# Comparing the Performance on the MI and SAT/PSAT for the Purpose of Monitoring Student Achievement

Michael McDonald and Yulei Pang

***Abstract*—This paper will discuss the correlation between the SAT and the Math Inventory Test. Many school districts adopted the Math Inventory as a tool to measure student growth from grades kindergarten through high school. The Math Inventory is a computer-administered test that gives students math problems spanning from counting to high school level math. When completed, the students are given a quantile measure, much like a Lexile score for reading skill. The purpose of this study is to figure out if success on the Math Inventory is a good indicator for performing well on the SAT. For most high schools around the United States, objectives and lessons are aligned with those of the SAT. The goal of high school teachers is for students to excel on the SAT so that they can go to college, which means the tests used in middle school should be aligned with that goal. If the Math Inventory is not, then it might not be a very good use of school time and resources. Data was analyzed from the 2017-2018 school year from ten different high schools in an urban school district to determine the correlation between Math Inventory score, and the math score/sub scores of SAT/PSAT. The value of the Pearson's correlation coefficient is used to suggest a fairly moderate positive relationship between these two variables.**

***Index Terms*—Math inventory, odds ratio, regression, SAT/PSAT.**

## I. Introduction

In 2016, Connecticut changed its high school accountability measure from the Smarter Balanced Assessment Consortium (SBAC) to the Scholastic Assessment Test (SAT) under the leadership of Governor Dannel Malloy. The SAT also satisfies Connecticut Public Act No. 15-238, which requires 11th graders to take a nationally recognized college readiness assessment, approved by the board of education that measures skills in mathematics, reading, and writing [1]. Another incentive for students to succeed on this test is that "the federal education department also approved a new program to allow the state to rate the schools 1-5 based on their performance" [2]. Test scores are one of the indicators for that rating system. Schools use other assessments to measure student achievement, one in particular that is administered in parts of Connecticut is the Math Inventory (MI).

This paper examined the correlation between the SAT and the MI, also known as the SMI (Scholastic Math Inventory). Some school districts in New England adopted the Math Inventory as a tool to measure student growth from grades kindergarten through high school. The Math Inventory is a computer-administered test that gives students varying math problems spanning from counting to high school level math. When completed, the students are given a quantile measure, much like a Lexile score for reading skill. The purpose of this study was to figure out if success on the Math Inventory is an indicator for performing well on the SAT. All schools in Connecticut want their students to succeed on the SAT since they want their students to be college and be career ready, and since the SAT is an accountability measure. This means that their objectives, lessons, textbooks and their other assessments should be aligned with the topics of the SAT. If the Math Inventory is not properly aligned, then it might not be a very good use of the schools' time and resources.

Manuscript received April 2, 2021; revised July 20, 2021. This work was supported in part by the CSU Research Grant.

The authors are with Southern Connecticut State University, New Haven, USA (e-mail: pangy1@southernct.edu).

## II. Background

### A. Alignment of Tests

The importance of this study is to see if the PSAT/SAT and the MI are aligned. When using multiple assessments to measure something, they must not only align with the standards and curriculum that go with the area of study, but they also need to complement each other. "Through the administration of assessments that are carefully aligned to standards and curriculum, educators are able to gain an understanding of how student learning is progressing. Like curriculum, assessments must be aligned to content and grade-specific standards, in order to assess whether or not a student has gained the knowledge, skills, and abilities described in the standards" [3].

### B. The Math Inventory

The most useful resource for learning about the Math Inventory itself was the technical guide provided by the creators of the program. Some of the main topics of the guide include the quantile framework, development of the test, administration, and validity. The SMI was developed in 2008 and launched in 2010. Throughout those years and beyond, there were multiple phases of studies and validity analyses that were completed in the year 2012. The SMI website provided some other documents that went into more detail on areas like the estimated growth, how the quantile framework was created, and stating the standard of what it means to be college and career ready.

### C. The Math Inventory

The most useful resource for learning about the Math Inventory itself was the technical guide provided by the creators of the program. Some of the main topics of the guide include the quantile framework, development of the test,

administration, and validity. The

SMI was developed in 2008 and launched in 2010. Throughout those years and beyond, there were multiple phases of studies and validity analyses that were completed in the year 2012. The SMI website provided some other documents that went into more detail on areas like the estimated growth, how the quantile framework was created, and stating the standard of what it means to be college and career ready.

### D. Research on Standardized Testing

Since there are not many studies done in this particular area, general view on standardized testing was looked into. The motivation of this study was to determine if it is worth each district's time and resources to give the Math Inventory test to students multiple times a year. A 2015 education survey by Paul DiPerna polled over a thousand people on topics like school choice, common core and standardized testing [4]. About 42% of the people questioned said that the amount of time spent on standardized testing is 'too high'. The category that had the highest percent of people claiming the testing time was too high was school parents, which was at 47% [4].

A similar study in Georgia by Sandra Cochrane wanted to see teachers' perceptions on the Common Core Standards, as well as standardized testing [5]. A majority of teachers indicated that they felt more pressure under the Common Core Georgia Performance Standards (CCGPS) than with the No Child Left Behind (NCLB), and that they were spending too much time preparing the students. Teachers in this survey also said that they "look unfavorably upon the amount of time students spend taking these tests" [5]. Cochrane however had a sampling issue with the 66 out of the 70 completed female educator responses. Not all groups lean towards believing educators spend too much time on standardized testing. In the survey by DiPerna, only 30% said that they think the time spent testing was too high.

A case in Alabama involved a test similar to the Math Inventory, where students were taking the STAR test, which is a formative computer adaptive assessment taken several times a year. A parent was concerned the district was spending too much time testing and requested she be exempt from the mandated test. The researchers of STAR found that "across all grades, STAR scores were statistically significantly related to the end of year scores as measured by the Florida Comprehensive Assessment Test". Instances of students and teachers protesting the standardized tests were then referenced. A school in New York had almost one third of its students withdrawn from school to avoid taking a new standardized test aligned with the Common Core [6]. In Seattle, teachers joined together against the standardized testing movement and boycotted the implementations of state exams in their district [7]. However, an OECD PISA focus study found that about 25% of 15-year-old students attend a school where mandatory standardized tests are never used, and 60% attend schools where these tests are used only one or twice a year. Using up too much time is not the only concern of all this testing. They also found that on average about 59% of students worry about taking a test, and 55% worry about a test even when they are well prepared [8].

There are many teachers, parents, and students who have a problem with all the standardized testing that occurs in their schools. In order to justify using standardized tests, they must be aligned with the goals of the school.

### E. Predicting with Standardized Tests

The purpose of this study was to determine if success on the MI is a good predictor of success on the SAT. The technical guide for the Math Inventory references a comparison to the SAT, but only shows the results, which can be seen in Figure 1, and there contains no explanation of the study. There were other studies of either predicting with standardized tests or using other sources to predict the performance of a standardized test. Barbara Flexer sought to predict achievement by comparing several variables to the students' end of year algebra grade [9]. These variables included IQ, math computation, math concepts, problem solving, and a prognosis test. Flexer found that, "Of the variables included in the study, an algebra prognosis test was identified as the best overall predictor of success in an eight-grade algebra course" [9]. There were some cases where standardized tests were helpful in predicting and others where they were not. A study was done to see if there is a relationship between students' result on the SAT and their ability to solve 'complex mathematical modeling problems'. The idea was that traditional tests don't usually show the full abilities of students and that the questions only contain lower level facts and procedures. In this study students worked in groups on interdisciplinary realistic problems they call MEA's. Their results were that the SAT was not a significant predictor of MEA performance of the individual students. Their evidence suggests "MEA's require capabilities different than those tapped by conventional standardized tests" [10].

An advantage to standardized testing is the ability to provide a lot of data and compare results with a large group of students. An interesting investigation shed some light on what educators might actually be learning about what students know. Pietro Di Martino suggested that standardized tests in general are all product and no process, and that time restrictions and lack of clarifying questions could result in students who may know the material, but still get the question wrong [11]. A 2017 study had teachers selecting test questions with low results that they thought their students should have done better on. They then gave the question to the students in an environment where they were allowed asked clarifying questions and had no time constraints. An example led the researchers to find out that it was a simple miscommunication in the word 'close' that led many students to get the answer incorrect. The question was designed to see if students were able to 'compare decimal numbers' and in fact, many students got the questions right when it was rewritten to overcome the linguistic difficulty [11].

Most of the studies had test scores as the predictor, but in a study by Angela Duckworth, IQ and self-control were tested to see if they could predict standardized test scores and an overall grade for the math class [12]. Intelligence was measured with the Wechsler Abbreviated Scale of Intelligence, and self-control was measured with a survey by the parents and teacher of the student. Duckworth found that,

"changes in report card grades from eighth grade to ninth grade were predicted by both self-control and IQ while longitudinal changes in standardized achievement test scores were predicted by just IQ" [12]. Flexer's study differs from these results, since they found all of their variables, except IQ, to have a significant correlation with an end of year grade. It is also important to note that these studies were about their success in middle and high school. A meta-analysis conducted by Nathan R. Kuncel and Sarah Hezlett looked at success in graduate school. After synthesizing studies, it was found that standardized tests, "are effective predictors of performance in graduate school, and combined with undergraduate grades can predict important academic outcomes beyond grades earned in graduate school" [13].

## III. Methodology

### A. Data Source

The analyzed data was of ten different high schools from the 2017-2018 school year in an urban school district in Connecticut. Math Inventory scores of students in both the fall and spring were provided, since the students took the test a couple of times throughout the school year. Also provided were the PSAT scores for the freshmen, sophomores, and juniors of each school with a breakdown of each score. For the juniors only, their SAT scores were also provided with a breakdown of sub scores. There are 885 freshmen, 832 sophomores, and 997 juniors, which combine to a total of 2,714 students in the total sample. Out of the 997 juniors, 827 of them are in the PSAT group, and 777 of them are in the SAT group. Students' PSAT scores were paired with their fall Math Inventory score and the SAT score with their spring MI to match up tests taken around the same time. This means that there were four groups analyzed. With each PSAT/SAT score, the student's math score, and sub scores were listed, if provided. The three sub scores are heart of algebra, problem solving and data analysis, and passport to advanced mathematics; however, freshmen do not get a passport to advanced mathematics sub score. Table I below (McDonald Data) provides an example of the data, which includes the school, PSAT scores/sub scores, and MI score. The 9th Graders table is a little smaller since the PSAT 8/9 does not contain a sub score for Passport to Advanced Mathematics. The PSAT/SAT Score, sub scores, and MI have their own range. The PSAT/SAT score ranges from 160-800, the sub scores range from 1-15, and the Math
Inventory ranges from -400 to 1540+.

TABLE I: Example of the Data Set

| Student | School | PSAT | Algebra | Problem Solving | Passport to Adv. Math | MI Fall 2017 |
|---|---|---|---|---|---|---|
| Student 1 | School | 390 | 4 | 7 | 6 | 460 |
| Student 2 | School | 440 | 7 | 8 | 7 | 710 |
| Student 3 | School | 400 | 6 | 5 | 7 | 695 |

### B. Methodology

The data was analyzed using the 'R' programming language, which is widely used among statisticians and data scientists for statistical computing and data analysis. R provides open source route to participate in linear and nonlinear modeling, classical statistical tests, classification, clustering, graphical techniques and so on. The research objective was to find the association between variables (Math Inventory, SAT/PSAT math score, and sub score of SAT/PSAT). Several independent variables (PSAT/SAT score/sub scores) and the dependent variable (MI score) were defined and used in this study. Some traditional and classical statistical tests: i.e., simple linear regression, residual analysis, odds ratio, were applied to determine association between variables, depending on the characteristics of the variable.

Linear regression is a basic and commonly used type of predictive analysis. The overall idea of regression is to examine two things:

1) Does predictor (MI) do a good job in predicting an outcome (dependent) variable (PSAT/SAT)?
2) Which variables in particular are significant predictors of the outcome variable, and in what way do they – indicated by the magnitude and sign of the beta estimates – impact the outcome variable?

Simple linear regression assumes that each group comes from an approximately normal distribution and that the variability within the groups is roughly constant. The normality was checked for the data set with a Q-Q plot, which is a graphing technique to see if two data sets come from the same distribution. A residual versus fitted plots the difference between the dependent variable values and the predicted ones. No discernable pattern with the points and equal variance amongst the positive and negative residuals are indicators that the linear model was the most appropriate.

### C. Threats to Validity

One threat to the validity of this study is that the 2017-18 school year was the first year the students are taking the MI. Their score on the MI might not be as accurate since they are experiencing the test for the first time, while they have been preparing and taking practice tests for the SAT/PSAT in previous years. However, students are taking the MI multiple times throughout the year so this problem isn't as urgent in the SAT/Spring MI data. Another threat to the validity is that students may not take the test seriously. To combat this problem on the MI, students are advised to try their hardest since the test will take longer if they simply guess. The test ends when the students are consistently getting questions right/wrong in certain areas and can accurately place them in a quantile [14].

## IV. Results

We found the association was slightly positive between MI and PSAT/SAT through all grades (grade 9-11) in the study (see Fig1). In this section, we will give a brief introduction of the linear regression model and apply it to analyze the relationship between MI and PSAT/SAT. Particularly, we will also use linear regression the explore the relationship between MI and each of the sub scores for the PSAT/SAT.

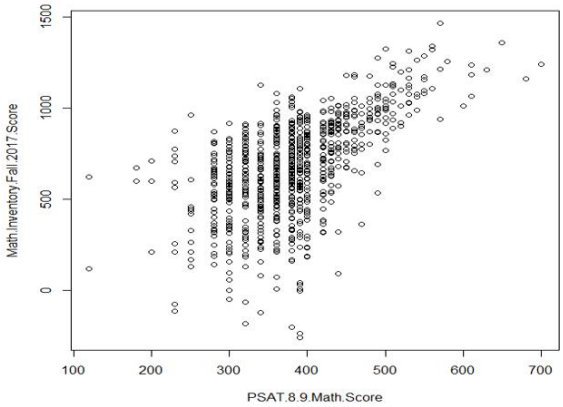


Fig. 1. Scatterplot of 9th graders MI vs PSAT 8/9 math score.

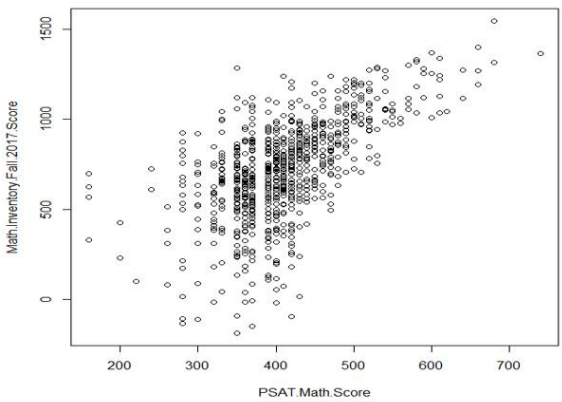


Fig. 2. Scatterplot of 10th graders MI vs PSAT math score.

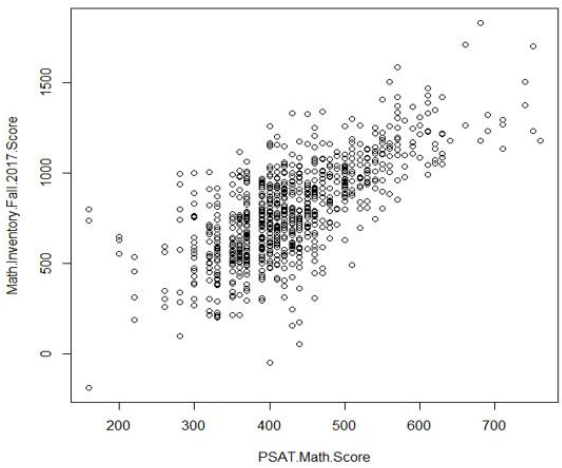


Fig. 3. Scatterplot of 11th graders MI vs PSAT math score.

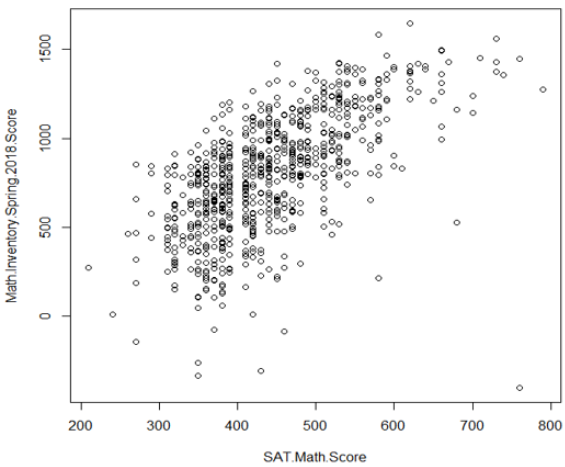


Fig. 4. Scatterplot for 11th graders MI vs SAT math score.

### A. A. Linear Regression Model

Linear Regression(LR) is a very simple approach for supervised learning. In LR, the function is linear equation and dependent variable can be expressed as a function of independent variable(s) in the form of

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{1}$$

$Y_i$ - PSAT/SAT

$X_i$ - MI

$\varepsilon_i$ - Random error term

We can use Ordinary Least Squares to estimate the values of the coefficients $\beta_0$ and $\beta_1$. The Ordinary Least Squares procedure seeks to minimize the sum of the squared residuals (see equation 2).

$$SSE = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \tag{2}$$

This provides the average squared error over all the data points. Therefore, this cost function is also known as the Mean Squared Error(MSE) function. LR algorithms aim to find the best values for $\beta_0$ and $\beta_1$, such that the MSE value settles at the minima.

To evaluate the performance of the model, we use coefficient of determination as our metric (see equation(3)).

$$r^2 = \frac{\text{explained variation}}{\text{total variation}} = \frac{SSR}{SST} \qquad 0 \le r^2 \le 1 \tag{3}$$

$$SSR = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2 \qquad df_R = 1$$

$$SST = \sum_{i=1}^{n} (Y_i - \bar{Y})^2 \qquad df_T = n - 1$$

### B. Performance of the Algorithm

R statistical Software (version 3.6.3; The R Foundation) was used for statistical analysis. The parameters of linear regression model are shown in Table II-V.

TABLE II: RESULTS FROM THE LINEAR REGRESSIONS OF MI AND PSAT 8/9 FOR 9TH GRADERS

| **10th Graders** | Coefficient | 95% Confidence Interval | P-Value | Multiple R-Squared |
|---|---|---|---|---|
| PSAT | 2.281 | (2.075, 2.487) | $2.2 \times 10^{-16}$ | 0.3617 |
| Algebra | 70.518 | (64.174, 76.861) | $2.2 \times 10^{-16}$ | 0.3645 |
| Problem Solving | 50.148 | (42.458, 57.837) | $2.2 \times 10^{-16}$ | 0.1649 |
| Passport to Adv. Math | 49.63 | (42.252, 57.014) | $2.2 \times 10^{-16}$ | 0.1735 |

TABLE III: RESULTS FROM THE LINEAR REGRESSIONS OF MI AND PSAT FOR 10TH GRADERS

| 9th Graders | Coefficient | 95% Confidence Interval | P-Value | Multiple R-Squared |
|---|---|---|---|---|
| PSAT 8/9 | 2.128 | (1.931, 2.324) | $2.0 \times 10^{-16}$ | 0.3382 |
| Algebra | 58.934 | (51.892,65.976) | $2.0 \times 10^{-16}$ | 0.234 |

TABLE IV: RESULTS FROM THE LINEAR REGRESSIONS OF MI AND PSAT FOR 11TH GRADERS

| 11th Graders (PSAT) | Coefficient | 95% Confidence Interval | P-Value | Multiple R-Squared |
|---|---|---|---|---|
| PSAT | 2.111 | (1.962, 2.26) | $2.0 \times 10^{-16}$ | 0.4837 |
| Heart of Algebra | 62.043 | (57.267, 66.819) | $2.2 \times 10^{-16}$ | 0.4407 |
| Problem Solving | 61.534 | (55.496, 67.573) | $2.2 \times 10^{-16}$ | 0.3266 |
| Passport to Adv. Math | 52.136 | (46.298, 57.975) | $2.2 \times 10^{-16}$ | 0.2713 |

TABLE V: RESULTS FROM THE LINEAR REGRESSIONS OF MI AND SAT FOR 11TH GRADERS

| 11th Graders (SAT) | Coefficient | 95% Confidence Interval | P-Value | Multiple R-Squared |
|---|---|---|---|---|
| SAT | 2.359 | 2.154 – 2.564 | $2.2 \times 10^{-16}$ | 0.3968 |
| Algebra | 78.028 | 70.915 – 85.141 | $2.2 \times 10^{-16}$ | 0.3743 |
| Problem Solving | 58.73 | 52.155 – 65.306 | $2.2 \times 10^{-16}$ | 0.284 |
| Passport to Adv. Math | 60.662 | 51.908 – 69.416 | $2.2 \times 10^{-16}$ | 0.1927 |

The tables shown above are broken into four groups based on grade level with the exception of the juniors having two groups. In the four data groups, there was a pattern amongst the scatterplots. The general positive trend can be seen in each case, although it not showing very strong correlation due to the points being spread out and forming a 'cloudy' look. Each scatterplot also has several outliers. The results in Tables II – V are similar as well. Each row of a table represents a linear regression with the Math Inventory. For each regression, the coefficient, a 95% confidence interval for the coefficient, p-value, and the multiple r-squared value are listed. The coefficient represents predicted increase in Math Inventory if the independent variable were increase by 1. For example, in Table V the SAT math score coefficient of 2.359 means that for every unit increase in the SAT math score, the predicted MI score goes up 2.359. The 95% confidence intervals show a range that contains a 95% chance of containing the true coefficient and not just the sample in the study. For example, there is a 95% chance that the true coefficient of the SAT math score lies between 2.154 – 2.564. The low p-values from each group indicate that you can reject the idea that the two variables are not related. A p-value of $2.0 \times 10^{-16}$ which is essentially zero is saying the odds of seeing these results between two unrelated variables is essentially zero. A multiple r-squared value tells us how much of the variation of the data can be explained by the model. The multiple r-squared of 0.3968 for SAT math score means that roughly 40 percent of the variation in the Math Inventory score can be explained by its linear model with the SAT math score. The sub score coefficients and confidence intervals are much higher than the PSAT/SAT since the range is lower. An increase in a sub score by one would correlate to a much higher increase in Math Inventory score compared to increasing a PSAT/SAT score by one.

Another comparison made between the two assessments was grouping the students into a 'Low' or 'High' group for each test based on their percentile. The data is shown below in Table VI. The top half performing students on the PSAT/SAT were labeled high and the bottom half were labeled low. This was repeated for the Math Inventory and then the number of students in each combination of groups were examined. For example, there were 326 9th Graders who finished in both 'Low" group for both the PSAT 8/9 and the Math Inventory and there were 117 9th Graders who finished in the 'Low' group for the PSAT 8/9 and the 'High' group for the Math Inventory.

## V. DISCUSSION

Overall the results from this study have shown that the Math Inventory is an appropriate test to use as an indicator of success on the mathematics used on the SAT. Results from the regressions, residuals, Q-Q plots, and odds ratios helped show that students who performed well on one test were likely to perform well on the other.

In regression analysis, p values and coefficients work together to show which relationships in the model are statistically significant and the nature of those relationships (positive or negative). The p values for each independent variable test the null hypothesis that the variable has no correlation with the dependent variable.

A low p value (<.05) indicates that the null hypothesis should be rejected. In other words, a predictor that has a low p value is likely to be a meaningful addition to that model. The coefficients for the PSAT/SAT math scores ranged from 2.111 to 2.359, which indicated a general positive trend in the data. This shows that a higher score on the PSAT/SAT correlates to a higher score on the Math Inventory. What is worth pointing out is the very consistent low p-values of about 2.2 x 10-16 for all of those tests which show significance in the coefficients for the models.

The r-squared values for the PSAT/SAT math score ranged from .3382 to .4837. Those values indicate that about 34 to 48 percent of the variation in the Math Inventory can be explained by the variation in the PSAT/SAT. In general, the higher the r-squared, the better the model fits the data. However even noisy, high-variability data can have a significant trend. In some fields, like the field that attempts to predict human behavior, it is entirely expected that the r-squared values will be low. Having that much of an impact on the variation of the variable is evidence that these two variables are positively correlated. Receiving those r-squared values should be viewed as high when considering the large sample size and the many variables that go into students taking both tests. These variables include students' attitude towards the PSAT/SAT or MI, motivation, and the differences amongst the tests. While researching about standardized tests, there were studies that looked into other

variables that had an impact on standardized test scores. Duckworth's study showed that IQ had a significant impact on the students' test score [12].

Odds rations (OR) are an important measure of the relative chance of an event of interest happening. The OR represents the odds that an outcome (high PSAT/SAT) will occur given a particular exposure (high MI), compared to the odds of the outcome occurring in the absence of that exposure.

- OR = 1 Exposure does not affect odds of outcome
- OR > 1 Exposure associated with higher odds of outcome
- OR < 1 Exposure associated with lower odds of outcome [15]

The more reassuring data came from the High Low comparison from both tests that can be seen in Table VI. The important statistic to take away from this table is the odds ratios, which are used to determine the probability of an event occurring across two groups [16]. In this case the event is placing in the top fifty percent of both the PSAT/SAT and the Math Inventory. The odds ratios ranged from 7.74 to 10.39, which means that students placed in the 'High' group for the PSAT/SAT are seven to ten times more likely to be placed in the 'High' Group for the Math Inventory. Cohen's Rule of Thumb dictates that a 4.3 is considered 'large' while the article by Chen and Cohen mentioned the numbers 5 and 6.71 as minimums of an indicative large effect size [16].

Linear regressions were performed between the Math Inventory and the sub scores to see if any of the sub scores had a more significant correlation to the Mi. The results from the study concluded that the Heart of Algebra had a stronger positive correlation the MI than Problem Solving and Data Analysis and Passport to Advanced Mathematics. Two indicators that help support this conclusion are the coefficients and r-squared values found in the regressions across the four groups. The coefficient for Heart of Algebra ranged between about 59 to 79. Problem Solving and Data Analysis had a coefficient range of 50 to 62 and Passport to Advanced Mathematics ranged from 50 to 61. These coefficients meant that improving the Heart of Algebra sub score by 1 correlated to a higher increase in MI score than increasing the other sub scores by 1. Across the four groups analyzed, Heart of Algebra had a higher range of r-squared values. The r-squared values of Heart of Algebra ranged from 0.234 to 0.4407 while Problem Solving and Data Analysis ranged from 0.1649 – 0.3266 and Passport to Advanced Mathematics ranged from 0.1735 to 0.2713. These results showed that the Heart of Algebra was able to explain more of the variation in MI scores.

There are a couple of ways that the correlation is higher than what the results have said. One reason is that the Math Inventory scores from the fall of 2017 and spring of 2018 were in the first year of implementing the MI. Students may not have gotten used to the format of the test, and may have needed a couple more attempts to really round out their appropriate score. Another reason to question the validity is that students may not take the test seriously. Some students will not try their hardest and would like to just get the test over with. The differences among the MI and SAT also could have had an impact on the correlation. Backes' study on the comparison between paper versus computer test did show a difference in students' test scores [17]. A recommendation for the future would be to compare MI and PSAT/SAT data when students have had more experience with the Math Inventory.

The main limitation of this study is the fact that this study was conducted in a single comprehensive cancer hospital. Multicentric-international studies would likely better reflect the impact of pandemic on molecular cancer imaging. Pandemic timeline was based on Connecticut data and may not reflect the results of the national or global course of the pandemic. Extreme sample bias may be confounded race analyses due to disproportionately high percentage of Caucasian population. Also, we used some open source packages in the R programming languages, where external threats may be introduced.

TABLE VI: RESULTS FROM LOW/HIGH GROUP COMPARISONS FOR EACH GRADE

| *Low* | *326* | *117* | *7.74* | *9th Grade* |
|---|---|---|---|---|
| High | 117 | 325 | | |
| **Low** | **306** | **110** | **7.74** | **10th Grade** |
| **High** | **110** | **306** | | |
| Low | 316 | 98 | 10.36 | 11th Grade (PSAT) |
| High | 98 | 315 | | |
| **Low** | **297** | **92** | **10.39** | **11th Grade (SAT)** |
| **High** | **92** | **296** | | |

## VI. CONCLUSIONS AND FUTURE WORK

In this paper we conducted a case study to assess the association between Math Inventory and SAT. Overall the results from this study have shown that the Math Inventory is an appropriate test to use as an indicator of success on the mathematics used on the SAT. The data from 10 different high schools from the 2017-2018 school year in an urban school district in Connecticut. It would be interesting to investigate whether the results can be applied to other states or not.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

Michael conducted the research and did the data analysis part. Dr. Pang was advising Michael to carry out and complete this project. All authors had approved the final version.

## ACKNOWLEDGMENT

This work is supported by Minority Recruitment and Retention Grants and Faculty Creative Activity Research Grants at Southern Connecticut State University.

## REFERENCES

[1] Connecticut Administration of the SAT Frequently Asked Questions for Superintendents. (2015). [Online]. Available: https://portal.ct.gov/-/media/SDE/Student-Assessment/SAT/sat_faqs_superintendents.pdf

[2] M. D. Andrew, "11th — Graders will avoid controversial smarter balanced test," *Hartford Courant 6 Aug. 2015*, Hartford Courant.

[3] Center on Standards and Assessments Implementation (CSAI), *WestEd, Standards Alignment to Curriculum and Assessment*, CSAI Update, 2018.

[4] DiPerna, Paul, and Friedman Foundation for Educational Choice, "2015 schooling in America survey: Perspectives on school choice, common core, and standardized testing. polling paper No. 24," *Friedman Foundation for Educational Choice*, 2015.

[5] C. Sandra and J. A. Cuevas, "Teacher perceptions of the common core performance standards," *Georgia Educational Researcher*, vol. 12, no. 1, 2015, p. 1.

[6] H. C. Javier and A. Baker, "A tough new test spurs protest and tears," *The New York Times*, 2013, p. A24.

[7] W. Teresa, "Standardized testing becomes the great divide in schools policy," *Los Angeles Times*, 2013.

[8] M. M. Mukaka, "Statistics corner: A guide to appropriate use of correlation coefficient in medical research," *Malawi Medical Journal: The Journal of Medical Association of Malawi*, vol. 24, no. 3, 2012, pp. 69-71.

[9] K. F. Barbara, "Predicting eighth-grade algebra achievement," *Journal for Research in Mathematics Education*, vol. 15, no. 5, Jan. 1984, pp. 352–60.

[10] K. Ozgul *et al*., "The relationship between students' performance on conventional standardized mathematics assessments and complex mathematical modeling problems," *International Journal of Research in Education and Science*, vol. 2, no. 1, Jan. 2016, pp. 239–252.

[11] D. M. Pietro and B.-F. Anna, "Beyond performance results: Analyzing the informational and developmental potentials of standardized mathematics tests," *For the Learning of Mathematics*, vol. 37, no. 3, Jan. 2017, pp. 39–44.

[12] D. L. Angela *et al*., "What 'no child left behind' leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades," *Journal of Educational Psychology*, vol. 104, no. 2, May 2012, pp. 439–451.

[13] R. K. Nathan and S. A. Hezlett, "Standardized tests predict graduate students' success," *Science*, vol. 315, no. 5815, 2007, pp. 1080–1081.

[14] R. Heri, "The comparison of accuracy scores on the paper and pencil testing vs. computer-based testing," *Turkish Online Journal of Educational Technology - TOJET*, vol. 14, no. 4, Oct. 2015, pp. 135–142.

[15] S. Magdalena, "Explaining odds ratios," *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, vol. 19, no. 3, 2010, p. 227.

[16] H. N. Chen *et al*., "How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies," *Communications in Statistics – Simulation and Computation*, vol. 39, no. 4, 2010, pp. 860-864.

[17] B. Ben and J. Cowan, "Is the pen mightier than the keyboard? The effect of online testing on measured student achievement," *Economics of Education Review*, 2018.


Copyright © 2021 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).

**Michael McDonald** received his M.S. degree in mathematical education from Southern Connecticut State University in 2019. He currently works as a math teacher at Glastonbury High School. He spent five years prior working at James Hillhouse High School in New Haven, Connecticut.

**Yulei Pang** received her M.S. degree in statistics and her Ph.D. degree in mathematics from Texas Tech University in 2012 and 2014, respectively. She is currently working as an associate professor in the Department of Mathematics at Southern Connecticut State University, US. Her current research interest includes applied mathematics and applied statistics.