Spring 2021

Thesis Is STEM Education Equal? How School and Student Characteristics Impact Student STEM Outcomes

Rafael Villa
rafael.villa@trincoll.edu

Trinity College

Is STEM Education Equal? How School and Student Characteristics Impact Student STEM Outcomes.

Senior Thesis

Rafael Villa

Educational Studies 400: Senior Research Seminar

Professor Daniel Douglas

22 December 2020

Abstract

There is a large disparity in current STEM education across racial and socio-economic status. The result is detrimental to these student populations, especially since STEM study and occupations are often used as an agent of economic mobility. So how do student and school characteristics in high school affect student outcomes in STEM? Using data from the HSLS:09, I used logistical regression analysis and cross tabs to find that student academic achievement and student SES were the biggest factor when predicting student performance and perception odds in STEM. Race and gender were also statistically significant characteristics when predicating student perception and performance outcomes. Surprisingly, school characteristics were not predictive of student performance and perception outcomes. The results show that mathematics education is not meritocratic, and schools need to do more to better foster student interest and success in STEM post-secondary for vulnerable student populations.

Introduction

High school educational experiences often make or break student success post-secondary. It is no secret that students from better schools have a greater chance at post-secondary success. Students from lower income and urban schools often do not have adequate resources and course rigor to do well or have the confidence to succeed post-secondary. In STEM education particularly, there is an idea that if students work hard, they will succeed, regardless of the school they go to or their race and socioeconomic status. STEM education is a catalyst for social and economic mobility for many lower income and students of color. But is the quality of their education the same as those from nonurban schools?

In my undergraduate experience, I noticed right away that some students were better prepared for college level mathematics courses than I and my peers were. Even though I took algebra II in high school, I struggled to barely get a C+ in the class and lost lots of sleep over the course. The stress was almost unbearable. I wondered why other students from suburban and boarding schools seemed to barely study for the course and do well? Was it me or was it mathematics being designed against me? At the time I strongly believed in the idea of meritocracy, since I did well in high school level mathematics. However, I pondered this a lot as I continued my degree in mathematics here at Trinity. I noticed the majority of my classes were either white and/or international students. I am the only Latinx senior math major, but I know I am not the only one who was interested in STEM degree from the get-go. I remember a lot of my peers, other students of color, expressing interest and a desire to pursue a degree in a STEM major, particularly in engineering. Many of them took calculus in high school as well, yet a lot of them ended up changing majors after Calculus I at Trinity. I wondered what deterred them and not other students.

Obviously, something is gate keeping students of color from wanting to continue majoring in STEM. At one point I even felt like changing my major, but I had a good Calculus II professor that convinced me I had what it took to stick with the program, and a good academic advisor as well. However, it is still an issue that so many students feel that they have to change their path because they figure STEM is not for them. The fact is that the major is overwhelmingly international and white students. It is not just mathematics, but other hard sciences, and engineering as well. Students of color and from lower income backgrounds feel discouraged to continue. It is important that students from all backgrounds feel they can succeed post-secondary in STEM. Students of color having limited access to STEM and being discouraged from continuing in STEM post-secondary is an issue of equity and is reflective that not enough is being done in their secondary education to prepare them for college level STEM course or to foster a sense of confidence in STEM. In general education is not fostering the success students of color post-secondary, and this is particular detrimental as STEM is often an opening into economic mobility for lower income and students of color.

This then raises the question of how do student and school characteristics in Highschool affect student outcomes in STEM? Throughout this paper, I will be analyzing this question. I begin this paper by referencing what current literature has to say about student STEM self-perception, racial barriers in STEM education and why students choose degrees in STEM. To answer this, I will be looking at data from the National center of Educational statistics. The study in question is the HSLS:09 which is a longitudinal study following high schoolers from 2009 onward to post-secondary. It is a relatively new source of data, so not much research has been done on it. After this, I will analyze the data with base level descriptive statistics and logistical regression analysis. I hypothesize that high school characteristics play a large role in student success and interest in STEM. Specifically, I believe school urbanicity and teacher certification play a large role in preparing students adequately for STEM post-secondary, both in student STEM perception and student performance.

Literature Review

Current literature on Mathematics education fails to holistically capture how high school characteristics impact student STEM outcomes. While much of the research focuses on how self-concept and self-efficacy influence student performance in Math and Science courses, others have highlighted the racial disparities in STEM course taking and STEM course offerings in schools. Additionally, there is research that highlights why students choose STEM majors and how their high school experiences impact this decision. Much of this past research hones in on specific school factors or students' demographic characteristics but does not tie these factors together when analyzing what impacts student self-concept and performance STEM outcomes. However, I will be considering specific high school and student characteristics such as race, gender, social economic status, and school location and its impact on student perception outcomes and performance STEM outcomes.

Math and Science Self Concept and Self Efficacy

Yeung and Marsh (1997) researched "the relations between academic self-concept and academic achievement, and more specifically, whether changes in academic self-concept lead to changes in academic achievement" (Marsh and Yeung, 41). To answer their research question, they used a sample of 605 catholic schoolboys from mostly working and middle class (though there were some upper-class students as well) from grades 7 to 10. Data collection took place over the course of three years, making the study longitudinal. Using regression analysis, they found "prior academic achievement affects subsequent academic self-concept." (Marsh and Yeung, 49). This was shown to be true despite what academic subject the model referred to. Interestingly enough, they also find that prior self-concept also influences academic achievement in the subjects analyzed (Marsh and Yeung, 50). This research shows that there is a strong relationship with the way students view themselves and how they perform in the classroom. However, Marsh and Yeung fail to consider how the race, ethnic background and gender can have an effect on their self-concept.

Parker et al (2014) also research how self-concept influences student mathematics achievement. However, in addition, they consider self-efficacy, a student's perceived ability to succeed academically, as a factor in student achievement. They state, "little research has juxtaposed these self-belief constructs as predictors of academic achievement and achievement related outcomes." (Parker et al, 30). The goal of their research was to compare how student self-concept and self-efficacy impact student achievement are related to each other. The research utilizes a longitudinal study of Australian young people (LSAY), which collected data on science, math and reading achievement tests, mathematics, self-concept and mathematics self-efficacy, to answer this question (Parker et al , 35). The data was analyzed regression analysis and controlled for the covariates: parental education and socio-economic status, year in school (hereafter grade), gender, and immigrant and indigenous status. Outcomes were measured by student enrollment in STEM courses, university enrollment and standardized test scores. They found that "achievement was strongly related to both self-beliefs and that controlling for prior achievement and a host of covariates, self-concept and self-efficacy were significant predictors of TER. Further, self-efficacy was found to be a significant predictor of university entry, and self-concept was a significant predictor of STEM course selection." (Parker et al, 42). Their results suggest that if students view themselves strong in a subject, it can influence if they pursue courses in that subject. However, these findings fail to take into account race and school urbanicity into account. The study is based in Australia, so it may not be so generalizable to the United states education system. However, the findings are still strong enough to show that there is a positive correlation among student self-concept and self-efficacy and success in STEM.

Moakler and Mikyong (2014) also delve into how math self-concept and self-efficacy impact choice of STEM major choice. The text "investigated confidence and demographic factors associated with the choice of a STEM major" (Moakler and Mikyong, 129). Specifically, the text focused on these questions:

1. How do background factors such as gender, minority status, parental socioeconomic status (SES), parents with STEM occupations, and academic preparation affect a STEM major choice?

2. How does academic confidence affect a STEM major choice?

3. How does mathematics confidence affect a STEM major choice?

Unlike other research, they take into account Parental education background and socioeconomic status well as data from the Cooperative Institutional Research Program (CIRP) of the Higher Education Research Institute (HERI) at the University of California, Los Angeles. The data was analyzed using logistical regression models while controlling for different variables to predict what factors influenced STEM major choice of students of different filter variables. The regression model predicted STEM major choice based on student confidence in STEM in high school. The study found “several positive indicators of STEM major choice: having parents with a STEM occupation, having higher SAT scores, having a higher high school GPA, having spent more hours studying or doing homework, being a minority (African American or Latina/o), having higher academic confidence, and having higher mathematics confidence.” (Moakler, 138). This research has great implications on how family background and student self-concept impacts their decisions in STEM, however it does not take into account how high school characteristics impact this choice. In addition, it does not go in depth in the role of race involved in STEM but does make mention of race as a factor. The main findings centralize on parent occupation and student scores.

<u>Racial Disparities in STEM Course Taking</u>

Kelly (2009) brings race to the forefront of their research and use data from the National Education Longitudinal Study of 1988 (NELS:88) to answer the following questions:

1. To what extent can differences in course taking among black and white students be attributed to differences in academic achievement or other factors that are associated with individual students, such as family background?

2. To what extent can the lower levels of academic course taking in mathematics among black students be explained by course-enrollment patterns at the schools that black students attend?

3. To what extent can lower levels of course taking be attributed to a contextual effect within integrated schools, whereby black students are disadvantaged in predominantly white schools?

4. Do inequalities in black-white course taking vary across school sectors?

Using multivariate regression analysis, they found that “the black-white gap in mathematics course taking is the greatest within integrated schools where black students are in the minority and cannot be entirely accounted for by individual-level differences in the course-taking qualifications or family backgrounds of white and black students.” (Kelly, 47). They also note that “there appears to be a connection between the racial composition of a school and the chances of black and white students enrolling in high-track mathematics courses.” (Kelly, 61). This is interesting as they note the racial disparities in an integrated high school setting, as well as segregated schoolsettings. This research supports mine as it is also analyzing a longitudinal study to make inferences on student STEM outcomes, particularly across racial lines.

Cogner et al (2009) also aimed to carefully examine different explanations for demographic disparities in advanced course-taking. The data used in the text was from high school student cohorts in Florida public schools and analyzed with regression analysis. They find that, when controlling for “pre high school characteristics”, the gap of which students take advance courses disappears. Meaning that the quality of education students receive before high school largely impacts the level of courses they will take in high school. In my research, I am analyzing how high school characteristics and student demographics and self-concept impact student STEM outcomes. This research shows that, school characteristics prior to high school impact student STEM courses enrollment in high schools and diminishes the gap across racial lines. The research focuses on data Florida high schools, which may not be as generalizable as the HSLS, which I am using. In addition, the study fails to take into account how these factors play a role in student STEM outcomes post high school. Nevertheless, the implications of this research show that more research needs to be done on the effects of school characteristics and race impact student STEM outcomes

Battery (2013) takes a different approach to analyzing racial disparities in mathematics education. They state “Oftentimes mathematics gets framed as neutral subject matter, devoid of culture, feelings, and based on a system of meritocracy. However, this field of mathematics education has a long history of giving access to students differentially, particularly based on the ideological construction of race.” (Battery, 332). Referencing prior research on tracking and color-blind ideology, they claim “that the mathematics curriculum, as an institution, is not neutral, but functions along the dominant racial ideologies in society.” (Battery, 333). Battery draws data from the “High School and Beyond (HSB) 1980 (and follow-ups), National Education Longitudinal Study (NELS) 1988 (and follow-ups), Education Longitudinal Study 2002 (and follow-ups), and Current Population Survey (CPS) 1972–2005.” (Battery, 342). Using logistical models they find that “the earning differentials attributable to mathematics education are appalling…” (Battery, 350). He goes on to elaborate that mathematics being treated neutrally consequently acts a gatekeeper trough race. This gatekeeping of quality mathematics education prevents students of color from gaining adequate background necessary to succeed or gain credit for college level mathematics. These findings suggest that mathematics education be viewed through a critical race lenses to ensure that students are not continuously stratified under the guise of meritocracy and test scores. While this research does not make implications on why students choose STEM or not, it does frame how race is a critical factor in what students are offered in STEM education and how that is valued by colleges and universities.

Why Students Choose STEM

Current research analyzing the HSLS09 brings a lot of the factors discussed earlier together. Saw (2018) investigates "the cross-sectional and longitudinal disparities in STEM career aspirations at the intersection of gender, race/ethnicity, and SES using the nationally representative High School Longitudinal Study of 2009 (HSLS:09)" (Saw, 525). In their analysis of the HSLS09, they reinforce the fact that students of color, girls and students with lower socioeconomic status are disinterested in pursuing and less represented in STEM careers and education. Although their study does not focus on hard STEM outcomes, the results show that more interest in STEM needs to be fostered in schools with marginalized students of color and lower socioeconomic status and that race/ethnicity, gender and socioeconomic status do in fact intersect with STEM aspirations. I will be building off this research to make implications on how, in addition to race/ethnicity, gender and socioeconomic status, school locale and student performance impact student interest in STEM and students actually pursuing degrees in STEM.

Data and Methods

In this study, I analyzed data from the Hight School Longitudinal Study of 2009 (HSLS09). The HSLS09 is a nationwide longitudinal survey that followed around 25 thousand high schoolers across the country in their perceptions of and assessment in mathematics. It also collected demographic data and detailed data on the schools that students attended. The initial base year survey was followed up with a second survey during the senior year, and then a third post-secondary. Student's parents, guidance counselors, teachers and school administrators were also surveyed. Overall, the study holistically captures data from high schools nationwide pertaining to student STEM experiences and performance.

For my research, I used a program called PowerStats on the National Center for Educational Statistics' website. PowerStats has the capability to create various tables and run some regression analyses based of the large list of variables from different Education studies. The HSLS09 study has a large list of variables available through PowerStats. The selected variables help me to answer the question: How do school characteristics and factors in Highschool affect student outcomes in STEM(Math) when analyzing survey data? Through PowerStats, I used percentage tables to create crosstabs. I also used the logistical regression tables to create predictive models of student STEM Outcomes. Much of the research I have referenced uses logistic regression analysis in the HSLS09, and other similar longitudinal studies, in order to make their implications on Student outcomes in STEM. In the HSLS09, there are two different definitions of STEM, the NSF and the SMART definition. For consistency sake, I use the NSF definition throughout the data analysis.

Despite Powerstats having access to the HSLS09 data, there are limitations to the extent of data available to analyze. Most of the variables I am analyzing are categorical variables, with socio-economic status being an exception to this, which means I had to utilize logistical regression analysis instead. Powerstats also did not give information on the sample size, meaning that all analysis had to focus on proportions, which were often times weighted, rather than any specific numbers of the sample population. Standard deviation was also not available through the mean table, which meant inferences were limited to results from the regression analysis and percentage tables. Overall, the sample was limited to categorical variables and important information about sample size and standard deviation was not available.

Using the software of PowerStats, I was able to create the tables I need to analyze the relationships among the independent and dependent variables. I used race as a filter variable in the study. Specifically, I filtered for Black, Latinx and white students. The proportion of ethnic backgrounds represent in the HSLS:09 was too small to be significant in the analysis. Below are the used in the analysis:

Table 1. Variable table

	Variable	Description
Independent	*X1LOCALE*	Indicates the locale (urbanicity) of the sample member's base year school.
	X1SEX	Indicates the student's sex/gender.
	X1SES	Indicates the sample member's socio-economic status.
	X1RACE	Indicates the student's race/ethnicity. Broken into the categories Black, White and Latinx students.
	X1TXMQUINT	Indicates the student's mathematics quintile score. The math quintile score is a norm-referenced measure of achievement.
	X1FREELUNCH	Indicates the percentage of students enrolled in the school who receive free or reduced price lunch.
	X1TMCERT	Indicates the math teacher's base year math teaching certification status by grade level and type of certification.
Dependent	*S2MPERSON1*	Indicates the extent to which respondents see themselves as a math person.
	S2SUREBA	Indicates how certain the respondent is that he/she will pursue a bachelor's degree.
	X2STU30OCC_STEM1	Indicates the STEM code 1 (sub domain) of the job the sample member expects or plans to have at age 30. Categorized as either STEM or Non-STEM.
	X4ENTMJSTNSF	Indicates whether the major the student was most seriously considering when first entering postsecondary education was a major supported by the National Science Foundation (NSF). Categorized as either STEM or Non-STEM.
	S4EVRATNDCLG	Indicates whether respondent attended any college or trade school by the end of February 2016.
	X4RFMJSTNSF	Indicates if the reference undergraduate degree/certificate's first major field of study is a major supported by the National Science Foundation (NSF). Categorized as either STEM or Non-STEM.
	X3THIMATH	Indicates the highest mathematics course taken. Categorized as either higher than Algebra II or at most Algebra II.
Filter Variables	*X1RACE*	Filtered for Black, White and Latinx students.

Of the variables above, only socioeconomic status is a continuous variable. The dependent variables listed consist of student performance, school characteristics, teacher characteristics variables and student demographics. I used the dependent variables to predict on both student perception outcomes and performance outcomes. Analysis was performed by creating cross tabs on all the variables listed against the race categories I used as a filter. Cross tabs provide information on descriptive statistics of the variables used. Specifically, the cross tabs give proportions of the sample population of race against each individual variable. However, they do not take into account the outside factors or other variables involved when crossing the variables. Because the cross tabs are a base level analyses, this means I also had to use logistic regression to make a stronger analysis of the data. The logistical regression does not generate traditional regression coefficients to make predictions. Instead, I have to rely on odds ratios to make inferences on the likelihood of certain outcomes based on my dependent variables. I did not have access to HSLS09's student assessment scores, only the student quintile scores, which divides of student assessment into five categories. Having access to student assessment scores as a predictor for mathematics perception outcomes and performance outcomes would have strengthen my analysis by comparing these outcomes to race. However, I still used the quintile scores to make inferences on student outcomes. Odds ratios were analyzed as statistically significant at $p<.05$ at least.

The Findings

Before beginning logistical regression analysis, I first constructed cross tables with all variables by race. Results are divided based on the two desired variable outcomes: perception outcomes and performance outcomes. Below, I begin with a cross table of race against the independent variables.

Table 2. Student- and School-Level Characteristics by Race

	White	Black	Latinx	All
Sex				
Male	50.9	47.3	50.1	50.4
Female	49.1	52.7	49.1	49.6
School Locale (urbanicity)				
City	20.9	43.3	46.7	31.9
Suburb	34.0	30.4	32.6	33.3
Town	15.0	8.9	6.1	11.8
Rural	30.1	17.4	14.5	23.0
SES				
Mean(SE)	0.2(0.02)	-0.3(0.03)	-0.5(0.03)	-0.1(0.02)
Race/ethnicity				
White	N/A	N/A	N/A	57.8
Black	N/A	N/A	N/A	16.2
Latinx	N/A	N/A	N/A	26.0
Mathematics quintile score				
First quintile	14.0	35.8	24.9	19.8
Second quintile	18.2	26.3	22.1	20.1
Third quintile	20.0	16.7	23.8	20.2
Fourth quintile	22.9	13.3	16.7	20.0
Fifth quintile	24.9	8.0	12.6	19.8
Math teacher's math teaching certification				
High School Certification	82.3	73.2	75.0	79.2
Not High School Certification	17.7	26.8	25.0	20.8
Grade 9 percent free lunch-categorical				
less than 20%	25.0	7.9	11.2	19.9
At least 20% but less than 60%	62.8	41.7	43.9	54.3
Greater than 60%	12.2	50.5	44.9	25.8

Unsurprisingly, White students have an average socioeconomic status of .2, with a standard error of .002, while Black and Latinx students have an average socioeconomic status of -.3 and -.5 respectively (where the mean is referenced from 0). We also see that there is race variation in math quintile scores; White students overwhelmingly dominate the Mathematics quintile score. 25% of White students are in the fifth quintile, compared to only 8% of Black students and 12% of Latinx students. At the same time, Black and Latinx students are disproportionately represented in the lowest quintile with Black students at 35.8% and Latinx students at 25%. Another interesting statistic is that 82.3% of White students in the sample have a High School certified mathematics teacher, compared to only 73.2% for Black students and 75% for Latinx students. Lastly, only 12.2% of White Students attend a school where more than 60% of students qualify for free or reduced priced lunch. In contrast, 50% of Black students and 45% of Latinx students attend a school where more than 60% of students qualify! It is also important to note that while 25% of white students attend a school where less than 20% of students qualify for free or reduced-price lunch, 8% of Black Students and 11.2% of Latinx students attend such schools. Thus, there is a large disparity in the economic status of students and the schools they attend across racial lines. At the same time, we see that white students in the study have outperformed Black and Latinx students. Perhaps there is a correlation, then on student socioeconomic status and student mathematics achievement. It is also the case that almost 50% of Black and Latinx students attend schools in the city, so this may impact the disparities seen in these proportions.

Table 3. Perception Outcomes by Race

	White	Black	Latinx	All
Teenager sees self as a math person				
Agree	44.5	44.3	39.6	43.6
Disagree	55.5	55.7	60.4	56.4
Certainty: will Pursue bachelor's degree				
Will	81.1	83.0	77.8	80.8
Won't	18.9	17.0	22.2	19.2
Considering STEM major upon postsecondary entry (NSF)				
STEM	34.5	30.3	32.1	34.0
Non-STEM	65.5	69.7	67.9	66.0
Student occupation at age 30: STEM				
STEM Occupation	9.6	6.1	7.2	8.6
Non-STEM Occupation	90.4	93.9	92.8	91.4

Despite these disparities noted in the sample population, there is hardly a difference when crossing race against perception outcomes across racial lines. In Table 2 we see that a similar proportion of White, Black and Latinx students view themselves as math people. At the same time, about a third of White, Black and Latinx students perceive themselves pursuing a bachelors in a STEM subject. Students across the board seem to view themselves capable of succeeding in STEM in the sample. However, in performance we see differences in students from the sample. By February 2016, 73.8% of White students were enrolled in some type of college, versus 62.1% of Black students and 66.1% of Latinx students. In addition, although the difference is not huge, 30% of White students have a STEM major as their first major while only a quarter of Black and Latinx students do. Lastly, there is about a 10% difference in the proportion of White students who have taken courses higher than Algebra II and Black students, and 15% difference with White and Latinx students. 58.2% of White students have taken a course higher than Algebra II in high school, while only 49.7% of Black students and 43.4% of Latinx students have. The ambition students have is not parallel to the outcome's students have in mathematics and STEM. These outcomes more closely align to the disparities noted earlier in student performance and socioeconomic status. It is plausible that socioeconomic status and initial performance contribute to this later difference in student STEM outcomes.

Table 4. Performance Outcomes by Race

	White	Black	Latinx	All
Highest level mathematics course taken/pipeline				
At most Algebra II	41.8	50.3	56.6	46.4
Higher than Algebra II	58.2	49.7	43.4	53.6
Ever attended college by the end of February 2016				
Yes	73.8	62.1	66.1	70.3
No	26.2	37.9	33.9	29.7
Reference degree first major - STEM (NSF)				
STEM	30.0	25.4	26.6	29.2
Non-STEM	70.0	74.6	73.4	70.8

Logistical Regression Analysis

The baseline odds of the regression analysis are a white male student who attends a suburban school that is "average" in the proportion of students who qualify for free or reduced-price lunch, come from a home with an average socioeconomic status and have a high school certified mathematics teacher. This student also has a math quintile score in the third (middle) quintile. What is surprising is that in the table 5, school locale is not statistically significant at $p<.05$. I assumed that school locale would predict the odds of student perception outcomes.

Female students, when holding other variables constant, have lower odds by a factor of .67 ($p<.001$), when compared to similarly situated male students, when predicting the odds of whether they see themselves as math people. In addition, Black students have higher odds, by a factor of 1.76, when compared to similar white students in predicting if they view themselves as math people. Students who score in the fifth quintile have higher odds by a factor of 3.08 ($p<.001$) when predicting the odds of whether students view themselves as math people at. At the same time, students in the lowest quintile have lower odds by a factor of .45 ($p<.001$) when predicting odds of whether they see themselves as math people. Race and gender seem to play a role in increasing the odds that a student see's themselves as a math person, holding other variables constant, However, we also see that student performance in the math quintile has higher odds for the students in the fifth quintile and lower for those in the lowest, exclusive or race and gender.

When predicting the certainty that a student will pursue a bachelor's degree, female students have higher odds by a factor of 2.12 ($p<.001$). Black students also have higher odds by a factor of 1.84 ($p<.001$) when predicting certainty that a student will pursue a bachelor's degree. In addition, Latinx students have higher odds by a factor of 1.35 ($p<.05$). This highlights that Black, Latinx and female students who are similar situated to the base case student, have higher odds of viewing themselves with the ability to pursue a bachelor's degree. Socioeconomic status increases the odds by a factor of 1.87 ($p<.001$) when predicting if a student will pursue a bachelor's degree. There then might be a relationship among race and socioeconomic status when predicting the odds that a student will pursue a bachelor's degree. Students from towns have lower odds by a factor of .69 ($p<.01$) while students from rural areas have lower odds by a factor of .79 ($p<.01$). In other words, students not from cities and suburban areas have lower odds of certainty of pursuing a bachelor's degree. Lastly, students in the lowest quintile have lower odds by a factor of .51 ($p<.001$) while students in the highest quintile have higher odds by a factor of 2.62 ($p<.001$). Of the other variables in the perception outcomes, Certainty of students pursing a bachelor's degree had the highest pseudo R^2, meaning it is more robust than the other outcome models. This does not imply, however that this is an objectively good model, as pseudo-R-squared is not as precise a measure as R-squared in a linear regression model.

Table 5. Predicted Odds of Perception Outcomes, Logistic Regression

	Teenager sees self as a math person	Certainty: will Pursue bachelor's degree	Student occupation at age 30: STEM	Considering STEM major upon postsecondary entry (NSF)
Baseline Odds	0.76	3.12	0.13	0.24
	(0.072)	(0.412)	(0.024)	(0.035)
Student's sex (ref. Male)				
female	0.67***	2.12***	0.33***	1.06
	(0.052)	(0.197)	(0.039)	(0.093)
Student's race/ethnicity-composite (ref. White)				
Black/African American	1.76***	1.84***	1.11	1.25
	(0.106)	(0.308)	(0.223)	(0.231)
Latinx	1.03	1.35*	0.82	1.149
	(0.230)	(0.204)	(0.141)	(0.152)
School locale (urbanicity) (ref. Suburb)				
City	1.00	0.95	1.05	1.00
	(0.097)	(0.149)	(0.161)	(0.131)
Town	0.93	0.69**	1.04	0.75*
	(0.105)	(0.077)	(0.166)	(0.087)
Rural	1.10	0.79*	0.80	0.92
	(0.095)	(0.093)	(0.128)	(0.126)
Socio-economic status composite	0.93	1.87***	1.20*	1.33***
	(0.049)	(0.141)	(0.100)	(0.084)
Mathematics quintile score (ref. third (middle) quintile)				
First (lowest) quintile	0.45***	0.51***	0.67	0.53**
	(0.058)	(0.079)	(0.191)	(0.121)
Second quintile	0.60***	0.68**	0.96	0.66*
	(0.061)	(0.080)	(0.205)	(0.107)
Fourth quintile	1.70***	1.77***	1.32	1.27
	(0.174)	(0.266)	(0.260)	(0.183)
Fifth (highest) quintile	3.08***	2.62***	1.88**	2.26***
	(0.312)	(0.476)	(0.347)	(0.261)
Math teacher's math teaching certification (ref. High School Certification)				
Not High School Certified	0.90	0.86	0.88	1.08
	(0.074)	(0.120)	(0.156)	(0.164)
Grade 9 percent free lunch-categorical (ref. At least 20% but less than 60%)				
Less than 20%	1.14	1.11	1.34*	1.06
	(0.102)	(0.175)	(0.181)	(0.127)
At least 60%	1.08	1.26	0.84	1.16
	(0.105)	(0.208)	(0.156)	(0.173)
N	8800	8800	8800	6900
Pseudo R^2	0.078	0.111	0.071	0.057

*p<.05 **p<.01 ***p<.001

When predicting whether students will consider a STEM major upon postsecondary entry, socioeconomic status has higher odds by a factor of 1.33 (p<001). This also implies that socioeconomic status positively increases the likely hood of considering STEM. In addition, the lowest quintile lowers odds by a factor of .53 (p<.001) while the highest quintile increases odds by a factor of 2.26 (p<.001) when holding other variables constant. Socioeconomic status and student prior performance as the statistically significant factors when predicting likely hood a student will consider majoring in a STEM major post-secondary.

Performance Outcomes

The first performance outcome is predicting odds of whether students took a course higher than Algebra II. Female students had higher odds by a factor of 1.36 (p<.001) when predicting the if they took mathematics courses higher than Algebra II. Latinx students had higher odds by a factor of ratio of 1.09 (p<.05), whereas the base case student had lower odds by a factor of 0.72, when predicting if they took courses higher than Algebra II. Socioeconomic status also had higher odds by a factor of 1.66 (p<.001), meaning students with average socioeconomic status had higher odds of taking a course higher than Algebra II. Students in the lowest quintile had lower odds by a factor of .38 (p<.001), while students in the highest quintile had higher odds by a factor of 4.23 (p<.001). The higher the quintile score, the more likely students are to have the odds of taking a course higher than Algebra II. The last predictor, which is rather surprising, is that students who attend a school where at least 60% of students have free or reduced priced lunch have higher odds by a factor of 1.54 (p<.05).

When predicting students' odds of attending college by the end of February 2016, female students have higher odds by a factor of 1.69 (p<.001), This is surprising as female students have lower odds of viewing themselves as math people. Latinx students had higher odds by a factor of

1.24 ($p<.05$), compared to the lower odds of the baseline student, when predicting if they attended college by the end of February 2016, which aligns with the odds of their certainty of attending college. Socioeconomic status had higher odds by a factor of 1.79 ($p<.001$) when predicting whether students have the odds of attending college by the end of February 2016. Student quintile scores also positively predicted whether students were enrolled in college by the end of February 2016. Students in the fifth quintile had higher odds by a factor of 2.39 ($p<.001$) whereas students in the lowest quintile had lower odds by a factor of ,54 ($p<.001$), which means they were negatively predictive of having odds of being enrolled in college. Lastly, students who attended schools where less than 20% of students qualified for free or reduced priced lunch had higher odds by a factor of 1.24 of being enrolled in college by the end of February 2016.

Table 6. Predicted Odds of Performance Outcomes, Logistic Regression

	Highest level mathematics course taken/pipeline	Ever attended college by the end of February 2016	Reference degree first major - STEM (NSF)
Baseline Odds	0.72	0.91	0.19
	(0.095)	(0.099)	(0.030)
X1 Student's sex (ref. Male)			
female	1.36***	1.69***	1.05
	(0.097)	(0.123)	(0.102)
X1 Student's race/ethnicity-composite (ref. White)			
Black/African American	1.65	1.18	1.38
	(0.245)	(0.147)	(0.264)
Latinx	1.09*	1.24*	1.13
	(0.158)	(0.117)	(0.149)
X1 School locale (urbanicity) (ref. Suburb)			
City	1.03	1.15	1.00
	(0.185)	(0.123)	(0.126)
Town	1.11	0.79*	1.00
	(0.272)	(0.072)	(0.151)
Rural	1.05	0.85*	0.94
	(0.167)	(0.064)	(0.121)
X1 Socio-economic status composite	1.66***	1.79***	1.34***
	(0.108)	(0.101)	(0.101)
X1 Mathematics quintile score (ref. third (middle) quintile)			
First (lowest) quintile	0.38***	0.54***	0.53**
	(0.052)	(0.062)	(0.127)
Second quintile	0.63***	0.78*	0.63*
	(0.062)	(0.085)	(0.116)
Fourth quintile	2.17***	1.44***	1.16
	(0.230)	(0.126)	(0.177)
Fifth (highest) quintile	4.23***	2.39***	2.01***
	(0.639)	(0.265)	(0.276)
X1 Math teacher's math teaching certification (ref. High School Certification)			
Not High School Certified	0.85	1.11	1.13
	(0.131)	(0.106)	(0.181)
X1 Grade 9 percent free lunch-categorical (ref. At least 20% but less than 60%)			
Less than 20%	1.27	1.24*	1.11
	(0.211)	(0.126)	(0.119)
At least 60%	1.54*	0.92	0.98
	(0.276)	(0.090)	(0.148)
N	9000	9000	6600
Pseudo R^2	0.151	0.101	0.051

*p<.05 **p<.01 ***p<.001

Surprisingly, the odds for female students are not significant when predicting whether their first major is a STEM major. Socioeconomic status, however, had higher odds by a factor of 1.34 ($p<.001$) when predicting the odds of students first degree is in STEM. In addition, students in the fifth quintile had higher odds by a factor of 2.01 ($p<.001$) when predicting the odds if a students' first major was a STEM major, while students in the lowest quintile had lower odds by a factor of .53 ($p<.01$).

Surprisingly, school urbanicity only factored in the odds that students were enrolled in college by the end of 2016, and not in other performance outcomes. Furthermore, teacher certification was not significant in any of the outcomes of the analysis. Percentage of students with free or reduced-price lunch also was found to be insignificant for the outcomes of if a student was enrolled in college by the end of February 2016 and if their first major was in STEM.

Discussion/Conclusion

To conclude, the focus of this research was to find out how do school and student characteristics affect student outcomes in STEM (math) when analyzing survey data from the HSLS:09. Specifically, I analyzed how school locale, teacher certification, percent of students on free or reduced priced lunch, base student achievement and demographic factors affect student perception outcomes and student performance outcomes. I hypothesized that school locale, prior student achievement, student race, teacher certification and percent of students on free or reduced-price lunch would predict strongest impact student STEM outcomes. Using cross tabs, I found that students across racial lines were equally likely to view themselves positively in math and their ability to go to college. However, the cross tabs highlighted a difference in student performance outcomes across racial lines, where white students unsurprisingly were enrolled in college and majored in STEM majors at larger rates than the students of color in the data. Through the logistical regression analysis, I found that the biggest factors to impact student perception outcomes included race, gender, socioeconomic status and math quintile scores. To my surprise, other factors of the school had no statistically significant impact on student perception outcomes. Similarly, the factors that predicted student performance outcomes were also race, gender, socioeconomic status and student math quintile scores. Although race was the least significant of the factors listed, despite the racial disparities in STEM outcomes, it was not as strong a predictor as SES and math quintile scores when predicting student performance.

From these findings we see that mathematics education has more underlying factors that predict student interest and performance than meritocracy. It is widely believed that students who work hard in STEM courses will do, and external factors are rarely taken into account. Battery states "When referring to students of color, gaps are framed as deficits, pathologizing the intelligence of students of color based on test scores, intelligence, and ability; the same is not done when White students have lower test scores." (Battery, 2). In his research, he finds that students of color are often not given the supports necessary to succeed as result of mathematics being framed for white students. Thus, student's low achievement in STEM courses is pinned on them and the solution to this is for students to work harder or be placed in lower tracks (Battery, 23). However, as I mentioned, my research finds that more factors are present when gaging student success in STEM. For instance, gender has quite a big impact on a student's perception outcomes as shown earlier. Yet, female students also have increased odds of taking higher level courses and being enrolled in college.

Furthermore, student prior performance and student SES are huge factors in predicting odds for both performance and perception outcomes. This implies that household factors play a large role into student academic self-perception and success, regardless of school factors. Similarly, Kelly finds, "the black-white gap in course taking in mathematics can indeed be explained primarily by differences in academic and family background upon entry to high school." (Kelly, 61). Student prior performance also makes the implication that current school factors do not play a strong role, but prior education does. In that same light, prior education performance can predict student post-secondary performance (Kelly, 66). The implications of these findings are that SES and gender paly a large role in interest of and performance in STEM.

However, despite the similar interest in STEM across racial lines, there is an obvious disparity in course taking and post-secondary major choice in STEM. The regressions found SES and prior student achievement to be the largest factors to predict student STEM outcomes. Perhaps there is a close relationship then with socioeconomic status and a student's quintile score then. Holmes, et al, find that "It appears that achievement at school in both reading and numeracy, while related to other non-significant variables such as SES and indigenous status, is a better predictor of interest in a STEM career than these demographic variables." (Holmes, et al, 11). In other words, they find that there is a relationship among student SES and prior student performance. This aligns with the results from logistical regression analysis.

It is important to recognize, that although these disparities exist, students are interested and confident in STEM, nonetheless. It is then necessary going forward to better foster and cultivate this interest. Failing to do so will continue the cycle of lack of representation of students of color and women in STEM majors and occupations. More research needs to be done on the pedagogical changes necessary to close these gaps and to ensure students achieve their full potential. In addition, this research highlights a need for schools to adjust and support students adequately who come from lower SES backgrounds, as they have less likely odds of performance and perception outcomes in STEM. More research can be done to gain understand of what these supports would look like.

In my research, I was limited by the amount of data of the HSLS:09 I had access to through the NCES. Because of this, statistical methods incorporated to reach my results were limited to logistical regression analysis and cross tabs. In future research, having access to the complete data file would strengthen the methods used. I would be able to run traditional regression analyses and gather other descriptive statistics to make stronger implications on why students choose STEM. This research also lacked a variable to take into account why student's choose majors in other disciplines and how that intersected the independent variables chosen for the study. In future research, it would be interesting to see how many students choose another major initially and then switch to STEM, or how many students choose a STEM major then switch to a different department. These results may be telling in what populations of students are more likely to change their major or course of study and why they make this change. In addition, future research may include an independent variable that measure's student immigration status (such as 1.5 or 2nd generation immigrant). Immigration status of students and families may be telling in what students choose to study and how they perceive themselves academically. Lastly, future research might take into account student anecdotal evidence through interviews to track themes as to why students make the choices they do or how their perception over time changes as they progress academically. Strong ethnographic research may bring deeper answers to issues of tracking and gate keeping in mathematics education.

Bibliography

A. Katrin Arens, Anne C. Frenzel & Thomas Goetz (2020) Self-Concept and Self-Efficacy in Math: Longitudinal Interrelations and Reciprocal Linkages with Achievement, The Journal of Experimental Education, DOI: 10.1080/00220973.2020.1786347

Charles, Maria. "Venus, Mars, and Math: Gender, Societal Affluence, and Eighth Graders' Aspirations for STEM." *Socius*, Jan. 2017, doi:10.1177/2378023117697179.

Conger, Dylan, et al. "Explaining Race, Poverty, and Gender Disparities in Advanced Course-Taking." *Journal of Policy Analysis and Management*, vol. 28, no. 4, 2009, pp. 555–576., doi:10.1002/pam.20455.

Dan Battey (2013) Access to Mathematics: "A Possessive Investment in Whiteness"[1], Curriculum Inquiry, 43:3, 332-359, DOI: 10.1111/curi.12015

Henriette Tolstrup Holmegaard, Lene Møller Madsen & Lars Ulriksen (2014) To Choose or Not to Choose Science: Constructions of desirable identities among young people considering a STEM higher education programme, International Journal of Science Education, 36:2, 186-215, DOI: 10.1080/09500693.2012.749362

Holmes, Kathryn, et al. "An Integrated Analysis of School Students' Aspirations for STEM Careers: Which Student and School Factors Are Most Predictive?" *International Journal of Science and Mathematics Education*, vol. 16, no. 4, 2017, pp. 655–675., doi:10.1007/s10763-016-9793-z.

Kelly, Sean. “The Black-White Gap in Mathematics Course Taking.” *Sociology of Education*, vol. 82, no. 1, 2009, pp. 47–69., doi:10.1177/003804070908200103.

Marsh, H. W., & Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology, 89*(1), 41–54. https://doi.org/10.1037/0022-0663.89.1.41

Moakler, Martin W., and Mikyong Minsun Kim. “College Major Choice in STEM: Revisiting Confidence and Demographic Factors.” *The Career Development Quarterly*, vol. 62, no. 2, 2014, pp. 128–142., doi:10.1002/j.2161-0045.2014.00075.x.

Philip David Parker, Herbert W. Marsh, Joseph Ciarrochi, Sarah Marshall & Adel Salah Abduljabbar (2014) Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes, Educational Psychology, 34:1, 29-48, DOI: 10.1080/01443410.2013.797339

Riegle-Crumb, Catherine, et al. “Who Wants to Have a Career in Science or Math? Exploring Adolescents' Future Aspirations by Gender and Race/Ethnicity.” *Science Education*, vol. 95, no. 3, 2010, pp. 458–476., doi:10.1002/sce.20431.

Sahin, Alpaslan, et al. “The Relationships among High School STEM Learning Experiences, Expectations, and Mathematics and Science Efficacy and the Likelihood of Majoring in STEM in College.” *International Journal of Science Education*, vol. 39, no. 11, 2017, pp. 1549–1572., doi:10.1080/09500693.2017.1341067.

Saw, Guan, et al. "Cross-Sectional and Longitudinal Disparities in STEM Career Aspirations at the Intersection of Gender, Race/Ethnicity, and Socioeconomic Status." *Educational Researcher*, vol. 47, no. 8, 2018, pp. 525–531., doi:10.3102/0013189x18787818.