metis associates
making a meaningful difference

120 Wall Street
21st Floor
New York, NY 10005
Tel: 212-425-8833
Fax: 212-480-2176
www.metisassociates.com

April 23, 2013

The Specialized High School Admission Test and High School Academic Achievement

Abstract

To determine the relationship between the Specialized High School Admission Test (SHSAT) and subsequent high school performance from the first two years of high school, the relationship between the SHSAT and three selected high school achievement metrics was examined. Five annual cohorts of 8th grade SHSAT takers were followed into their first two years of high school to obtain their performance on measures such as GPA, Regents examinations and Advanced Placement (AP) examinations.

Correlation and regression analyses pointed to a strong positive predictive relationship between the SHSAT and high school achievement. These findings are consistent across all cohorts – as well as for all cohorts combined – irrespective of the academic rigor of the high schools ultimately attended by students. Of particular note were the strong relationships observed between the SHSAT and the mathematics and science subject areas. Further, strong relationships were also exhibited between the SHSAT and achievement on college level metrics for history as measured by the AP exams. The consistency of these results over the observed five-year period suggests that the results were not spurious and accurately reflected the true nature of the relationship between the SHSAT exam and subsequent high school achievement.

Introduction

Each year more than 20,000 students sit for the Specialized High School Admission Test (SHSAT) as the sole admissions criterion to NYC's specialized high schools. These students are competing for approximately 3,800 seats – in other words, fewer than 20% of the test takers will gain acceptance to these schools. The NYC DOE has asked Metis Associates to conduct a predictive validity study to determine the extent to which the SHSAT is a reliable predictor of future academic success in high school.

Data from the SHSAT and select high school academic outcome metrics (described below) were compiled and analyzed for a seven year period from school year 2005-06 to school year 2011-12 comprising five annual cohorts of 8th grade SHSAT takers from fall 2005 to fall 2009. Five cohorts of test-takers were selected to ensure that any connection found between the SHSAT and the selected outcome metrics was consistent over time.

Analyses point to a consistent strong positive predictive relationship between the SHSAT and high school academic achievement for each cohort as well as all cohorts combined. This relationship exists even when controlling for the clustering effects of students within schools and

irrespective of the type of high school that is ultimately attended by the SHSAT takers. Described in the sections below are the selected high school academic outcome metrics, methodology and findings that have led to these conclusions.

Outcome Metrics

A number of outcomes were considered for this study as potential indicators of academic success in high school – scores on New York State Regents exams, scores on Advanced Placement (AP) exams, and grade point averages (GPAs). To minimize contamination from "hidden" biases, such as motivation or maturation, it was desirable to select high school outcomes obtained in close temporal proximity to the administration of the SHSAT – i.e., during the first and/or second years of high school. While GPAs are typically used in studies such as the one presented herein, Metis and the NYC DOE also selected Regents and AP examinations as outcomes due to their continued use as high school achievement metrics and known standard, psychometric properties. The specific Regents and AP exams used as outcomes were based on a review of the actual numbers of SHSAT test-takers who also sat for each of the various subject area exams.

Along with the examinations, a review of grade point averages from the first, second and combined first and second years of high school was undertaken to determine their suitability as outcome metrics. As described later, hierarchical linear modeling (HLM) techniques were employed to control for possible school effects that may have influenced obtained GPAs or standardized examination results.

Grade Point Averages (GPAs)

As shown in Appendix A, grade point equivalents associated with "standard courses" range in value from 0 (F) to 4.3 (A+), while grade point equivalents for AP courses (i.e., college level courses) range from 0 to 5. Grade point averages were generated for each student by multiplying the numerical value shown in Appendix A by the credit value of each course, adding these products and dividing by the total number of credits taken. The following example shows the GPA calculation for a student taking four courses in a particular semester:

Table 1. Example Grade Point Average (GPA) calculation

Course	Type	Letter Grade Received	Grade Point Equivalent (GPE)	Number of Credits	Value (GPE X Credits)
Course 1	Standard	B+	3.50	3	10.50
Course 2	AP	B+	3.75	3	11.25
Course 3	AP	A	4.30	4	17.20
Course 4	Standard	A	4.00	3	12.00
			TOTALS	13	50.95
				GPA	3.92

In the example above, the student took two standard courses (worth three credits each), and two AP courses (a three-credit and a four-credit course), thereby accumulating 13 credits. His/her letter grades were A's and B+'s. Using the values contained in Appendix A, a grade point equivalent (GPE) was obtained for each letter grade. The GPE was then multiplied by the number

of credits associated with each course, yielding the value shown in the right-most column above. The sum of these values (50.95) was then divided by the total credits (13), to yield a GPA of 3.92.

Population

Metis was provided with rosters for all 8th grade students who sat for the SHSAT exam for a five (5) year period between fall 2005 to fall 2009 (i.e., SY2005-06 to SY2009-10). These data were checked for duplicates, cleaned, processed and stored in a local database. Student records from each cohort of students were then matched to the NYCDOE public school audited registers, high school course deck, Regents examination results and Advanced Placement (AP) examination results for the immediate two consecutive years following the exam. For example, the 2009 SHSAT cohort (SY2009-10) was matched to high school records from SY2010-11 and SY2011-12.

Within each cohort, a flag was generated for each student to denote their acceptance to one or more of the eight (8) specialized schools for which the SHSAT is the admission requirement[1]. Two groups of students were then identified – accepted and not accepted – to explore the possibility of either creating a subgroup analysis or limiting the outcome analyses to only those students accepted to specialized schools. Metis had been concerned that any subgroup analysis or limitation to accepted students may result in an inadequate amount of variance in outcomes due to their performance on the SHSAT.

Empirical evidence confirmed Metis's suspicions – looking simply at the students who were accepted to a specialized school would not be appropriate because the admission criterion creates an artificial "floor" for the exam which does not permit the full range of performance to be explored (see Appendix B). Since this study is concerned with the predictive validity of the full exam (not just the top scores) Metis conducted *all analyses* presented herein on the **entire population** of 8th grade students sitting for the SHSAT.

Analyses

Once data quality was established, Metis undertook several different analyses of the data to determine the relationship between the SHSAT and the selected outcomes. First, a series of descriptive tables were generated for each cohort and all cohorts combined to gain a better understanding of the SHSAT cohorts, missing data patterns and number of students for whom outcome data were available. For the purposes of the current study, Metis decided to use a criterion of *at least 500* students with outcomes for all analyses as it represents approximately 2.5% of the population that sits for the SHSAT and should result in stable correlation coefficients less affected by "noise" variance that may be associated with a smaller number of observations.

[1] While there are nine (9) specialized high schools in New York City, admission to Fiorello H. LaGuardia High School of Music & Art and Performing Arts is not based on the SHSAT results.

After reviewing the descriptive tables, scatterplots, correlation matrices and linear regressions were developed to gain a better understanding of the relationship between the SHSAT and each of the outcome metrics. These were disaggregated by year to determine whether any observed relationships between the SHSAT and outcomes were consistent over time.

Finally, to control for possible school experience confounds, additional rigorous analyses with hierarchical linear modeling (HLM) procedures were conducted based on all cohorts combined. To further control for possible fixed effects of school type, schools were classified into three categories: specialized (admission to which depends on the SHSAT), selective[2] and other. This school type indicator was also included as an additional predictor at the school level to confirm the results based on the SHSAT predictor only.

Comparative Outcomes

Grade Point Averages. As illustrated in the table below displaying the GPAs for all students sitting for the SHSAT over the five year period, students accepted to specialized high schools (i.e., those who scored at or above the cut-point on the SHSAT) fared better than the students who were not accepted. For example, it can be seen that the GPAs for accepted students ranged from 3.036 (1st Year) to 3.104 (2 year aggregate), while the GPAs for students not accepted ranged from 2.299 to 2.396. Standard deviations (SDs) for the accepted group were also smaller than those for the not accepted group, reflective of a smaller spread of values around higher mean GPA performance for accepted students.

Table 2. Grade Point Averages for all 8th grade students sitting for the SHSAT from 2005 to 2009

Metric	Accepted			Not Accepted			Total		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
1st Year GPA	19,298	3.036	0.694	79,345	2.387	0.894	98,643	2.514	0.897
2nd Year GPA	19,102	3.059	0.725	77,734	2.299	0.950	96,836	2.449	0.959
2 year Aggregate GPA	19,344	3.104	0.688	80,466	2.396	0.896	99,810	2.533	0.904

Regents Examinations. Table 3 presents the mean scores for the Regents examinations taken in the first two years of high school for all students sitting for the SHSAT over the period of the study. The mean scores for accepted students ranged from 82.59 to 93.41, while the mean scores for students not accepted ranged from 68.69 to 79.16. Note that average scores on subsequent Regents exams are also higher for accepted students than for students not accepted into a specialized high school. As with GPAs, smaller standard deviations were noted for the accepted group, suggesting smaller distributions of scores around better mean Regents performances for accepted students.

Table 3. Mean score on first two years of Regents examinations for all 8th grade students sitting for the SHSAT from 2005 to 2009

[2] The list of selective schools was provided to Metis by the NYC DOE (see Appendix C). Schools were selected for this list based on three criteria: 1. Applicants are screened for grades between 80 and 100 in core subject areas; 2. Demand for the school is higher than citywide demand; and 3. The school is within the top quartile of high schools in incoming proficiency.

Subject	Accepted			Not Accepted			Total		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Algebra	3,986	85.98	5.94	44,103	74.79	9.32	48,095	75.72	9.60
Biology	15,684	89.50	5.58	65,503	77.05	10.15	81,190	79.45	10.64
Chemistry	17,758	82.59	9.28	32,502	69.52	10.96	50,261	74.14	12.13
Earth Science	†			22,336	69.08	14.08	22,576	69.25	14.13
English	3,958	89.24	7.20	19,605	78.67	11.93	23,563	80.45	11.95
Geometry	10,766	87.85	7.60	34,170	73.45	12.20	44,937	76.90	12.84
Global Studies	18,558	91.91	6.69	67,075	77.75	13.93	85,637	80.82	13.99
Math A	4,145	93.41	4.83	14,511	79.16	11.52	18,656	82.32	11.98
Math B	4,630	86.59	9.82	5,561	71.97	15.74	10,191	78.61	15.23
Physics	2,375	85.28	9.96	3,777	68.69	14.74	6,152	75.09	15.39
Trigonometry	6,418	86.80	11.46	8,269	69.50	17.72	14,687	77.06	17.55
US History	543	90.02	7.82	10,735	76.65	14.82	11,278	77.29	14.84

† Not reported – N < 500.

Advanced Placement (AP) Examinations. While all AP exams were reviewed for the study, only two exams – both in social studies – have sufficient Ns to report. Further, as so few students sit for AP exams in the first year of high school, analyses only focus on year 2 AP exams. This is not unexpected since AP exams are more typically taken in the third and fourth years of high school.

For the two exams with sufficient Ns, table 4 shows the same pattern observed with the Regents and GPA outcomes: students accepted to specialized high schools generally fare better – on average – than students who are not accepted. For these two exams, mean scores for accepted students are over one point higher (on the five-point AP grading scale) than for those students who were not accepted into a specialized high school. Unlike the GPAs and Regents exams, standard deviations for the accepted and not accepted groups were roughly equivalent although this was likely an artifact of the discrete five point scale upon which AP examinations are scored.

Table 4. Mean score on second year Advanced Placement examinations for all 8th grade students sitting for the SHSAT from 2005 to 2009

Subject	Accepted			Not Accepted			Total		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
European History	986	3.53	1.17	925	1.85	1.05	1,911	2.72	1.40
World History	4,785	3.64	0.97	4,743	2.40	1.02	9,528	3.02	1.17

Across all selected outcomes, students accepted to specialized high schools based on their performance on the SHSAT appear to outperform those who are not accepted[3]. While this could be an artifact of the high schools the students ultimately attend, it points to the possibility of a positive relationship between performance on the SHSAT and subsequent high school outcomes. The standard deviations for GPA and Regents exams seem to support this, as there appears to be less variability around higher mean performance for the accepted group than the not accepted group.

[3] This is also true for each individual annual SHSAT cohort. Comparative tables for each SHSAT cohort as well as for separate first and second year Regents results yield the same trends as those displayed in tables 1 to 3.

Correlations/Linear Regressions

As described below, generated correlation coefficients confirm positive relationships between the SHSAT and the chosen high school performance metrics. This is true for metrics from the first year of high school, second year of high school or both the first and second years combined.

Per standard conventions in behavioral statistics almost all of the calculated correlation coefficients could be classified as strong (between 0.40 and 0.69), with a few that could be best characterized as very strong (0.70 or above). It is important to note that very strong correlations are rare in social sciences due to the complicated nature of human behavior.

Further, because the relationships between the SHSAT scores and the selected outcomes have been determined to be strong, it is not surprising that the SHSAT was found to be a statistically significant predictor of all studied outcomes in the linear regression models. In linear regressions, the value of r^2 (the coefficient of determination) indicates the proportion of the variation explained in an outcome by a given predictor (the SHSAT in this study). It is not uncommon in education to find "meaningful" statistically significant predictors with relatively small r^2 values[4]. For example, the raw correlation coefficient reported for the combined revised Scholastic Achievement Test (SAT) and first year college GPA in an initial predictive validity study conducted by the College Board[5] was 0.35, yielding a coefficient of determination (r^2) of only 0.123 (i.e., explaining 12.3% of the variation in first-year GPA).

ALL values of r and r^2 presented in tables 5a through 5e below are statistically significant with a p-value < 0.01. The tables present the correlation coefficients (r) for each relationship tested, the proportion of variation in the outcome explained by the SHSAT (r^2) and the number (N) of students with both SHSAT and outcome metric results.

Table 5a presents the observed relationships between the SHSAT and Grade Point Average (GPA) metrics:

- As suggested earlier, the SHSAT correlates well with GPAs with correlation coefficient values nearing 0.5, thereby explaining approximately 25% of the variation in the metric.
- As will be shown below, while the GPA correlations with the SHSAT are strong, they are not as strong as those observed with the standardized examinations. Metis suspects that this is a function of the subjectivity that is inherent in course grading.

Table 5a. Summary of Correlations and Linear Regressions – 8th Grade SHSAT and Grade Point Average (GPA) Metrics

Metric	2005 SHSAT	2006 SHSAT	2007 SHSAT	2008 SHSAT	2009 SHSAT	Total SHSAT

[4] Since r^2 by definition is the mathematical square (a number multiplied by itself) of the correlation coefficient (r), strong relationships between variables do not necessarily produce large values for r^2 (e.g., when r=0.500, r^2=0.250).

[5] *Validity of the SAT for Predicting First-Year College Grade Point Average.* College Board Research Report No. 2008-5. The College Board, New York, 2008.

Table 5a. Summary of Correlations and Linear Regressions – 8th Grade SHSAT and Grade Point Average (GPA) Metrics

Metric		2005 SHSAT	2006 SHSAT	2007 SHSAT	2008 SHSAT	2009 SHSAT	Total SHSAT
1st Year GPA	r	0.470	0.473	0.458	0.463	0.452	0.462
	r^2	0.221	0.224	0.210	0.214	0.204	0.213
	N	19,479	18,255	19,353	20,917	20,645	98,649
2nd Year GPA	r	0.484	0.489	0.470	0.477	0.463	0.475
	r^2	0.234	0.239	0.221	0.228	0.214	0.226
	N	19,029	17,911	19,042	20,499	20,370	96,841
2 Year Aggregate GPA	r	0.494	0.500	0.481	0.491	0.480	0.487
	r^2	0.244	0.250	0.231	0.241	0.230	0.237
	N	19,742	18,479	19,554	21,110	20,941	99,816

Table 5b presents the observed relationships between the SHSAT and 1st year Regents examinations:

- Of the twelve (12) Regents examinations for which 500 or more students had records in the first year of high school, the relationship with the SHSAT can be characterized as statistically significant, positive and strong.
- The strongest relationship between the SHSAT and first year Regents is observed with the Math A exam (2007, r=0.747).
- Overall, the strongest relationships are observed between the SHSAT and year 1 mathematics (Geometry, Math A, Trigonometry) and science (Biology, Physics, Chemistry) subject areas.
- The SHSAT is also a significant predictor of all Regents exams, with the strongest predictive value for Math A (2007, r^2=0.558).
- The strongest predictive value of the SHSAT overall in year 1 is observed with Biology and Math A.

Table 5b. Summary of Correlations and Linear Regressions – 8th Grade SHSAT and 1st Year Regents Examinations

SUBJECT		2005 SHSAT	2006 SHSAT	2007 SHSAT	2008 SHSAT	2009 SHSAT	Total SHSAT
Algebra	r		0.641	0.595	0.552	0.574	0.587
	r^2		0.411	0.354	0.305	0.329	0.345
	N		9,449	10,585	10,710	11,152	41,894
Biology	r	0.699	0.711	0.699	0.708	0.680	0.684
	r^2	0.489	0.506	0.489	0.501	0.462	0.468
	N	13,276	13,165	14,074	14,817	14,438	69,769
Chemistry	r	0.537	0.614	0.664	0.608	0.636	0.603
	r^2	0.288	0.377	0.441	0.370	0.404	0.364
	N	648	523	862	1,023	1,122	4,178
Earth Science	r	0.579	0.524	0.551	0.548	0.557	0.548
	r^2	0.335	0.275	0.304	0.300	0.310	0.300
	N	581	668	802	1,154	1,181	4,386

Table 5b. Summary of Correlations and Linear Regressions – 8th Grade SHSAT and 1st Year Regents Examinations

SUBJECT		2005 SHSAT	2006 SHSAT	2007 SHSAT	2008 SHSAT	2009 SHSAT	Total SHSAT
English	r					0.452	0.468
	r^2					0.204	0.219
	N	†	†	†	†	551	1,625
Geometry	r			0.665	0.686	0.64	0.663
	r^2			0.442	0.471	0.410	0.440
	N			5,524	7,579	7,709	20,811
Global Studies	r	0.532	0.576	0.449	0.506	0.539	0.513
	r^2	0.283	0.332	0.202	0.256	0.291	0.263
	N	619	732	686	723	665	3,425
Math A	r	0.699	0.738	0.747			0.725
	r^2	0.489	0.545	0.558			0.526
	N	4,639	2,708	899			8,246
Math B	r						0.507
	r^2						0.257
	N	†	†	†	†		532
Physics	r				0.600	0.625	0.676
	r^2				0.360	0.391	0.457
	N	†	†	†	675	656	2,632
Trigonometry	r						0.695
	r^2						0.483
	N				†	†	624
US History	r	0.566	0.571	0.565	0.541	0.532	0.552
	r^2	0.320	0.326	0.319	0.293	0.283	0.305
	N	1,187	1,182	1,488	1,665	1,801	7,323

† Not reported – N < 500.

Table 5c presents the observed relationships between the SHSAT and 2nd year Regents examinations:

- Of the twelve (12) Regents examinations analyzed for the second year of high school, again the relationship with the SHSAT can be characterized as statistically significant, positive and strong.
- The strongest relationship between the SHSAT and second year regents is observed with the Math A exam (2006, r=0.769). Likewise, the greatest predictive value of the SHSAT is observed with Math A (2006, r^2=0.591).
- Overall, the strongest predictive relationships are observed between the SHSAT and year 2 Global Studies and Chemistry exams.

Table 5c. Summary of Correlations and Linear Regressions – 8th Grade SHSAT and 2nd Year Regents Examinations

SUBJECT		2005 SHSAT	2006 SHSAT	2007 SHSAT	2008 SHSAT	2009 SHSAT	Total SHSAT
Algebra	r	0.486	0.459	0.496	0.529	0.513	0.502

Table 5c. Summary of Correlations and Linear Regressions – 8th Grade SHSAT and 2nd Year Regents Examinations

SUBJECT		2005 SHSAT	2006 SHSAT	2007 SHSAT	2008 SHSAT	2009 SHSAT	Total SHSAT
	r^2	0.236	0.211	0.246	0.280	0.263	0.252
	N	663	2,796	2,608	3,327	3,132	12,522
Biology	r	0.581	0.557	0.618	0.600	0.601	0.582
	r^2	0.338	0.310	0.382	0.360	0.361	0.339
	N	3,096	2,821	2,651	2,984	2,746	14,295
Chemistry	r	0.644	0.658	0.657	0.647	0.653	0.648
	r^2	0.415	0.433	0.432	0.419	0.426	0.420
	N	8,837	8,769	9,168	9,740	9,773	46,286
Earth Science	r	0.515	0.490	0.502	0.518	0.476	0.492
	r^2	0.265	0.240	0.252	0.268	0.227	0.242
	N	3,180	3,460	3,579	4,281	4,048	18,547
English	r	0.632	0.569	0.585	0.518	0.521	0.553
	r^2	0.399	0.324	0.342	0.268	0.271	0.306
	N	2,689	2,908	3,755	6,359	6,648	22,359
Geometry	r		0.585	0.551	0.497	0.504	0.529
	r^2		0.342	0.304	0.247	0.254	0.280
	N		7,842	8,150	8,332	8,585	32,908
Global Studies	r	0.650	0.633	0.632	0.625	0.646	0.633
	r^2	0.423	0.401	0.399	0.391	0.417	0.401
	N	15,706	15,191	16,355	17,922	17,447	82,621
Math A	r	0.674	0.769				0.689
	r^2	0.454	0.591				0.475
	N	9,772	1,088				10,862
Math B	r	0.587	0.561	0.623			0.576
	r^2	0.345	0.315	0.388			0.332
	N	4,570	4,456	697			9,729
Physics	r	0.571	0.483	0.561	0.579	0.540	0.546
	r^2	0.326	0.233	0.315	0.335	0.292	0.298
	N	655	611	728	683	894	3,571
Trigonometry	r			0.599	0.610	0.579	0.596
	r^2			0.359	0.372	0.335	0.355
	N			5,531	6,987	7,158	19,675
US History	r	0.593	0.613	0.589	0.551	0.541	0.575
	r^2	0.352	0.376	0.347	0.304	0.293	0.331
	N	952	961	889	926	1,078	4,806

Table 5d presents the observed relationships between the SHSAT and first two years of Regents examinations combined:

- Table 5d presents the Regents examinations taken in the first two years of high school for the SHSAT cohorts. As with the separate first and second year results, the relationships found are statistically significant, positive and strong.

- Biology (r=0.682), Geometry (r=0.650), Chemistry (r=0.645) and Global Studies (r=0.630) exhibit the overall strongest relationship with the SHSAT over the first two years of high school.
- The SHSAT has the strongest predictive value for the Biology, Math A, Geometry and Chemistry Regents (all $r^2 > 0.400$).
- All else being equal, the SHSAT appears to be a strong predictor of performance on the Regents examinations taken within the first two years of high school.

Table 5d. Summary of Correlations and Linear Regressions –8th Grade SHSAT and First Two Years of Regents Examinations Combined

SUBJECT		2005 SHSAT	2006 SHSAT	2007 SHSAT	2008 SHSAT	2009 SHSAT	Total SHSAT
Algebra	r		0.639	0.592	0.567	0.584	0.593
	r^2		0.408	0.350	0.321	0.341	0.352
	N		10,977	12,030	12,514	12,579	48,095
Biology	r	0.689	0.700	0.706	0.706	0.679	0.682
	r^2	0.475	0.490	0.498	0.498	0.461	0.465
	N	15,934	15,253	16,270	17,111	16,626	81,190
Chemistry	r	0.638	0.656	0.656	0.643	0.651	0.645
	r^2	0.407	0.430	0.430	0.413	0.424	0.416
	N	9,465	9,252	9,980	10,725	10,840	50,261
Earth Science	r	0.522	0.494	0.510	0.523	0.498	0.502
	r^2	0.272	0.244	0.260	0.274	0.248	0.252
	N	3,716	4,061	4,316	5,335	5,149	22,576
English	r	0.629	0.567	0.565	0.513	0.515	0.546
	r^2	0.396	0.321	0.319	0.263	0.265	0.298
	N	2,795	3,065	3,983	6,653	7,067	23,563
Geometry	r			0.675	0.648	0.635	0.650
	r^2			0.456	0.420	0.403	0.423
	N			13,397	15,542	16,000	44,937
Global Studies	r	0.647	0.630	0.628	0.621	0.643	0.630
	r^2	0.419	0.397	0.394	0.386	0.413	0.397
	N	16,256	15,829	16,966	18,561	18,025	85,637
Math A	r	0.685	0.746	0.748			0.714
	r^2	0.469	0.557	0.560			0.510
	N	14,016	3,739	901			18,656
Math B	r	0.589	0.563	0.592			0.575
	r^2	0.347	0.317	0.350			0.331
	N	4,727	4,638	787	†		10,191
Physics	r	0.649	0.628	0.624	0.603	0.610	0.620
	r^2	0.421	0.394	0.389	0.364	0.372	0.384
	N	1,028	1,049	1,183	1,349	1,543	6,152
Trigonometry	r				0.616	0.585	0.600
	r^2				0.379	0.342	0.360
	N				7,187	7,501	14,687
US History	r	0.555	0.575	0.543	0.527	0.517	0.539
	r^2	0.308	0.331	0.295	0.278	0.267	0.291

Table 5d. Summary of Correlations and Linear Regressions –8th Grade SHSAT and First Two Years of Regents Examinations Combined

SUBJECT		2005 SHSAT	2006 SHSAT	2007 SHSAT	2008 SHSAT	2009 SHSAT	Total SHSAT
	N	1,975	1,970	2,208	2,462	2,663	11,278

† Not reported – N < 500.

Table 5e presents the observed relationships between the SHSAT and 2nd year Advanced Placement (AP) examinations:

- As observed with the Regents examinations, the correlation coefficients derived for the AP exams are positive, strong and statistically significant.
- The SHSAT also appears to be a good predictor for AP history examinations taken in the second year of high school, accounting for approximately one-third to two-fifths of the variation in AP World History exam results.

Table 5e. Summary of Correlations and Linear Regressions – 8th Grade SHSAT and 2nd Year Advanced Placement Examinations

SUBJECT		2005 SHSAT	2006 SHSAT	2007 SHSAT	2008 SHSAT	2009 SHSAT	Total SHSAT
European History	r						0.686
	r^2						0.471
	N	†	†	†	†	†	1,911
World History	r	0.573	0.585	0.634	0.626	0.586	0.602
	r^2	0.328	0.342	0.402	0.392	0.343	0.362
	N	1,349	1,509	1,997	2,176	2,497	9,528

† Not reported – N < 500.

As stated before, the correlation and linear regression analyses point to a strong predictive positive relationship between the SHSAT and the selected indicators of high school performance. In the next section, it will be determined whether these relationships are influenced by the clustering effects as well as the types of schools (i.e., specialized, selective or other) that are ultimately attended by the students who sit for the SHSAT.

Hierarchical Linear Models (HLMs)

In educational settings, students are nested within schools. It is widely accepted that students attending the same school share many common, educationally relevant experiences and therefore tend to be more similar to each other in terms of academic performance than they are to students in a different school. In other words, student academic outcomes collected from the same school may be related to each other more so than student outcomes from different schools.

When data have nested structures, estimated standard errors obtained by a conventional linear regression could be smaller than actual standard errors due to a failure to account for the

correlated observations within the same cluster.[6] Hierarchical linear modeling (HLM) techniques were developed to analyze data with a hierarchical or nested structure and correct the dependency of observations within a cluster (i.e., the clustering effects). HLMs were therefore conducted in the current study to verify the findings from the linear regressions described above.[7]

The fixed effects of the SHSAT predictor for **ALL** outcome variables presented in Tables 6a – 6e below are positive and statistically significant with a *p*-value < 0.001. The tables summarize the HLM results, including an R-square type measure[8] that indicates the amount of student level variance explained by the SHSAT predictor, the school sample size and the student sample size. According to the R-square type measure, the SHSAT predictor explains a substantial amount of student level variance in all models controlling for the clustering effects. The amount of variance explained ranges from 0.145 (1st year GPA) to 0.428 (1st Year Physics Regents exam). In addition, the conclusions remain the same when the type of school (i.e., specialized, selective or other) is also included as a predictor at the level 2 (school level) model.

Table 6a summarizes the HLM results for Grade Point Average (GPA) metrics:

- Based on the selected R-square type measure, the SHSAT explains between 15% (1st Year GPA) and 18% (2 Year Aggregate GPA) of the student-level variance in GPA.
- Similar to the results based on conventional linear regressions, the SHSAT serves as a statistically significant positive predictor for GPA.

Table 6a. Summary of HLM Results – Grade Point Average for All 8th Grade Students Sitting for the SHSAT from 2005 to 2009

Metric	R-square Type Measure for SHSAT[8]	School N	Student N
1st Year GPA	0.145	490	98,262
2nd Year GPA	0.161	589	96,822
2 Year Aggregate GPA	0.181	596	99,413

Table 6b summarizes the HLM results for 1st year Regents examinations:

- For all twelve (12) of the Regents examinations analyzed for the first year of high school, the SHSAT is found to be a significant positive predictor.
- The SHSAT has the strongest predictive value for Physics (0.428) based on the selected R-square type measure. It shows strong predictive value for Math A and Biology (0.388 and 0.355 respectively) as well.

[6] The underestimation of standard errors may cause an increase in the probability of a Type I error, which would indicate a higher possibility of concluding that obtained results are significant, even though they may not be so.
[7] Specifically, a two-level random intercept HLM was used to analyze the data, with the only predictor (i.e., the SHSAT) included at the level 1 (student level) model.
[8] There is no measure in HLM that is perfectly analogous to the r^2 in linear regressions. Several pseudo- r^2 measures have been used in the field to provide the prediction indices. For the purposes of the current study, the Snijders and Bosker's (1999) version of R_1^2 (the proportional reduction of prediction error variance at level 1 for a model) is generated and reported in Tables 6a - 6e.

Table 6b. Summary of HLM Results – First Year Regents Examinations for All 8th Grade Students Sitting for the SHSAT from 2005 to 2009

Subject	R-square Type Measure for SHSAT[8]	School N	Student N
Algebra	0.289	442	41,862
Biology	0.355	409	69,735
Chemistry	0.301	167	4,176
Earth Science	0.298	246	4,374
English	0.183	118	1,625
Geometry	0.292	338	20,802
Global Studies	0.298	148	3,424
Math A	0.388	294	8,244
Math B	0.229	74	532
Physics	0.428	32	2,632
Trigonometry	0.398	68	624
US History	0.312	159	7,321

Table 6c summarizes the HLM results for 2nd year Regents examinations:

- For all twelve (12) of the Regents examinations analyzed for the second year of high school, the SHSAT again serves as a statistically significant positive predictor.
- The SHSAT explains at least 20% of the student level variance in the second year Regents performance, with the strongest predictive value for Math A (35%).

Table 6c. Summary of HLM Results – Second Year Regents Examinations for All 8th Grade Students Sitting for the SHSAT from 2005 to 2009

Subject	R-square Type Measure for SHSAT[8]	School N	Student N
Algebra	0.206	505	12,495
Biology	0.320	498	14,266
Chemistry	0.251	372	46,260
Earth Science	0.240	399	18,525
English	0.246	418	22,351
Geometry	0.223	435	32,858
Global Studies	0.311	470	82,526
Math A	0.350	350	10,857
Math B	0.198	234	9,724
Physics	0.237	165	3,571
Trigonometry	0.197	332	19,670
US History	0.340	303	4,806

Table 6d summarizes the HLM results for the first two years of Regents examinations combined:

- The analysis of the Regents examinations taken in the first two years of high school combined for the five cohorts also show findings that are consistent with the separate first

and second year analysis results: clearly, the SHSAT is a statistically significant positive predictor for all Regents outcomes.

- The predictive value of the SHSAT is strongest for Math A and Biology (both R-square type measure estimates > 0.360).

Table 6d. Summary of HLM Results – First Two Years of Regents Examinations Combined for All 8th Grade Students Sitting for the SHSAT from 2005 to 2009

Subject	R-square Type Measure for SHSAT[8]	School N	Student N
Algebra	0.314	520	48,038
Biology	0.361	522	81,127
Chemistry	0.250	373	50,233
Earth Science	0.252	420	22,542
English	0.236	419	23,555
Geometry	0.276	429	44,900
Global Studies	0.310	473	85,541
Math A	0.373	368	18,649
Math B	0.189	240	10,186
Physics	0.283	167	6,152
Trigonometry	0.202	316	14,683
US History	0.327	324	11,276

Table 6e summarizes the HLM results for 2nd year Advanced Placement (AP) examinations:

- The SHSAT also serves as a significant positive predictor for AP history examinations taken in the second year of high school, accounting for approximately 30% of the student-level variation in the two AP history exam performances.

Table 6e. Summary of HLM Results – Second Year Advanced Placement Examinations for All 8th Grade Students Sitting for the SHSAT from 2005 to 2009

Subject	R-square Type Measure for SHSAT[8]	School N	Student N
European History	0.301	42	1,911
World History	0.298	123	9,528

With the adjustment of apparent clustering effects, the rigorous HLM results uniformly confirm that the SHSAT is a statistically significant positive predictor of all selected measures of academic success in high school, which is consistent with the findings based on conventional linear regressions. Despite the differences in estimation methods of the two analytic approaches and calculations of the amounts of variance explained, the SHSAT dependably shows meaningful predictive value for all outcomes analyzed.

Appendix A – Course Letter and Number Grade Point Equivalent Conversion Table

Letter grades	Number grades		Standard Grade Point Equivalent	Advanced Placement Grade Point Equivalent*
	Start	End		
A+	98	100	4.30	5.00
A	95	97	4.00	4.30
A-	93	94	3.75	4.00
B+	88	92	3.50	3.75
B	85	87	3.00	3.50
B-	83	84	2.75	3.00
C+	78	82	2.50	2.75
C	75	77	2.00	2.50
C-	73	74	1.75	2.00
D+	68	72	1.50	1.75
D	65	67	1.00	1.50
D-	60	64	0.00	0.00
F	55	59	0.00	0.00
F-	10	54	0.00	0.00
E+	98	100	4.30	5.00
E	95	97	4.00	4.30
E-	93	94	3.75	4.00
G+	88	92	3.50	3.75
G	85	87	3.00	3.50
G-	83	84	2.75	3.00
S+	78	82	2.50	2.75
S	75	77	2.00	2.50
S-	73	74	1.75	2.00
N+	68	72	1.50	1.75
N	65	67	1.00	1.50
N-	60	64	0.00	0.00
ME	95	100	4.00	4.30
MA	85	94	3.00	3.50
MT	75	84	2.00	2.50
MP	65	74	1.00	1.50
MB	55	64	0.00	0.00

* Additional weight added to course grade for Advanced Placement courses

Adapted from the NYC Department of Education High School Academic Reference Guide - January 2013.

Appendix B – Relationship between the SHSAT and outcome metrics for those accepted/not accepted to specialized high schools

The plots below depicting the relationship between the 2009 SHSAT and the 2011 Algebra Regents are provided to illustrate how limiting the study to students accepted to a specialized high school produces inappropriate conditions to conduct correlational analyses due to restriction of range.

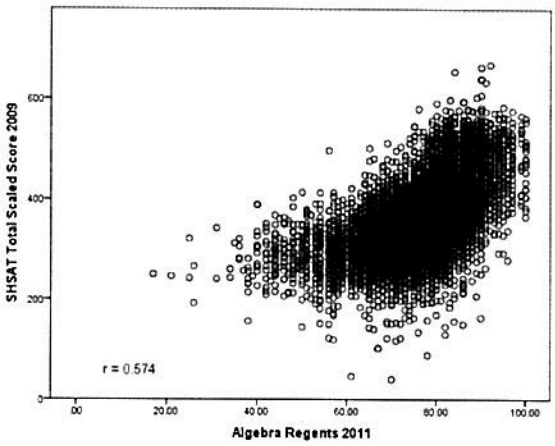

Figure B1. This plot represents the entire 2009 SHSAT cohort, irrespective of acceptance to a specialized high school. The relationship can be characterized as positive and strong, with a Pearson's r correlation coefficient of 0.574. Note that most of the SHSAT scaled scores range between 200 and 600.

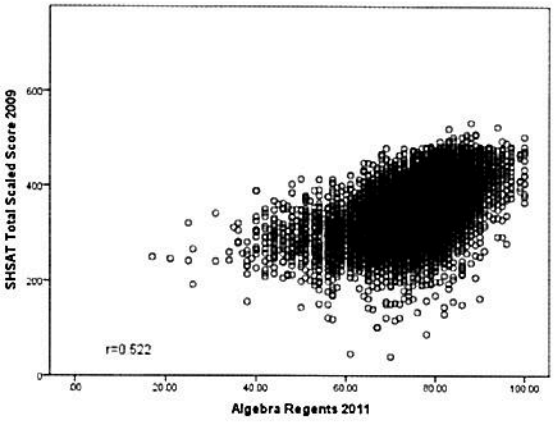

Figure B2. This plot represents the individuals in the 2009 SHSAT cohort that *were not accepted* to a specialized high school. Again, the relationship can be characterized as positive and strong (r=0.522). Note that the majority of the SHSAT scores now range between 200 and 500 while the range of Algebra Regents scores remains relatively unchanged.

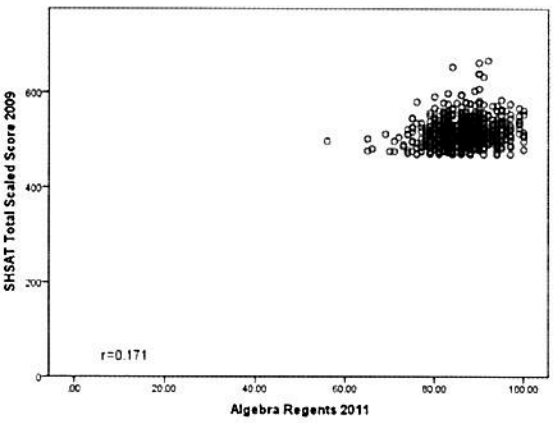

Figure B3. This plot represents the individuals in the 2009 SHSAT cohort that *were accepted* to a specialized high school. The relationship here can be best characterized as negligible (r=0.171). This is primarily due to the restriction in range of SHSAT scores, which now fall largely between 450 and 600. Note that the data plotted herein correspond to the upper right quadrant in the first plot. This is akin to a close-up of a particular feature without the benefit of seeing the entire picture.

As the scatter plots clearly show, the **entire** SHSAT cohort must be included in the study to best determine the predictive value of the SHSAT. The SHSAT will not accurately predict performance only for those accepted to specialized high schools because the admission criterion creates an artificial "floor" for the exam which does not permit the full range of performance to be explored.

Appendix C – List of Specialized and Selective Schools

DBN	School Name
Specialized Schools	
10X445	The Bronx High School of Science
14K449	The Brooklyn Latin School
13K430	Brooklyn Technical High School
05M692	High School for Mathematics, Science and Engineering at the City College
10X696	High School of American Studies at Lehman College
28Q687	Queens High School for the Sciences at York College
31R605	Staten Island Technical High School
02M475	Stuyvesant High School
03M485	Fiorello H. LaGuardia High School of Music & Art and Performing Arts
Selective Schools	
01M450	East Side Community School
03M417	Frank McCourt High School
02M439	Manhattan Village Academy
04M610	Young Women's Leadership School
13K595	Bedford Academy High School
02M413	School of the Future High School
17K543	Science, Technology and Research Early College High School at Erasmus
24Q610	Aviation Career & Technical Education High School
02M376	NYC iSchool
02M414	NYC Museum School
17K590	Medgar Evers College Preparatory School
04M435	Manhattan Center for Science and Mathematics
03M541	Manhattan/Hunter Science High School
02M411	Baruch College Campus High School
02M412	NYC Lab School for Collaborative Studies
28Q680	Queens Gateway to Health Sciences Secondary School
02M418	Millennium High School
05M362	Columbia Secondary School
01M539	New Explorations into Science, Technology and Math High School
22K535	Leon M. Goldstein High School for the Sciences
03M479	Beacon High School
30Q580	Baccalaureate School for Global Education
02M416	Eleanor Roosevelt High School
24Q299	Bard High School Early College Queens
01M696	Bard High School Early College
25Q525	Townsend Harris High School

Selective School List based on the following criteria:

1. Applicants are screened for grades between 80 and 100 in core subject areas
2. Demand for the school is higher than citywide demand; and
3. The school is within the top quartile of high schools in incoming proficiency.