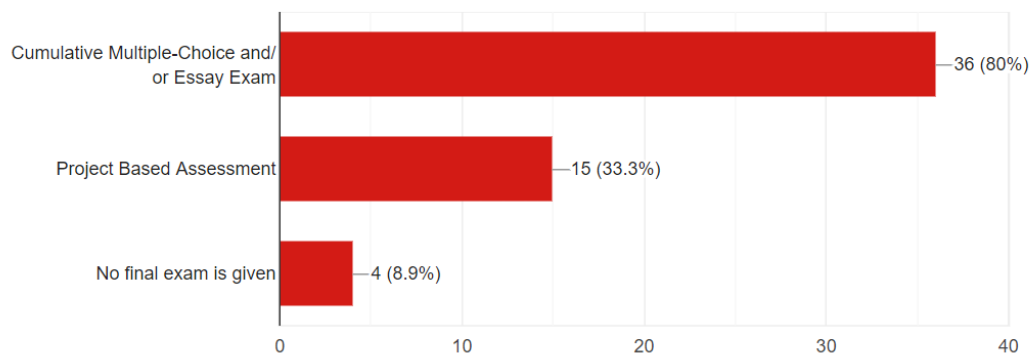
The survey established that most participating teachers at the time of the survey

were using a cumulative multiple-choice and/or essay exam (Figure 3). Teachers were given the following options to choose from: Cumulative Multiple-Choice and/or Essay Exam, Project Based Assessment, or No Final Exam. Teachers were able to select multiple options since they could teach more than one course with a different final exam type. Of the 45 respondents, 80% indicated that they used the traditional final exam. Four teachers responded that no final exam was given, however all four also selected either a project-based or cumulative exam as well.

All the teachers of English, math, social studies, science and world language selected cumulative multiple-choice and/or essay exams. Of those teachers, three math teachers also selected no final exam for their elective courses and four teachers (three English and one world language) selected project-based assessment for their elective courses. All remaining nine teachers who selected project-based assessment were elective teachers. This provided context to the responses for the high school specific Likert statements.

Figure 3

Question 1 Results: I teach a course that uses the following type of final exam



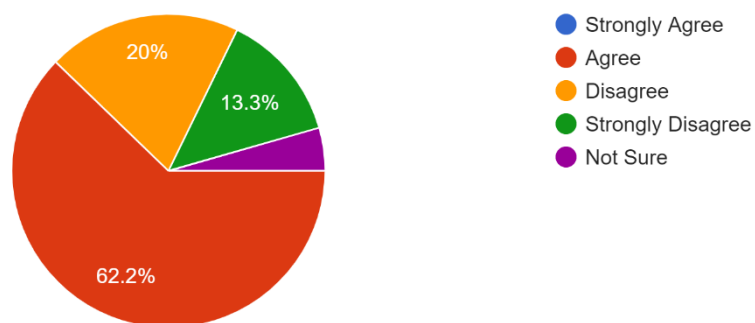
The first category of Likert scale questions included eight statements specific to

each individual teacher's final exam practice at Peters Township High School at the time of the survey. The second category of Likert scale questions included six statements related to general beliefs about final exams. While the Likert responses collected quantitative data, two open ended questions were included to collect qualitative data. This allowed teachers to provide more detail or expand their thoughts. The Likert scale statements required each participant to indicate their level of agreement or disagreement. Their choices included: Strongly Agree, Agree, Disagree, Strongly Disagree, and Not Sure.

The first Likert scale question showed that most teachers agreed that they used final exam data to determine if students mastered the course content (Figure 4). While no teachers strongly agreed with the statement, 62.2% agreed with it. However, 20% of teachers disagreed and 13.3% strongly disagreed. Math teachers were evenly split while social studies and English teachers mostly agreed with the statement. Of the elective teachers, 70% agreed, all of whom said they used project-based assessments.

Figure 4

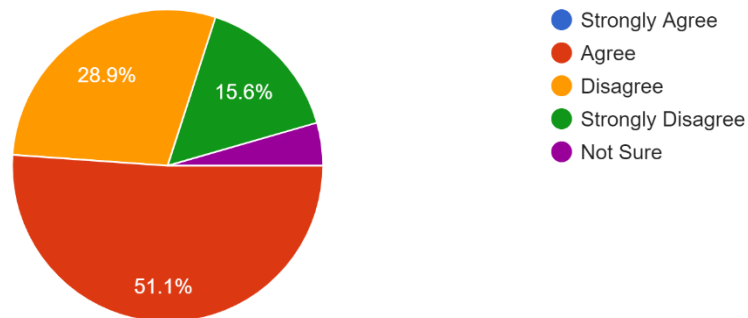
Question 2 Results: I use final exam data to determine if students have mastered the content.



Responses to the next question revealed a divide when teachers were asked if they used final exam data to evaluate curriculum (Figure 5). While 51.1% stated that data was used, no teachers strongly agreed with the statement and 44.5% either disagreed or strongly disagreed with the statement. The split was consistent across all content areas except social studies which showed 5 of the 8 teachers agreeing. Two teachers marked that they were unsure.

Figure 5

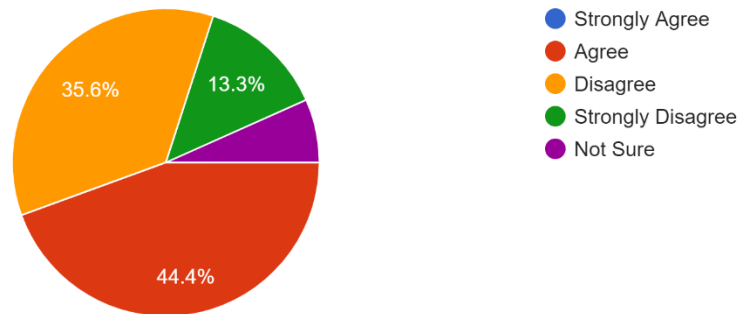
Question 3 Results: I use final exam data to evaluate my curriculum.



Similarly, teachers were divided on whether the final exam grade was an accurate measure of student learning for their classes. Nearly half of them disagreed with the statement (Figure 6). In total, 44.4% of teachers felt that their final exam grade was an accurate measure of student learning, while 48.9 % felt that it was not. This included 35.6% disagreeing and 13.3% strongly disagreeing. Three teachers indicated that they were not sure. This discrepancy was consistent across all departments, including electives.

Figure 6

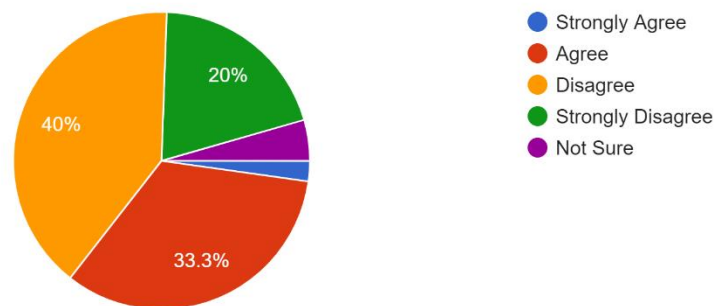
Question 4 Results: The final exam grade is an accurate measure of student learning in my class.



An overwhelming number of teachers felt that the final exam grade was not consistent with how students performed on assessments throughout the course term with 60% disagreeing or strongly disagreeing with the Likert question: The final exam grade is consistent with how students perform on assessments throughout the course term (Figure 7). Only 33.3% of teachers felt that the end of year assessment was consistent with prior performance and one teacher strongly agreed. The social studies department was the only group that had more teachers agree than disagree.

Figure 7

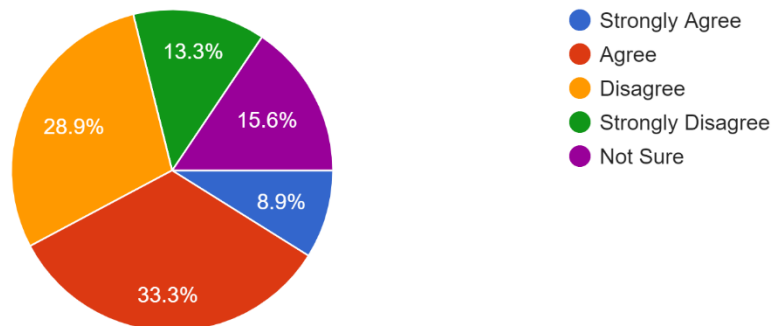
Question 5 Results: The final exam grade is consistent with how students perform on assessments throughout the course term.



The same trend continued with the next statement: The final exam prepares students for college-level cumulative assessments. Teachers were equally divided on whether the final exam prepares students for college-level cumulative assessments (Figure 8). Of the 42.2% of respondents who agreed with the statement, 8.9% selected strongly agree. Of the 42.2% who disagreed, 13.3% selected strongly disagree. Most respondents who selected the not sure option were elective teachers. English and social studies teachers leaned slightly more towards agree while science teachers more frequently disagreed. Math teachers were equally divided.

Figure 8

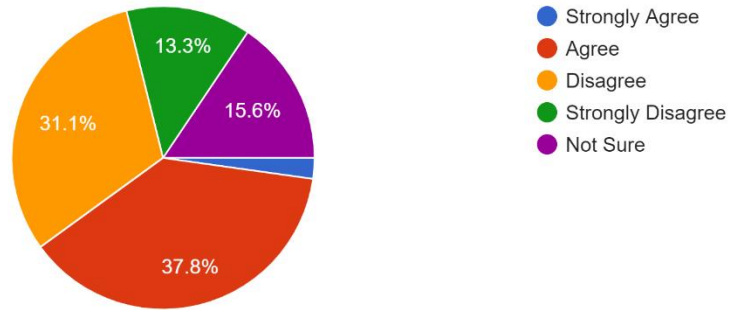
Question 6 Results: The final exam prepares students for college-level cumulative assessments.



Teacher perceptions on whether the final exam calculation gave students the opportunity to improve their grade followed the same trend, but with slightly more disagreeing (Figure 9). While 40% of teachers agreed or strongly agreed, 44.4% of teachers disagreed or strongly disagreed. English and social studies teachers agreed more frequently than math and science teachers. Most teachers who were unsure also came from social studies.

Figure 9

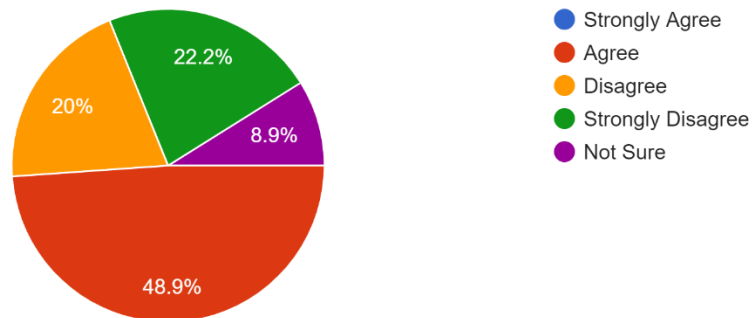
Question 7 Results: The final exam calculation allows students who are behind a chance to improve their grade.



The last question that focused specifically on the final exam process at Peters Township High School illustrated the continued split in perceptions. In this question, teachers were asked if the final exam calculation was fair and consistent to all students. Slightly more teachers believed it was (Figure 10). While 48.9% agreed with the statement, 42.2% disagreed or strongly disagreed. Most departments were divided on their responses except for science. Of the science teachers, 71% agreed. Four teachers indicated that they were unsure.

Figure 10

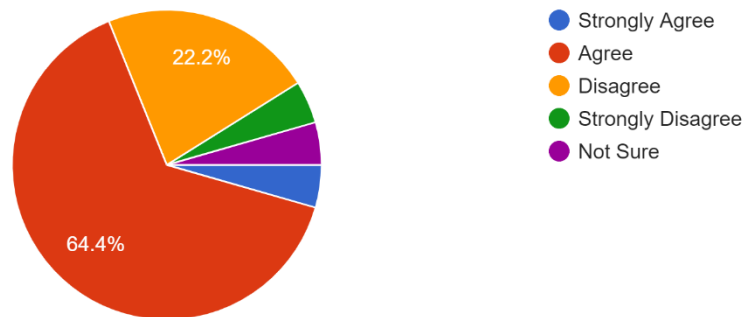
Question 8 Results: The final exam calculation is fair and consistent to all students.



The next series of questions included six Likert statements designed to gauge general teacher beliefs about final exams. The first statement was: Having a final exam holds students accountable for their learning. Most teachers (68.8%) believed that having a final exam does hold students accountable for their learning (Figure 11). All the science teachers and 83% of the math teachers agreed. Over 60% of both English and social studies teachers also agreed. Of the 24.4% who disagreed or strongly disagreed, most were elective teachers. Two teachers indicated that they were not sure.

Figure 11

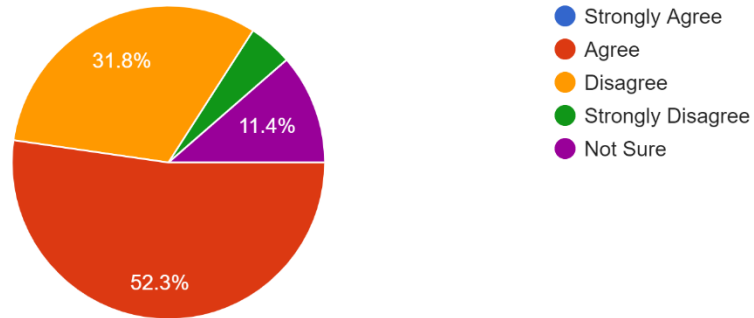
Question 9 Results: Having a final exam holds students accountable for their learning.



When asked if studying for a final exam increased student retention of their learning, a slight majority, 52.3%, of teachers agreed (Figure 12). No teachers strongly agreed and 36.3% of teachers disagreed or strongly disagreed. Five teachers, making up 11.4% of responses, were not sure. In both math and social studies, 75% of teachers believed studying for the exam increased retention, while in English 50% of teachers disagreed with 25% unsure. Most elective teachers also disagreed.

Figure 12

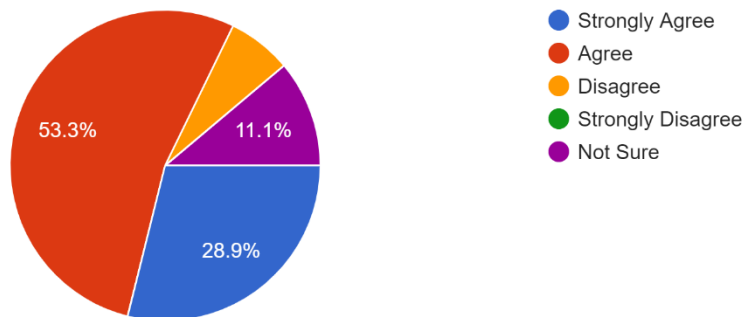
Question 10 Results: Studying for a final exam increases student retention of learning.



An overwhelming majority of teachers (82.2%) supported the position that students need experience studying for a final exam as preparation for college (Figure 13). Only three teachers disagreed, and five teachers were unsure.

Figure 13

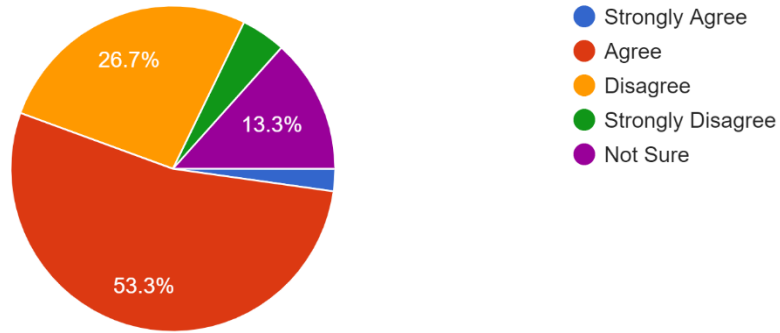
Question 11 Results: Students need experience studying for a final exam in order to be prepared for college.



Most teachers also agreed that final exams provide teachers with important data on student learning (Figure 14). By department, 71% of science, 67% of math, and 63% of social studies agreed. English was evenly split. The electives had 50% who disagreed but 20% who were unsure.

Figure 14

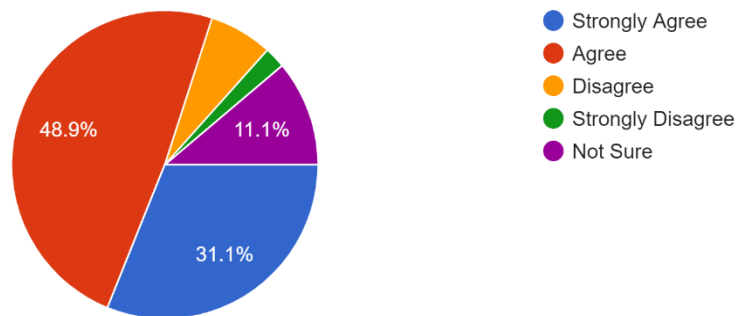
Question 12 Results: Final exams provide teachers with important data on student learning.



The next question elicited the largest strongly agree response of any of the Likert statements. Teachers overwhelmingly responded that they prefer multiple summative assessments throughout the course term rather than one final cumulative exam (Figure 15.) Only four teachers (two in English, one in social studies, and one in science) disagreed or strongly disagreed with 11.1% unsure.

Figure 15

Question 13 Results: I prefer multiple summative assessments throughout the course term instead of one final cumulative exam.

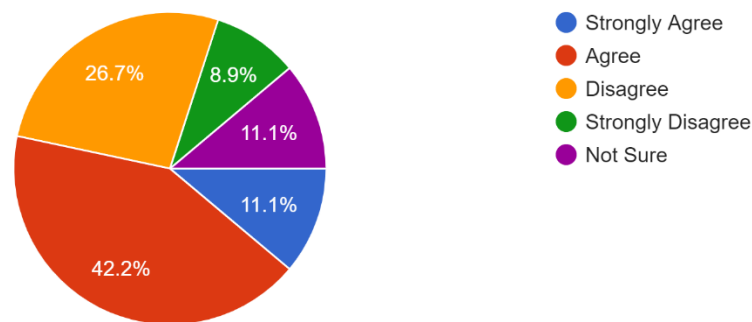


For the last general statement about final exams, most teachers responded that end

of course assessments can cause undue stress on students (Figure 16). Of the 53.3% of affirmative responses, 42.2% selected agree and 11.1% selected strongly agree. Even so, 26.7% of teachers disagreed and 8.9% strongly disagreed. More teachers from English and the electives responded affirmatively with half and both math and social studies evenly split. Five teachers were not sure.

Figure 16

Question 14 Results: An end of course assessment can cause undue stress on students.



Following the Likert statements, teachers were asked two open-ended questions. There was no limit put on their responses and it provided important qualitative data about teacher perceptions.

The first open-ended question was: What would be your preferred method of assessing student learning for your course? Most responses could be categorized into three general areas: using one or two cumulative exams, using project-based assessments, or giving no final exams.

The most popular sentiment expressed by teachers was to have multiple cumulative assessments with either a midterm and final or two semester finals. Several teachers stated that cumulative exams should only test one semester of content and others

added that the weight of the exam should impact a student's overall grade. The following comments are from the 18 teachers who preferred multiple cumulative assessments:

- I would prefer no final exam extend past content that was taught over a semester long period – students will never, in college, take an exam that covers a year.
- I prefer meaningful cumulative exam at the end of each semester. So, for a full year course, midterm and final. This prepares students for more rigorous preference-based assessment. Also, it makes a student more accountable for their learning. Finally, if students consistently perform poorly on a final it forces a teacher to better evaluate themselves.
- Midterm and final exam to help students prepare for cumulative testing at the collegiate level. Each exam would consist of multiple choice and free response questions. The average of the two exams worth 15-20% of the course grade.
- I think it would be helpful to have a midterm and a final that really counts toward their final grade of the course. This would help prepare them for college.

Of the 18 teachers who preferred two cumulative assessments, six were math, five were science, four were social studies, and one each from English, technology, and world language.

Eleven teachers wrote that they preferred a final cumulative exam. Some of these responses were not specific to one yearlong final or having multiple assessments and

some responses just said “testing.” Those teachers included six math teachers, one English teacher and four social studies teachers. Comments included:

- A final is a good review assessment to ensure that students can connect unit concepts that they learned into one cohesive understanding of course material.
- I prefer to have regular assessment built into my course throughout the year. This allows me to highlight specific content in order to make sure that students have learned important concepts before moving on to a new unit. A final is a good review assessment to ensure that students can connect unit concepts that they learned into one cohesive understanding of course material.
- The final has to mean something. The way grades are calculated right now with the final is "a joke" (student words not mine). "There's no point in studying." "I'm just turning in a blank form." This does not prepare kids for anything in the future.

There were 13 teachers who preferred cumulative project-based assessments over a traditional final exam. Not surprisingly, most of the elective teachers felt this with seven of the ten teachers represented in this category. Five English teachers also suggested project-based assessments along with one science teacher. Teachers said things like:

- For my class, a writing portfolio would be a better assessment of student knowledge, critical thinking, and reasoning.
- Projects, multiple assessments throughout the course, journaling.

- Ongoing assessments and/or project-based assessment.
- Cumulative project-based assessment.
- Unit Exams/Class Assessments/Practical Assessments.
- Application questions or problem scenario that requires fundamental content knowledge in order to solve.

Two teachers suggested not requiring final exams. One wrote, “Get rid of finals – or make them impactful to every student grade. Even if only in the fourth nine weeks.” Another teacher stated, “No required final or midterm. At most 2 midterms decided on a course-by-course basis.” These teachers were from math and technology.

The second open-ended question asked: What recommendations, if any, do you have for giving final exams? Teacher feedback was sorted into three categories: determination of final grades, frequency of cumulative testing, and instructional practices. Two teachers said they had no recommended changes and two others had responses that did not fall into any of the categories. One teacher suggested eliminating finals and the other thought they were good practice for students, but not a good measure of their learning.

Most teachers recommended changes to how final grades are determined. Twenty-five responses mentioned either changing the exam weight or switching to a percentage-based grading system. All content areas were represented with math having the most teachers at seven, followed by five in English, four in social studies, three in science, and three in electives.

Of the teachers who want changes to how grades are determined, 20 focused on changing the weight of the final exam. Comments included the following:

- Our students do not take our final exams as seriously as they should. They are much weightier and important in a college level course, so there might be some false sense of reality with our final exams.
- Students should not be able to get a zero on an exam with no effect on their final grade.
- They should be worth a higher percentage to make students accountable.
- Either increase the weight of the final exam or make it a part of the course grade. The way Peters Township computes grades, multitudes of student's grades cannot be impacted by any final exam. For B's? The student earns a B. The course grade cannot be impacted. The student has 3 A's and one B? Student needs to get a C or better (relatively easy if the student has that grade). On the other hand, a number of students may have two A's and two B's. In that case, the student would have to devote HOURS of time to study – to make certain the final grade is achieved. So, there is a mix of students with no impact, low impact, and HIGHEST impact. To devote days to this – and in non-Covid years – sometimes a week or more with senior finals, everyone finals, and then core finals for remaining students seems foolish. I feel discouraged from adding additional work during that time out of class, so it decreases the amount of class time.
- The current grading system does not make sense. If we are giving finals, they should have some real weight. If we say that finals are

important, students should not be able to skip it and not have that impact their grade.

- The exam needs to count regardless of quarter grades. It needs to be weighted more (or the average of the midterm and final needs to be weighted more) so that students with 4 A's must study for the exam. Currently, only students that have 2 of one grade and 2 of another grade are the ones really placing effort into study for the cumulative test.

There were five teachers who suggested using a percentage-based grading system. These teachers represented English, math, social studies, and world language. Comments included:

- I am not objected to the idea of the cumulative final exam, but I do object to HOW it is weighted and administered. If we operated on a straight percentage scale, then the final exam would actually mean something, students would take it more seriously, better prepare for it, and possibly retain it.
- Changing the final grade calculations to focus on percentages rather than a 'final exam' rule.
- If the grades were percentage driven and there was a midterm and a final, I think that we could get accurate reflections of mastery of content and skills. These tests must be coupled with performance based assessments.

Eight teachers recommended more frequent cumulative testing in the form of semester finals with some suggesting semester grading. Respondents included two teachers each from math, social studies, and electives, and one each from English and science. Recommendations included the following:

- Transitioning to a system that records semester grades and holds students accountable for semester final exams.
- I think it would be helpful to have a midterm and a final that really counts toward their final grade of the course. This would help prepare them for college.
- Any exam should never cover more material than is learned in a 5 month period. That lines up with the college timeline.

The remaining comments centered on instructional practices such as test design, student preparation, and teacher communication. Seven teachers, two each from English, science, and social studies, and one elective teacher, suggested changes to what teachers were testing and how they were messaging the exams. Comments included:

- In the subject I teach, I'm sometimes annoyed with my colleagues who ask specific content questions for a final exam. It's not remembering the content that matters – it's the skills the kids gain. Cramming non-essential content makes me feel bad for the students. Teachers could be better educated about the role/purpose of the final exam.
- Final exams should address the big ideas/core body of knowledge for this course, include multiple assessment item types (multiple

choice, problem solving, short answer, and a project/performance component), and have a variety of DOK level questions with the majority being DOK 3.

- It is important that the instructor sets the tone that the final exam is to be taken seriously and provides quality review-type activities and guides (not study guides that are identical to the final exam, but that give the students a measure of the quality of the material and rigor at which they need to apply what they have learned.
- Students must be taught how to study for these exams.

Some teacher's made specific recommendations like, "If possible, ask your students how they would like to receive the final exam for the class," and "Having alternative methods to assess certain students can be beneficial. For example, utilizing a one-on-one discussion-based assessment may be a better fit for some students than taking a traditional exam."

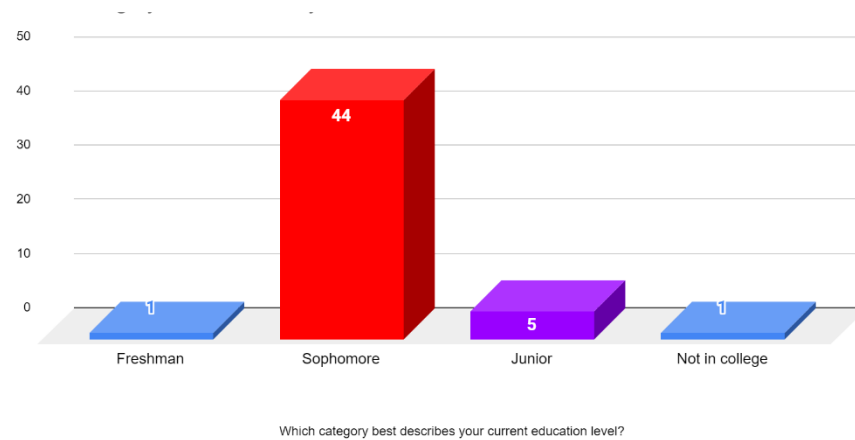
Alumni Survey Results

A survey was sent to 347 students who graduated from Peters Township in 2019 with 65 completing the survey and consenting to participate in this study. After review of the responses, 13 were removed as duplicates and one was removed as unusable, leaving 51 viable responses. The survey contained five Likert scale questions and two open-ended questions. Likert scale questions were formed as statements that allowed each participant to indicate their level of agreement or disagreement. There choices included: Strongly Agree, Agree, Disagree, Strongly Disagree, and Not Sure. The first three statements focused specifically on student experiences in high school and the next two

statements on student experiences in college. The open-ended questions allowed participants to provide additional details or suggestions. Respondents were also asked to identify their area of study in college and their current level in college. Most respondents identified as second year college students with 44 sophomores participating. Of the remaining respondents, five said they were juniors, one a freshman, and one was not in college.

Figure 17

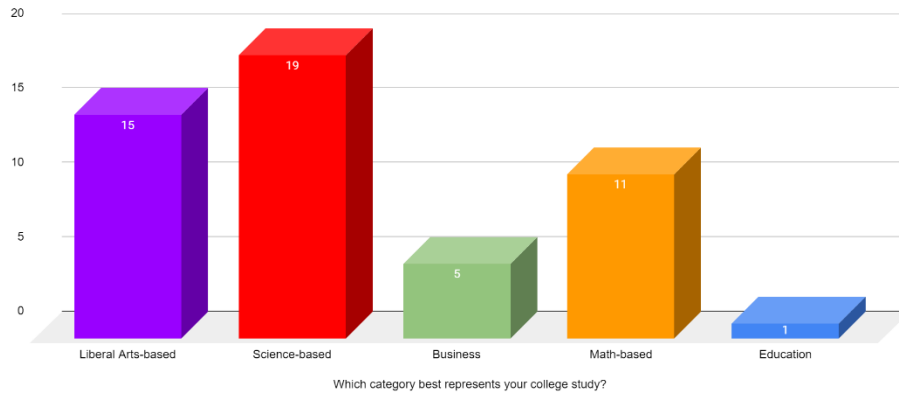
Education Level of Alumni Respondents



Science studies were the most frequently identified course of study with 19 responses, followed by liberal arts with 15 and math with 11. Five alumni identified business and one identified education.

Figure 18

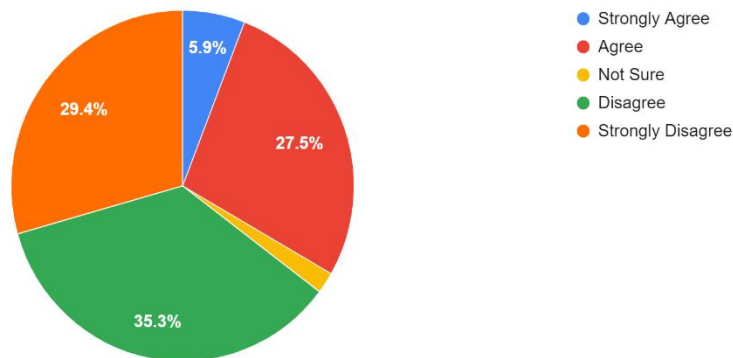
Category of College Study of Alumni



On the first Likert statement, nearly 65% of alumni disagreed with the statement: Taking final exams in high school prepared me for college-level testing (Figure 19). Nearly half of those that disagreed with the statement had selected strongly disagree. Approximately 33% of respondents felt that high school final exams helped prepare them for college. Science majors made up most of the negative responses with 14 strongly disagreeing or disagreeing, followed by 11 liberal-arts majors, five math majors, and three business majors.

Figure 19

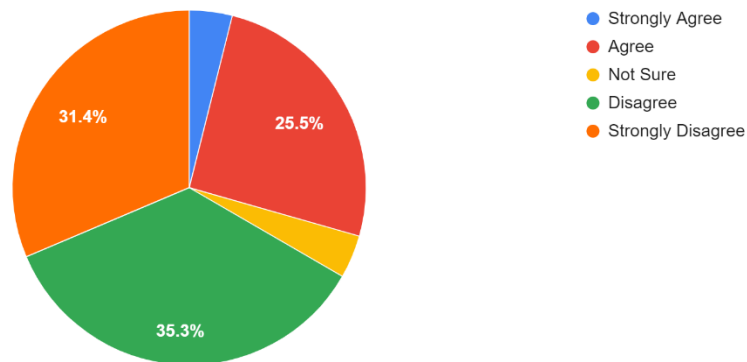
Question 1 Results: Taking final exams in high school prepared me for college-level testing.



On the second question, alumni were asked if the final exams they took in high school were similar to the exams they take in college (Figure 20). Opinions were stronger on this response with nearly 67% indicating that final exams were not similar, 31.4% of which strongly disagreed. Only 29.4% of alumni agreed or strongly agreed. Responses were evenly dispersed across course of study, though science had the most with seven.

Figure 20

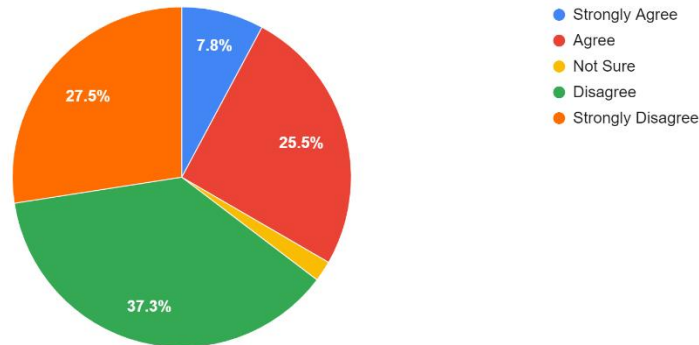
Question 2 Results: The final exams that I took in high school were similar to the exams I take in college.



For the last question related to high school experiences, participants were asked if their final exams accurately measured their learning (Figure 21). Following the same trend as the previous two questions, nearly 65% disagreed that final exams were an accurate learning measure. Only 33.3% agreed with a small number who selected strongly agree. One participant was unsure.

Figure 21

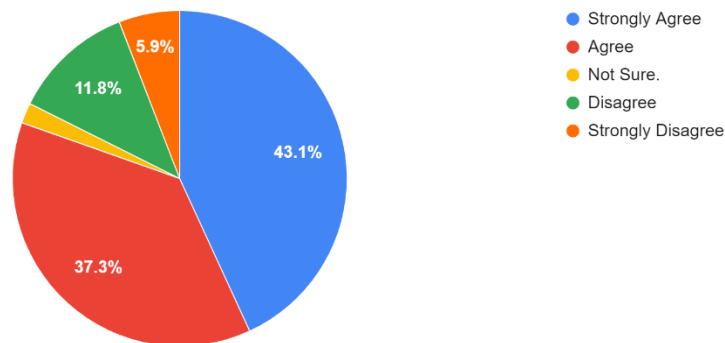
Question 3 Results: My final exam grades in high school were an accurate measure of my learning in the class.



The next two questions were meant to gather information regarding current college practices. An overwhelming majority of alumni responded in agreement with the first statement: My college classes rely on cumulative exams (midterms and finals) to assess my learning (Figure 22). Of the 80.4% who agreed, 43.1% strongly agreed. Less than 18% disagreed.

Figure 22

Question 4 Results: My college classes rely on cumulative exams (midterms and finals) to assess my learning.

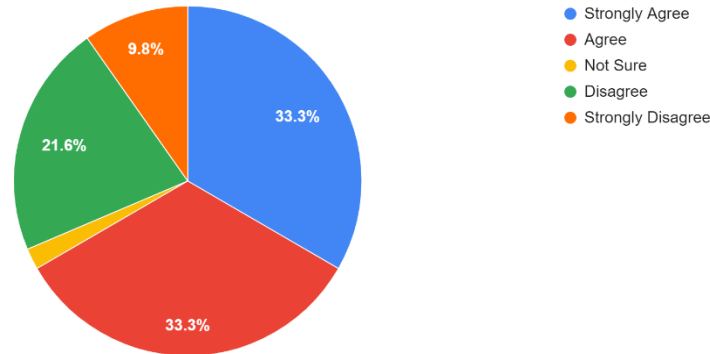


The final Likert statement asked alumni if their college classes relied on project-based assessments to which 66.6% agreed (Figure 21). It was equally split between agree

and strongly agree. Less than 32% of participants responded that they were experiencing project-based assessments.

Figure 23

Question 5 Results: My college classes rely on project-based assessments to assess my learning.



The two open-ended questions provided qualitative data for this research study. These questions allowed alumni to describe their experiences with assessments in their college courses and to give feedback on what they wished they had learned in high school. Responses were expected to vary based on the collegiate course of study, the level of course taken in high school, and the college or university attended.

The first open-ended question asked: What surprised you most about how you are assessed in your college classes? Responses were categorized into the following areas: exam weight, assessment rigor, project-based/writing assessments, college expectations, and college preparation.

A frequently referenced difference between high school and college was the weight of exams. One participant said,

I was surprised at the weight of the tests in respect to my final grade in the class.

In most classes, I will have a midterm worth 20% of my grade and a final worth 25%. This is very different to what I experienced in high school.

Another wrote, "Exams are worth a much larger percentage of my total grade than I anticipated." In total, 14 participants focused on the increased weight of college exams in determining their final course grade. There were five alumni in each content area of liberal-arts and math, three in science, and one in business.

One respondent took it a step further and referenced the difference in assessment rigor,

For my STEM courses, there are usually only a total of 3 exams for the whole semester and they matter a LOT. This was something that blindsided me. The types of questions on these exams were also foreign to me. They obviously weren't information-recall and were application-based questions, but the application-based questions I had in high school were pale in comparison.

Others made similar comments about assessment rigor such as, "There are still points and grades in college, but the grades are more determined by how well the professor sees you actively understand and apply the material." Another said, "The questions are worded and asked differently." One former student compared the rigor to high school tests,

The tests are very similar to tests taken in high school. Material is learned more fast paced and the questions are deeper and thought provoking. I was surprised that a lot of professors give out partial credit and actually do care about how their students do.

Overall, there were six references to rigor with respondents equally representing liberal arts, math, and science.

Multiple respondents said they have experienced an increased focus on writing and project-based assessments. For example, one alumnus said,

Almost all of my college courses pretty exclusively do papers for midterm/finals.

This could be because I'm an English major, and even when I'm not taking English, I'm taking history or language courses for which papers work well. But I was very surprised by how much writing REALLY mattered. Every midterm for every class was a long paper; every final was an even longer one.

Four respondents, mostly in the science track, talked about the project-based focus of their courses. "Many of my classes have projects and labs in place of formal exams in order to apply knowledge." Another said, "I am surprised at how few homework assignments I am given. In addition, most of my finals are now project-based."

Additional comments were made about expectations at the college level that were unanticipated. Students wrote that they were surprised by "How little time you get in between assignments and exams to study" and "the amount of work you have to actually put in to do well." There were four comments about unanticipated expectations.

Two alumni felt well prepared for college. One in liberal-arts writing, "I'm glad that I was prepared to face the stress and anxiety of taking an exam without any sort of aid or help but rather just my knowledge." A student studying science said, "They (high school classes) were sometimes more difficult than things I've seen in college, which really helped me out when I got there."

Two alumni stated that their college experiences were better than expected. “We are given much more tools to study from than high school. We were told we wouldn’t be given study guides in college, and I find that most of my professors provide them,” one alumnus studying business wrote. A student studying education said, “Many professors still grade attendance, participation, quizzes and homework. It’s not just a few exams to determine your grade.”

The second open-ended question asked: What do you wish you had learned in high school to be better prepared for college? Responses to this question demonstrated more consistent common themes focused on writing, study skills, and instructional and/or assessment practices.

Six responses were specific to writing. Most of these were from liberal-arts majors with two from math. Comments included:

- I wish I had learned better writing practices. In general, the type of essay I was taught in high school is seldom used (5 paragraph style). I only had one teacher throughout my four years of high school that never required 5 paragraphs. In general, I think the exams I took in high school (except AP exams I took) were not at all a reflection of exams in higher ed.
- I wish I had learned better writing and researching skills.
- How to write actual 3–5 page papers and that college is nothing like high school teachers say it’s like.
- Writing essays, taking tests similar to college level, and being challenged more.

- I think a stronger emphasis on research writing...I have learned a lot about research writing/academic writing simply by doing it so often, but I felt kind of foolish in my first few weeks at school – I think our education in high school is very rigidly “5 paragraph, intro/conclusion” - esque which does not at all reflect how most college professors approach writing. We should learn more about writing theses (that aren’t three-pronged!)

Regarding test-taking or study skills, 18 respondents expressed that they wished they had been exposed to more rigorous testing and taught studying techniques and strategies. Additionally, some recommended improved instructional practices. Science majors dominated this mindset with nine responses, three from math, five from liberal arts and one from business. Comments included:

- I wish we have more cumulative exams because finals didn’t matter in high school.
- How to study effectively for cumulative exams.
- I wish our finals carried a little more importance in high school regardless of the class or the grades we had, this would have helped me learn to study and prepare for college finals.
- In college classes, final exams are required in order to get a “good” grade. In PTHS, final exams were sometimes unnecessary because the result would not change the final grade.
- I wish I learned better test-taking skills.

- I wish I had learned how to effectively study concepts and take general notes instead of memorizing things and copying most of the textbook into my notes.
- How to study vs what to study.
- Figure out what study methods work best for students and teach them how to study.
- I wish I had learned actual learning skills to help set me up for success in college.

Other alumni focused on general life skills like time and money management and how to live independently. Of the seven responses in this category most were science with two from business and one from math.

Two alumni would like to see more rigor in assignments and more projects. One wrote, “High school has an insane amount of busy work. College assignments take more effort, but you have more time to complete them and less assignments in a semester.” While another said, “Semester projects instead of just multiple choice questioned (Scantron) exams.”

Finally, two respondents indicated they were well prepared like the alumnus who said, “Nothing. I still recall things I’ve learned in 8th grade in my college writing classes.”

Grading Data Analysis

The purpose of the grading analysis was to determine the relationship between the final exam grade and its correlation to quarter grades. Stored grades for nine required English and math courses from the 2017-2018 and 2018-2019 school years were used for

this study. A review of this data using Excel and RStudio, a statistical programming software, sought to answer the following questions:

1. What percentage of final exams had an impact on final grades?
2. How often did a final exam increase or decrease a final grade?
3. Are there any differences in patterns due to course content or academic level?
4. How often did a final exam grade align with the average of the quarter grades?

Each course was analyzed independently than combined to determine overall impact. Within each course, a line was created for each student record. This included the four quarter grades, the final exam grade, and the end of year final grade. There were 1,326 student course grades analyzed.

The calculation used at Peters Township for final grades at the time of this study used quality points of each quarter grade and the final exam grade. Each quarter was a standalone grade that was doubled in the final calculation. The final exam quality point was not doubled. Table 1 shows the calculation as listed in the student handbook.

Table 1

PTHS Final Grade Calculation

FYI: How to calculate Final Grades For Transcripts

						How is FV calculated for you if your grades are in on time
						↓
Term of course	Qtr 1 weight	Qtr 2 weight	Qtr 3 weight	Qtr 4 weight	Finalex weight	FV in PS (Final Average for transcripts)
Semester 1 S1	2 X letter grade equivalent of the percent	2 X letter grade equivalent of the percent			1 X letter grade equivalent of the percent	$\frac{[(2 \times Q1) + (2 \times Q2) + FX]}{\text{divide by } 5}$ =FV printed as a letter grade
Year	2 X letter grade equivalent of the percent	2 X letter grade equivalent of the percent	2 X letter grade equivalent of the percent	2 X letter grade equivalent of the percent	1 X letter grade equivalent of the percent	$\frac{[(2 \times Q1) + (2 \times Q2) + (2 \times Q3) + (2 \times Q4) + FX]}{\text{divide by } 9}$ =FV printed as a letter grade
Semester 2 S2			2 X letter grade equivalent of the percent	2 X letter grade equivalent of the percent	1 X letter grade equivalent of the percent	$\frac{[(2 \times Q3) + (2 \times Q4) + FX]}{\text{divide by } 5}$ =FV printed as a letter grade

A=4, B=3, C=2, D=1 for calculating FV. Appropriate weights are applied to Honors and AP Courses after calculating FV (Final Average.)

To determine how many students were impacted by the final exam grade, each quality point average possible was examined to determine how many quality points would be needed on the final exam to move a grade up to the next letter grade or keep the grade the same. The first step was to look at what quality points were needed to achieve each letter grade. Table 2 shows the progression from percent to letter grade to quality point. The “Final Grade Value” column shows the number of quality points needed at the end of the school year to achieve the corresponding letter grade. Students who have received four “A” s would have 32 quality points prior to taking the final exam. Regardless of their grade on the exam, they would still receive an “A” in the class.

Table 2*Quality Point Conversion Table*

Percent	Letter Grade	Quality Point Value (QPV)	Sum of Quarter QPVs	Final Grade Value
90 - 100	A	4	14 - 16	32 - 36
80 - 89	B	3	10 - 13	23 - 31
70 - 79	C	2	6 - 9	14 - 22
60 - 69	D	1	2 - 5	5 - 13
59 & below	F	0	0 - 1	0 - 4

A review of each grade band revealed that only students within certain quality point values could be impacted by the final exam grade as shown in Table 3. There are six scenarios that require a student to achieve a minimum grade on the final exam to maintain their grade. A student with a “D” average and five quality points would need an “A” on the final exam to improve his/her grade to a “C.” That is the only scenario that allows for improvement. For students with an “A” average but only 14 Quality Points, an “A” on the final exam would be needed to maintain the grade.

Table 3*Quality Points Impact on the Final Grade*

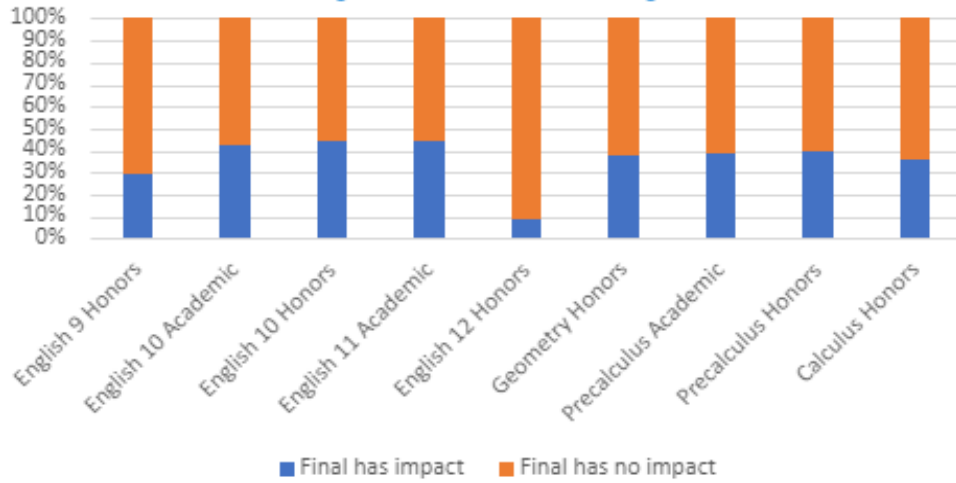
Letter Grade	QPV4	QPV4x2	Needed to Maintain	Needed to Improve	Decrease
A	16	32	0	--	--
A	15	30	2	--	1, 0
A	14	28	4	--	3,2,1,0
B	13	26	0	--	--
B	12	24	0	--	--
B	11	22	1	--	0
B	10	20	3	--	2,1,0
C	9	18	0	--	--
C	8	16	0	--	--
C	7	14	0	--	--
C	6	12	2	--	1,0
D	5	10	0	4	--
D	4	8	0	--	--
D	3	6	0	--	--
D	2	4	1	--	0
F	1	2	0	--	--
F	0	0	0	--	--

Of those 1,326 course grades used in this study, 817 could not mathematically be impacted by the final exam (Figure 24). All 817 students could have received a zero on the final exam and still maintained the average of their four quarter letter grades for their final grade. The remaining 509 students could have been impacted by the final exam grade. These students had one of the following quality point totals prior to the final exam: 15, 14, 11, 10, 6, 5 or 2. There were 19 students who could have increased a letter grade if they had gotten an “A” on the final exam, but the final exam could not hurt their grade. In all 19 cases students' grades remained the same with no increases. The remaining 490 students had to achieve a certain score on the final exam to maintain the grade average of

the quarters as noted in Table 3. All but 210 of those students achieved the grade needed on the final exam to maintain their grade.

Figure 24

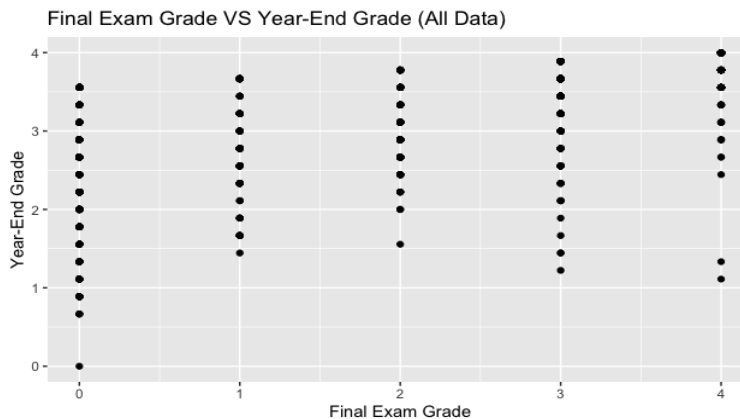
Final Exam Impact Per Course



A basic scatterplot of final exam grades and year end grades for all data shows a slight positive linear relationship between final exam grades and final year end grades (Figure 25). The scatterplot shows a large cluster of year-end grades are 2s and above (grades of C and higher), while the final exam grades are more equally distributed across all quality points, with a significant number of grades in the 1 and 0 range.

Figure 25

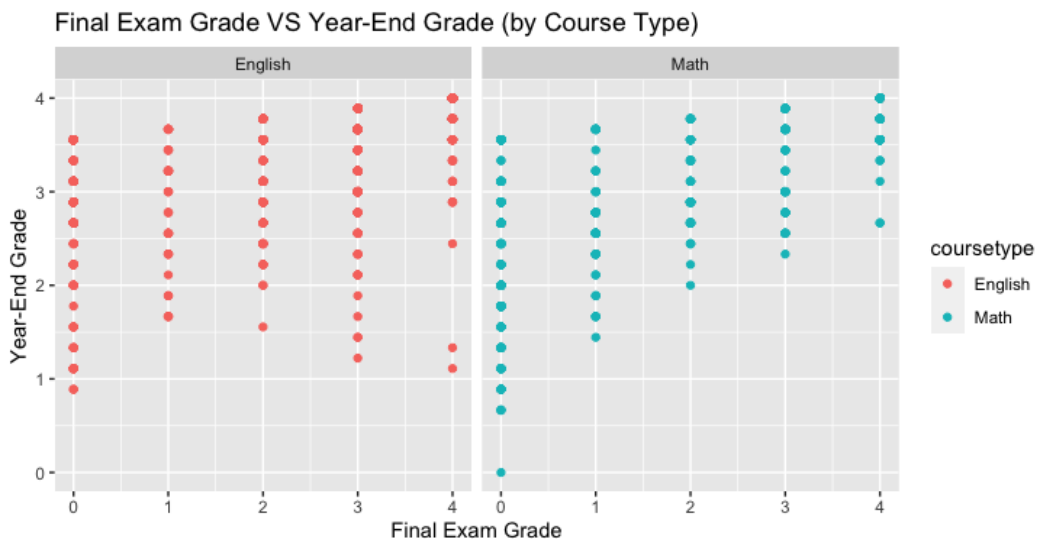
Final Exam Grade vs. Year-End Grade



The correlation coefficient of final exam grades and year end grades is .615. The closer a correlation gets to 1, the closer two items are connected. A .615 is a moderate indicator of a positive relationship. This represents the relationship for the nine English and math courses combined. There were small differences between English and math courses (Figure 26). Math tended to be more linear, though both subjects had a significant number of students receiving a D or F on the final exam, even when their year-end grade was a 3 or 4.

Figure 26

Total English and Math Scatterplot

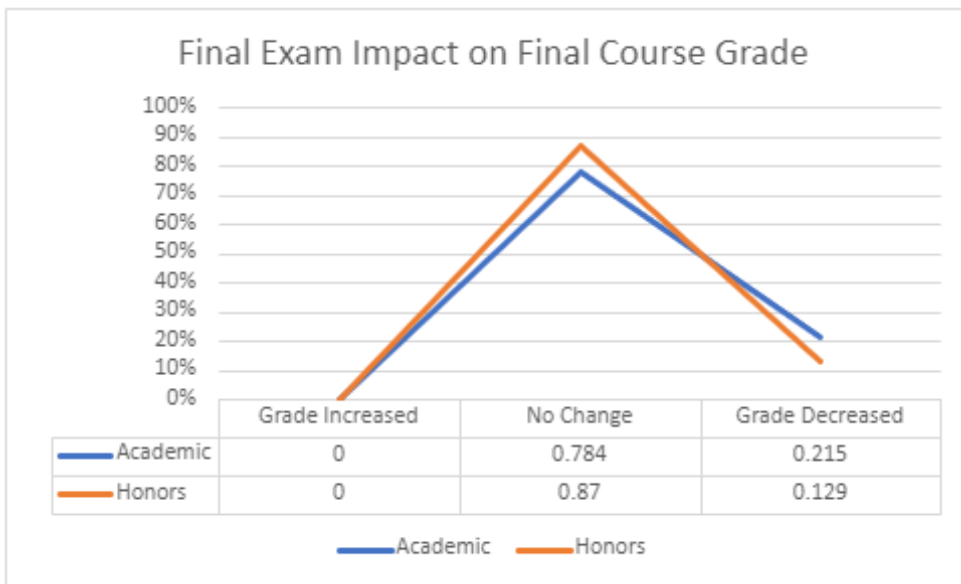


Of the 210 grades that decreased as a result of the final exam, 122 were from English and 88 from math. One more English class was used in this study which explains the higher number. Percentages were nearly identical for both English and math with approximately 16% of English grades decreased and 15% of math grades decreased after the final exam. No student grades increased after the exam but 84% of English students and 85% of math students maintained the average of their quarter grades.

A bigger discrepancy was noted between academic and honors level courses (Figure 27). Since there were more grades included for honors level courses in this study, percentages were used to illustrate a comparison more accurately. Honors students were more likely to maintain their grade and less likely to decrease their grade on the final exam. Only 12.9% had a decrease in course grade whereas for academic level students, 21.5% had a decrease in their course grade.

Figure 27

Final Exam Impact on Final Course Grade



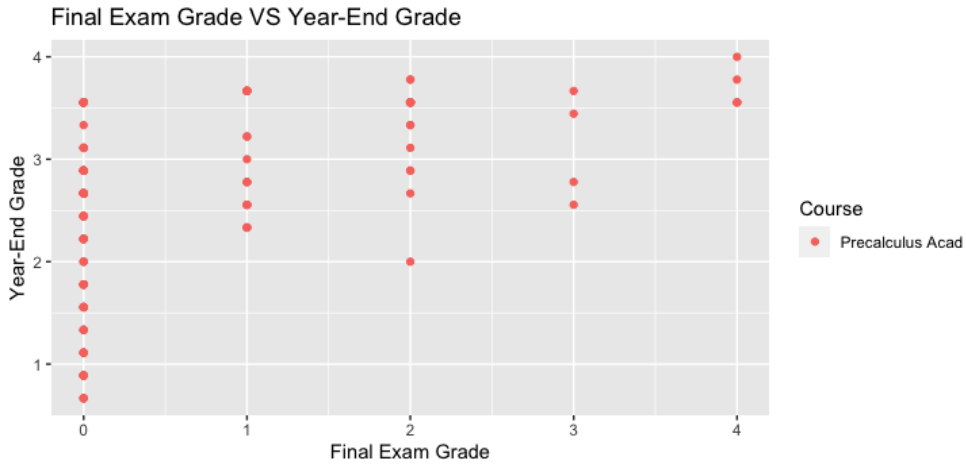
Two sets of courses were used in this study that had equivalent academic and honors levels, Precalculus and English 10. A review of just these four courses showed the discrepancy, though it was stronger in math than English.

For Precalculus Academic, the correlation was .457 with 79% of grades staying the same and 21% of grades decreasing after the final exam (Figure 28). There was a limited relationship between final grades and final exam grades. Grades on the final exam

were clustered in the 0 and 1 range indicating that most students failed or received a low grade on the final exam.

Figure 28

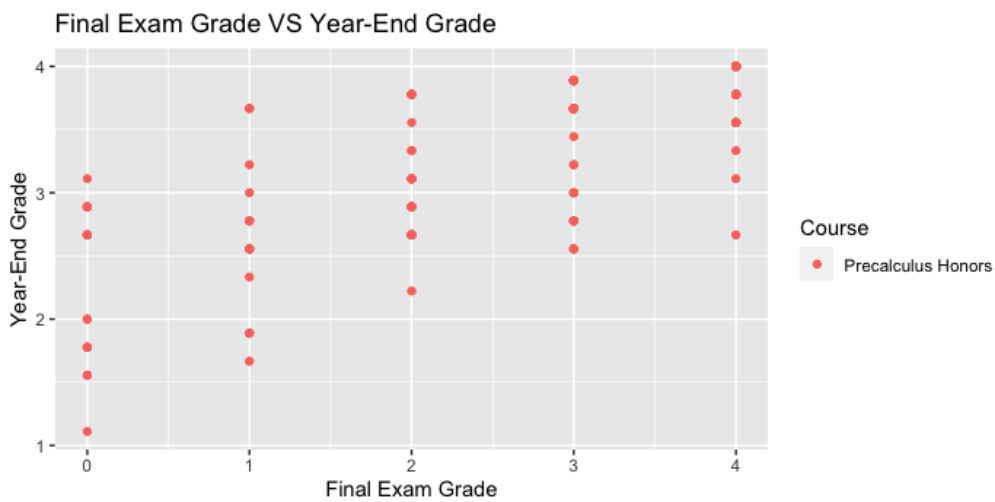
Precalculus Academic Scatterplot



For Precalculus Honors, there was a stronger correlation of .737 and only 4% of students had a decrease in their grade. Most students, 96%, not only maintained their grade, but scored more consistently in the 3 and 4 range as shown on the scatterplot below (Figure 29).

Figure 29

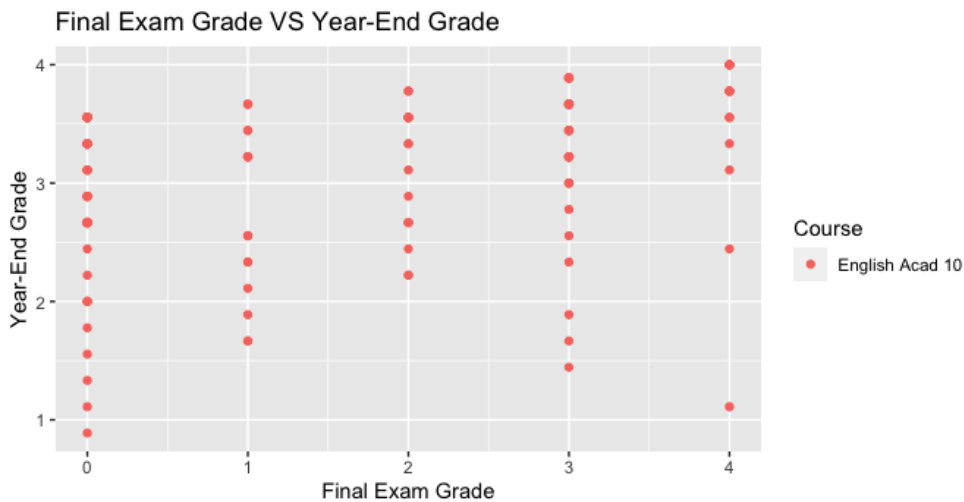
Precalculus Honors Scatterplot



The English courses had the same trend, but it was less pronounced. The correlation for English 10 Academic was .288 while English 10 Honors was .576. Only 78% of English 10 Academic students maintained their grade and 21.7% decreased their grade (Figure 30).

Figure 30

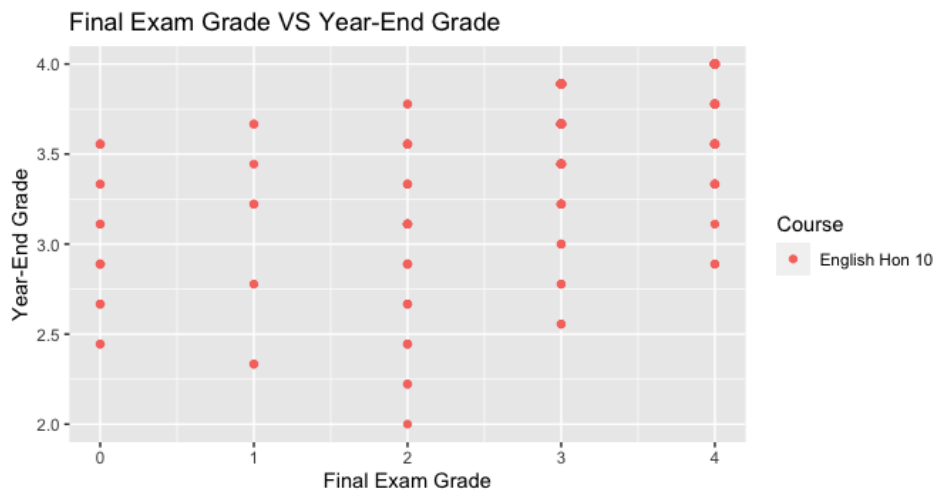
English Academic 10 Scatterplot



For English 10 Honors, 16% of student grades decreased and 84% maintained their grade. More final exam grades fell in the grade bands of 2, 3 and 4 (Figure 31).

Figure 31

English Honors 10 Scatterplot



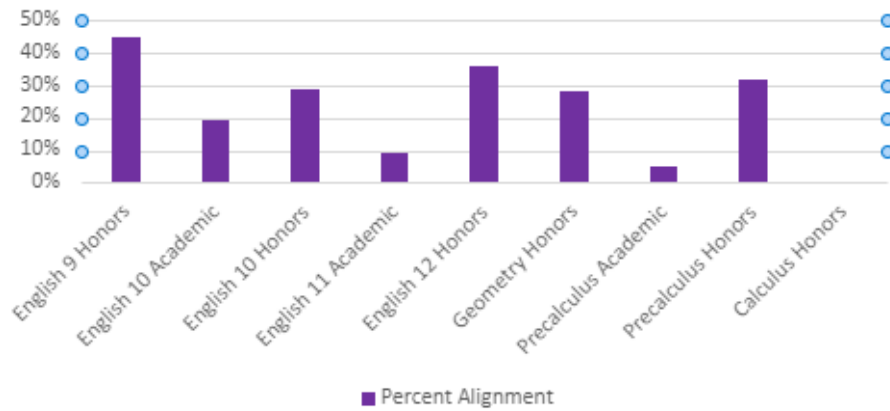
Both sets of courses provided evidence that the academic level was a factor in the amount of impact a final exam grade was on the final course grade.

In review of the first three grading questions, all were answered. Final grades were impacted less than 16% of the time by the final exam. In each case, the final grade decreased. Only 1.4% had the opportunity to improve their grade, but this did not occur for any of the students. Though there was no difference between content areas, there was a difference between academic level.

The final question asked how well the final exam grade aligned with the average of the quarterly grades. To many students and teachers, the average of the four quarters is representative of each student's achievement. As such, it would be expected that the final exam grade be similar to the quarter grades. Each course was reviewed to see how aligned student quarter grades were to their final exam grade (Figure 35). The course with the most alignment was English 9 Honors with 45.2%. English 12 Honors had 35.6% alignment and Precalculus Honors had 32%. On the low end, Calculus I Honors had zero alignment followed by Precalculus Academic which had 5%. It is important to note that students who take Precalculus Academic typically take Calculus I Honors as the next course in the math sequence which may explain the similar pattern in each course's data.

Figure 32

Percentage of Final Exam Grades Aligned with Quarter Grades



With all the courses combined, only 23.6% of final exam scores aligned with the how students performed throughout the course as reflected by quarterly grades.

Discussion

This study utilized a mixed method approach to data collection using surveys and an analysis of historical grades. Both quantitative and qualitative data was gathered through Google Form surveys to teachers and alumni. The teacher survey was used to identify perceptions about the exam process at Peters Township High School as well as determine the value teachers place on having final exams. The alumni survey was designed to identify current assessment trends in higher education through their experiences and to elicit their perceptions of the final exam process at Peters Township High School. Included in the teacher and alumni surveys were two parallel questions that provided a comparison between teacher and alumni perceptions.

The use of the grading data from historical grades provided additional quantitative data that was used to find the statistical impact of the final exam grade on student course

grades and to determine if there were notable trends in student performance on the final exams. This information was used to compare with both teacher and alumni perceptions. This section outlines the results of the triangulation of the qualitative and quantitative analysis from the three data sources to answer the four research questions.

The first two research questions were answered from the teacher survey responses. The research questions are:

1. What are teachers' perceptions of the final exam?
2. What are teachers' perceptions of the reliability of the final exam as a measure of student learning?

Results revealed that teachers were generally divided on the current process used and the reliability of the final exam as a measure of student learning, but they were more consistent in their perceptions of the practice of administering final exams. There were only two areas that the teachers showed a moderate agreement, but these were contradictory. When asked if they use final exam data to determine if students had mastered the content, 62.2% said they did. Nearly the same percentage of teachers also responded that the final exam grade was not consistent with how students perform on assessments throughout the course. Assuming assessments throughout the course were also measuring content mastery, it would be expected that students would perform similarly on a comprehensive exam. This discrepancy may be evident of what teachers believed they should be doing versus their actual practice.

Teacher responses also revealed that only half of them used final exam data to evaluate their curriculum and less than half believed it was an accurate measure of student learning. One of the parallel questions on the alumni survey was, "My final

exams grades in high school were an accurate measure of my learning in a class.” Results showed that 64.8% of alumni did not believe the final exams were an accurate measure of their learning. Nearly 46% of teachers agreed with the alumni, but 44.4% of teachers thought it was an accurate measure. Even though alumni responded more strongly than teachers, combined responses indicated that both groups had doubt about the effectiveness of the final exam as accurate measure of student learning.

Teachers were also divided on whether the final exam prepared students for college-level assessments. This was the second parallel statement on the alumni survey, “Taking final exams in high school prepared me for college-level testing.” Responses from teachers and alumni revealed a greater disconnect between their perceptions. There were 22.5% more alumni than teachers that did not believe final exams prepared them for college assessments. A notable finding was that 15.6% of teachers were unsure. This provides evidence that teachers may not be confident in their understanding of college assessment trends.

Lastly, teachers remained split on whether the final exam calculation allowed students to improve their grades and was fair and consistent to all students. With less than a 50% response rate for both agree and disagree selections, a number of teachers marked not sure. Results indicated that teachers may not have a strong understanding of the mathematical impact of final exam grades.

When questioned about the use of final exams as a practice, unrelated to the process at Peters Township, teachers' perceptions were more consistent. Teachers overwhelmingly believed that administering final exams holds students accountable for their learning and prepares them for college. Most also believed that studying for a final

exam increases retention and provides important data on student learning. The number of not sure responses indicated that more professional development on this topic would be needed on assessment practices.

Qualitative data showed that most teachers prefer using multiple assessments throughout a course term with a cumulative assessment that tests no more than a semester worth of content. Also supported by the qualitative data from the alumni surveys was a strong belief that the final exam grade did not have an impact on student overall grades. Additional results were that teachers favored increasing the weight of the exam to make it more meaningful and they wanted to provide more skilled-based or writing assessments. Most of the anecdotal responses from the alumni survey supported an increased focus on final exams, skill-based assessments, and research-type writing assessments.

The third research question provided an opportunity for quantitative analysis of grading data to see if it aligned with general beliefs from the teacher and alumni surveys. The question is:

3. What is the impact of final exam grades on students' grades?

The data analysis showed that out of the 1,326 course grades analyzed, only 38.4% could have been impacted by the final exam. Only 1.4% of students had a shot at improving their grade to the next letter grade, but these were struggling students with "D" grades who needed an "A" on the final. Of the remaining 490 students, 57% were able to do well enough on the final exam to maintain their grade. This achievement level needed looked different for students even if they had the same average grade going into the final exam. For example, a student with an "A" average could have quality points of 14 or 15.

To maintain the “A” grade, a student with 15 needed to get a “C” on the final while a student with a 14 had to get an “A” on the final.

The 210 students whose grades dropped as a result of the final exam made up less than 16% of all students. In 44 of those cases, teachers overrode the grading calculation to give the student the higher grade. Most often this was for students with the “A” average who needed an “A” on the final but earned a “B.” This occurred almost exclusively in English courses, with only one in math.

Based on the nine courses analyzed, the impact of the final exam on student grades was minimal, impacting only 38% of all students. While 210 grades decreased after the final exam, 21% of those were changed by the teacher. This decreased the overall impact of the final exam to less than 13% of all grades being impacted by the final exam. This data supported the beliefs of both teacher and alumni that the final exam has a minimal impact on student grades.

Teacher survey results showed that 40% of teachers believed that the final exam calculation allowed students a chance to improve their grade, however the grading data illustrated the opposite with only 19% able to improve. Nearly 50% of teachers believed that the final exam calculation was fair and consistent, but this was not supported by the data that showed only 38% of students had to pass the final exam. In addition, the data showed that students with the same average grade at the end of the year experienced different expectations in performance on the final exam to maintain their grade. Both teachers and students commented in the survey responses that final exam grades do not matter for most students.

Data from the alumni survey was used to answer the final research question:

4. How does the current final exam structure compare to current assessment trends at the collegiate level?

According to alumni, final exams at Peters Township High School did not compare to those taken in their college courses in both weight and rigor. Cumulative exams in the form of midterms and finals were reported to be used in colleges by 80% of participants. College students also reported being assessed often through projects and research papers. While several respondents felt well-prepared for their college-level courses, most said they would have benefitted from more rigorous cumulative assessments in high school. Many commented that high school finals were not valued and did not provide them with the experience of studying. They also noted differences in the rigor of exam questions, with college questions requiring more application of content and less memorization.

This data can be used to help teachers better understand the demands on students as they proceed into higher education. Because student high school experiences vary based on the course level, content area, and teacher skill, some students most likely felt more prepared than others. However, a strong emphasis on improving student study skills and writing would benefit all students.

Summary

A significant amount of data was collected and analyzed for this study to develop a complete understanding of the final exam process from both teacher and student perspectives. The results of the surveys revealed teacher inconsistencies and misconceptions which demonstrates a need for additional professional development. While teachers favored having a cumulative testing plan to increase student

accountability and better prepare them for college, they did not feel strongly about maintaining the current system. They were also unsure of current collegiate practices and expectations.

Former students demonstrated strong feelings about the ineffectiveness of the current structure and provided reasonable suggestions for areas of improvement. Some of these suggestions could be implemented even without changing the grading system, like adding more research papers. Both teachers and alumni agreed that the weight of the final exam was a problem that should be addressed.

The data analysis from historical grades affirmed the beliefs of both teachers and students that the final exam is not valued and has little impact on student grades. As a result, any data provided to teachers from the final exams would not be useful for evaluating curriculum or instruction. Furthermore, it was clear from the comparison of teacher survey responses and grade analysis that many teachers did not fully understand the final exam calculation and its ultimate impact on student grades.

Lastly, both the quantitative and qualitative survey data supported increasing student exposure to rigorous cumulative assessments, but this should also include teaching students how to prepare for the assessment. Because student collegiate experiences varied based on area of study, having the same type of assessment in all courses would not fully prepare all students. Incorporating more application-based cumulative projects and research papers, however, could fill the gap that many alumni reported.

The information from this study provided necessary guidance that will be used by teachers and administrators to identify areas of focus so that next steps can be developed.

These will include improvements in teacher practices that can have an immediate impact and broader decisions about the grading practices and use of final exams across all courses at Peters Township High School.

CHAPTER V

Conclusions and Recommendations

Since the creation of the first high schools in America, the mission of secondary education has been to prepare students for societal success. While the expectations of society have shifted over the years, the pressure on high schools to equip their graduates with the knowledge and skills to transition into adulthood has remained constant. What that means in 2021 is being college and career ready. The focus in high schools throughout America has been on preparing students for college, though how that looks can vary greatly from state to state and school to school. One consistency, however, has been the use of standardized testing to measure this. From the state-mandated Keystone exams to the SAT, there remains a heavy reliance on test scores to determine not only individual student futures but also to drive decisions about curriculum and instruction for schools. Among the efforts by high schools to increase rigor to prepare students for the demands of both standardized testing and college-level assessments has been the adoption of cumulative final exams as end of year assessments. While this practice is widely in use in high schools, there is surprisingly little research on it.

Whether in high school or college, a cumulative final exam can impact student grades, sometimes significantly. Because grades are a measure of student learning, final exams need to be reliable and part of a clear cumulative testing plan (McDermott et al., 2014). Final exams have been used at Peters Township High School in nearly all courses following an established process. This process was not a policy established by the school board, but rather a practice that has been in place for many years. Guidelines are given to teachers that includes following the testing calendar and using common assessments for

the same course. A testing schedule is used at the end of the first semester for semester course finals and at the end of the school year for full year finals. This calendar is meant to prevent students from having too many tests on the same day. Some guidance is given to teachers on best practices for question design, length, and test reliability. The final exam grade is factored into each student's end of year grade as either 1/9 for a full year course or 1/5 for a semester course. As described in Chapter 4, the final course grade is determined using a unique grading system that averages quality points rather than percentages. This grading system and the use of the final exam has caused confusion and frequent discussion by teachers, students, and parents. Questions about the efficacy and fairness of the final exam even led to a school board committee that made recommendations for changes. Though no changes were adopted, the debate has continued.

This study intended to take a more in-depth look at the use of final exams as a practice and enquire into the process used at Peters Township High School. As a first step it made sense to quantify the perceptions of teachers about the use of final exams at Peters and then compare those to their overall beliefs about cumulative assessments. Teachers create the assessments, prepare students to take them, and grade them. It was important to see how invested they were in the process. Next, seeking input from alumni provided genuine feedback on their experiences taking final exams in high school and how those finals compared to their assessment experiences in college. This presented evidence of current college practices. The third and one of the most important steps in this research process was to review actual grading data. This illustrated the reality of how students have performed on the final exams and how that performance has impacted their

grades. Lastly, all the data pieces were analyzed for alignment with previous research on cumulative assessments in the literature review.

The foundation of this study was the four research questions that formed the lens through which all data was filtered and analyzed. The results of this study highlight areas that need to be addressed to have an effective cumulative assessment plan. Decisions will need to be made about whether to continue with year-end final exams at Peters Township or not. The research indicated a continued desire to use final exams, but with changes. Next steps have been created based on the outcomes of this study to give guidance for reviewing the current process and developing a clear and consistent cumulative assessment plan. This chapter will present the conclusions of this study including application of the results and fiscal implications, limitations of the study, and recommendations for future research.

Conclusions

This study provided a considerable amount of information to begin the process of reviewing and revising the final exam practices at Peters Township High School. The review of literature provided context with which to compare survey results and grading data and the four research questions created a framework for the data collection. The first research question addressed the perceptions of teachers towards the purpose of the final exam. The second question concentrated on teacher perceptions of the reliability of the final exam as a measure of student learning. The third question focused on actual student data to determine what the impact of the final exam grade was on students' overall grades. The fourth and final question solicited feedback from alumni on their experiences at the collegiate level. Within these questions, teachers and alumni were encouraged to

share their opinions and offer suggestions. As a result of this study, it is recommended that Peters Township High School create an assessment committee to review these findings and design a comprehensive assessment plan that addresses identified deficiencies with the current system and meets the current needs of all students. This may include identifying immediately actionable changes to improve the current process and a more holistic plan that would include broader changes. This section outlines the planning steps and recommendations based on the research and findings of this study.

To understand why an assessment committee and planning steps are recommended, it is important to review the key findings of this study. The first part of this research involved identifying teacher perceptions of the final exam with the intent of discovering prevailing final exam beliefs and practices only to find there was not a consistent belief about the Peters Township final exam process. In fact, support for and confidence in the current system was low. Teachers did find common ground, however, on the value of giving a final exam. They believed the exams would hold students accountable for their learning, prepare them for college, and increase student retention of learning. This aligns with research that has proven both at the collegiate and secondary school level that a test on studied material promotes learning and retention as well as encourages ownership of learning and creates a strong basis on which to build (McDermott et al., 2014). Identifying why a final exam or other cumulative assessment should be used is important to developing a sound plan. A 2013 study identified three important reasons for giving finals. These included:

1. It motivates students to learn the course content and potentially seek further study.

2. It increases long-term retention of content after the course ends.
3. Final exam test questions indicate important content to students (Glass et al., 2013).

Unfortunately, the Peters Township final exam system evaluated in this study did not appear to meet these expectations. Many teachers and most alumni said that the final exams at Peters did not matter to students because it did not impact their grades. The analysis of grading data supported these claims and yielded one of the most significant findings of this study. The final exam did not have value for most students. Planning for and administering the final exams takes a considerable amount of time for teachers with little to no reward. Most student scores on the final exam were not representative of their classroom performance and not impactful on their final grades. Therefore, teachers were not able to use that data to evaluate student learning and students were not getting a realistic experience of studying for a cumulative assessment. The number of students who received an “F” on their final exam but still achieved a grade of “A” for the year was striking. Of the students who had an average quarterly grade of an “A,” only 16% of them scored an “A” on the final exam while nearly 10% of them received an “F” on the final. A closer look at the calculation showed that it did not improve any students grade and it was nearly impossible with the current final exam calculation to do so. The impact on their grades was not enough to encourage students to study and do their best, therefore minimizing any benefits to taking a final exam. The benefits of being held accountable for their learning and increasing their retention were lost. In college, however, alumni reported that cumulative assessments mattered a great deal, and many felt unprepared for

this reality. Therefore, an assessment plan will need to address how to improve student study skills for high-stakes assessments and increase the value placed on the final exam.

Teachers and alumni responses showed agreement with the findings of the literature review that a final exam should be a measure of student learning. But most alumni and many teachers felt that the current exam system did not accurately measure student learning. Based on the large discrepancy of final exam grades and end of year grades this assumption was most likely true. Therefore, the perception that the final exam as designed is not a reliable measure of student learning needs to be addressed. If teachers are expected to invest time and energy into creating a reliable and authentic assessment, they must have confidence in the final exam process. The most frequently given explanation for this phenomenon in both surveys was exam weight. Teachers and alumni agreed that the exam weight was not enough to give the final exam any real value. Considering that of the 1,326 students grades analyzed, less than half of them could have been impacted by the final exam this seems to be true. The data analysis showed that for some, the final exam had significant value while for most it had none.

The current weight of the final exam for a full year course is 11%. With the system of quality points that is used, it appears that the weight of the final exam has created reverse inflation. This is a term used by Franke to explain when an exam weight will only cause a final grade to stay the same or go down (2018). Instead, a final exam weight should allow for student growth. Franke's study identified three problems to consider when determining weight:

1. Too little weight decreases student motivation to retain information.

2. A ceiling effect can occur if the weight does not allow for improvement, causing an inequality in opportunity for students.
3. Too heavy a weight can cause high stress and diminishing returns (Franke, 2018).

These guidelines provided a good starting point for review of the weighting at Peters Township. Among the considerations moving forward should be whether the weight needs to be the same for each course, if students who have hit the grade ceiling can opt out, or whether the final exam should be decoupled from final grades. Decoupling final exams was a solution presented by Franke that essentially removed the final exam as part of the final grade calculation. The idea is that students who would take the final exam would have the points added to their grades as “extra,” allowing more students to improve their grades (Franke, 2018). Any recommendations made for adjusting the weight of the final exam will need to be part of a larger school community discussion, but the committee should provide recommendations. Changing the weight of the exam may only resolve one issue, however. To realize the full benefit of using final exams, more work will need to be completed to build comprehension of the benefits of final exams to enhance learning and prepare students for collegiate assessments.

The final research question relied heavily on alumni feedback to identify collegiate practices. One thing that teachers and alumni agreed on was that final exams in high school should prepare students for college level assessments. They were divided on whether the final exams given at Peters Township were achieving that aim. Teachers were also unclear if the types of final exams they were giving compared to those taken by students in college. Alumni made it quite apparent that the high school final exams were

not comparable to their college exams. These findings were similar to the study conducted at Terre Haute South Vigo High School in which the same disconnect occurred between teacher and alumni views about college preparation (Hall & Ballard, 2000). The qualitative data from alumni provided valuable insight into what graduates viewed as college-readiness skills. Some of their comments were very specific to what they needed in high school to be successful in college including, “writing essays, taking tests similar to college level, and being challenged more” and “how to study versus what to study.” Alumni recommended increased focus on writing, study skills, assessment rigor, and application-based activities/projects. This supports findings during the review of literature in which college professors identified college-readiness as students’ prior knowledge, study habits, confidence, and motivation (Myers & Myers, 2007).

Other studies revealed that high school teachers and college professors often have different definitions of college-readiness. For example, in a study that defined being college ready as completing the first year without remediation being needed, professors defined college-ready as students having a solid grasp of basic skills whereas high school teachers focused on higher-order thinking skills (Woods et al., 2018). Without additional training or information, when planning for instruction that they believe is aligned to collegiate practices, teachers may automatically default to their own college experiences. This could cause teachers to miss opportunities to embed more relevant college-preparatory practices that would benefit students. The committee will need to develop a working definition of college-readiness, including what skills are most valuable. In a high school that sends over 90% of students to post-secondary education, it is imperative that

college-readiness be defined and understood by all educators and students. Collaborating with local colleges or universities is recommended to begin this process.

One of the promising outcomes of the teacher survey was the overwhelming support for diversifying their assessment practices and using multiple cumulative assessments. Their comments suggest a desire for more writing, application of content, and project-based assessments. Teacher recommendations often fell along content lines, but that also aligns with how college

students responded on the alumni survey. More heavily content-laden courses relied on traditional final exams while science and arts courses often relied on project-based or research-based assessments. Diversifying the high school practice would expose students to multiple types of rigorous assessments. Furthermore, teachers overwhelmingly said they preferred to administer summative assessments throughout the course term. Studies support the use of low-stakes testing, regardless of format, to prepare students for cumulative assessments (McDaniel et al., 2007). The committee should consider the research on the testing effect as a strategy to develop intentional testing plans for each course.

Beginning this action research project, it was surprising how little research existed on the use of final exams at the high school level when there exists considerably more research on their use in college. However, much of the college research can and should be applied to the high school setting. Building teacher understanding of cumulative testing research on student motivation, distribution of learning, and retention will help to maximize the effectiveness of giving final exams. Treating a final exam, the same as a quiz or unit test will not yield the same benefits. There is a clear difference between

testing in the classroom setting to measure student comprehension of learned content and final exams that are meant to encourage students to remember content learned over a span of time (Szpunar et al., 2007). Several alumni responded in the survey that high school focused on memorization while college assessments focused on application of knowledge. Professional development should be explored by the committee to assist teachers in developing their own assessment plans and ensure best practices are employed with a school-wide plan. Any assessment plan should be developed with a full understanding of the testing effect, distribution of learning, retention, and the psychology of testing.

Application of the Results

As the creators of assessments and communicators to students, teachers are essential to the success of any assessment plan. If they are unsure of the purpose or benefits of a cumulative assessment, their students will be too. Educators need to be transparent with their students about the process, including how testing will help students retain valuable information for future use (Khanna et al., 2013). Before they can accomplish this, however, they need to understand it. Defining the role of the teacher is the first step in developing a plan. The assessment committee should be made up of teachers representing all departments and include representatives from the counseling department and administration.

There are two options to create the committee, the first is to seek volunteers from all teachers and the second is to create a subcommittee from department facilitators. Once the committee is formed, a recommended sequence of steps will guide their work. These

steps are based on the findings of this research study which can be used to guide their work:

Step 1: Review Research and Findings

Step 2: Educate Staff

Step 3: Define College-Readiness

Step 4: Identify Opportunities for Improvement or Change

Step 5: Create an Assessment Plan

Step 6: Determine Exam Weight

Step 7: Present Recommendations and Solicit Feedback

Step 8: Make Revisions and Determine Next Steps

The first step recommended is a thorough review of the data, research, and findings. The committee should consider areas that need further study and prioritize the areas that need to be addressed. The next step is educating staff to ensure all teachers have a complete understanding of the current grading system, exam weight, and goal of the assessment committee. For the third step, the committee should develop a working definition of college-readiness. A review of alumni feedback and findings from the literature review will provide a starting point for their work. Collaborating with local colleges and universities offers an opportunity to ensure the definition aligns with college expectations. Reviewing current practices that support the college-readiness definition is the next step. This process may require additional input from teachers. Questions for the committee to consider include: Where are students already building college-readiness skills, and how can those practices be more intentional? Once these are established, the

committee needs to identify new opportunities for skill development and cumulative assessment practices.

The fifth step will be the most significant and lengthy process as it will address the specific needs identified in this study. Creating an assessment plan should be thoughtful and deliberate and based on current research. The plan should consider different cumulative assessment practices including the use of research papers, projects, and traditional multiple-choice exams. The committee will need to determine which courses should use a cumulative assessment and which type of assessment should be used. Additionally, they should consider the frequency and timing of testing to control over-testing and student stress. For teachers, the committee will need to identify if any professional development would be needed to implement the assessment plan.

Once a plan is developed, the weight of the exam needs to be reviewed with recommendations for improvements or changes. This will require a significant analysis of the impact on student grades of any changes to the current system. The committee should also consider enhancements to the current system if changes are not possible. The next step for the committee is to present their plan for consideration and solicit feedback for possible changes. It is recommended that the committee separate their recommendations into two categories: those that can be implemented immediately and those that require school board approval. Finally, the committee will reflect on the presentation and feedback to identify revisions or redirection. Next steps will be developed for either implementation of recommendations or further presentations to a broader committee.

While not included in the steps, the committee should consider other factors that would impact the success of an assessment plan. The stress level of students is an area

explored briefly in this study. While research suggests that final exams do not overstress students, teachers perceive that it does. Recognizing the stress factor can be addressed by coordinating an assessment plan to avoid conflict between departments and with high-stakes tests like AP, Keystones, and SAT/ACT. Additionally, clear expectations for teachers will need to be created with a follow-up plan.

Fiscal Implications

There were no fiscal implications to this study that needed to be addressed. All data analysis programs and survey tools were used at no cost. Potential future fiscal implications may include professional development time and training. In-service time is already factored into the teacher school year and may be used for committee work and data review. With the implementation of the Canvas Learning Management Platform, there may be additional tools considered by the assessment committee that would require no additional costs.

Limitations

Several limitations should be considered when reviewing the outcomes and recommendations of this study. First, only teacher and alumni perceptions were collected and reported. Additional stakeholders including parents, school board members, and administrators may have differing perceptions that will need to be considered as part of any future changes. Second, while alumni perceptions were fairly consistent, other variables may have impacted their responses including academic level of courses taken in high school and the type of college they are attending. In addition, the alumni sample was small. A larger sample may provide additional insights not already expressed. Third, interpretation of the qualitative data was limited by the responses given. Some teachers

expanded their thoughts to clarify their ideas, while others gave short simple responses. There may have been more consensus on some topics than was indicated in this report. In addition, this study focused on only the final exam and did not factor in other assessment practices. These assessments may also play a greater role in the overall assessment plan. Lastly, the grading system was not evaluated as a whole, only as it related to the impact of the final exam. A comparison of percentage grading system to the quality point grading system was not done as it was not the focus of this study. Future recommendations from the assessment committee may require a more specific analysis of the grading system.

Recommendations for Future Research

The use of testing to evaluate student learning, curriculum, and college-readiness has a solid footing in education. Though the Covid-19 Pandemic has led to many colleges relaxing their SAT and ACT requirements, it is unlikely that these will go away entirely. Studies have shown that these standardized assessments are still the best predictor of college success (Koretz & Langi, 2017). Knowing that testing is a critical part of student study means that research on this topic needs to continue.

As a result of studying the perceptions of teachers and alumni towards final exams and research on cumulative assessments, several other topics should also be explored that could enhance the value and reliability of assessing students. These include assessment development and evaluation, grading systems, the role of the teacher, and writing assessments.

This study did not review teacher training on assessment practices which would include creating authentic assessments and using assessment data. Because the reliability

of a final exam is crucial to its effectiveness as a learning tool, it is recommended that more research be conducted on classroom assessment practices including how teachers evaluate the efficacy of their assessments. Furthermore, some teachers had reported using the data to make decisions about curriculum. They were not asked if decisions were made in collaboration with other teachers, what they looked for in the data, or what decisions they have made. Additional staff surveys or interviews could identify additional areas that require in-service training.

Within this study, five teachers showed preference for switching to a percentage-based system of grading. Many schools and colleges calculate grades using percentages to determine letter grades. At Peters Township High School, a hybrid system is used. This has created confusion for teachers, students, and parents since a percentage based system is used prior to coming to the high school. When they calculate the percentage, sometimes the letter grade does not match their expectations. Future studies should investigate the use of a percentages compared to a quality point grading system to determine how each system impacts student performance. This must take into consideration, however, that students who know and understand the current grading system may change their behavior under a different grading system.

A study on the role of the teacher in the assessment process was not discussed in this study but additional research is recommended. Specific to this research should be the perceptions of the role of the teacher by both the teacher and the students, and teacher's comfort level and ability to create quality cumulative assessments. Clarification of the teacher role along with feedback from student perspectives would assist in creating a stronger assessment culture.

An unexpected finding of this study was the lack of preparation to write research papers reported by alumni. In fact, six alumni made comments such as, “I wish I had learned better writing and researching skills.” Several alumni commented on the strict five paragraph writing protocol that they were taught at Peters Township. Several teachers also commented on preferring writing assignments for cumulative assessments. Further study on the use of writing assignments in Peters Township is warranted based on this feedback to ensure that the writing instruction and opportunities in the curriculum are aligned with expected standards for college preparation.

Summary

Students graduating from high school in America today are expected to be “College and Career Ready.” This educational mission has become embedded in nearly all state standards for public schools. Quantitative measures for high schools are used to evaluate their effectiveness at preparing students for post-secondary pursuits. These measures are often how schools are evaluated in the public eye. As a result, schools have taken steps to improve student preparation for college and career. This focus includes ensuring students are exposed to a rigorous course of study and that students perform well on college-readiness tests and state-mandated end of course assessments.

As a high performing high school, Peters Township has always placed considerable effort on ensuring that its graduates that continue to college are well prepared. The high school provides a rigorous course of study with numerous college-preparatory courses. Teachers have participated in professional development aimed at improving instructional practices and expanding educational opportunities for students. In addition, a focus on increasing student performance on the SAT and ACT has led to the

use of the PSAT in grades 9, 10 and 11. Essentially, all teacher practices have been aimed at college and career preparation, including the use of final exams in all courses.

Nearly all courses at Peters Township High School use an end of course assessment. The core areas of English, math, science and social studies give an 80-minute assessment that includes multiple-choice and open-ended questions. Elective courses have the option of giving a traditional final exam or a project-based assessment. Students in an Advanced Placement class can opt out of taking the final exam if they take the AP test. A considerable amount of instructional time is used for course review in preparation for the final exam and administration of the final exam. For year-long classes, the final exam is a separate grade that accounts for 11% of a student's grade. For a semester course, the final exam accounts for 20% of a student's grade. Complaints and confusion about the final exam have led many in the school community to question its value, especially with the additional demands of state testing.

If the original intent of the final exam was to provide students with the experience and practice of taking a cumulative assessment and to hold them accountable for their learning, then the research indicates it is not functioning as intended. The school can consider eliminating giving final exams altogether, but the results of this study identify a need to continue the practice with improvements to build student college-readiness skills and increase their retention. By combining the findings from the review of literature with the findings of this study, an assessment committee will have the ability to make recommendations for a comprehensive assessment plan that address the college-readiness needs of Peters Township Students

References

- Bleske, R. A., Zeug, N., & Webb, R. M. (2007). Discrepant performance on multiple-choice and short answer assessments and the relation of performance to greater scholastic aptitude. *Assessment & Evaluation in Higher Education*, 32(2), 89 - 105.
- Burns, D. J. (2004, November/December). Anxiety at the time of the final exam: Relationships with expectations and performance. *Journal of Education for Business*, 80(2), 119 - 124.
- Caves, K., & Balestra, S. (2018). The impact of high school exit exams on graduation rates and achievement. *Journal of Educational Research*, 111(2), 186-200.
<https://doi.org/10.1080/00220671.2016.1226158>
- D'Agostino, J. V., & Bonner, S. M. (2009). High school exit exam scores and university performance. *Educational Assessment*, 14(1), 25 - 37.
- Franke, M. (2018). Final exam weighting as part of course design. *Teaching & Learning Inquiry*, 6(11), 93 – 103.
- Gannon, K. (2018, November 30). What is the purpose of final exams, anyway? *Chronicle of Higher Education*, 65(13), 1.
- Gewertz, C. (2019, March 3). *What tests does each state require?* Education Week.
<https://www.edweek.org/ew/section/multimedia/what-tests-does-each-state-require.html>

- Glass, A. L., Ingate, M., & Sinha, N. (2013). The effect of a final exam on long-term retention. *The Journal of General Psychology, 140*(3), 224 - 241.
- Hall, J., & Ballard, L. (2000). School/university collaborative inquiry: Determining the effectiveness of secondary final exam practices. *Contemporary Education 71*(4), 49.
- Hunt Institute. (2016, January 14). *The update: ESEA reauthorization*. <https://hunt-institute.org/resources/2016/01/the-update-esea-reauthorization-every-student-succeeds-act/>
- Kerdijk, W., Cohen-Schotanus, J., Mulder, B. F., Muntinghe, F. L. H., & Tio, R. A. (2015). Cumulative versus end-of-course assessment: effects on self-study time and test performance. *Medical Education, 49*(7), 709 - 716.
- Khanna, M. M., Brack, A. S. B., & Finken, L. L. (2013). Short- and long-term effects of cumulative finals on student learning. *Teaching of Psychology, 40*(3), 175 – 182.
- Koretz, D., & Langi, M. (2017). Predicting freshman grade-point average from test scores: Effects of variation within and between high schools. *Educational Measurement: Issues and Practice, 37*(2), 9 - 19.
- Lott, E. (2020, May 12). *2020 school guide rankings: Southwestern Pennsylvania's top districts*. Pittsburgh Business Times. <https://www.bizjournals.com/pittsburgh/news/2020/05/12/2020-school-guide-rankings-southwestern.html>
- Lysne, S. J., & Miller, B. G. (2017). A comparison of long-term knowledge retention between two teaching approaches. *Journal of College Science Teaching, 46*(6), 100 – 107.

- McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, *19*(4/5), 494 - 513.
- McDermott, K. M., Agarwal, P. K., D'Antonio, L., Roediger, H. L., & McDaniel, M. A. (2014). Both multiple-choice and short-answer quizzes enhance later exam performance in middle and high school classes. *Journal of Experimental Psychology*, *20*(1), 3 – 21.
- Mertler, C. A. (2019). *Introduction to educational research*. Sage Publications, Inc.
- Myers, C. B., & Myers, S. M. (2007). Assessing assessment: The effects of two exam formats on course achievement and evaluation. *Innovative Higher Education*, *31*(4), 227 - 236.
- National Center for Education Statistics. (2017). *State high school exit exams, by exam characteristics and state: 2017*.
https://nces.ed.gov/programs/statereform/tab2_10.asp
- National Education Association. (2020, June 25). *History of standardized testing in the United States*. <https://www.nea.org/professional-excellence/student-engagement/tools-tips/history-standardized-testing-united-states>
- Pennsylvania Department of Education. (2020, October). *Future ready index*.
<https://futurereadypa.org/>
- Perkins-Gough, D. (2008). Unprepared for college. *Educational Leadership*, *66*(3), 88 - 89.

Petrowsky, M. (1999). *The use of a comprehensive multiple-choice final exam in the macroeconomic principles course: An assessment* (ED427830). ERIC.

<https://files.eric.ed.gov/fulltext/ED427830.pdf>

Ramirez, G., & Beilock, S. L. (2011, January 14). Writing about testing worries boosts exam performance in the classroom. *Science*, 331(6014), 211 – 213.

Roderick, M., Nagaoka, J., & Coca, V. (2009). College readiness for all: The challenge for urban high schools. *Future of Children*, 19(1), 185 - 210.

Sadler, P. M., & Tai, R. H. (2007). Advanced placement exam scores as a predictor of performance in introductory college biology, chemistry and physics courses. *Science Educator*, 16(2), 1-19.

Szpunar, K. K., McDermott, K. B., & Roediger III, H. L. (2007). Expectation of a final cumulative test enhances long-term retention. *Memory and Cognition*, 35(5), 1007 - 1013.

U.S. Bureau of Labor Statistics. (2020, April). *College enrollment and work activity of recent high school and college graduates study*.

<https://www.bls.gov/news.release/hsgec.nr0.htm#:~:text=Following%20are%20some%20highlights%20from,percent%20and%2069.8%20percent%2C%20respectively>

U.S. Congress, Office of Technology Assessment. (1992). *Testing in American schools: Asking the right questions* (ED340770). ERIC.

<https://files.eric.ed.gov/fulltext/ED340770.pdf>

U.S. Department of Education. (1983, April). *A nation at risk*.

<https://www2.ed.gov/pubs/NatAtRisk/risk.html>

U.S. Department of Education. (2003, October). *From there to here: The road to reform of American high schools*.

<https://www2.ed.gov/about/offices/list/ovae/pi/hsinit/papers/history.pdf>

U.S. News & World Report. (2021). *Best high schools rankings*.

<https://www.usnews.com/education/best-high-schools/pennsylvania/districts/peters-township-sd/peters-township-high-school-17190>

Warren, J. R., & Kulick, R. B. (2007). Modeling states' enactment of high school exit examination policies. *Social Forces*, 86(1), 215 – 229.

Woods, C. S., Park, T., Hu, S., & Bertrand Jones, T. (2018). How high school coursework predicts introductory college-level course success. *Community College Review*, 46(2), 176 - 196.

Appendices

Appendix A

Survey Disclosure Letter to Teachers

Dear Teachers,

I am conducting a study for my doctoral program at California University of Pennsylvania titled: The Perception of the Final Exam as Preparation for College-Level Assessments and its Reliability as a Measure of Student Understanding of Course Content. The goal of this action research project is to analyze the reliability of the current final exam process and to determine if it aligns with current practices at the post-secondary level to affirm its use or to make recommendations for improvement.

You are invited to participate in this study as a teacher who has firsthand experience the final exam process in your classroom. The study involves completing a survey that includes basic demographic information and questions about your experience and perceptions of the final exam process. Your participation is completely voluntary and you may withdraw from the study at any time. Your responses will be anonymous and no identifying information will be collected or included in the study. The survey will take 7 to 10 minutes to complete. If you do not wish to participate in this study, please do not complete this survey. If you would like to participate in the study, please complete the survey using the link below. By clicking on the survey, you are giving consent to participate.

Your participation in this research will provide important information about the final exam process and assist in developing recommendations for consideration for the school district. All responses will be electronically stored with access by the researcher only. If

you have questions about this research, please contact Mrs. Lori Pavlik at 412 417-1698 or pav0968@calu.edu or California University of PA Faculty Committee Chair Dr. Kevin Lordon at Lordon@calu.edu.

Approved by the California University of Pennsylvania Institutional Review Board. This approval is effective 09/10/20 and expires 09/09/21.

Thank you for your time and participation.

Sincerely,

Lori A. Pavlik

Appendix B**Teacher Survey & Consent to Participate**

Please answer the following questions based on your experience with the current final exam structure at Peters Township High School. By clicking submit at the end of this survey, you are consenting to participate in this study.

1. I teach a course that uses the following type of final exam assessment:
 - Cumulative Multiple-Choice and/or Essay Exam
 - Project Based Assessment
 - No final exam is given

2. I use final exam data to determine if students have mastered the content:
 - Strongly Agree
 - Agree
 - Disagree
 - Strongly Disagree
 - Not Sure

3. I use final exam data to evaluate my curriculum.
 - Strongly Agree
 - Agree
 - Disagree
 - Strongly Disagree
 - Not Sure

4. The final exam grade is an accurate measure of student learning in my class.
 - Strongly Agree

- Agree
 - Disagree
 - Strongly Disagree
 - Not Sure
5. The final exam grade is consistent with how students perform on assessments throughout the course term.
- Strongly Agree
 - Agree
 - Disagree
 - Strongly Disagree
 - Not Sure
6. The final exam prepares students for college-level cumulative assessments.
- Strongly Agree
 - Agree
 - Disagree
 - Strongly Disagree
 - Not Sure
7. The final exam calculation allows students who are behind a chance to improve their grade.
- Strongly Agree
 - Agree
 - Disagree
 - Strongly Disagree

- Not Sure

8. The final exam calculation is fair and consistent to all students.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree
- Not Sure

General Final Exam Questions: Please answer the following questions regarding the general use of a final or summative exam in a high school setting.

9. Having a final exam holds students accountable for their learning.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree
- Not Sure

10. Studying for a final exam increases student retention of learning.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree
- Not Sure

11. Students need experience studying for a final exam in order to be prepared for college.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree
- Not Sure

12. Final exams provide teachers with important data on student learning.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree
- Not Sure

13. I prefer multiple summative assessments throughout the course term instead of one final cumulative exam.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree
- Not Sure

14. An end of course assessment can cause undue stress on students.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

- Not Sure

Open-Ended Questions:

15. What would be your ideal method of assessing student learning for your course?

16. What recommendations, if any, do you have for giving final exams?

Demographic Questions:

1. I teach in the following content area:

- English
- Math
- Science
- Social Studies
- World Language
- Fine Arts (Art, Music)
- Technology (BCIT, Media, Technology Education)
- Health/PE

2. Years of teaching experience

- 0 – 2 years
- 3 – 5 years
- 6 – 10 years
- 11 – 15 years
- 16 – 20 years
- 21 – 25 years
- Over 25 years

Appendix C

Survey Disclosure to Alumni

Dear Peters Township Graduate,

I am conducting a study for my doctoral program at California University of Pennsylvania titled: The Perception of the Final Exam as Preparation for College-Level Assessments and its Reliability as a Measure of Student Understanding of Course Content. The goal of this action research project is to analyze the reliability of the current final exam process and to determine if it aligns with current practices at the post-secondary level to affirm its use or to make recommendations for improvement.

As a student who has first-hand information about how assessments are used at the college level and experience with the high school final exam process, you are invited to participate in this study. The study involves completing a brief survey that includes basic demographic information and questions about the college assessments and the final exam process. The survey will take 7 to 10 minutes to complete. Your participation is completely voluntary and you may withdraw from the study at any time by stopping the survey. Your responses will be anonymous and no identifying information will be included in the study. If you do not wish to participate in this study, please do not complete this survey. If you would like to participate in the study, please complete the survey using the link below. By clicking on the survey, you are giving consent to participate.

Your participation in this research will provide important information about the final exam process and assist in developing recommendations for consideration by the school district.

All responses will be electronically stored with access by the researcher only. If you have questions about this research, please contact Mrs. Lori Pavlik at 412 417-1698 or pav0968@calu.edu or California University of PA Faculty Committee Chair Dr. Kevin Lordon at Lordon@calu.edu.

Approved by the California University of Pennsylvania Institutional Review Board. This approval is effective 09/10/20 and expires 09/09/21.

Thank you for your time and participation.

Sincerely,

Lori A. Pavlik

Appendix D**Alumni Survey & Consent to Participate**

Please answer the following questions based on your experience with the testing at the collegiate level and final exams at Peters Township High School. By clicking submit at the end of this survey, you are consenting to participate in the survey.

1. Taking final exams in high school prepared me for college-level testing.
 - Strongly Agree
 - Agree
 - Disagree
 - Strongly Disagree
 - Not Sure

2. The final exams that I took in high school were similar to the exams I take in college.
 - Strongly Agree
 - Agree
 - Disagree
 - Strongly Disagree
 - Not Sure

3. My final exam grades in high school were an accurate measure of my learning in the class.
 - Strongly Agree
 - Agree
 - Disagree

- Strongly Disagree
 - Not Sure
4. My college classes rely on cumulative exams (midterms and finals) to assess my learning.
- Strongly Agree
 - Agree
 - Disagree
 - Strongly Disagree
 - Not Sure
5. My college classes rely on project-based assessments to assess my learning.
- Strongly Agree
 - Agree
 - Disagree
 - Strongly Disagree
 - Not Sure

Open-Ended response

1. What surprised you most about how you are assessed in your college classes?

2. What do you wish you had learned in high school to be better prepared for college?

Demographic Information

3. What level best categorizes your current educational level?
 - Freshman in College

- Sophomore in College
 - Junior in College
 - Senior in College
 - College Graduate
4. Which category best represents your college study?
- Math-based
 - Science-based
 - Liberal Arts-based
 - Other: _____

Appendix E**IRB Approval**

**Institutional Review Board
California University of Pennsylvania
Morgan Hall, 310
250 University Avenue
California, PA 15419
instreviewboard@calu.edu
Melissa Sovak, Ph.D.**

Dear Lori,

Please consider this email as official notification that your proposal titled “The Perception of the Final Exam as Perception for College-Level Assessments and it's Reliability as a Measure of Student Understanding of Course Content” (Proposal #19-078) has been approved by the California University of Pennsylvania Institutional Review Board as submitted.

The effective date of approval is 9/10/20 and the expiration date is 9/09/21. These dates must appear on the consent form.

Please note that Federal Policy requires that you notify the IRB promptly regarding any of the following:

- (1) Any additions or changes in procedures you might wish for your study (additions or changes must be approved by the IRB before they are implemented)
- (2) Any events that affect the safety or well-being of subjects
- (3) Any modifications of your study or other responses that are necessitated by any events reported in (2).
- (4) To continue your research beyond the approval expiration date of 9/09/21 you must file additional information to be considered for continuing review. Please contact instreviewboard@calu.edu

Please notify the Board when data collection is complete.

Regards,

Melissa Sovak, PhD.
Chair, Institutional Review Board