Entrance Examination Scores as a Predictor of Academic Success at the University of Rhode Island.

July 2007

Office of Institutional Research

Introduction

A resolution by the Faculty Senate directing the Admissions Advisory Committee to investigate the possibility of making SATs optional for admission to the University prompted a request for the Office of Institutional Research to conduct a study of "the effectiveness of the SAT and ACT in predicting success at URI including data on gender, race, and major." This report is a summary of the findings.

Entrance examination policy is varied among the thousands of institutions of higher education in the United States. The National Center for Educational Statistics collects data for public and private, not-for-profit, four-year, degree-granting, Title-IV participating institutions. A summary table (Appendix A) indicates that as of 2004 admissions test scores are required at 1395 (63%) and recommended at 112 (5%) of the 2218 schools with available data. At the remaining schools provision of test scores is either not applicable (27%) or not required/recommended (5%). Public institutions (75%) are more likely to require scores than are private ones (58%).

At the University of Rhode Island questions about the validity and usefulness of admission testing are not being raised for the first time. Eighty years ago University president Howard Edwards (1926) in his annual report to the Board of Managers stated that *"the so-called intelligence tests as applied to high school graduates are a delusion"* and *"In its efforts to measure the imponderable, it seeks to emulate the physicist both in method of procedure and in decimal accuracy of results. Under its analysis the power of the human will vanishes, error becomes a constant quantity, repentance is a myth, conversion to new ideals of effort is a fairy tale."* Coincidentally, in the same year Edwards offered his opinion the forerunner of the Scholastic Aptitude Test was first administered to high school students. Appendix B contains a brief history of the test.

In 1958 admissions tests were required at URI for the first time. Nine years later the Office of Institutional Research investigated the "reliability" of SAT scores and high school class rank as a "predictor of academic success or failure." That study (Office of Institutional Research 1967) concluded that low test scores possessed limited usefulness in predicting academic dismissal and cautioned against placing too much reliance on them for the purpose of admission (Appendix C).

Dr. Roberta Zwick, professor of Education at the University of California, Santa Barbara, has written a white paper for the National Association for College Admission Counseling (Appendix D) that provides a comprehensive summary of standardized testing (both SAT and ACT) including current research on its usefulness. It is intended "to help admissions officers, counselors and other stakeholders better understand the role of standardized tests in undergraduate admissions." She finds that current research indicates SAT tests have some predictive validity of first-year grade point average (GPA), especially when combined with other measures such as high school GPA.

Procedure

Two challenges arise with the studies of this type: the first is subjective and the second is operational. In order to predict "success" an acceptable definition of the criterion must be established. Success could be viewed as completion of a degree, but current standards for the measurement of graduation rates allow up to six years from initial enrollment. During such an extended time period, the influence of numerous factors unrelated to academic ability can mask the predictive strength of test scores. A better measure is the cumulative grade point average (GPA) because it quantifies the level of success course-by-course and term-by-term on a standardized basis for all students. Even students who ultimately do not complete a degree can be included in the analysis. Furthermore, as students progress through academic careers its precision continually is refined making it a more accurate indicator of accomplishment. Finally, as a quantitative variable, GPA can be correlated with numeric test scores more precisely than a qualitative one.

The second challenge for this study is limitation in scope. While it is assumed that SAT scores to some degree measure ability to perform college level work, research has shown that other factors are associated significantly with academic success and earning a bachelor's degree (Adelman 2006). Examination of multiple variables and their relationship to SAT scores, although worthwhile, exceeds the goal of this study.

Data

Student data were extracted from the University's student information system for all degree-seeking undergraduates and first professional students (six-year Pharmacy PMD program) enrolled in the fall 2006 semester. Variables include SAT scores (critical reading and mathematics) and demographic data (gender, race/ethnicity, residency,) supplied on applications for admission. Cumulative grade point average, academic level (i.e., class year), and cumulative earned credits, administrative unit, and discipline area of major (U.S. Department of Education 2002) were obtained about one month after the conclusion of the semester. Only data from the most recent year were analyzed because information from previous years is unlikely to differ substantially.

In this analysis, students without SAT scores were excluded. While most students submit SAT scores, some do not: transfers with 24 or more college credits, students admitted through the College of Continuing Education, and applicants submitting ACT scores as an alternative to the SAT. Although ACT scores can be converted to the SAT scale, the tests are not precisely equivalent and the conversion has not been applied here. Unfortunately, the number of students submitting ACT scores alone is too small to allow a separate analysis that is meaningful. For various reasons some students also may not have received grades when end-of-term data were acquired. Most would be students who withdrew or took leave of absences during their first semester. Only records with GPAs have been included.

Analysis

As a starting point, basic statistics were determined for critical reading, mathematics, and composite SAT scores as well as cumulative GPA. These consist of the count, mean, standard deviation, 25th percentile, 50th percentile, and 75th percentile values for the total number of students with scores (n=8796) and reported grades (n=12,131). Besides totals, statistics also were generated for the group categories of gender, race/ethnicity, residency, academic level, cumulative credits level, administrative unit, and discipline area. See Appendix E - Tables 1 and 2 for results.

The next step created distributions (count and percentage) of SAT scores in 50-point ranges for the same group variables (Appendix E - Tables 3 and 4) and in 0.50-point ranges for cumulative GPA (Appendix E - Tables 5 and 6). Source data consisted only of the unit records containing both SAT and GPA information (n=8796).

Analysis of variance of the cumulative GPA by SAT Composite score ranges was performed for each of the groups (Appendix E - Table 7). Values of the R-square statistic for all groups were collected into a table for comparison. This statistic measures on a scale from 0 (none) to 1 (all) the fraction of the total variance in the data that is explained by fitting a linear model – in this case, the part of the variation in cumulative GPA that can be attributed to the influence of SAT scores, or more accurately, the ability to do college level work that is quantified by SAT scores. A high R-square value, for example, would occur with most students who scored high on the SAT also earning high a GPA and most who scored low earning lower GPAs.

Statistical calculations were performed with JMP Statistical Discovery Software (SAS Institute, Inc. SAS Campus Drive, Cary, NC 27513).

Results

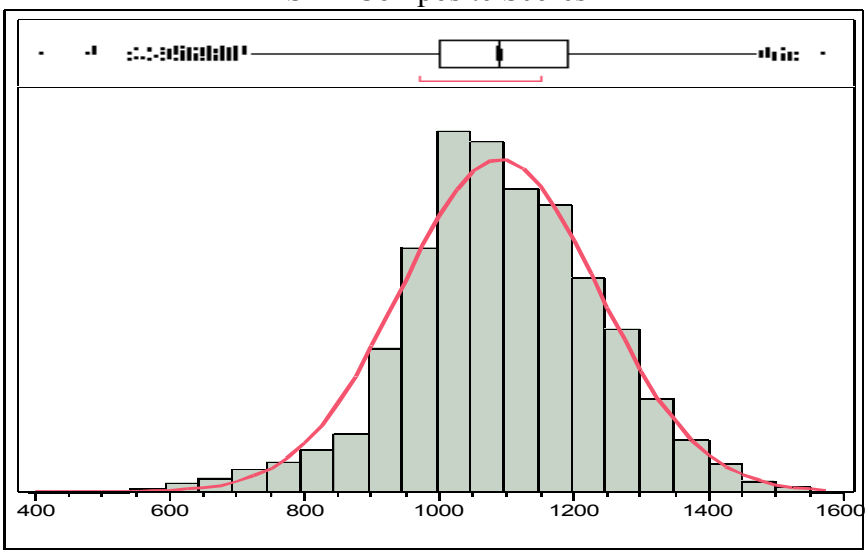
The basic statistics for the SAT Composite show a mean value of 1090 for all students submitting scores. When split by gender, male student have somewhat higher mean scores (1116) than females (1069). Mean scores for Caucasian, Non-resident Alien, and students not reporting race/ethnicity lie above the overall mean. Regional students have the highest mean scores and instate students the lowest. Mean scores rise with increasing class level and accumulated credits. The colleges of Pharmacy and Engineering have students with the highest scores and CCE and Nursing have the lowest. Philosophy, Mathematics, and Engineering are the disciplines with strongest scores; the weakest scores belong to Human Sciences disciplines. Similar patterns are found for the Critical Reading and Mathematics scores. All of these results are neither extraordinary nor unexpected, having been seen in the past.

Basic statistics for cumulative GPA show a mean of 2.88 for all students, with females having a 2.97 and males 2.76. Caucasian, Non-resident Alien, and students not reporting race/ethnicity have GPAs above the overall mean. International and Regional students achieved average grades well above the mean while Instate and Out-of-state are within 0.01 of a point on either side. GPA rises with class year and cumulative credits. The highest average GPAs belong to the

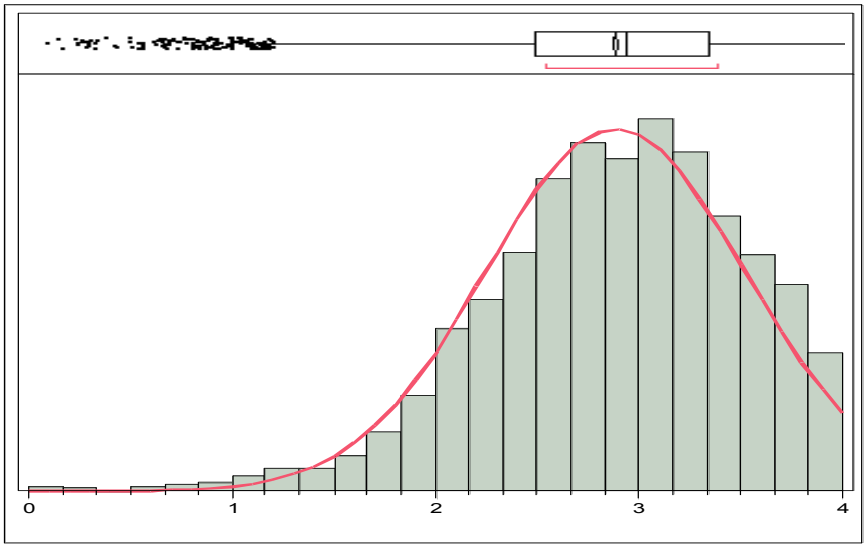
Colleges of Pharmacy, UC-Pharmacy, and Nursing and the discipline areas of Philosophy, Health, and Architecture. The weakest mean GPAs are found with Undecided/Undeclared majors and all University College groups except Pharmacy. The patterns again are familiar and unremarkable.

The both SAT Composite scores and cumulative GPAs have distributions that are close to but not statistically normal. Minor amounts of skewness are found toward the left-side tails and the shapes of the curves are slightly flattened.

SAT Composite Scores

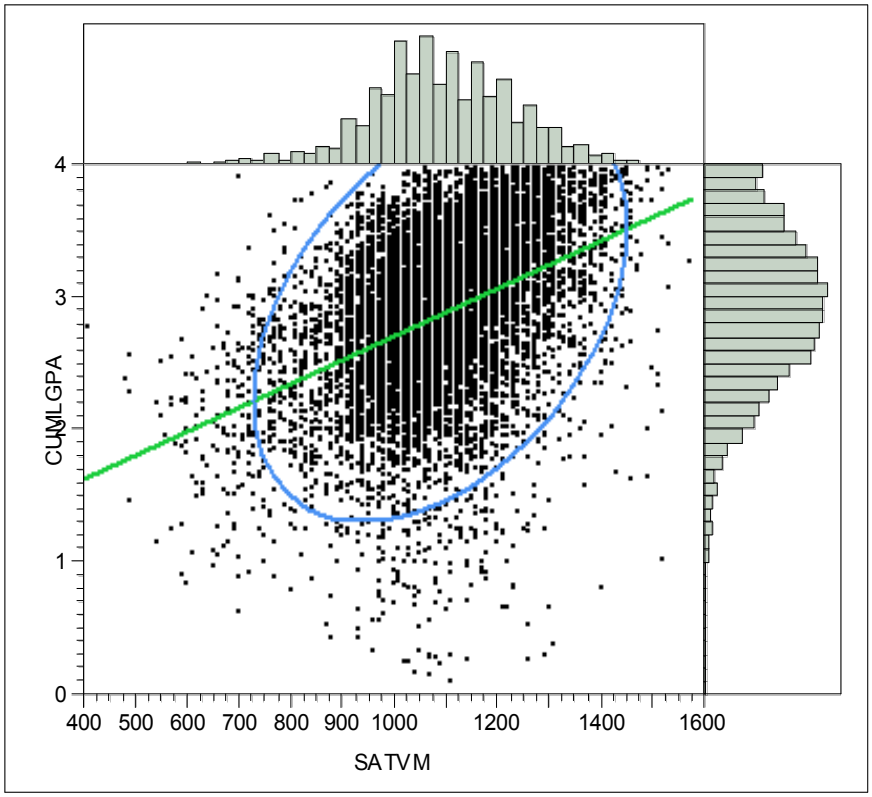


Cumulative GPA



A bivariate scatterplot with a straight-line fit and .95 density ellipse show that the two variables are poorly correlated. The equation of fit is: CUMLGPA = 0.9113484 + 0.0018005*SATVM

Similar results are obtained for the Critical Reading and Mathematics components taken separately.



The large degree of spread in GPA values for each of the SAT values means that the test scores are a weak predictor of cumulative GPA. Using the R-square statistic, less than 18% of the variance in GPA is explained by the SAT (see table on next page). When the demographic groups are evaluated, in most cases only about one-quarter or less of the variance is accounted for. In the few cases (Non-resident Alien, Native American, and International students) where R-square exceeds the still low value of 0.5, the number of students is less than 30 in each group and the results should be accepted with caution.

Within this set of generally low R-square values some tendencies are noticeable. Females exceed males in the composite as well as the critical reading and math tests. Asian students show values twice as large as the overall number. SAT scores are the least predictive of GPA for out-of-state students than for the other residency groups. Predictive strength is slightly better for the middle years (sophomore-junior) than for either freshman or senior and the same pattern holds true for accumulated credit level. Interestingly, for students with over 120 credits R-square is higher than the other ranges. Among the administrative colleges, GPAs for University College Nursing students are the most predictable while for Pharmacy there is almost no association between the variables. When parsed by discipline area, about one-third of the cumulative grades for students majoring in Philosophy, Mathematics, Computer Science, Architecture, History, Health, Human Sciences, and Languages can be explained by their SAT scores.

Proportion of GPA variance explained by SAT score

All Undergraduates and PharmD enrolled in Fall 2006 as of 7 Feb 2007

	Number of Students		R-square		
Variable	Total	Analyzed	VERB	MATH	COMB
All Students	12131	8796	0.1538	0.1249	0.1769
Female	6877	4977	0.2036	0.1908	0.2503
Male	5254	3819	0.1270	0.1300	0.1655
Hispanic	541	380	0.1332	0.1613	0.2063
Non-resident Alien	36	20	0.5157	0.5841	0.7893
African-American	586	376	0.0920	0.0754	0.1216
Asian/Pacific Island	317	221	0.2864	0.3225	0.3504
Caucasian	9110	6753	0.1132	0.0846	0.1339
Native American	41	22	0.4627	0.2709	0.5329
Unknown	1500	1024	0.2061	0.1401	0.2124
In-state	7228	4606	0.1860	0.1552	0.2122
Out-of-State	4375	3793	0.1055	0.0786	0.1265
International	50	28	0.3768	0.5746	0.5708
Regional (NEBHE)	478	369	0.1717	0.1752	0.2211
Freshman	3971	3635	0.1352	0.1086	0.1533
Sophomore	2634	2176	0.1612	0.1462	0.2047
Junior	2723	1855	0.1987	0.1356	0.2167
Senior +	2803	1130	0.1406	0.0894	0.1492
0-30 credits	2988	2779	0.1281	0.0941	0.1383
31- 60 credits	2605	2225	0.1687	0.1681	0.2198
61- 90 credits	2851	2056	0.1813	0.1261	0.2035
91-120 credits	2510	1498	0.1588	0.1003	0.1686
over120 credits	1177	238	0.2089	0.1770	0.2588

Proportion of GPA variance explained by SAT score

All Undergraduates and PharmD enrolled in Fall 2006 as of 7 Feb 2007

	Number of Students		R-square		
Variable	Total	Analyzed	VERB	MATH	COMB
AS	1688	988	0.1520	0.1124	0.1718
BUS	799	586	0.1360	0.1189	0.1828
CCE	534	14	0.2261	0.1606	0.2734
ELSCI	1080	682	0.1549	0.1531	0.2038
ENGR	495	315	0.0888	0.0938	0.1364
HSS	902	586	0.2117	0.1399	0.2400
NURS	414	199	0.1544	0.1602	0.2347
PHARM	354	165	0.0539	0.0414	0.0716
UC_AS	1931	1700	0.1263	0.0806	0.1416
UC_BU	611	570	0.1235	0.0716	0.1298
UC_EL	529	464	0.1268	0.1248	0.1603
UC_EN	458	418	0.1648	0.1776	0.2222
UC_HS	1040	936	0.1877	0.1381	0.2145
UC_NU	280	242	0.2383	0.2107	0.3223
UC_PH	197	194	0.0258	0.0746	0.0777
UC_UN	819	737	0.1289	0.0971	0.1587
1-Agriculture	262	187	0.2016	0.1824	0.2700
3-Environment Sciences	279	203	0.1648	0.1296	0.2309
4-Architecture	71	48	0.3065	0.3708	0.3565
5-Ethnic Studies	11	4	--	--	--
9-Communications	1012	764	0.1147	0.0903	0.1461
11-Computer Science	114	74	0.2396	0.3861	0.3655
13-Education	1069	865	0.1622	0.0967	0.1701
14-Engineering	890	672	0.1367	0.1711	0.1932
16-Languages	116	86	0.1647	0.2853	0.3175
19-Human Sciences	739	467	0.2464	0.2398	0.3220
23-English	304	174	0.1316	0.1248	0.2117
24-General Studies	19	0	--	--	--
26-Biological Sciences	668	512	0.1755	0.2016	0.2492
27-Mathematics	54	33	0.2594	0.2626	0.3775
38-Philosophy	40	27	0.1124	0.3010	0.4346
40-Physical Sciences	213	151	0.2163	0.2377	0.2824
42-Psychology	620	426	0.1517	0.1241	0.1618
44-Public Administration	6	5	--	--	--
45-Social Sciences	514	357	0.2082	0.1000	0.1921
50-Arts	374	273	0.1035	0.1448	0.1772
51-Health	1676	1073	0.2753	0.2969	0.3394
52-Business	1421	1102	0.1223	0.0941	0.1548
54-History	178	93	0.3317	0.1834	0.3419
Undeclared Major	1476	1196	0.1267	0.0736	0.1350
Unknown Major	5	4	--	--	--

Conclusions

The dream of those involved with admissions testing has been to devise a highly reliable instrument to predict how likely it is that a student will be able to perform college level work. Yet after eighty years of refinement such an instrument still does not have the desired precision and remains only one of several tools used in the process of admitting students into college.

The results of this study indicate that "the effectiveness of the SAT and ACT in predicting success at URI" is relatively small. R-square values that can be interpreted as a measure of the "contribution" of SAT scores to cumulative GPA generally lie in the lowest quartile of the range. The reason for this outcome is that a student's academic achievements are determined by many factors in addition to those reflected by a single written examination. Entrance examinations provide some insight to potential future success, but do not have high reliability as general forecasters of GPA. This is not to say, however, that in certain areas when combined with other data, the predictive ability might not be better.

In spite of its limited ability to predict individual grade point averages, the admissions test is the only instrument that is nationally standardized. Unlike high school grades and class rank which have problems of comparison and meaning, the SAT is carefully controlled. It is used with some justification by governmental agencies, evaluating entities, and commercial enterprises as a means to assess the selectivity of institutions. In aggregate, when comparing average scores to retention and graduation rates for a suite of institutions, the SAT has a high positive correlation to these metrics of institutional success.

References

Adelman, C 2006. *The toolbox revisited: paths to degree completion from high school through college*. U.S. Dept. of Education. Washington, D.C.

Edwards, H. 1926. Report of the president of the college in *Bulletin of Rhode Island State College.* Volume 21. Oxford Press, Providence.

Office of Institutional Research. 1967. *A study of the scholastic aptitude tests and high school class standings as predictors of academic success at the University of Rhode Island.* Report No. 22. Kingston, R.I.

U.S. Department of Education. 2002. *Classification of instructional programs: 2000 edition*, NCES 2002-165. Washington, D.C.

Zwick, R. 2007. *College admission testing.* (NACAS white paper) National Association for College Admission Counseling. Alexandria, VA. Retrieved from http://www.nacacnet.org/MemberPortal/ProfessionalResources/Research/Testing.htm

Appendices

A. Frequency distribution of admission test score policy types at four-year, not-for-profit, degree-granting, Title-IV participating institutions. Source: U.S. Department of Education, National Center for Educations Statistics, Integrated Post-secondary Educational Data System.

B. Brief timeline history of the SAT. Source: Public Broadcasting System. 1999. *Secrets of the SAT*. FRONTLINE documentary website. http://www.pbs.org/wgbh/pages/frontline/shows/sats/

C. 1967 report by the URI Office of Institutional Research on the SAT and high school class rank as a predictor of academic success.

D. National Association for College Admission Counseling (NACAS) white paper on college admission testing.

E. Data tables.

- E-1. SAT score basic statistics
- E-2. Cumulative GPA basic statistics
- E-3. Distribution of SAT score counts in 50-point ranges
- E-4. Distribution of SAT score percents in 50-point ranges
- E-5. Distribution of cumulative GPA counts in 0.50-point ranges
- E-6. Distribution of cumulative GPA percents in 0.50-point ranges
- E-7. Analysis of variance of cumulative GPA by SAT Composite score ranges

Appendix A.

Admission Test Score Policy in Fall 2003 or Fall 2004
Four-Year, Not-For-Profit, Degree-Granting, Title-IV Participating Institutions

SOURCE: National Center for Education Statistics,
Integrated Postsecondary Education Data System (IPEDS) Dataset Cutting Tool
http://www.nces.ed.gov/ipeds/find_data/data_cutting_summary.asp

	Private	Public	Total
Required	905	490	1395
Recommended	91	21	112
Neither required nor recommended	111	5	116
Not applicable	454	137	591
Policy unknown	4	0	4
Total Institutions	1565	653	2218

	Private	Public	Total
Required	57.8%	75.0%	62.9%
Recommended	5.8%	3.2%	5.0%
Neither required nor recommended	7.1%	0.8%	5.2%
Not applicable	29.0%	21.0%	26.6%
Policy unknown	0.3%	0.0%	0.2%
Total Institutions	100.0%	100.0%	100.0%

Appendix B.

Brief timeline history of the SAT. Source: Public Broadcasting System. 1999. *Secrets of the SAT*. FRONTLINE documentary website. http://www.pbs.org/wgbh/pages/frontline/shows/sats/

1900 Formation of the College Board

Formation of the College Entrance Examination Board in 1900. This organization, set up by presidents of 12 leading universities administers admissions tests. The purpose is to standardize the admissions process administratively and to force New England boarding schools to adopt a uniform curriculum. In 1901, the first College Boards were conducted. Exams were done in essay format for specific subject areas. Sample some questions from the 1901 test.

1905 Invention of the IQ

A French psychologist, Alfred Binet, is credited with inventing the first IQ test, a test that could measure one's intelligence. Binet's intent was to identify slow learners by determining their mental ages.

World War I Experimentation with army IQ Test

During the first World War, IQ testing advances greatly when Robert Yerkes, a Harvard professor, persuaded the army to let him administer IQ tests to nearly two million recruits. Yerkes wants to use tests to choose officer candidates and help the IQ movement build up a record of statistical evidence. The resulting Alpha and Beta tests mark the first time an IQ test has mass results. The goal of IQ testers is to select the most intelligent people of society, not necessarily to reform education.

1923-1926 Carl Brigham invents the SAT

Carl C. Brigham, who worked with Yerkes on the Army IQ tests, publishes a book, A Study of American Intelligence, on the results. Brigham's book analyzes the findings by race and concludes that American education is declining "and will proceed with an accelerating rate as the racial mixture becomes more and more extensive."

Around this time, Brigham also administers his own objective version of the Army test to Princeton freshmen as well as to applicants to Cooper Union, an all-scholarship technical college in New York City. The College Board then puts him in charge of a committee to develop a test that could be used by a wider group of schools. This test becomes the SAT. In 1926 the SAT is administered to high school students for the first time. Sample some questions from 1926 test.

1933 Conant and the use of SAT for scholarship programs

James Conant is appointed president of Harvard. His assistants, Henry Chauncey and Wilbur Bender are given the task of figuring out a way to select public school students for a Harvard

scholarship program. They travel to Princeton in the end of 1933, where they meett Brigham. Starting in 1934, the SAT is used to select students for Harvard scholarships. A year later, Harvard begins requiring all candidates to take the SAT.

In the same year, IBM machines descended from the Markograph are used to score tests for the NY State Regents and Providence, Rhode Island public schools. By the end of the thirties, the SAT was used as a scholarship test for all Ivy League schools.

1943 The SAT, a tool of the meritocracy

On January 24 Brigham dies at age 52.His death removed the main obstacle for the testing field to be more cohesive. On April 2, the Army-Navy College Qualifying Test is administered to at least 316,000 high school seniors all over the country proving that standardized multiple-choice tests can be given to a mass group. In May, Conant publishes the third in a series of articles for the Atlantic Monthly entitled "Wanted: American Radicals." He wrote that the American radical "believes in equality of opportunity, not equality of rewards."

1948 Creation of Educational Testing Service

Educational Testing Service (ETS), the new testing agency, officially opened for business in Princeton on Jaunary 1st. Henry Chauncey is president and Conant is made chairman of the board. The same year, before ETS is even chartered, a branch office was established in Berkeley, California. Chauncey's hope is to initiate relations with the University of California and get them to adopt the SAT as a requirement.

1952

In 1952, the current structure of questions for the verbal section of the SAT is established: reading comprehension, analogies, antonyms, sentence completion questions.

1957

In 1957, the number of students taking the SAT every year passes half a million.

1959

A new testing organization is formed, American College Testing (ACT) becoming ETS' leading rival.

1960

The University of California system begins requiring applicants to take the SAT, becoming ETS' biggest client.

Appendix C.

1967 report by the URI Office of Institutional Research on the SAT and high school class rank as a predictor of academic success.

Appendix D.

National Association for College Admission Counseling (NACAS) white paper on college admission testing.

Appendix E.

Data tables.

Appendix E-1.

SAT score basic statistics

Appendix E-2.

Cumulative GPA basic statistics

Appendix E-3.

Distribution of SAT score counts in 50-point ranges

Appendix E-4.

Distribution of SAT score percents in 50-point ranges

Appendix E-5.

Distribution of cumulative GPA counts in 0.50-point ranges

Appendix E-6.

Distribution of cumulative GPA percents in 0.50-point ranges

Appendix E-7.

Analysis of variance of cumulative GPA by SAT Composite score ranges