

CASE STUDY: LAMPETER-STRASBURG HIGH SCHOOL (PA)

Exceeding Predictions on the Algebra I Keystone Exam

In the early 2000s, a teacher in Lampeter-Strasburg High School in Pennsylvania developed the software Get More Math to improve long-term retention of mathematics through adaptive, individualized mixed review. The refinement of this software in the classroom coincided with the development and release of the Keystone Exams, which were intended to measure student growth and mastery of core subjects. The first testing year for the Keystone Algebra I Exams was the 2012-2013 school year. Students using Get More Math in the Lampeter-Strasburg teacher's classroom surpassed the expectations for likelihood of passing the Keystone Exam as predicted by the Pennsylvania Value-Added Assessment System.

- ▶ *From 2013 to 2016, 248 out of 265, or 93.6%, of the students using Get More Math in this classroom exceeded their predicted score on the Keystone Exam.*
- ▶ *Of the 265 students who took the Keystone Exam, only 123 were predicted to pass. However, with Get More Math, 214 of those students attained a passing score—73% more than expected.*
- ▶ *The mean score on the Keystone Exam of the students using Get More Math was 1534, which was a statistically significant ($p < 0.0001$) average improvement of 38 points over the mean predicted score of 1496.*
- ▶ *The growth in score from a predicted value of 1496 to an actual value of 1534 leads to a Cohen's d effect size of 0.98, which is considered a large effect size.*

In addition, the average growth index for the teacher using Get More Math far exceeded the norm.

TABLE OF CONTENTS

Introduction.....	2
School Profile.....	3
Implementation.....	3
Results.....	3
Conclusion.....	7
Appendix A.....	8
Appendix B.....	10
Appendix C.....	12
Appendix D.....	13
References.....	14

Introduction

Mathematics teaching and learning has come under scrutiny in the last several years, particularly in the United States. On the 2018 PISA, students from the United States ranked 37th out of 78 educational systems assessed in mathematics (Schleicher, 2019, p. 7). Educational experts have long debated the merits and challenges of conceptual versus procedural understanding, individualized instruction, and how students learn best. Two evidence-based strategies that address long-term retention and student learning are retrieval practice and differentiation.

Retrieval practice is simply recall of knowledge over time, which reinforces learning by pulling information out of the brain rather than trying to improve retention by re-listening, rereading, or re-observing. For mathematics in particular, Kim et al. (2013) posited that mastering cognitive skills benefits from distributed practice (p. 31). When retrieval practice is interleaved, spaced, and varied, learners mimic how we experience life, which makes our memory stronger (Brown et al., 2014, p. 66). Brown et al. (2014) further explained our misunderstanding of massed practice—most people insist that repetitively practicing a concept with a single focus will help us to master it (p. 47). However, research in cognitive science does not support this idea. While educators often perceive distributed retrieval practice as an assessment tool, it functions as a learning tool that produces “desirable difficulty during learning” (Roediger & Karpicke, 2006, p. 254).

While incorporating appropriate retrieval practice and other learning strategies is possible without technology, there are ways that computers can address complex issues and analyses that are beyond the capacity of most teachers. Get More Math is a program that provides adaptive, individualized, cumulative practice for mathematics students to increase long-term retention. Get More Math leverages research-based retrieval practice using spacing and interleaving. The software determines which skill a student should attempt founded on data about the student’s strengths, weaknesses, and time since last practicing each skill.

Another way Get More Math addresses complex student needs is through content differentiation. Get More Math is not an instructional tool, so it allows the teacher to deliver the appropriate instruction for students while the software dynamically provides practice problems that are tailored to each student’s needs. The program intelligently selects problems for a student from a set of teacher-assigned skills based on each student’s past successes and failures, providing each student with a differentiated experience. Little et al. (2009) explained, “the value of differentiation [is] to respond to student readiness... [and provide] opportunities for all students to work with tasks that challenge them” (p. 42). Get More Math also scaffolds individual skills to meet students at the appropriate level. This multi-level approach to individualization of practice keeps students in their zone of proximal development. Murray and Arroyo discussed accessing a student’s zone of proximal development as integral in making learning efficient and effective (p. 749).

Get More Math was created by a mathematics teacher in the early 2000’s and was refined for over a decade in the classroom before being released to a set of pilot schools in Pennsylvania. This document will discuss the results that the first school using Get More Math experienced when the creator of the software implemented it in his classroom.

School Profile

Lampeter-Strasburg School District is in central Lancaster County in Pennsylvania, serving the townships of Lampeter and Strasburg and the historic borough of Strasburg. For the years discussed in this study, L-S High School graduated about 250 students per year. According to recent data from the National Center for Education Statistics, the high school had an enrollment of 972 students in the 2020-2021 school year (NCES, 2022). Of those students, approximately 17% were minority or two or more races. The high school is a Title I school with about 21% of the student body economically disadvantaged and receiving free or reduced-price lunches.

Implementation

The creator of Get More Math began teaching high school math in 1996. Beginning with his first year, he experimented with strategies to improve long-term retention. He built the first version of Get More Math software in 2004 and continued to refine it in his classroom through 2016.

In the 2012–2013 school year, the Keystone Algebra I Exam was mandated by the state of Pennsylvania for all students as an end-of-course assessment upon completing Algebra I. That, along with the existing Pennsylvania Value-Added Assessment System (PVAAS) used to measure student growth, provided a robust environment in which to measure the growth and achievement of the students using Get More Math software.

Get More Math was utilized in two different levels of Algebra I for ninth grade students; most students took the course for a full year of block scheduling, while about 20% took it for one semester of block scheduling. For this document, those students are not disaggregated into subgroups. However, the results of this study show significant growth across different ability levels, which are typically demarcated by predicted scores generated by the PVAAS model. The data from the Algebra I classes discussed in the study span the Keystone Exam testing years from 2013 to 2016.

Results

The results of using Get More Math were significant for student growth. To measure growth, Pennsylvania uses the Pennsylvania Value-Added Assessment System, or PVAAS, to make predictions about student scores. According to the Pennsylvania Department of Education, non-partisan researchers have called the PVAAS approach one of the most reliable approaches in measuring student growth (Pennsylvania Department of Education, 2021a). In order to determine an estimated likelihood of passing the Algebra I Keystone Exam, the PVAAS model uses all appropriate, available prior state assessment scores for an individual and then creates a predicted score based on the profile of all students who had similar previous scores (Pennsylvania Department of Education, 2021b). The state calculates the probability of a student scoring greater than or equal to the cut score for proficiency based on the predicted score and the associated standard error for that student profile.

The first notable measure of growth for L-S students was in the number of students who were predicted to pass the Keystone Algebra I Exam versus the number of students who passed (see Figure 1). From 2013 to 2016, only 123 out of 265 students using Get More Math with this teacher were predicted to pass, or fewer than half. However, 214 ended up passing the exam—over 80%, which is notably higher than the typical Pennsylvania pass rate of 60% to 70%.

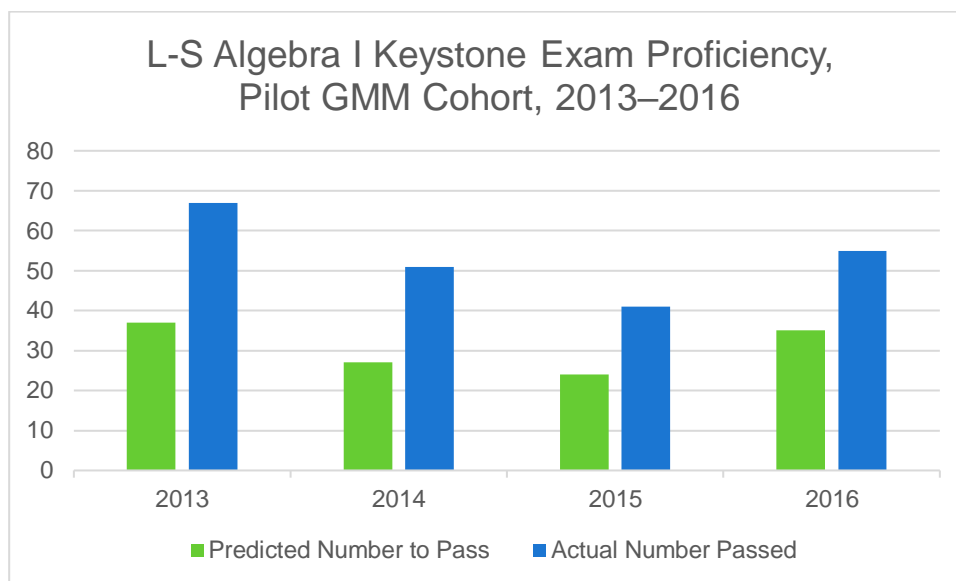


Figure 1: Keystone Exam Proficiency, L-S Algebra I Pilot GMM Cohort, 2013-2016

While not all students attained a rating of proficient or advanced, most of the students who did not still showed growth. Of the 265 students in the 2013 to 2016 study, 248 students scored better than they were predicted to score, an astounding 94%.

In Figure 2 below, the results of the 2016 PVAAS data are displayed for each individual student (for raw data, see Appendix A). Predicted results are shown by red and blue bars, while actual results are shown by the additional green bars, except for the two students who scored lower than predicted. Bold horizontal lines indicate the threshold ratings of Basic, Proficient (1500—the minimum passing score), and Advanced.

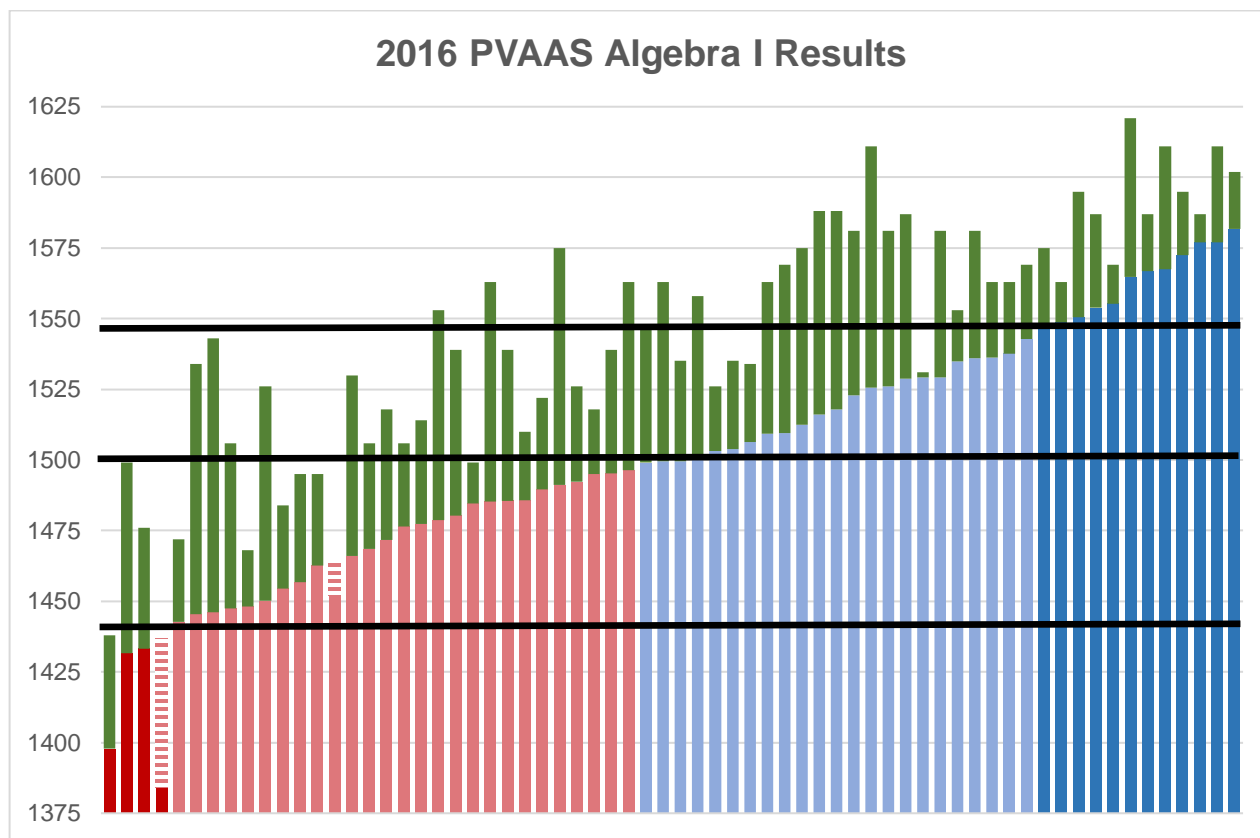


Figure 2: Keystone Algebra I Exam Predicted & Actual Scores, L-S Algebra I, 2016 (see Appendix B for 2013–2015 graphics)

From 2013 to 2016, the average predicted score for students in these Algebra I classes was 1495.9. The average actual score for those students was 1534.2. This is a statistically significant improvement ($p < 0.0001$, see Appendix C) of 38.3 points over the predicted average. Using Cohen's d , with pooled standard deviation, the effect size of this average score increase is 0.98 (see Appendix C), which meets the criteria for a large effect size.

The growth in scores for this L-S teacher utilizing Get More Math was not exclusive to low-achieving students. The scatterplot in Figure 3 plots the predicted scores vs. the actual scores for the 265 students in this study. The line of best fit is also plotted with these points ($r = 0.795$). The slope of 0.96, slightly less than 1, for the line of best fit, indicates that growth was consistent across all levels of achievement, with an inclination towards increased growth for low-achieving students. The line $y = x$ is also plotted on the graph to show the expected line of best fit when the actual score matches the predicted score for each student.

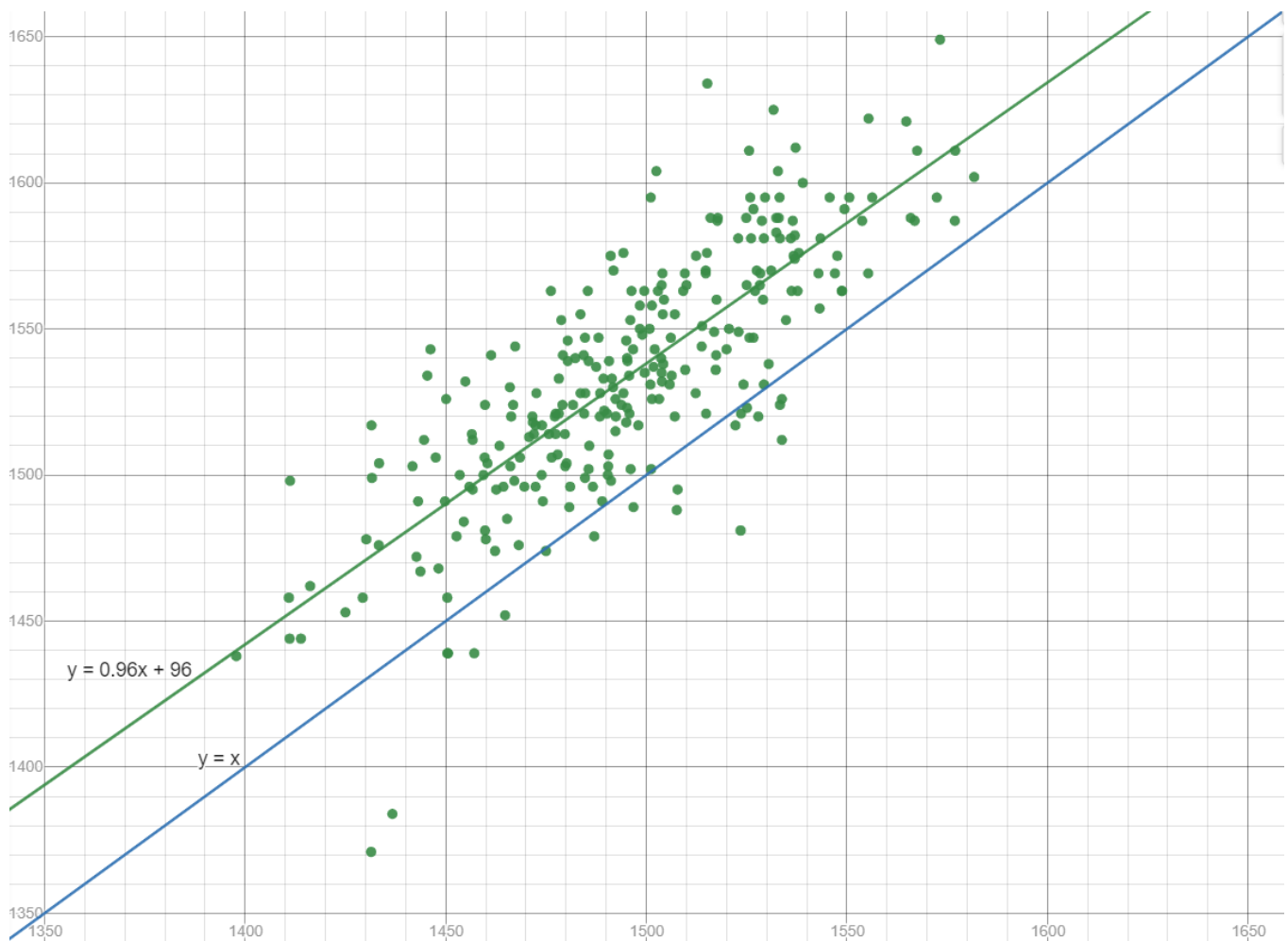


Figure 3: Predicted Score vs. Actual Score Scatterplot, L-S Algebra I, 2013-2016

The Pennsylvania Value-Added Assessment System distills growth measures down to one number, the average growth index, which can be compared across schools and districts. The growth index is also rolled into a three-year composite score. For this L-S classroom, the three-year composite growth index was 16.41 for 2014 to 2016. This is not an average of those three years but a compilation of the data across those three years. The growth index of each individual year for this teacher from 2014 to 2016 is shown below in Figure 4; over this time period the indices hovered between 7 and 13.

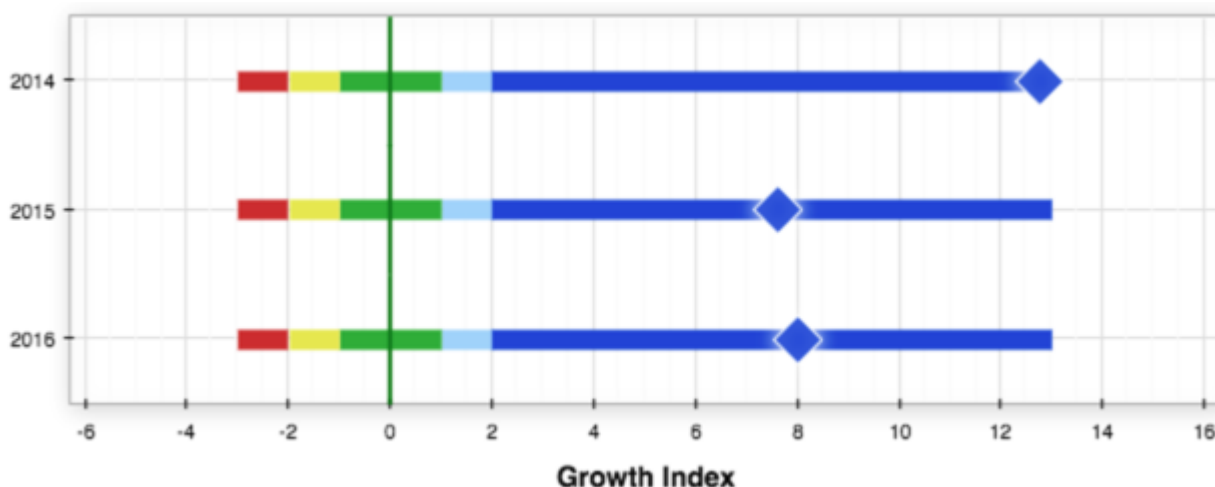


Figure 4: PVAAS Average Growth Index, L-S Algebra I, 2014-2016

The average growth index is a measurement of how many standard errors away from a growth measure of zero a set of students has scored (Pennsylvania Department of Education, 2021b, p. 20). Thus, scores between negative one and one are considered to have met the growth standard, as they are within one standard error of zero, showing evidence of no difference between the actual achievement and the expected achievement. Average growth indices greater than two are considered exceptionally good, as they are greater than two standard errors from a growth measure of zero.

Conclusion

The impetus behind Get More Math was one teacher's recognition that traditional math teaching techniques did not maximize his students' long-term retention. By utilizing technology and implementing Get More Math software, his students were able to show statistically significant growth over their predicted scores on the Keystone Algebra I Exam. This growth led to Lampeter-Strasburg High School achieving the third highest ranking in the state with their average growth index in 2016 (see Appendix D).

Appendix A: Predicted & Actual Keystone Algebra I Scores, L-S Algebra I, 2013–2016

(Values in bold text indicate actual scores that were greater than predicted scores. Green cells indicate proficient scores, while blue cells indicate scores that were not proficient but still exceeded the predicted score.)

Predicted	Actual
1581.7	1602
1577	1611
1576.9	1587
1573.2	1649
1572.4	1595
1567.5	1611
1566.9	1587
1565.9	1588
1564.8	1621
1556.3	1595
1555.4	1622
1555.3	1569
1553.8	1587
1550.6	1595
1549.4	1591
1548.8	1563
1548.7	1563
1547.6	1575
1547	1569
1545.7	1595
1543.4	1581
1543.2	1557
1542.9	1569
1539	1600
1538	1576
1537.7	1563
1537.2	1612
1537	1582
1537	1574
1536.7	1575
1536.5	1587
1536.2	1563
1536	1581
1534.8	1553
1533.8	1526
1533.8	1512
1533.3	1581
1533.3	1524
1533.2	1595
1533	1588
1532.8	1604
1532.4	1588
1532.4	1583

1531.7	1625
1531.1	1570
1530.5	1538
1529.6	1595
1529.3	1581
1529.3	1531
1529.1	1560
1528.8	1587
1528.4	1569
1528.3	1565
1527.9	1520
1527.5	1570
1527.1	1563
1526.7	1591
1526.7	1547
1526.1	1581
1525.9	1595
1525.7	1547
1525.6	1611
1525.1	1523
1525	1565
1524.9	1588
1524.2	1531
1523.6	1521
1523.5	1481
1523	1549
1522.9	1581
1522.2	1517
1520.6	1550
1520	1543
1517.8	1588
1517.7	1587
1517.5	1560
1517.4	1541
1517.3	1536
1516.9	1549
1516	1588
1515.2	1634
1515.1	1576
1514.9	1521
1514.8	1570
1514.8	1569
1513.9	1551
1513.8	1544
1512.4	1575

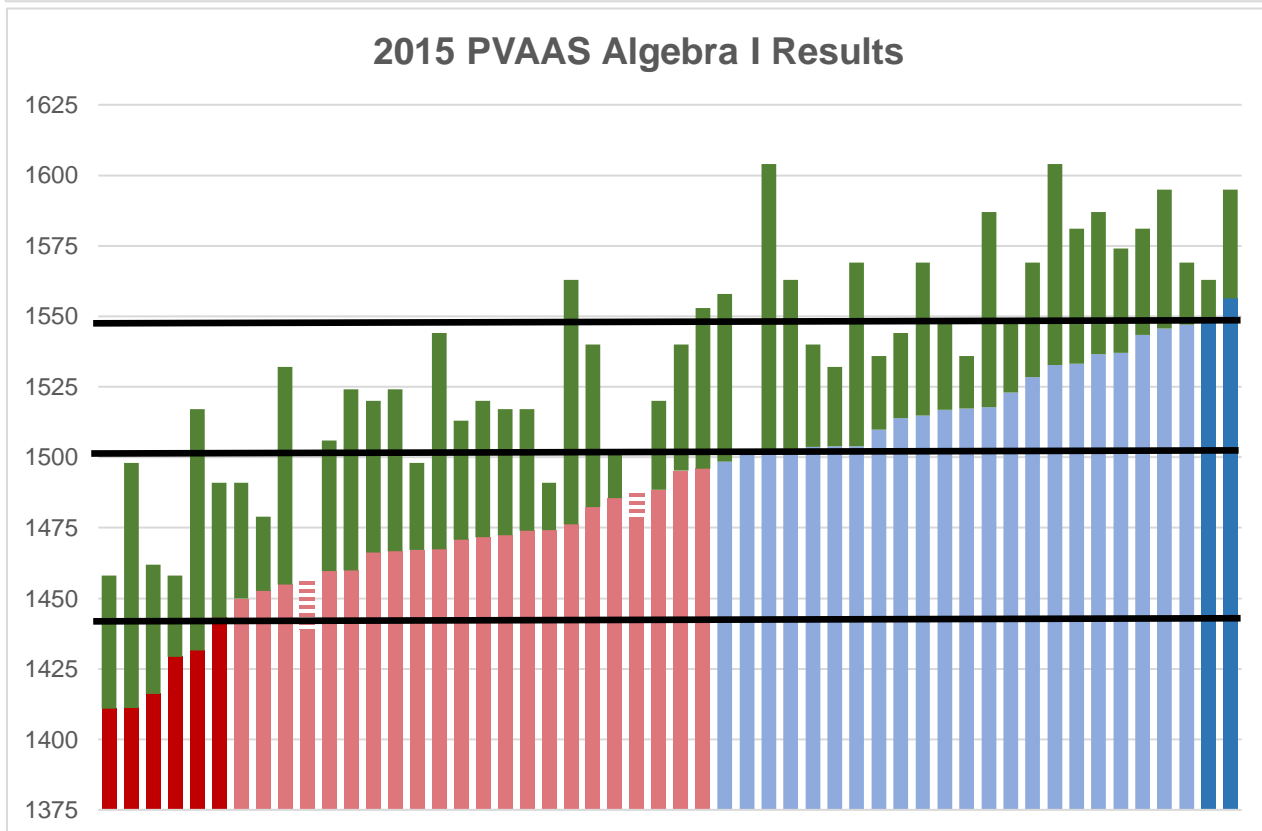
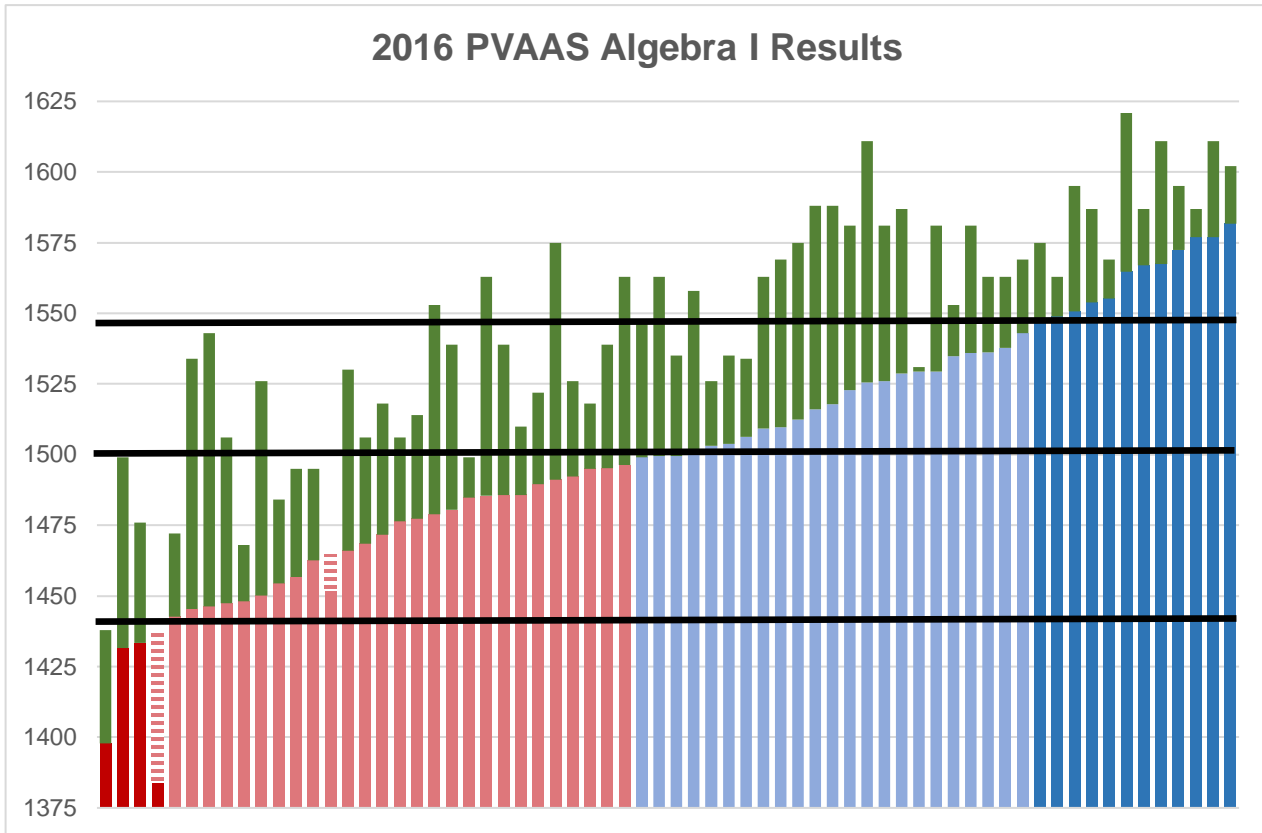
1512.3	1528
1510	1565
1509.7	1536
1509.6	1569
1509.2	1563
1507.8	1495
1507.6	1488
1507.1	1555
1507.1	1520
1506.3	1534
1506.1	1547
1505.8	1531
1504.4	1560
1504.2	1538
1504.1	1555
1504	1569
1503.9	1532
1503.8	1565
1503.8	1535
1503.7	1540
1503.2	1526
1502.9	1563
1502.5	1604
1502.1	1543
1501.8	1537
1501.4	1558
1501.4	1526
1501.2	1502
1501.1	1595
1501	1531
1500.8	1550
1499.6	1535
1499.5	1563
1499	1548
1498.4	1558
1498.4	1550
1498	1517
1496.8	1489
1496.7	1543
1496.3	1563
1496.1	1502
1496	1553
1495.8	1521
1495.7	1534
1495.3	1539

1495.2	1540
1495.2	1523
1495	1546
1495	1518
1494.3	1576
1494.3	1528
1493.8	1524
1492.4	1520
1492.3	1526
1492.3	1515
1491.8	1570
1491.7	1530
1491.4	1533
1491.2	1498
1491.1	1575
1490.7	1539
1490.6	1507
1490.5	1503
1490.4	1500
1490.2	1521
1489.5	1522
1489.3	1533
1489	1491
1488.5	1528
1488.4	1520
1488.1	1547
1487.5	1537
1487	1479
1486.7	1496
1485.8	1510
1485.6	1539
1485.6	1502
1485.4	1563
1484.8	1528
1484.7	1547
1484.7	1499
1484.5	1521
1484.4	1541
1483.6	1555
1483.6	1528
1482.3	1540
1481.7	1524
1481	1496
1480.8	1489

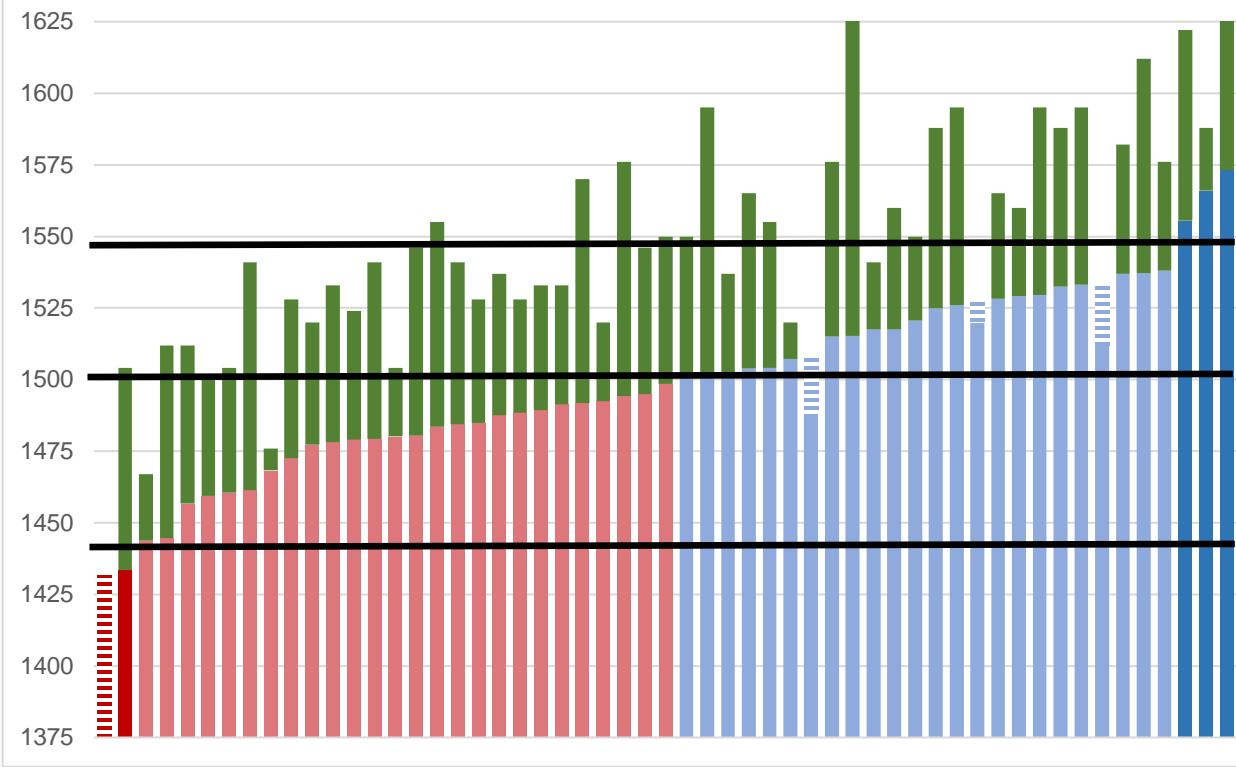
1480.4	1546
1480.4	1539
1480.1	1504
1479.8	1503
1479.7	1514
1479.2	1541
1479.1	1524
1478.8	1553
1478.2	1533
1478.2	1521
1477.9	1507
1477.4	1521
1477.4	1514
1477.2	1520
1476.4	1506
1476.2	1563
1475.7	1514
1475	1474
1474.2	1491
1474	1517
1473.9	1500
1472.6	1528
1472.4	1517
1472.4	1496
1472	1514
1471.7	1518
1471.6	1520
1470.8	1513
1469.6	1496
1468.5	1506
1468.2	1476
1467.3	1544
1467.1	1498
1466.8	1524
1466.3	1520
1466.1	1503
1466	1530
1465.3	1485
1464.8	1452
1464.4	1496
1463.4	1510
1462.6	1495
1462.3	1474
1461.3	1541

1460.4	1504
1460	1478
1459.8	1524
1459.8	1481
1459.7	1506
1459.4	1500
1457.1	1439
1456.7	1512
1456.7	1495
1456.5	1514
1455.9	1496
1454.9	1532
1454.5	1484
1453.5	1500
1452.7	1479
1450.6	1439
1450.4	1458
1450.4	1439
1450.1	1526
1449.8	1491
1448.2	1468
1447.5	1506
1446.2	1543
1445.4	1534
1444.6	1512
1443.7	1467
1443.1	1491
1442.7	1472
1441.7	1503
1436.7	1384
1433.4	1504
1433.3	1476
1431.6	1499
1431.5	1517
1431.4	1371
1430.2	1478
1429.3	1458
1425	1453
1416.2	1462
1413.9	1444
1411.2	1498
1411.1	1444
1410.9	1458
1397.8	1438

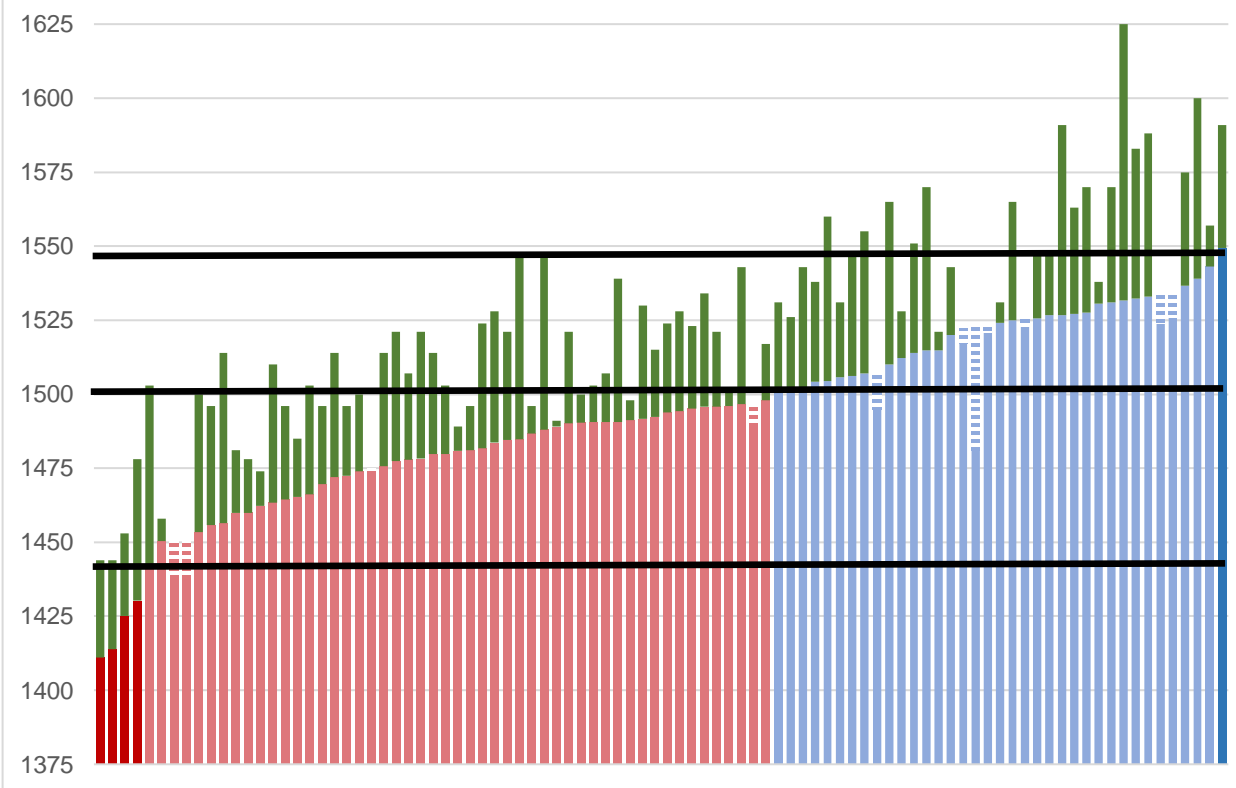
Appendix B: PVAAS Growth Results, L-S Algebra I, 2013–2016



2014 PVAAS Algebra I Results



2013 PVAAS Algebra I Results



Appendix C: Statistical Tests

Actual Scores of Get More Math Students vs. Predicted Scores

Summary statistics:

Column	n	Mean	Variance	Std. dev.	Std. err.
predicted	265	1495.90	1239.610	35.20810	2.16281
actual	265	1534.23	1813.386	42.58387	2.61590

Two sample T hypothesis test:

μ_1 : Mean of actual

μ_2 : Mean of predicted

$\mu_1 - \mu_2$: Difference between two means

H_0 : $\mu_1 - \mu_2 = 0$

H_A : $\mu_1 - \mu_2 > 0$

(without pooled variances)

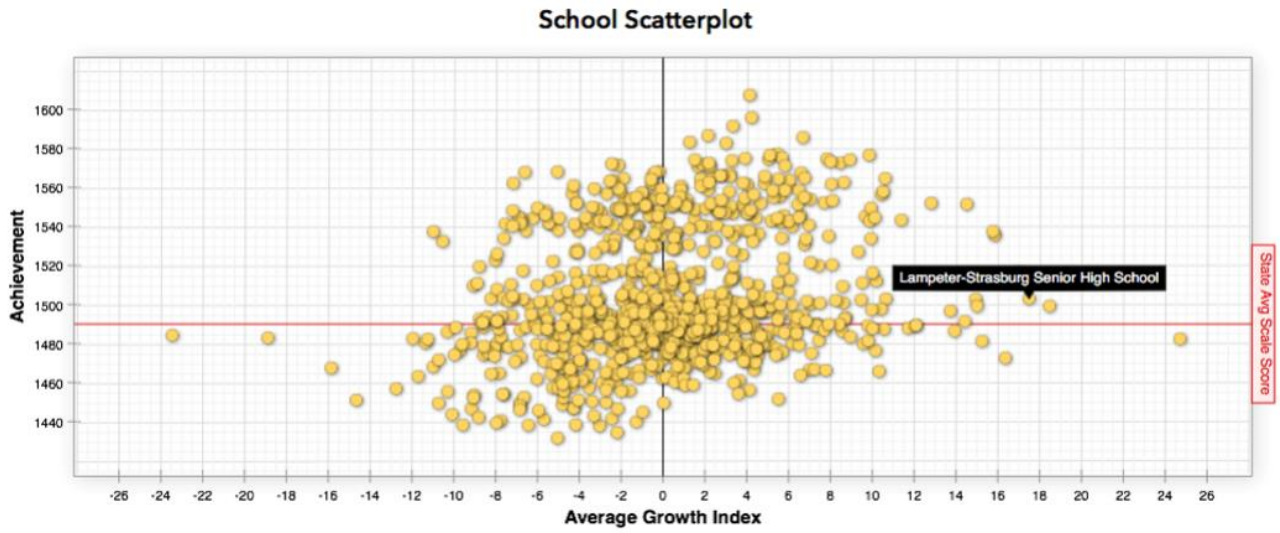
Hypothesis test results:

Difference	Sample Diff.	Std. Err.	DF	T-Stat	P-value
$\mu_1 - \mu_2$	38.328679	3.3942221	509.98681	11.292331	<0.0001

Effect Size for Actual Scores of Get More Math Students vs. Predicted Scores

$$d = \frac{38.328679}{\sqrt{\frac{42.58387^2 + 35.2081^2}{2}}} = 0.981 > 0.8$$

Appendix D: PVAAS Achievement vs. Average Growth Scatterplot, 2016



2016 Achievement vs. Average Growth Index, Pennsylvania Schools

References

- Brown, P.C., Roediger, H.L., III, & McDaniel, M.A. (2014). *Make it stick: The science of successful learning*. The Belknap Press of Harvard University Press.
- Kim, J.W., Ritter, F.E., & Koubek, R.J. (2013). An integrated theory for improved skill acquisition and retention in the three stages of learning. *Theoretical Issues in Ergonomics Science*, 14(1), 22-37. <http://dx.doi.org/10.1080/1464536X.2011.573008>
- Little, C.A., Hauser, S., & Corbishley, J. (2009). Constructing complexity for differentiated learning. *Mathematics Teaching in the Middle School*, 15(1), 34-42. <https://www.jstor.org/stable/41182949>
- Murray, T. & Arroyo, I. (2002, June 2). *Toward measuring and maintaining the zone of proximal development in adaptive instructional systems*. International Conference on Intelligent Tutoring Systems, 749-758. San Sebastian, Spain. 10.1007/3-540-47987-2_75
- National Center for Education Statistics. (n.d.). *School directory information – Lampeter Strasburg SHS*. Retrieved March 22, 2022, from <https://nces.ed.gov/ccd/>
- Pennsylvania Department of Education. (2021a). *What should you know about PVAAS?* <https://www.education.pa.gov/Documents/K-12/Assessment%20and%20Accountability/PVAAS/WhatShouldYouKnowAboutPVAAS.pdf>
- Pennsylvania Department of Education. (2021b). *PVAAS methodologies: measuring growth & projecting achievement*. <https://www.education.pa.gov/Documents/K-12/Assessment%20and%20Accountability/PVAAS/Methodology/PVAASMethodologies.pdf>
- Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>
- SAS Institute, Inc. (2022). *PVAAS scatterplots*. <https://pvaas.sas.com/scatterplot.html?as=d&aj=d&w4=116&ab=dD&x9=1&xp=2999&yb=1&wD=-1&x7=2>
- Schleicher, A. (2019). *PISA 2018: Insights and interpretations*. Organisation for Economic Cooperation and Development.