ABSTRACT

IMPACT OF STANDARDIZED TEST PERFORMANCE ON SUCCESS IN
INTRODUCTORY COLLEGE PHYSICS CLASSES

Michael Gattone, MS
Department of Physics
Northern Illinois University, 2019
Michael Eads, Director

This research project looks to analyze the ability of the ACT and other standardized tests to predict student success in introductory physics classes at Northern Illinois University (NIU). Admissions data from NIU was collected by the physics department from a total of 4262 students enrolled in introductory physics courses during the Fall 2012 semester until the Spring 2017 semester. This data includes academic level, gender, race/ethnicity, composite ACT score, high school GPA, and final letter grade. In particular, the composite ACT score of students was directly compared with their final letter grade from the student's introductory physics class. The student sample was then grouped according to each student's academic level, gender, and by whether or not the physics class required calculus. Each group of students then had their ACT score correlated with their final class grade to determine if the correlation was stronger for specific subgroups of students.

As a result of this research, there is a significant ($p < 0.001$) correlation between composite ACT score, and grade earned. However, the data has a large amount of variance, with students who received 14's on their ACT earning A's and students with perfect scores on the ACT earning F's. As such, the findings show that there exists a high amount of unexplained variability within the data.

NORTHERN ILLINOIS UNIVERSITY
DE KALB, ILLINOIS

August 2019

IMPACT OF STANDARDIZED TEST PERFORMANCE ON SUCCESS IN

INTRODUCTORY COLLEGE PHYSICS CLASSES

BY

Michael Gattone

A THESIS SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

MASTER OF SCIENCE

DEPARTMENT OF PHYSICS

Thesis Director:
        Michael T. Eads

TABLE OF CONTENTS

## LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Every year public school teachers and administrators face increasing pressure to have students meet academic performance benchmarks. Standardized tests such as the ACT and SAT are primarily used to mark this achievement. As such teachers and administrators are demanding higher test scores and high school students all across the country attempting to reach these demands are at dangerously high levels of stress. According to the American Psychological Association (APA), eighty-three percent of teens surveyed cited school as a source of stress, sixty-nine percent cited, "getting into a good college or deciding what to do after high school," and sixty-five percent cited, "financial concerns for their family" ("Stress in America", 2014). These results suggest that school and testing is the source for a large portion of today's students' stress. Abeles highlights this sentiment, describing the emphasis on testing as out of control and calling today's students, "the most tested generation in history" (Abeles, 2015). Abeles argues that this excessive focus on testing has negative consequences, including placing unhealthy stress on students.

In order to process the thousands of applicants, colleges are turning towards the ACT and SAT to rank students against each other. In order to earn admission, schools and families are pouring more of their time and money into preparing for standardized tests, such as the ACT and SAT. According to Mathew Chingos schools across America spent a cumulative $1.7 Billion on

standardized testing (Chingos, 2012). With an investment that large, it is the aim of this study to determine if the ACT and other standardized tests are a valid measure of performance in introductory college physics courses.

## Definition of Problem

Administrators at the high school level are pushing more and more for standardized test preparation which is costing schools billions of dollars. Illinois spent a total of $19,011,148 during the years 2007-2011 on primary assessment contracts alone, this is not including spending on assessments not included in state's primary assessment contracts nor does it include state-level spending on assessment-related activities that are not contracted out (Chingos, 2012). With schools and families spending more money on these services, it is imperative that the ACT and other standardized tests act as valid benchmarks for college readiness.

For students who are looking to pursue a physics degree, the ACT is lacking in a significant representation of physics material ("How Physics is Tested", 2018). According to their official ACT study guide, all the information you need to answer the questions will be presented in the passages. "There will be no questions such as, 'what is the second law of thermodynamics?', unless that answer is to be found explicitly in the given passage" ("How Physics is Tested", 2018). It stands to reason then, that for these specific students, standardized tests are especially poor indicators of potential student success. If a student wishes to go into a physics-related major, or even take an introductory physics course, is the ACT an appropriate indicator of potential success?

This study will compare college students' academic performance against their ACT scores when entering college. These students' academic achievement is defined by the grade they earned in these introductory courses.

## Significance of Study

The intent of this study is to analyze the effectiveness of standardized entrance exams for predicting different levels of success in first year college physics courses. To gauge student success, every student enrolled in a physics class as a freshman had their final grade in the class recorded. The five classes recorded were PHYS 273, 253, 211, 210 and 150. Over 4000 student grades were recorded and referenced to their ACT scores that they reported to NIU administration.

## Research Questions

I. Does ACT score correlate with performance in introductory college physics courses?

II. Are there specific groups of students that have a stronger correlation between ACT score and performance college physics?

## Limitations of Study

This study will investigate the performance of thousands of students; however, the data has been collected over the course of five years. In that time the classes have been taught by different instructors. This will naturally lead to slight variations in the way the content is

delivered.  However, these classes being introductory classes, the content of what is taught has remained primarily the same over the sample time period, and the grading guidelines remained the same at the university.

It is the goal of the ACT to be a measure of educational achievement. The purpose of this project is to determine if that goal was achieved by comparing ACT scores to grades earned in introductory physics courses. However, college GPA, and therefore the grades earned in these introductory physics classes, does not solely depend on education achievement. There are a number of non-cognitive factors that play a part in earning a specific grade. Such personal factors include: Getting enough sleep, class attendance, conformity and motivation (Goldman & Widawski, 1976). Different non-cognitive factors will be discussed.

CHAPTER 2

LITERATURE REVIEW

Effect on Students & Society

The ACT and other standardized testing have two major impacts on society. First is the stress it places on students and teachers. In Simpson's study, it was shown that testing leads to a narrowing of curriculum, cheating scandals, and more importantly poor physical and mental wellbeing (Simpson, 2016). These effects are the heart of this study and speak to how important a force the ACT has become, but even more so, the effect it has on students has become increasingly worrisome. Simpson states that students often face high levels of stress from pressure related to college admissions and the demands of school more generally. This deteriorating of students' mental and physical wellbeing is due to the culture of stress that is built because of the ACT and other standardized tests (Simpson, 2016).

The American Psychological Association conducted a survey in 2014 on teenagers (ages: 13-17) to understand how teens experience stress. The teens that responded reported stress levels far above what the APA considered healthy. Further analysis showed that on average, teens reported higher levels of stress than adults ("Stress in America", 2014). Eighty-three percent of teens surveyed cited school as a source of stress, sixty-nine percent cited, "getting into a good college or deciding what to do after high school," and sixty-five percent cited, "financial concerns for their family" ("Stress in America", 2014). The findings show that school and

excessive testing has negative consequences most importantly putting an unhealthy amount of stress on teenagers. Abeles highlights this sentiment, describing the emphasis on testing as out of control and calling today's students, "the most tested generation in history" (Abeles, 2015). These two results show that standardized testing comes with a high cost to students' wellbeing.

The other major impact of standardized testing comes from the financial burden it places on states across America. With the pressure from government programs such as No Child Left Behind, or Every Student Succeeds Act, the amount of spending by each state has seen dramatic increases over the past ten years. Matthew Chingos's study provides a picture of the significant amount of financial burden put on states, individual school districts, and families. It is shown in his report that in 2012 on average states spent an estimated $1.7 billion dollars on just testing alone (Chingos, 2012). To put this amount in perspective, the national education budget at the time was $77.4 billion. Spending on testing since 2012 has only increased while the education budget has decreased over the past seven years ("President's FY 2012 Budget Request", 2011). Combine that with the fact that families and school districts are spending thousands more on ACT prep classes and it becomes clear that ACT testing results in significant financial investment from society.

While this study is not expanding on the results from Simpson, the APA, and Chingos, their investigations provide context for this study. The results from these studies are not to say the ACT and other standardized tests should be done away with, but rather that we must weigh both the disadvantages of standardized testing along with the benefits.

Standardized Testing

A college education is a necessity. Having a college degree has been shown to result in higher income and being affected less by economic instability. Today's employees are required to have college degrees more and more. The ACT and other standardized tests are being used frequently as measures of academic achievement in terms of college readiness. In Diana Iborg's study, it was measured on a general scale of whether utilizing more standardized testing leads to increased college readiness on a national scale. It was concluded that there existed a relationship between increased testing and college enrollment (Iborg, 2014). Iborg's study sparked the need for more specific testing as seen in this study.

Iborg's study examined the effects of increased ACT testing on college enrollment and readiness for students graduating in 2010. State testing percentages were obtained from the 2010 ACT Profile Reports for each state in the sample. College enrollment information was obtained from the Digest of Education Statistics. College readiness was measured by the percentage of students who met or exceeded the ACT benchmark scores in English and Mathematics. However according to Zinth in the article "Defining college readiness" some potential drawback of using the ACT for measuring college readiness are that students are typically tested in 11th grade. This means that student's progress through their senior year without taking any rigorous classes, thus leaving them ill-prepared for an introductory level class (Zinth, 2012).

Results of Iborg's study revealed a relationship between increased testing and college enrollment, especially in Caucasian/White students, although the degree to which testing was

responsible for this increase was unclear. Strong relationships were found between increased testing and college readiness in both English and Mathematics. (Iborg, 2014).

A negative correlation was found between the percentage of students who took the ACT and the percentage of students who met benchmark scores (Iborg, 2014). This is interesting because it suggests that there may be some other underlying variables that lead to this correlation. The correlation coefficients were low indicating that the degree of correlation was such that no cause could be attributed (Iborg, 2014). Her findings are in line with the purpose of this study. A low correlation coefficient indicates a high amount of variability, which could indicate that there are other variables that attribute to the correlation, or rather that no cause can be attributed.

CHAPTER 3

METHODOLOGY

This study is intended to quantitatively determine the impact that higher ACT scores have on students' grades in introductory physics courses. This chapter discusses the design and methodology of the study including the source of the data and statistical analysis methods. Demographic data has been collected since Fall 2012 including the following categories: age band, major, academic level, class section, gender, race/ethnicity, SAT score, composite ACT score, high school GPA, NIU GPA, and final letter grade. Table 1 - Variable PropertiesTable 1 below indicates the properties of the variables analyzed.

*Table 1 - Variable Properties*

| Variable | Continuous/Discrete | Range |
|---|---|---|
| **Age Band** | Discrete | 17-50 |
| **Major** | Discrete | N/A |
| **Academic Level** | Discrete | Freshmen - Senior |
| **Class Section** | Discrete | 150 – 273 |
| **Gender** | Discrete | N/A |
| **Race/Ethnicity** | Discrete | N/A |
| **SAT score** | Continuous | 740 - 1120 |
| **Composite ACT Score** | Discrete | 13-35 |
| **High School GPA** | Continuous | 1.55 – 4.71 |
| **NIU GPA** | Continuous | 1.011 – 4.000 |
| **Final Letter Grade** | Discrete | F - A |

Description of Sample and Population

This study investigates student performance in one of five introductory physics courses. Northern Illinois University offers thirteen introductory physics classes. In this study students' grades and demographic data were collected from students enrolled in PHYS 150, 210, 211, 253, and 273. The data from each class was not always available for each semester. Table 2 below outlines which semester each section of class data was collected.

*Table 2 - Number of students per class per semester*

| | 2012 | 2013 | | 2014 | | 2015 | | 2016 | | 2017 |
|---|---|---|---|---|---|---|---|---|---|---|
| Class | Fall | Spring | Fall | Spring | Fall | Spring | Fall | Spring | Fall | Spring |
| 150 | | | | | | | | | 61 | 53 |
| 210 | | | | | | | 130 | 126 | 129 | 117 |
| 211 | | | | | | | | | 78 | 122 |
| 253 | 154 | 130 | 178 | 139 | 187 | 145 | | | 180 | 131 |
| 273 | 74 | 125 | | | | | | | 96 | 170 |

According to the undergraduate catalog PHYS 150, titled "Physics", is a class designed for students who have not taken physics in high school. PHYS 210 and 211 are titled "General Physics I" and II respectively. These two classes are typically taken in sequence and do not require calculus. Similarly, PHYS 253 and 273 are typically taken in sequence and are titled "Fundamentals of Physics I: Mechanics" and "Fundamentals of Physics II: Electromagnetism". These classes, however, do require calculus. Students enrolled in PHYS 253/273 typically go on to pursue engineering degrees.

These classes were chosen as they are often the first physics classes taken by Northern Illinois University students. The data was initially collected by the physics department to improve the curriculum of these introductory classes. These five classes are important as they are frequently the first classes taken by a student wishing to pursue a degree in STEM.

This study investigated a total of 4262 students across 52 individual sections, of which 24.5% were female, and 75.5% were male. The student sample was predominately white (57.7%), followed by Hispanic (13.2%) and then African American (12.1%). This study will not focus on the ethnic background of the students however it is given to establish the educational context. Table 3 below shows the breakdown for each class.

*Table 3 - Demographic data sorted by class*

| Class | Male | Female | White | Black | Hispanic | Asian | Other |
|-------|------|--------|-------|-------|----------|-------|-------|
| **150** | 76.5% | 23.5% | 42.0% | 24.4% | 21.0% | 9.2% | 3.4% |
| **210** | 50.6% | 49.2% | 58.6% | 13.0% | 10.8% | 8.3% | 9.3% |
| **211** | 46.8% | 53.2% | 59.0% | 10.7% | 15.6% | 11.7% | 3.0% |
| **253** | 83.6% | 16.4% | 59.1% | 10.6% | 13.0% | 12.1% | 5.2% |
| **273** | 83.8% | 16.2% | 59.2% | 7.3% | 15.0% | 13.2% | 5.3% |

To preserve the privacy of the individuals in the study, the specific academic performance of the students will not be included in this paper. In addition, all personally identifying information has been removed from the data. Student data was only examined in aggregate.

The student sample was then divided into groups in order to investigate whether a specific portion of the student sample was more subject to having their ACT score be an indicator of college success in physics. The sample study was grouped by academic level, gender, and by whether or not the class required calculus. The reasoning for these choices was based primarily on the researcher's experience as a teacher, and the generally accepted thought is that the smarter the student is, the harder they will "try" on the test resulting in a better representation of your potential success in college. This study hopes to collect more information on these correlations. The last variable analyzed was the high school GPA of the student before entering college and how that reflects on their success in college physics.

Statistical Analysis

Analysis and all graphical representations of the data were done with IBM SPSS (IBM, 2019); this was due to the large number of data points used in this study.

Student entrance ACT scores and class enrollment numbers will be summarized using histograms. Histograms will count the number of students to achieve a score, or how many are enrolled in a specific class within a range (bin), they will give the reader an opportunity to see the distribution of scores.

When comparing ACT entrance scores to the grade earned in the specific class, binned scatter plots will be used to represent the trend in that relationship. The darker the color on the dot, the more data points at that specific value. Along the *x*-axis of the plots represents the grade earned in the class organized to match the standard GPA scaling, where 4.0=A, 3.7=A-, 3.3=B+,

3.0=B, 2.7=B-, 2.3=C+, 2.0=C, 1.0=D, 0.0 = F.  It is important to note that NIU did not

implement the use of plusses or minuses in their grading until 2014. By scaling the pluses and

minuses according to common GPA scale, each of these +/- grades should still carry the same

weight.

One of the primary ways of comparing relationships in this study is the use of a

correlation coefficient. The correlation coefficient ($R$-Value) is a measure of the strength and

directions of a linear relationship between two variables. It is calculated by finding the sum

product of the standardized values of each of the two variables, divided by the degrees of

freedom.

$$R = \frac{\sum(Z_{yi}Z_{xi})}{(n-1)}$$

Where $Z_y$ and $Z_x$ are the standardized means for each variable and $n$ is the total number of data

points in the paired data set.

The null hypothesis for these comparisons will be that within the student sample any

observed trend between the two variables is due to random variation. A two-tailed significance

($p$-value) will be calculated as we do not know the direction of the trend and this method is more

conservative than the one-tailed $p$-value. If the $p$-value is less than 0.05 then the two variables

will be considered linearly related, thus rejecting the null hypothesis.  The standard 95%

confidence interval (CI) will be used for reporting testing as well, it shows the range where there

is a 95% chance that the 95% CI contains the true population mean (Devore, 2004). As it relates

to this study, $p$-value of less than 0.05 indicates that the observed trend is not likely to happen by

random chance and that there exists a statistically significant relationship between the ACT

composite score and class grade. Conversely, a large *p*-value suggests that changes in ACT score are not associated with changes in class grade.

Finally, when looking at the high school GPA, a multivariable linear regression was used. Through SPSS the coefficients for both HSGPA and Composite ACT score were compared with the grade earned in the class. The assumptions of using this linear regression are that the variables are continuous. When multiple groups need to be compared for statistical analysis, analysis of variance (ANOVA) will be used (Devore, 2004). This method of statistical analysis can be implemented when groups have a similar variance. The result of the ANOVA analysis is referred to as *F*-statistic, which is a measure of the variability in the scores between the conditions compared. The *F*-statistic will be reported with the degrees of freedom for the variable and uncertainty in parenthesis. *F*-statistic depends on the mean square of both the variance among the group ($MS_{Group}$), and the residual variance ($MS_{Residual}$).

$$F = \frac{MS_{Group}}{MS_{Residual}}$$

If the null hypothesis was true, the *F*-statistic should be close to 1. A high *F*-statistic indicates that you can reject the null hypothesis. In this study, a high *F* value would indicate that there exists a significant relationship between the ACT composite score and class grade.

CHAPTER 4

RESULTS

Descriptive Statistics

The study contains a total of 4262 students enrolled in the courses Physics 150, 211, 210, 253, 273 from the years 2012 to 2017 in either fall or spring semesters at NIU. Of the sample, 26.2 % are enrolled in physics classes that required calculus, and 73.8% are enrolled in classes that do not.

Of those students, only 3142 have their entrance ACT scores recorded. This is not uncommon, as community college graduates are not required to submit their ACT scores. In addition, a few students submitted SAT scores instead of their ACT. The students are then broken down by academic level. The sample contains 1356 Freshmen, 1147 Sophomores, 1135 Juniors, and 540 Seniors. Of the 4262 students, 1043 are females, 3218 are males and 1 is unknown. It should be noted that gender is not a binary classification, but this study is limited to the options of male and female in the available data.

Figure 1 below is a histogram showing the composite ACT scores of all the students enrolled in an introductory physics class.
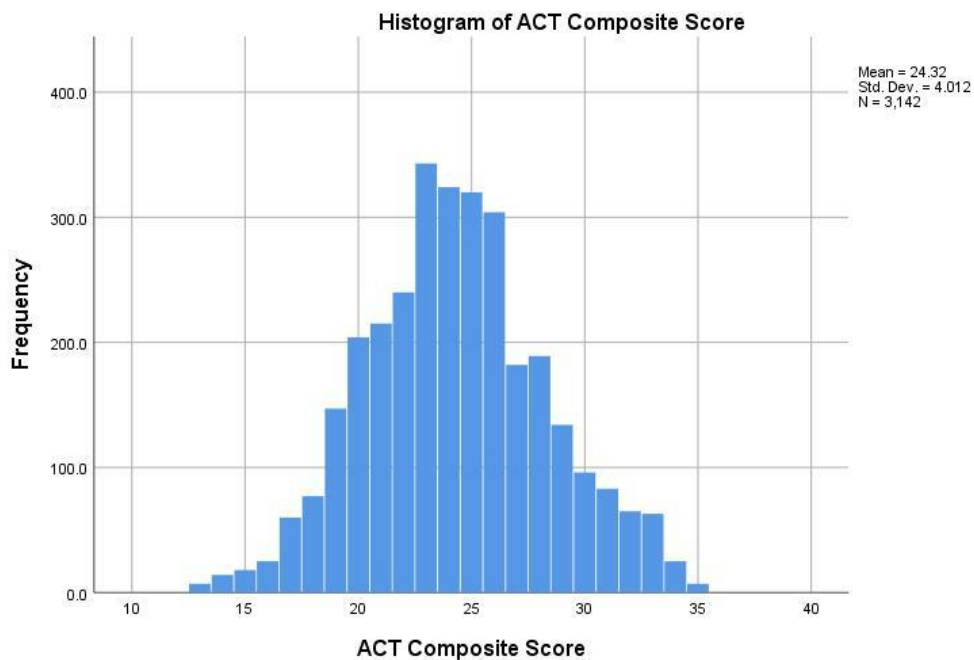
*Figure 1 - Histogram of composite ACT score.*

The data shows that there is a sharp decline in numbers of students with scores below 19, this is because students with those low scores are not typically admitted to NIU. This data is a great example of the strong emphasis that colleges put on standardized scores, according to National Center for Educational Statistics of the students who applied with an ACT score of less than 19, less than 25% of students were accepted ("Northern Illinois University Statistics", 2018).

Comparing this data to the grades earned in their physics course results in the scatterplot shown in Figure 2 below.
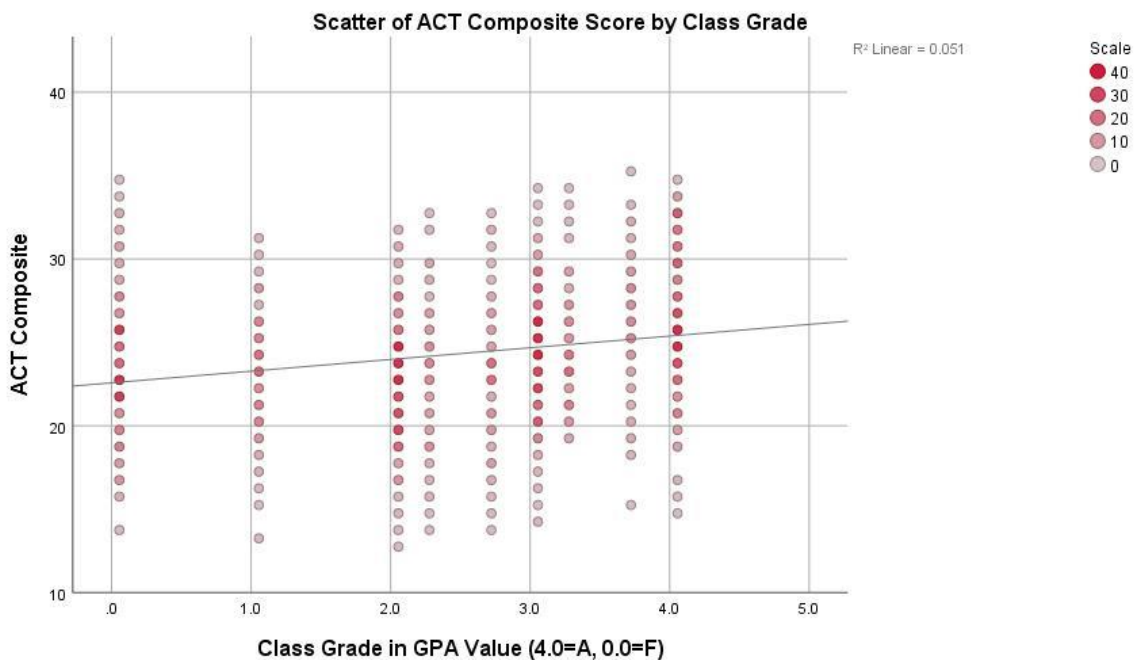
*Figure 2 - ACT Composite vs Grade in Class for All Students*

The data shows a statistically significant ($p < 0.001$) but weak (R = 0.225) correlation between the two values. This implies that while there is a valid trend, the amount of variation is significant. This trend indicates that there are other variables that might better explain this correlation. Therefore, the student sample was then further broken down for comparison between three major variables. The variables investigated in the study are Academic Level, Inclusion of calculus in the Curriculum, and Gender. By splitting the data, we are able to investigate this same relationship between ACT Composite Score and Class Grade in these smaller subgroups and determine if there are other variables that are weighted more heavily in this relationship, thus explaining the low *p*-value.

Inclusion of Calculus

The sample study consisted of students enrolled across five different classes; PHYS 150, 210, 211, 253 and 273. Figure 3 below is a histogram of the student enrollment by class. Each bar on the histogram is further split showing the academic level of students enrolled in each course.
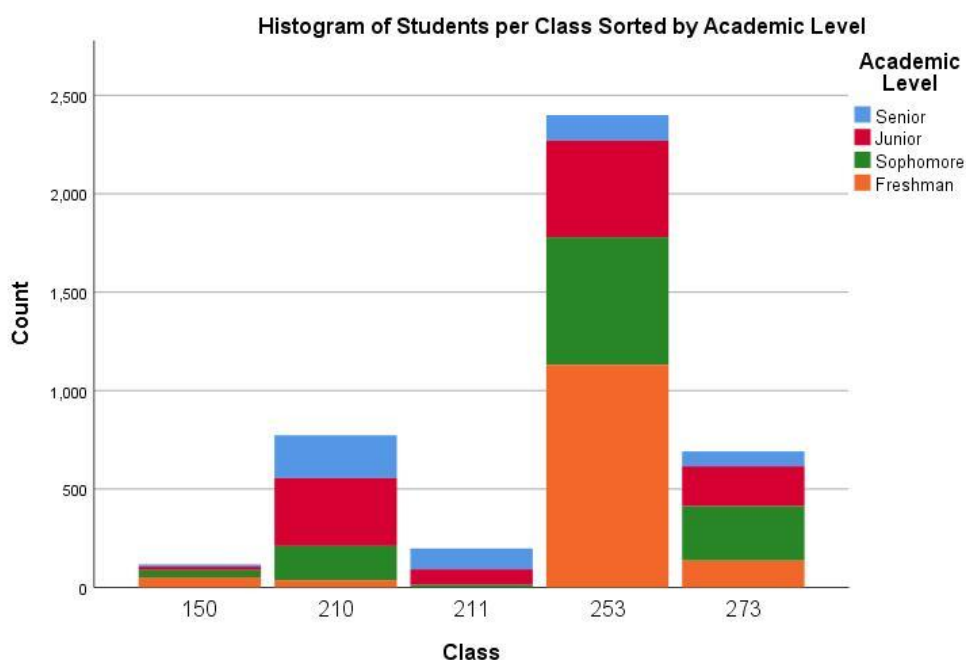


*Figure 3 - Histogram of Student Class*

Classes PHYS 150, PHYS 210 and PHYS 211 are structured in such a way as to not require calculus. According to the NIU undergraduate class catalog, PHYS 150 has no prerequisite classes and 210/211 require MATH 155 or equivalent. According to the undergraduate catalog at NIU, MATH 155 is titled "Trigonometry and Elementary Functions". It is a class reviews exponential and logarithmic functions, trigonometry, and complex numbers,

but not calculus. Therefore, these three classes may require trigonometric or algebraic background, but no calculus. PHYS 253 and PHYS 273 do have the prerequisite of MATH 229 and MATH 230, respectively. These two math classes are titled "Calculus I" and "Calculus II". Therefore, students enrolled in either of those classes must have taken some calculus course. The two groups are additionally distinct in that the students enrolled in PHYS 273 and 253 typically pursue a degree in engineering or physical science whereas students enrolled in the other three classes do not. The sample students were then split to investigate whether or not the inclusion of calculus in the curriculum affects the correlation seen above.

Figure 4 below is a scatter plot of student grade earned in class vs their entrance composite ACT score for the 3146 students enrolled in classes that do not require calculus (PHYS 150, PHYS 210, and PHYS 211).
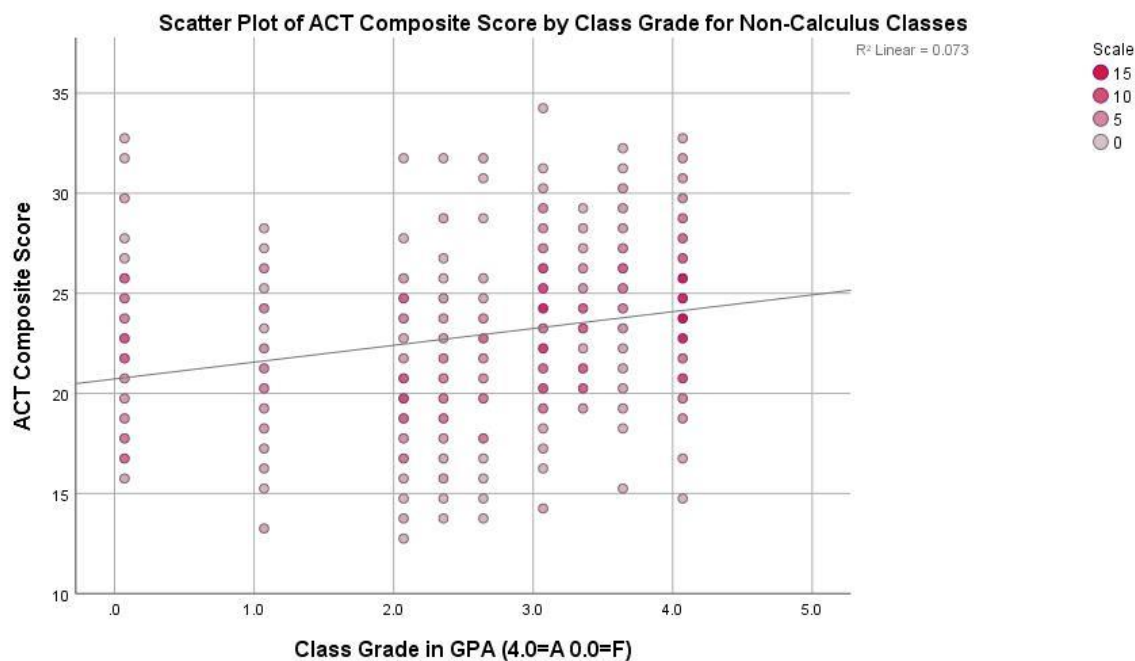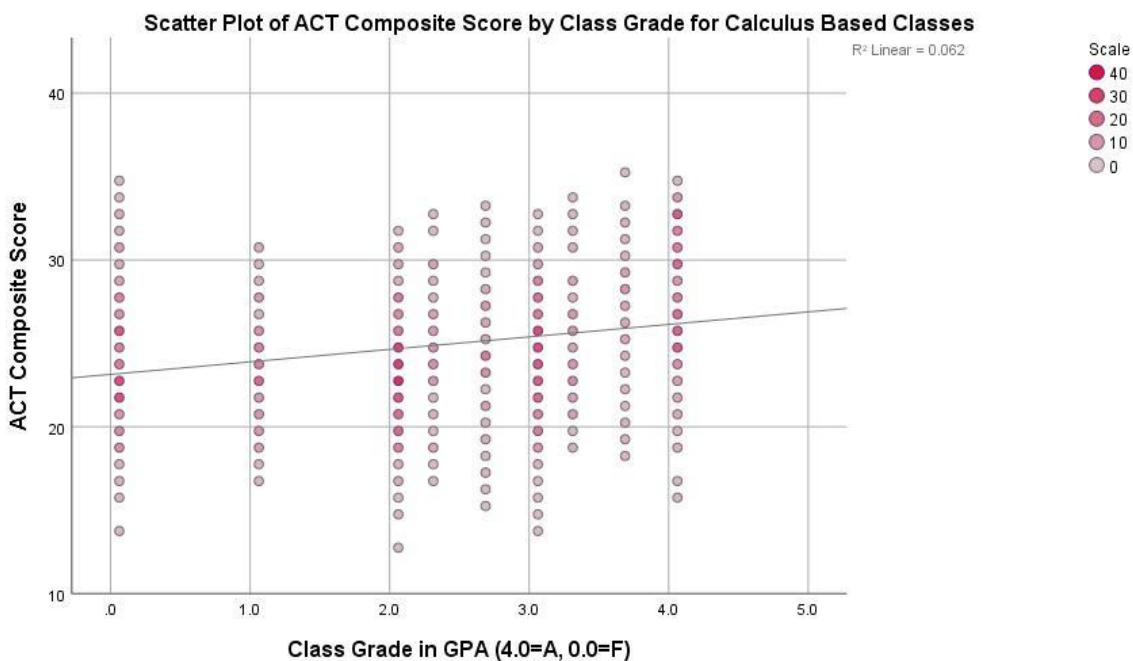


*Figure 4 - ACT Score vs Class Grade for non-calculus classes*

While there are significantly fewer data points than the total, there is a slightly stronger correlation between course grade and ACT score. The *p*-value was the previous result ($p <$ 0.001), and now the correlation constant is larger (R = 0.271). This indicates that students who are enrolled in a non-calculus introductory physics class at NIU have slightly lower variance in whether or not their ACT score will indicate their college success in that course.

Figure 5 below is a scatter plot of student grade earned in the class versus their entrance ACT score for the 1115 students enrolled in classes that require calculus.



*Figure 5 - ACT Score vs Class Grade for calculus based classes*

This subgroup had only half as many data points but still had a similar trend as the non-calculus group. The *p*-value for this group was still less than 0.001, and thus a statically significant trend is observed. However, the correlation coefficient had a similar value to the original data set (R=0.249). An interesting note is that both subgroups had a larger correlation

coefficient than the coefficient of the combined data. This phenomenon is known as Simpson's Paradox. As described by Blyth, "Simpson's paradox is a paradox in which a statistical trend appears to be present when data are segmented into separate groups of data but disappears (or reverses) when the data is considered as a whole." (Blyth, 1972). This indicates that there is a stronger relation between composite ACT score and class grade when looking at each group individually, however, when taken as a whole set of data, the correlation decreases.

## Academic Level

The next variable chosen to split the data set was the academic level. The expectation is that as students progressed further and further from high school the student's ACT score will have less of an effect, while other non-cognitive variables will become more important. However, the opposite trend was observed.

The student sample was then broken by academic level. The following is a histogram of the breakdown of the different levels in the student sample:
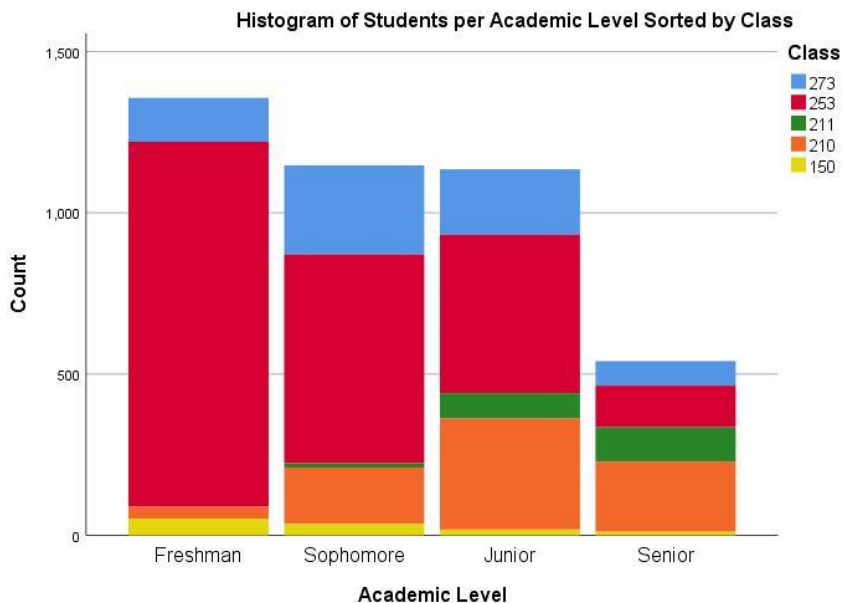
*Figure 6 - Histogram of student academic level*

The majority of the student sample are freshmen. This is because the classes were selected such that an entry level student, such as a freshman, would be interested in taking the course. However, it is not uncommon for other academic levels to also take these introductory courses.

The following four figures show the composite ACT score vs class grade for each of the four academic levels in the respective introductory class.
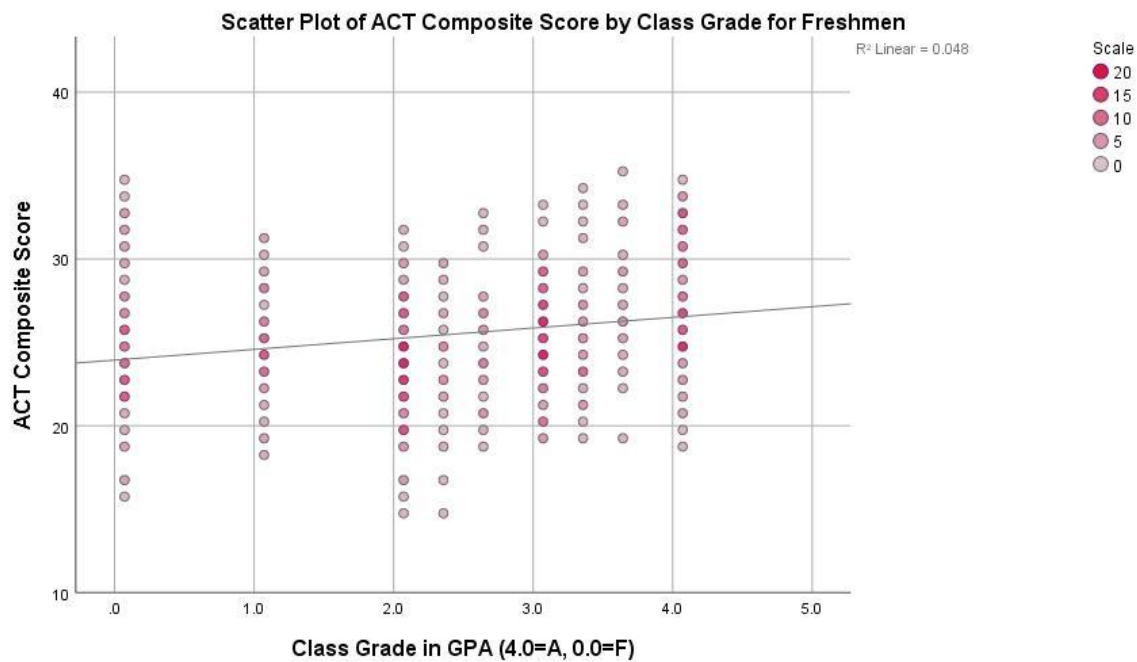
*Figure 7 - ACT score vs class grade for freshmen*
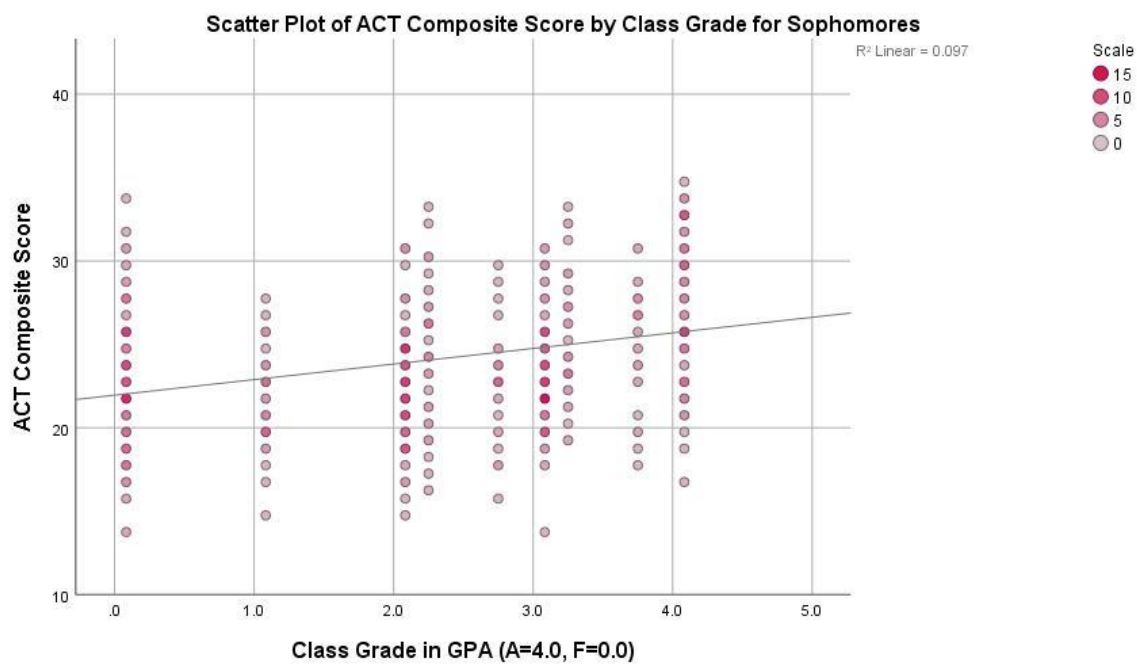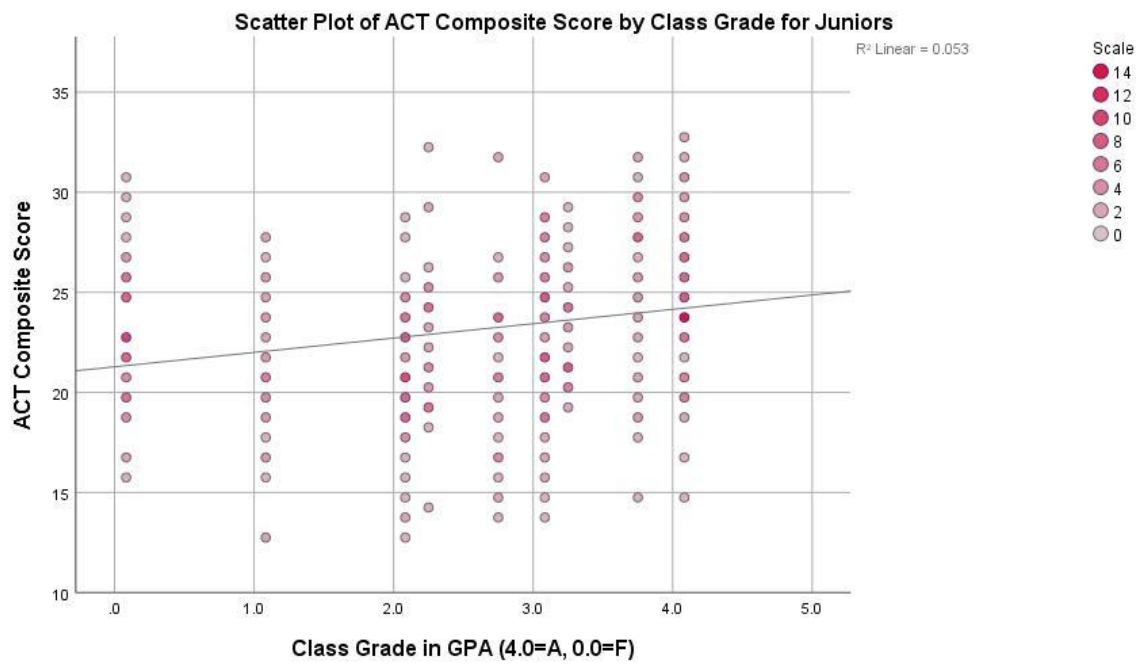


*Figure 8 - ACT score vs class grade for sophomores*
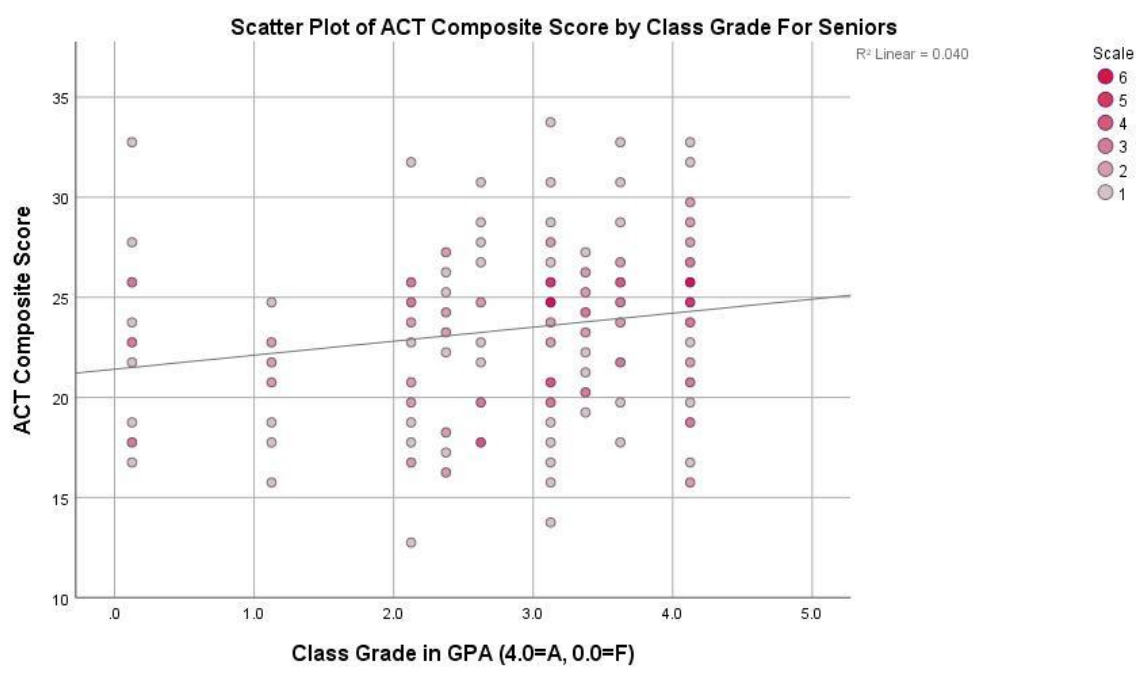
*Figure 9 - ACT score vs grade for juniors*



*Figure 10 - ACT Score vs class grade for seniors*

*Table 4 - Statistical breakdown for each academic level*

| Academic Level | $p$-Value | R - Value |
|---|---|---|
| Freshmen | $p < 0.001$ | 0.219 |
| Sophomores | $p < 0.001$ | 0.311 |
| Juniors | $p < 0.001$ | 0.230 |
| Seniors | $p < 0.001$ | 0.200 |

Table 4 - Statistical breakdown for each academic level above shows a summary of the statistical breakdown for each academic level. Much like the rest of the data, all of these subgroups had a significant trend between the ACT composite score and grade earned in the class. However, sophomores had a higher correlation constant.

Gender

The sample was split into gender. Figure 11 below is a histogram showing the numeric breakdown of the student sample into gender:
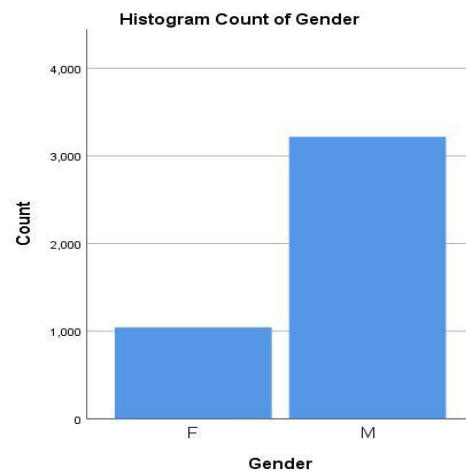


*Figure 11 - Histogram of student breakdown by gender*

While NIU enrollment is nearly equally distributed, 50% male and 50% female ("Northern Illinois University Statistics"*,* 2018), there were over three times more males enrolled across the five introductory physics courses. This is a phenomenon that is seen across many collegiate institutions. This study will not focus on why this imbalance exists, but rather if it leads to a stronger correlation between ACT scores and college success.

Figure 12 below shows the score distribution for the 1043 females enrolled in introductory physics courses.
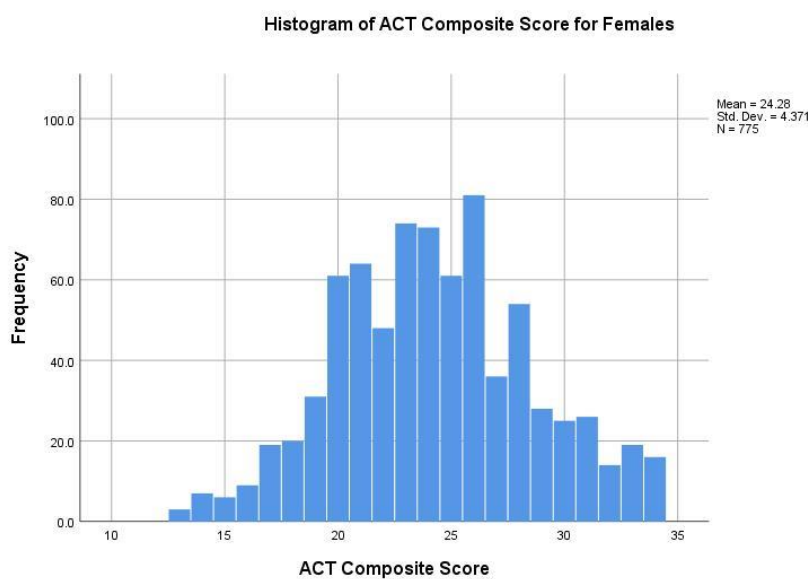


Histogram of ACT Composite Score for Females

Mean = 24.28
Std. Dev. = 4.371
N = 775

*Figure 12- ACT score distribution for females*

Figure 13 below shows the score distribution for the 3218 males enrolled in introductory physics courses.
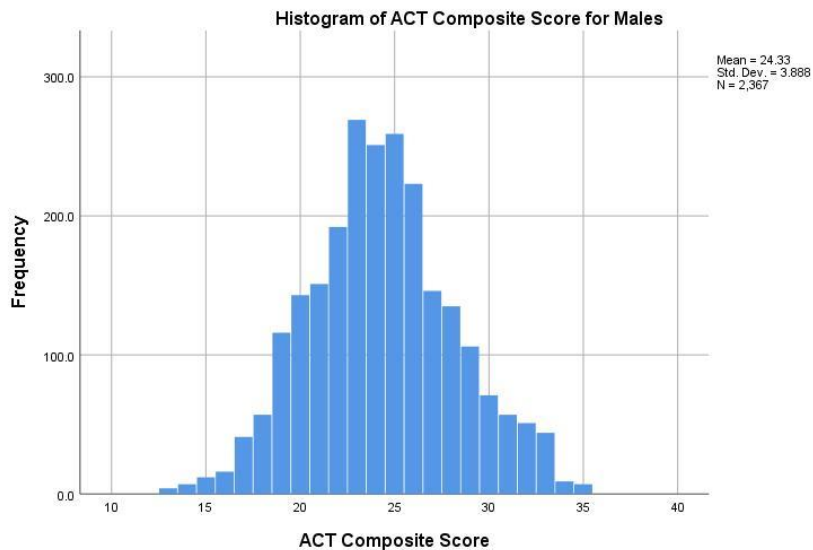
*Figure 13 - ACT score distribution for males*

Figures Figure 14 and Figure 15 below shows the relationship between ACT score and final class grade for each gender.
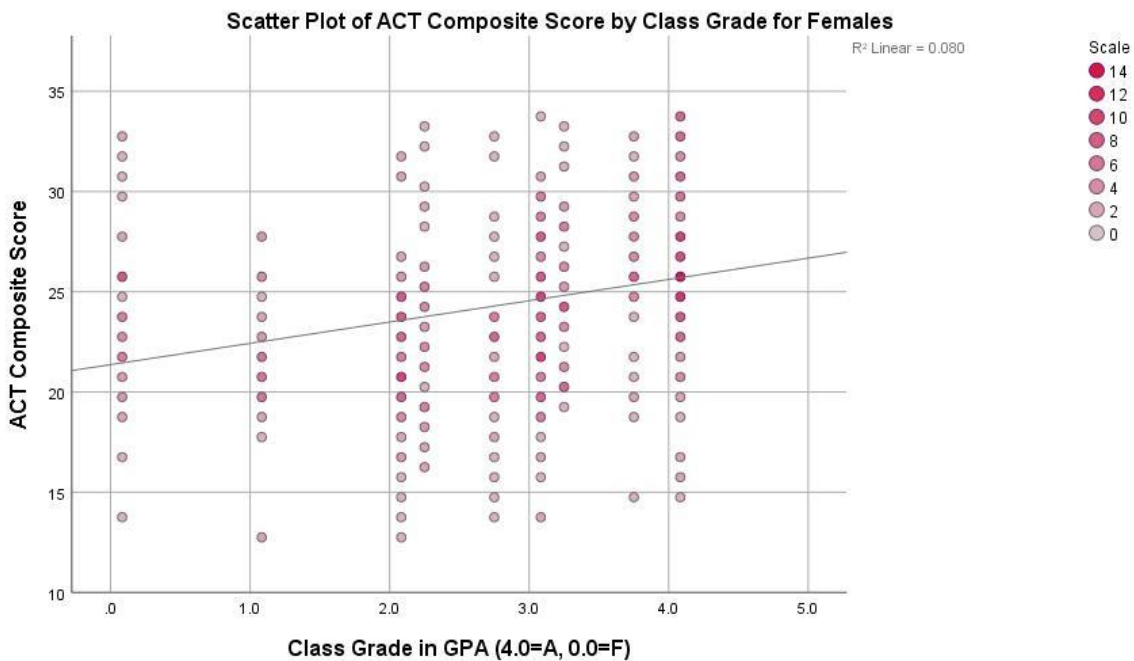


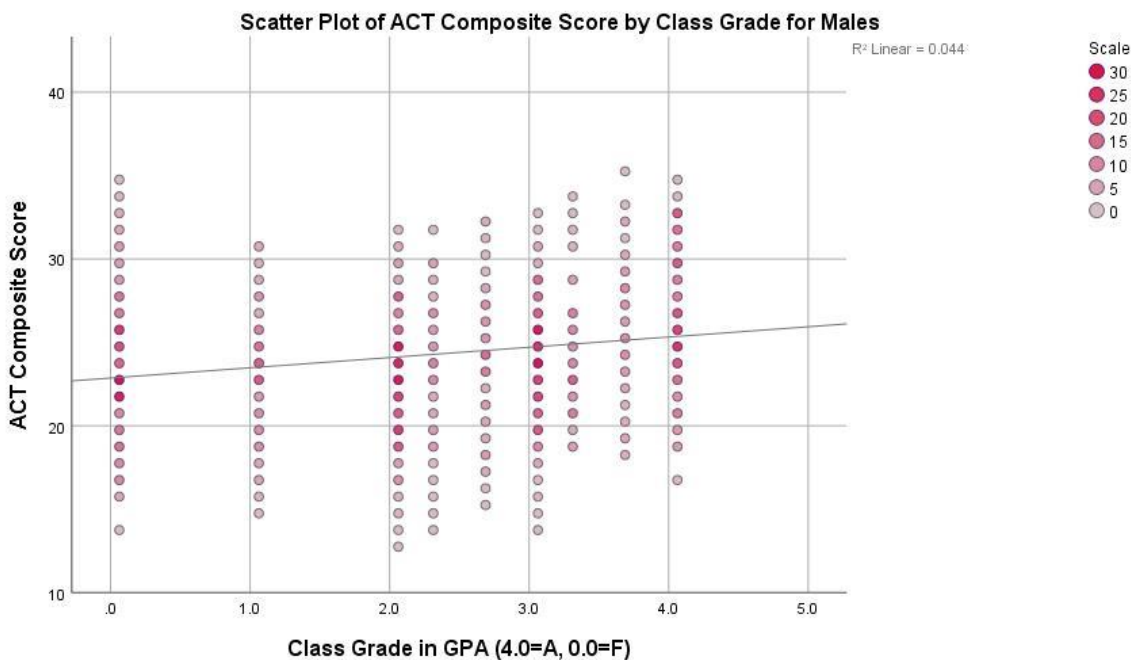*Figure 14 - ACT score vs grade for females*

*Figure 15 - ACT vs class grade for males*

Again, both genders show a similar trend of high significance ($p < 0.001$), but equally high variability ($R_{females} = 0.283$, $R_{males} = 0.210$). The two graphs shown above imply that while the ACT might be a better predictor for success for females, both genders have significant variation within this data.

## High School GPA

The last variable analyzed was the high school GPA of the student before entering college and how that reflects on their success in college physics. Recent trends have college admissions relying less and less on standardized testing and more on high school experiences and GPA (Furuta, 2017). In this study, high school GPA data was only collected from students enrolled in their physics class from fall 2016 and on. Therefore, the number of data points for high school GPA is significantly less than the rest of this study.

The figure below shows a histogram of the sample student's GPA in high school. Only 1086 students had their GPA value available.
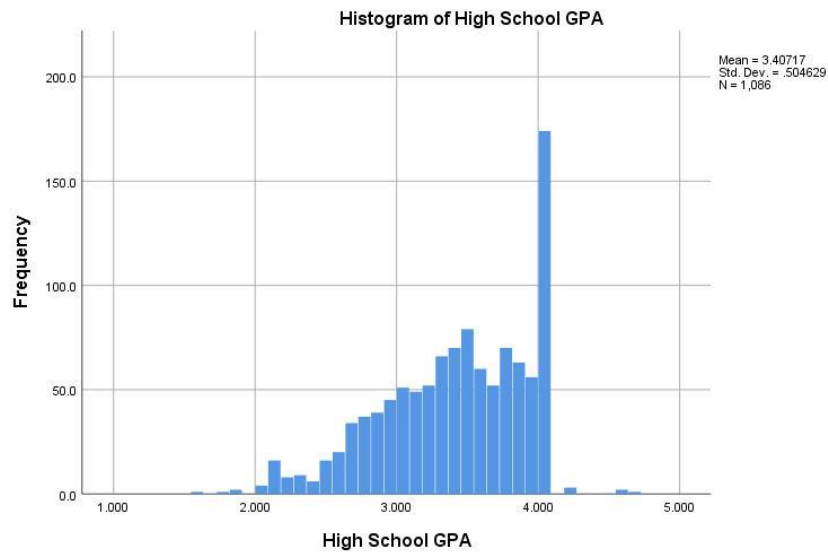


*Figure 16 - Histogram of student high school GPA*

There is a large spike at the 4.0 GPA because 16% percent of students admitted to NIU have a GPA of 4.0 or higher ("Northern Illinois University Statistics", 2018). As these are weighted GPAs there are also a few students who had GPAs of over 4.0.

The figure below represents this GPA data when plotted against class grade.
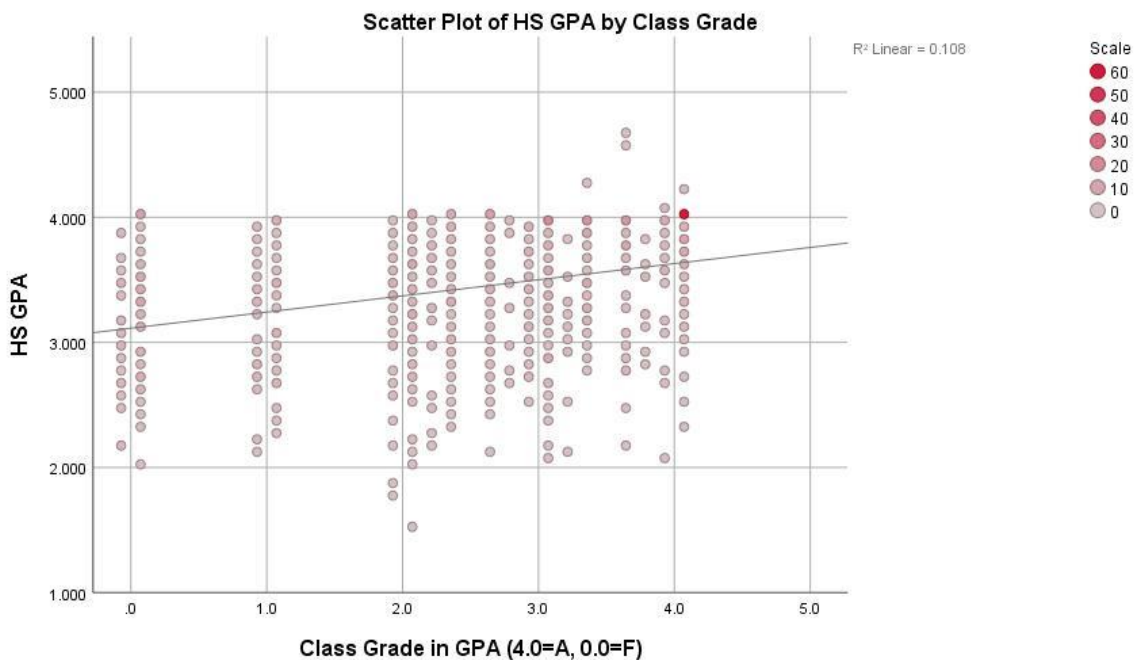
*Figure 17- High school GPA vs grade*

The *p*-value was $< 0.001$, and the correlation constant $R = 0.329$. This correlation here is significantly higher than that of the ACT Composite Score. This indicates that on average GPA has a higher correlation with a higher grade in introductory physics courses, than ACT score. However, similar to the previous data, there were students with low high school GPA scores that earned high grades and vice versa.

To further compare composite ACT score to high school GPA a multiple regression was performed to predict class grade from ACT score and high school GPA. These variables statistically significantly predicted class grade, $F(2, 854) = 54.825$, $p < 0.001$, $R = 0.337$. Both variables added statistically significantly to the prediction, $p < 0.05$. However, ACT composite score had a *p*-value of 0.029, whereas GPA was had a *p*- value less than 0.001. Even with the

combination of all three additional variables, the linear *R*-value is still well below an acceptable amount.

ACT Math Subscore

In addition, the subscores of the ACT were also recorded, however, the subscore that was believed to have the most correlation with student's grades in physics actually had a significantly lower subscore. Figure 18 below is a histogram showing the ACT Math subscore distribution for the 1256 students enrolled in the five introductory courses.
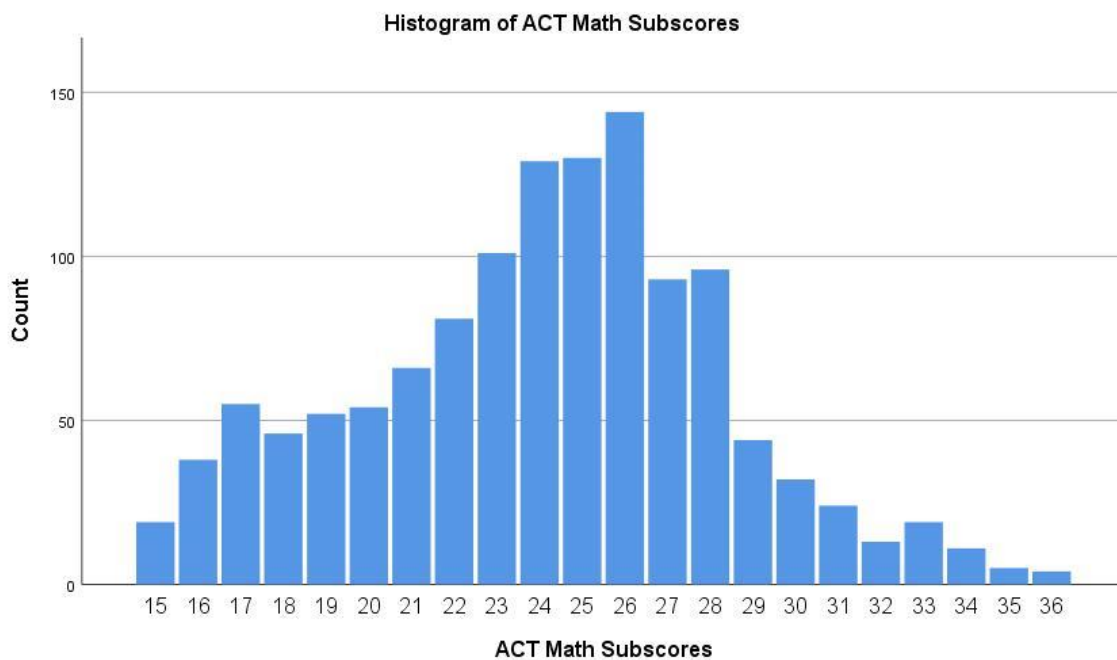


*Figure 18 - Histogram of ACT math subscore*

When comparing the ACT Math Subscore to grade earned in these introductory classes the correlation coefficient was significantly lower than that of the composite score. below shows the ACT Math subscore when plotted against class grade.
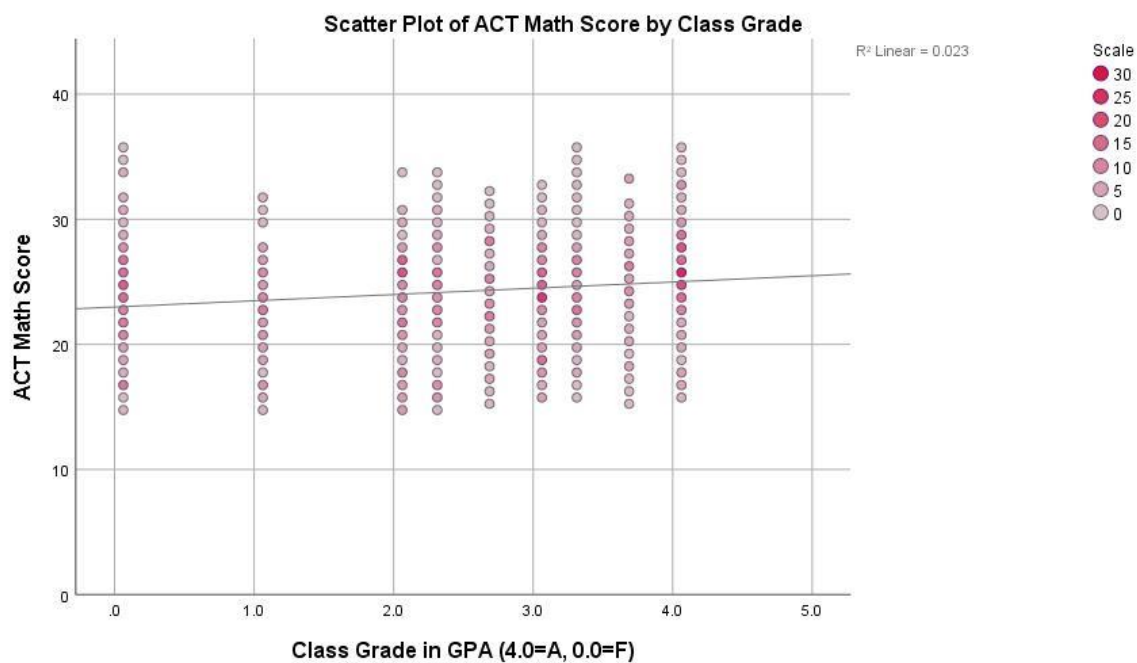


*Figure 19 - ACT math subscore vs class grade*

The *p*-value for this correlation was less than 0.001. However, the correlation coefficient was much lower (R = 0.151). This data implies that the composite ACT score is less of an indicator than the ACT math subscore for anticipating success in college introductory physics.

..

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

Summary of Findings

This study investigated the impact that ACT scores have on determining student success. The study used statistical analysis of ACT scores correlated with student success measured in class grade. It is important to note that, while the *p*-value was low for the correlations between composite ACT score and grade earned, in this study, the data has a large amount of variance, with students who received 14's on their ACT earning A's and students with near perfect scores on the ACT earning F's. The following conclusions based on the evidence presented for the student sample can be taken from this study:

1. There is evidence that a relationship between ACT Score and class grade exists, however, there is a significant amount of variability in the trend.

2. Even when adding multiple variables to the analysis, the correlation coefficient does not reach a value that would show a significant correlation. Thus the data contains an inherently higher amount of unexplained variability.

These results are far from conclusive and definitely require further study. With the correlation coefficients so small, it is hard to say conclusively either way. The amount of variance seen in the data is too much warrant the significant emotional and financial pressure put on students and families. The findings in this study explain the trend observed by Jared Furuta's

study about colleges transitioning to "Test-Optional Admissions". Especially in light of the 2019

testing scandal, more colleges are transitioning to not requiring SAT or ACT tests in the

admission process. This trend is consistent with the low correlation coefficients found in this

study.

The following recommendations for further research are recommended based on the

findings of this study:

1. Conduct a study collecting standardized test data other than the ACT (SAT, PARCC,

   etc.). This will give a broader view

2. Conduct a study that collected more non-cognitive variables specific to each student

   (amount of sleep, study habits, etc.). The purpose of this would be to investigate if other

   variables not explored in this study can explain the variability.

3. Collect High School GPA data for more than two years. Requesting GPA data for all five

   years of the study would make it comparable with the rest of the data.

4. Only track classes that have the same professor such that you can ensure that each class

   and section is being taught as close to identically as possible.

The following recommendations for improving practice are recommended based on the

findings of this study:

1. Conduct a survey of each professor for each section, and categorize each professor's

   teaching practice. Instructor effect cannot be eliminated entirely, at least the impact on

   grades can be documented and incorporated into the analysis.

2. Collect data for non-introductory courses. While this study focused on where the ACT would have the greatest predictive impact, it would be good to also see how it has on overall success in college, rather than just introductory courses.

3. Investigate how ACT score relates to graduation rates. This would be beneficial to the findings of this study because it would be an additional measure of college success and highlight other non-cognitive variables that influence students' success.

This study provides insight into how best to predict collegiate success in physics for high school students. The recommendations are based on the findings of this study and are offered to improve college selection of students based on potential success in physics.

REFERENCES

Abeles, V. (2015). *Beyond measure: Rescuing an overscheduled, overtested, underestimated generation.* (Rep.). Simon & Schuster.

Blyth, C. R. (1972). On Simpson's Paradox and the Sure-Thing Principle. *American Statistics Association, 67*, 364-366.

Chingos, M. M. (2012). *Strength in Numbers: State Spending on K-12 Assessment Systems* (Rep.). The Brookings Institution.

Devore, J. L. (2004). *Probability and statistics for engineering and the sciences*. Thomson-Brooks/Cole.

Furuta, J. (2017). *Sociology of Education* (Vol. 90, pp. 236-254) (United States, US Department of Health and Human Services). SAGE Publications.

Goldman, R. D., & Widawski, M. H. (1976). A Within-Subjects Technique for Comparing College Grading Standards: Implications in the Validity of the Evaluation of College Achievement. *Educational and Psychological Measurement, 36*(2), 381-390. doi:10.1177/001316447603600217

How Physics is Tested on the ACT Science Test. (2018, June 29). Retrieved from https://www.kaptest.com/study/act/how-physics-is-tested-on-the-act-science-test/

Iborg, D. H. (2014). *Examining the Effects of ACT Assessment of High School Graduates on College Enrollment and College Readiness* (Doctoral dissertation, Lindenwood University, 2014). UMI.

Northern Illinois University Statistics. (2018). Retrieved from https://nces.ed.gov/collegenavigator/?id=147703

President's FY 2012 Budget Request for the U.S. Department of Education. (2012, February 13). Retrieved from https://www2.ed.gov/about/overview/budget/budget12/index.html

Simpson, C. (2016). *Effects of Standardized Testing on Students' Well-Being* (Unpublished master's thesis). Harvard Graduate School of Education.

SPSS Software. (n.d.). Retrieved from https://www.ibm.com/analytics/spss-statistics-software

Stress in America. Are teens adopting adults' stress habits? (2014). Retrieved from http://www.apa.org/news/press/releases/stress/2013/stress-report.pdf

Zinth, J. (2012). Defining college readiness: Where are we now, and where do we need to be? The progress of education reform. *Education Commission of the States, 13*(2), 1-7.