Analyses of the Predictive Validity of the SAT® and High School Grades From 1976 to 1985

Rick Morgan

College Board Report No. 89-7

ETS RR No. 89-37

College Entrance Examination Board, New York, 1989

Rick Morgan is an associate measurement statistician at Educational Testing Service, Princeton, New Jersey.

Acknowledgments

The author wishes to thank ETS internal reviewers William Angoff, Nancy Burton, Neil Dorans, Daniel Eignor, Charles Lewis, Gary Marco, Nancy Petersen, Len Ramist, and Howard Wainer for their comments on various drafts of this paper. External reviewers Lloyd Bond, Bert Green, Rex Jackson, Lyle Jones, Robert Linn, Lorrie Shephard, and John Tukey provided many helpful suggestions. Warren Willingham deserves mention for coordinating the review process.

The author also wishes to thank Karen Carroll, Nancy Feryok, Bruce Kaplan, and Napolean Oleka for their data analyses and also Ted Blew and Susan Bryce for graphics. Charles Lewis deserves special thanks for being ever present with a smile, an idea, and a formula. While many individuals have been mentioned, the author is responsible for any shortcomings in the report.

The College Board is a nonprofit membership organization committed to maintaining academic standards and broadening access to higher education. Its more than 2,600 members include colleges and universities, secondary schools, university and school systems, and education associations and agencies. Representatives of the members elect the Board of Trustees and serve on committees and councils that advise the College Board on the guidance and placement, testing and assessment, and financial aid services it provides to students and educational institutions.

Additional copies of this report may be obtained from College Board Publications, Box 886, New York, New York 10101. The price is $7.

Printed in the United States of America.

CONTENTS

Foreword 1

Abstract 3

Introduction 3

Method 4
Population of Data 4
Method of Analysis 5
Grouping Variables 5

Results 5
Grouping Variables 8

Summary 12

References 12

Appendix A. 13
Yearly Estimates 13
Standard Errors of Yearly Estimates 13

Appendix B. 14
Number of Colleges for Each Pairwise Set of Years 14

FIGURES

1. 10-Year Estimates of the Correlation of SAT with F-GPA Corrected for Restriction of Range 8

2. 10-Year Estimates of the Correlation of SAT with F-GPA Corrected for Restriction of Range 11

3. 10-Year Estimates of the Correlation of SAT with F-GPA Corrected for Restriction of Range 11

4. 10-Year Estimates of the Correlation of SAT with F-GPA Corrected for Restriction of Range 12

TABLES

1. Correlations of SAT and High School Record with F-GPA 4

2. Estimates of the Yearly Means with Associated Standard Errors and Slopes of the Best-Fitting Lines of the Means for Predictor and Criterion Variables 6

3. Estimates of the Yearly Standard Deviations with Associated Standard Errors and Slopes of the Best-Fitting Lines of the Standard Deviations for Predictor and Criterion Variables 6

4. Estimates of the Yearly Correlations for SAT Scores and High School Record with F-GPA with Associated Standard Errors and Slopes of the Best-Fitting Lines for the Correlations. 7

5. Estimates of the Yearly Correlations for SAT Scores and High School Record after Multivariate Correction for Restriction of Range of SAT-V, SAT-M, and High School Record with Associated Standard Errors and Slopes of the Best-Fitting Lines for the Correlations 7

TABLES *(continued)*

6. Estimates of the Average Raw Regression Weights for SAT-V and SAT-M and the Incremental Validity and Partial Correlation after Multivariate Correction for Restriction of Range with Associated Standard Errors and Slopes of Best-Fitting Lines 9

7. Yearly Estimates within College Categories of Multiple Correlation of SAT Scores and High School Record with F-GPA along with Associated Standard Errors and Slopes of the Best-Fitting Lines for the Correlations 9

8. Yearly Estimates within College Categories of Multiple Correlation of SAT Scores and High School Record with F-GPA after Multivariate Correction for Restriction of Range with Associated Standard Errors and Slopes for the Best-Fitting Lines for the Corrected Correlations . . 9

9. Yearly Estimates within College Categories of Correlation of High School Record with F-GPA with Associated Standard Errors and Slopes of the Best-Fitting Lines for the Correlations 10

10. Yearly Estimates within College Categories of Correlation of High School Record with F-GPA after Multivariate Correction for Restriction of Range with Associated Standard Errors and Slopes of Best-Fitting Lines 10

11. Yearly Estimates within College Categories of Multiple Correlation of SAT Scores with F-GPA with the Associated Standard Errors and the Slopes of the Best-Fitting Lines for the SAT Correlations . 10

12. Yearly Estimates within College Categories of Multiple Correlation of SAT Scores with F-GPA after Multivariate Correction for Restriction of Range with Associated Standard Errors and Slopes of the Best-Fitting Lines 11

FOREWORD

The College Board's *Guidelines on the Uses of College Board Test Scores and Related Data* urge institutions to validate data used in the selection process regularly, generally at least every three years. Although studies validating the SAT, as well as other admissions criteria, have been conducted by the College Board and individual institutions for more than 60 years, the creation of a free Validity Study Service (VSS) in 1964 has encouraged many more colleges to conduct such studies. These studies analyze the relationship between admissions criteria (generally SAT scores and either high school class rank or grade-point average) and college performance (generally GPA at the end of the freshman year).

The VSS computes correlation coefficients and related statistics to describe these relationships. While the data from each institutional validity study is primarily intended for use by the individual institutions in formulating and evaluating their admissions policies and selection criteria, the collective data from hundreds of validity studies provide a rich data base for research. Numerous analyses and summaries of these data have been published over the years, with the most recent summary (1964 through 1982) appearing in *The College Board Technical Handbook for the Scholastic Aptitude Test and Achievement Tests* (1984). In the report that follows, Rick Morgan focuses on data collected for freshmen enrolling at selected participating institutions in 1976 through 1985.

In reviewing any summary of validity data, it is important to be aware of the tremendous diversity of colleges and universities in the United States. While averages compiled from many different studies provide a useful supplement to assist colleges in interpreting their own institutional data, it would be misleading to assume that such summaries can be generalized across all institutions. Furthermore, even within many institutions, there is considerable variation across academic disciplines. An additional caveat applies to the data derived from the Validity Study Service since participation is voluntary and the institutions that choose to conduct such studies are not necessarily representative of the more than 3,000 two- and four-year colleges and universities in this country.

Nonetheless, an examination of summarized data over a period of time may make it possible to identify trends that may be affecting the high school to college transition. The admissions process is dynamic, as it accommodates changes in the preparation and expectations of prospective students as well as changes in institutional goals and practices. In recent years, there have been many changes that may have influenced the admissions process. Some of these changes are based on public policy or institutional goals, such as increased emphasis on diversification of enrollment, changes in college "core" requirements, strengthening of academic standards, and increased attention to remediation and retention. Other influences may be more subtle, such as the declining numbers of high school graduates, changes in financial aid policies, freshmen course selection, and grade inflation—to name just a few. Any of these factors might well affect the relationship between the predictors used in admissions decisions (such as SAT scores and high school record) and college grades, often in systematic ways.

In fact, over the years in which VSS data have been compiled, the average correlation between SAT scores and freshman grades has fluctuated, in part because of the different mix of institutions participating in each annual cycle of validity studies, but also probably because of other influences on the admissions process. An example of the latter is that, in recent years, a number of colleges have been enrolling students representing a relatively narrow range of abilities. Whether caused by student self-selection, sophisticated marketing and recruitment activities, or college admissions decisions, a restricted range of abilities of enrolled students results in lower correlation coefficients. This particular phenomenon was noted earlier with the 1964-1982 data in the *Technical Handbook,* and the data were reanalyzed to account for the effects of restriction of range.

In this study, Morgan examines the most recent 10-year period during which the average correlation coefficients between the SAT and college grades have shown a small, but consistent, downward trend. In order to take into account the mix of different institutions in each year of validity data, Morgan conducted most of his analyses on a subgroup of institutions that had conducted validity studies in at least two different years during this period. Variations in the range of student abilities at these institutions were accounted for by the use of multivariate restriction of range adjustment. Because self-reported high school record data generally yield slightly lower correlations than actual high school record data, colleges that had used self-reported data in their study also were excluded from most of the analyses in this report. The results of these analyses, after accounting for the restriction of range of abilities and excluding self-reported high school grades, indicate that the average SAT correlations changed from .51 in 1976 to .47 in 1985. (There was less change during this period for private colleges, small colleges, and more selective colleges.) While the change is not large, any systematic trend calls for further examination.

Several additional studies are currently in progress to further explore a variety of validity issues. These studies are intended not only to better understand factors that may be related to the most recent trends, but also to provide colleges with a framework within which they might understand fluctuations in the specific correlations they observe at the institutional level. As suggested in this report, some of the factors that warrant further exploration are the characteristics of students included in each validity study, the characteristics of predictor information (such as test scores and high school record) and the characteristics of the college grade-point average.

In reviewing this report, it becomes evident that there are a variety of ways to summarize and analyze validity data. This study employs sophisticated procedures to summarize the results of validity studies conducted by different kinds of colleges over time. Of particular interest is Morgan's analysis of pairwise differences to measure changes within colleges during the decade covered by this study. This research represents a valuable contribution to validity research and will undoubtedly form the basis for the development of other methods to provide baseline

validity information for admissions officers, as well as researchers and other interested educators. For example, many colleges use self-reported data in their validity studies and large numbers of studies, therefore, are not included in this report. This report also excludes studies that were conducted on subgroups of students, such as males and females separately, different racial/ethnic groups, and different academic disciplines. Ways to incorporate these studies in regular summaries of VSS data might be addressed in future research.

During the past year, a number of changes have been made to the Validity Study Service. Procedural changes, such as the option to provide college data on floppy disks and a quarterly processing cycle, should increase participation. While the basic analysis that relates test scores and high school record information to college GPA remains the same, a number of additional features have been added. For example, an option to choose SAT analyses based on latest, highest, or average SAT scores is now available.

Colleges and universities that have not conducted validity studies within the past few years should consider doing so. The resulting data will provide useful information for evaluating and modifying admissions policies, and enable admissions officers to make better selection decisions. In addition, more current summaries of validity data can be developed to assist all educators in better understanding the variety of factors and influences that affect the transition from high school to college.

Gretchen Wyckoff Rigol
Executive Director, Access Services
The College Board

New York
August 1989

ABSTRACT

Since 1964, colleges have been sending data to the College Board Validity Study Service in order to determine the degree to which measures used in admissions predict college performance. These studies have allowed for the monitoring of general trends in the relationship of SAT scores and high school grades with freshman grades. Beginning in the early-1970s and continuing to the mid-1970s, the observed average correlations of SAT scores with freshman grades increased dramatically. In the decade that followed, the strength of this relationship gradually declined. In order to determine to what extent the observed variations in the average correlations might be associated with the composition of colleges conducting validity studies each year, validity data were analyzed for the enrolling classes of 1976 to 1985. Estimates of change, based on a repeated measures estimation approach utilizing data from colleges conducting more than one predictive validity study within the 10-year period, and correcting for the restriction of range in SAT scores and high school record, indicated a general linear trend from 1976 to 1985. The correlations of SAT scores with first-year college grades declined from about .51 to about .47, for a change of approximately .04. There was less change during this period for private colleges, small colleges, and more selective colleges.

INTRODUCTION

Many colleges conduct studies to determine the extent to which measures used in admissions predict college performance. Beginning with the enrolling class of 1964, the College Board established the Validity Study Service (VSS), to which colleges may send data for an analysis of the predictive validity of college admissions measures. The analysis of the data, conducted for the College Board by the Educational Testing Service (ETS), provides least squares regression estimates for sets of predictors of freshman grade point average (F-GPA). Following the completion of all the requests for data analysis for an enrolling year, average multiple correlations of SAT scores and high school grade information with F-GPA are computed for all studies in which SAT-Verbal, SAT-Mathematical, and a measure of high school record were used to predict F-GPA for the entire freshman class. These average correlations are the estimates of predictive validity for each enrolling year.

Two major problems with drawing conclusions about trends in predictive validity using these yearly averages are the potential unrepresentativeness of the group of academic institutions conducting validity studies in any given year and the changes in the composition of this group from year to year. Consequently, yearly fluctuations in the average validity coefficients could result from different colleges choosing to participate in different years, as well as from several other factors. These other factors include 1) changes over time in the characteristics of students choosing to attend college, 2) changes in college admissions practices, 3) changes in the test, and 4) changes in grading practices or in the meaning of the criterion.

Yearly average VSS validity data taken from Ramist (1984) and more recent data made available by the VSS program are summarized in Table 1. The table shows the size of the average yearly multiple correlation of SAT-V, SAT-M, and high school record with F-GPA, the average yearly multiple correlation of SAT-V and SAT-M with F-GPA, and the average correlation of high school academic record with F-GPA for the enrolling classes of 1964 to 1985. The observed data suggest a steady gradual decline in the predictability of college performance from the optimal weighting of high school record and SAT scores (the multiple correlation). The trend in average yearly correlations for high school record with freshman grades is generally parallel to the multiple correlation. However, it does not necessarily follow that these observed data represent true changes in predictive validity, because the participating colleges change each year. Also, one cause for these trends can be traced to the 1972 introduction of the option of using student-reported high school data from the Student Descriptive Questionnaire (SDQ). Prior to 1972, colleges provided actual high school record information. Because self-reported data generally have lower correlations, averages based solely on the studies supplying actual high school record information are also shown in Table 1. As expected, correlations involving college-reported high school record are higher, with the average correlation of F-GPA with high school record leveling off in the last 10 years.

Two distinct trends in the average yearly correlation of SAT scores with freshman grades are present in Table 1. Beginning in the early 1970s and continuing to the mid-1970s, the average correlations of SAT scores with freshman grades increased. In the years that follow, this relationship returned to levels approximating the levels of the mid-1960s. In assessing all the correlations, it should be noted that the average number of studies on which the yearly averages are based has increased. From 1964 to 1969, the correlations are based on an average of 53 studies per year, while the average number of studies is 96 for the period from 1970 to 1975, and 165 for the years following 1975. The effect of the increase in the number of colleges on the correlations is not known.

Because the fluctuations in college participation in the VSS could be influencing the trends in average correlations presented in Table 1, this study was designed to explore these trends after removing effects resulting from differential college participation. Specifically, this study investigates the predictive validity of the SAT and measures of high school record from 1976 to 1985. This 10-year period begins with data from the peak years for the average correlation of SAT scores with F-GPA and ends with the most current set of data.[1] Most of the emphasis in

1. Because of processing changes associated with the introduction of a new Student Descriptive Questionnaire in the fall of 1985, participation in VSS was more difficult than usual for colleges conducting validity studies for the enrolling classes of 1986 and 1987. Consequently, the number of studies conducted for those years was not sufficient for analysis. Beginning with the enrolling class of 1988, the relatively easy option of utilizing data from the SDQ for validity studies is once more available to colleges.

Table 1. Correlations of SAT and High School Record with F-GPA

	All Available VSS Data[1]			Studies Using Actual High School Record		
Year	Multiple Correlation[2]	SAT Correlation[3]	High School Record[4]	Multiple Correlation	SAT Correlation	High School Record[4]
1964	.60	.44	.52	.60	.44	.52
1965	.60	.40	.54	.60	.40	.54
1966	.58	.39	.52	.58	.39	.52
1967	.57	.40	.52	.57	.40	.52
1968	.59	.42	.52	.59	.42	.52
1969	.57	.41	.51	.57	.41	.51
1970	.55	.39	.49	.55	.39	.49
1971	.54	.38	.48	.54	.38	.48
1972	.56	.41	.49	.56	.41	.49
1973	.57	.43	.49	.56	.43	.49
1974	.58	.47	.50	.58	.46	.50
1975	.56	.44	.48	.56	.45	.48
1976	.56	.45	.47	.57	.45	.50
1977	.56	.44	.49	.56	.43	.49
1978	.54	.42	.46	.56	.42	.49
1979	.55	.42	.47	.57	.42	.50
1980	.54	.41	.46	.56	.42	.48
1981	.54	.38	.47	.56	.39	.50
1982	.54	.40	.47	.58	.41	.52
1983	.53	.39	.46	.55	.40	.49
1984	.51	.37	.45	.54	.38	.48
1985	.52	.37	.47	.56	.38	.50

1. From Ramist (1984).
2. Average multiple correlation of SAT-V, SAT-M, and high school record.
3. Average multiple correlation of SAT-V and SAT-M with F-GPA.
4. Average correlation of high school record.

this study is on the SAT, although an understanding of overall predictive validity trends requires analysis of the trends involving high school record as a predictor of F-GPA. In order to analyze within-college trends, the study uses data from colleges that conducted more than one predictive validity analysis in the 10-year period. In the first part of the study, a pairwise difference estimation technique is used to provide estimates of trends in several measures of predictive validity. The second part of the study applies the same analysis to groups of colleges defined by various college characteristics in order to investigate whether the predictive validity trends are related to college characteristics.

As a result of the pairwise estimation technique, differential yearly participation in VSS is much better controlled than the observed data found in Table 1. The pairwise difference estimation technique, however, can not account for fluctuations in correlations that may result from other factors, such as changes in the composition of students attending participating colleges or from changes in the grading or the composition of courses taken by the students during their freshman year. Nor can the results of the study be generalized to colleges that did not participate in VSS.

METHOD

Population of Data

Because the study focuses on within-college trends, the data set included only colleges that conducted analyses in at least 2 of the 10 years under study. The study sample included only analyses that had at least 75 complete student records and were based on the entire freshman class of the college (rather than a subset based on gender, ethnicity, or major). The selected studies utilized SAT-V, SAT-M, and one of two college-reported measures of high school record (high school grades on a 0-4 scale and class ranks converted to a 20-80 scale[2]) to predict F-GPA, which was also limited to a 0-4 scale. In order not to underestimate the relationship of F-GPA with high school record, the data set excluded self-reported high school record information. These constraints on the data permitted a more accurate analysis of predictive validity change over time than was provided in Table 1. The sample includes 222 colleges that conducted 778 different validity studies over the 10-year period.

For each of the 778 analyses included in this study, the mean, standard deviation, and correlations among the three predictor variables and F-GPA were extracted from the VSS data base. These data were merged with college demographic information provided from the 1987 College Board Research Tape C, which contains information detailing student composition, academic programs, and campus life for both foreign and domestic colleges.

2. Class rank information is converted by VSS to a normally distributed 20-to-80 scale. The scale has a mean of 50 and a standard deviation of 10 for all high school seniors.

Method of Analysis

An optimal way of estimating trends in predictive validity would be to examine only those colleges that have conducted analyses in each of the 10 years under study. However, the number of such colleges was too small to produce meaningful validity estimates. An alternative is to examine changes over time in the results of colleges with multiple analyses. The repeated measures estimation procedure outlined by Goldman and Widawski (1976) was used to produce estimates for each of the 10 years, based on pairwise estimates of change. In this procedure estimates of the average difference in measures between a pair of years were produced for each of the 45 nonredundant pairwise sets of years (10 years taken 2 at a time). These pairwise difference estimates were based only on colleges with analyses for both of the years in the pair. For example, the 1976-1985 pairwise change estimates are the mean change for all colleges with analyses in 1976 and 1985. In order to produce a deviation estimate for any year, all pairwise estimates involving the year, including the zero difference of the year with itself, were averaged. A deviation estimate for a year is analogous to a main effect deviation estimate in analysis of variance. These deviation estimates were then added to a mean over all years to provide yearly estimates. Appendix A provides formulas used to obtain the yearly estimates and their associated standard errors.

The pairwise difference estimation technique produced yearly estimates of the means and standard deviations of the predictor variables and F-GPA to permit an examination of the trends of these measures on predictive validity results. Both raw correlations and correlations corrected for the multivariate restriction of range SAT-V, SAT-M, and high school record[3] as well as regression weights, were also estimated. The incremental validity, which is the difference between the multiple correlation of F-GPA with SAT scores and high school record and the correlation of F-GPA and high school record alone, was estimated.The partial correlation of the SAT with F-GPA, after partialling out the high school record was also estimated.

3. The multivariate restriction of range formula was supplied by Charles Lewis (personal communication, April 1989). The formula is equivalent to the formula for explicit selection found in many sources, such as *Theory of Mental Tests* by H. Gulliksen (pp. 165-166). Reference correlations and standard deviations were based on the data from the VSS data base for the enrolling class of 1985. As a result, the corrected correlations refer to the population of enrolling students, rather than the population of those who took the SAT. The reference group had average scores for both SAT-V and SAT-M approximately 40 points higher than the population of students administered the SAT in 1985. The reference group was also less variable, with SAT-V and SAT-M standard deviations approximately 10 points less than the population of students administered the test. The assumed population standard deviation for those colleges supplying high school GPA was 100 for SAT-V, 107 for SAT-M, and .63 for high school GPA. The assumed correlation was .60 for SAT-V with SAT-M, .38 for SAT-V with high school GPA, and .45 for SAT-M and high school GPA. The assumed population standard deviation for those colleges supplying high school rank was 103 for SAT-V, 107 for SAT-M, and 7.9 for high school rank converted to a 20-80 scale. The assumed correlation was .63 for SAT-V with SAT-M, .37 for SAT-V with high school rank, and .42 for SAT-M and high school rank.

Grouping Variables

The second part of the study examined the predictive validity within three categorizations of colleges. Validity estimates using the repeated measures procedure were derived for the following categories:

Type of college control:	Public colleges
	Private colleges
Selectivity:	Colleges with SAT V+M at or above 950
	Colleges with SAT V+M below 950
Freshman class size:	Colleges with 750 or more freshmen
	Colleges with fewer than 750 freshmen.

When the categorization of a college changed across years, the data from non-matching category years were not used in the analysis. For example, if a college had a SAT V+M mean of 940 in 1978, 960 in 1982, and 970 in 1983, the data from 1978 were not used in the analysis.

RESULTS

The average SAT-V score for the analyses in the study was 455, while the mean SAT-M was 493. The average within-college standard deviation was 86 for SAT-V and 90 for SAT-M. The number of pairs of analyses was 1,408. The pair of years with the most analyses was 1981-1982, with 49 colleges conducting analyses utilizing college-reported high school record information both years. The pair of years with the fewest analyses was 1976-1985, with 14 colleges. The median number of colleges for the pairs of years was 30. Appendix B provides the number of colleges for each pair of years for all analyses in this study.

The tables to follow present estimates of validity-related measures for each of the 10 years. The standard errors, which provide measures of the precision of these estimates, are shown in parentheses. Most of the focus of the interpretation of the data is on the linear trends provided by slope coefficients. These slopes refer to the linear regression lines characterizing year-to-year change. The size of the standard errors for the slopes provide a measure of the precision of the slope. The absolute size of the slope is important, with slopes close to zero indicating little or no trend. Slopes that differ from zero by a few standard errors and whose values represent meaningful change are interpreted as suggestive of a trend. Slopes that differ from zero by less than a standard error and whose values represent minimal change are interpreted as representing no systematic trend. Slopes that meet neither criterion are interpreted with caution. Multiplying the slope coefficient by nine provides an estimate of the amount of change in a measure for the 10-year period.

Table 2 presents the estimates for the means of SAT-V, SAT-M, high school rank (converted to a 20-80 scale), high school GPA (0-4), and F-GPA for each of the 10 years. The table shows no systematic trends in the means.

Table 3 shows the yearly averages of the within-college standard deviations for SAT-V, SAT-M, high school rank, high school GPA, and F-GPA. The negative slope for the within-college standard deviation estimates of SAT-V, SAT-M, and F-GPA indicate lower standard deviations for the mid-1980s than for the mid-1970s. Little systematic change occurred in the standard

Table 2. Estimates of the Yearly Means with Associated Standard Errors and Slopes of the Best-Fitting Lines of the Means for Predictor and Criterion Variables

Year	*SAT-V*	*SAT-M*	*High School Rank*	*High School GPA*	*F-GPA*
1976	460(4.4)	497(5.1)	58.2(.54)	3.03(.04)	2.59(.03)
1977	459(4.0)	495(4.8)	58.1(.43)	3.04(.04)	2.58(.02)
1978	459(4.1)	495(4.5)	58.1(.41)	3.05(.04)	2.60(.02)
1979	455(3.8)	491(4.4)	58.2(.34)	3.03(.04)	2.61(.02)
1980	452(3.9)	490(4.3)	57.9(.38)	3.04(.03)	2.58(.02)
1981	451(4.0)	491(4.3)	58.0(.38)	3.05(.03)	2.59(.02)
1982	454(3.7)	490(4.6)	58.2(.42)	3.03(.03)	2.58(.02)
1983	452(4.0)	491(4.3)	58.3(.35)	3.01(.03)	2.55(.02)
1984	454(4.0)	494(4.6)	57.9(.44)	3.02(.04)	2.55(.03)
1985	458(4.1)	499(4.6)	58.0(.35)	3.02(.04)	2.57(.02)
Slope	−.60	−.14	−.01	−.002	−.005
S. E. Slope	.35	.44	.06	.004	.003

Table 3. Estimates of the Yearly Standard Deviations with Associated Standard Errors and Slopes of the Best-Fitting Lines of the Standard Deviations for Predictor and Criterion Variables

Year	*SAT-V*	*SAT-M*	*High School Rank*	*High School GPA*	*F-GPA*
1976	89(.91)	93(1.25)	7.5(.10)	.51(.01)	.73(.01)
1977	88(.91)	91(.89)	7.3(.13)	.52(.02)	.73(.01)
1978	86(.86)	89(.95)	7.4(.11)	.52(.02)	.71(.01)
1979	87(1.13)	90(.98)	7.5(.14)	.53(.02)	.71(.01)
1980	86(.95)	89(1.02)	7.5(.12)	.52(.02)	.70(.01)
1981	85(.86)	88(.99)	7.3(.16)	.50(.01)	.69(.01)
1982	85(.85)	88(1.25)	7.5(.15)	.51(.01)	.68(.01)
1983	85(.84)	89(.86)	7.5(.12)	.52(.01)	.70(.01)
1984	86(.99)	89(.94)	7.4(.12)	.52(.02)	.71(.01)
1985	86(.97)	89(1.20)	7.5(.15)	.51(.01)	.69(.01)
Slope	−.33	−.35	.002	−.000	−.004
S. E. Slope	.08	.11	.015	.002	.001

deviations in either of the two measures of high school record. The reduction in the average standard deviation for SAT-V and SAT-M would be expected to result in lower correlations of SAT scores with F-GPA. The increased homogeneity, with regard to SAT scores of the students within the colleges in this study, has not resulted from increased homogeneity in the total test-taking population. Based on data from the College Board (1976 and 1985), in the population of SAT examinees, the standard deviation of SAT-V increased from 110 to 111, while the standard deviation of SAT-M decreased from 120 to 119 between 1976 and 1985. The reduction of the standard deviations for colleges in this study might be a result of different recruitment practices that identify a more homogenous group of applicants, increased emphasis on the SAT in college admissions, or self-selection in the colleges that students apply to and choose to enroll. Reductions in the standard deviation of F-GPA could result in lower correlations of any variable with F-GPA and lead to lowered average regression weights for predictors of F-GPA.

The yearly averages for the correlations of the SAT and high school record with F-GPA are presented in Table 4. High school rank and high school grade information have been combined. The table indicates no systematic trend in the correlation of high school record with F-GPA. The multiple correlation of SAT scores and high school record with F-GPA shows a per year slope of -.0021, which may not be different from zero. The multiple SAT correlation shows a linear trend with a slope of -.0068, the linear trend for the SAT-M correlations has a slope of -.0051, and the slope for SAT-V is -.0071. The correlation estimates for SAT-V and SAT-M are at very similar levels, ranging from .39 to .32. The repeated measures estimates of change are smaller than are the change estimates based on the observed correlations found in Table 1.

The yearly estimates for the correlations adjusted for restriction of range are presented in Table 5. The restriction of range adjustment controls for year-to-year changes in the distributions of SAT scores and high school record for the colleges being

Table 4. Estimates of the Yearly Correlations for SAT Scores and High School Record with F-GPA with Associated Standard Errors and Slopes of the Best-Fitting Lines for the Correlations

Year	*SAT-V*	*SAT-M*	*Multiple SAT*	*High School Record*	*Multiple Correlation*
1976	.39(.01)	.38(.01)	.45(.01)	.52(.01)	.58(.01)
1977	.37(.01)	.38(.01)	.44(.01)	.51(.01)	.57(.01)
1978	.37(.01)	.38(.01)	.44(.01)	.50(.01)	.57(.01)
1979	.36(.01)	.36(.01)	.42(.01)	.50(.01)	.56(.01)
1980	.35(.01)	.37(.01)	.42(.01)	.49(.01)	.56(.01)
1981	.35(.01)	.33(.01)	.40(.01)	.51(.01)	.56(.01)
1982	.36(.01)	.35(.01)	.41(.01)	.51(.01)	.57(.01)
1983	.34(.01)	.36(.01)	.41(.01)	.50(.01)	.56(.01)
1984	.32(.01)	.35(.01)	.40(.01)	.48(.02)	.55(.01)
1985	.32(.01)	.32(.01)	.38(.01)	.52(.01)	.57(.01)
Slope	−.0071	−.0051	−.0068	−.0009	−.0021
S. E. Slope	.0015	.0012	.0014	.0018	.0012

Table 5. Estimates of the Yearly Correlations for SAT Scores and High School Record after Multivariate Correction for Restriction of Range of SAT-V, SAT-M, and High School Record with Associated Standard Errors and Slopes of the Best-Fitting Lines for the Correlations

Year	*SAT-V*	*SAT-M*	*Multiple SAT*	*High School Record*	*Multiple Correlation*
1976	.45(.01)	.46(.01)	.51(.01)	.57(.01)	.64(.01)
1977	.45(.01)	.46(.01)	.51(.01)	.57(.01)	.64(.01)
1978	.45(.01)	.46(.01)	.51(.01)	.56(.01)	.64(.01)
1979	.43(.01)	.44(.01)	.49(.01)	.56(.01)	.62(.01)
1980	.44(.01)	.46(.01)	.50(.01)	.55(.01)	.63(.01)
1981	.44(.01)	.44(.01)	.49(.01)	.58(.01)	.65(.01)
1982	.44(.01)	.45(.01)	.50(.01)	.58(.01)	.65(.01)
1983	.43(.01)	.44(.01)	.49(.01)	.56(.01)	.63(.01)
1984	.42(.01)	.44(.01)	.48(.01)	.56(.02)	.63(.01)
1985	.41(.01)	.42(.01)	.47(.01)	.58(.01)	.64(.01)
Slope	−.0038	−.0033	−.0040	−.0005	−.0010
S. E. Slope	.0013	.0010	.0012	.0015	.0011

studied. The restriction of range adjustment also makes colleges with different degrees of selectivity more comparable, so that averages over colleges are more meaningful. The table indicates no systematic trends in the multiple correlation of SAT scores and high school record with F-GPA or in the correlation of high school record with F-GPA. Although a decline still exists, the trends for the SAT correlations in the unrestricted population are not as steep as those found in the restricted samples. The multiple SAT correlation has a negative slope of -.0040. SAT-M has a slope of -.0033, while SAT-V has a slope of -.0038. These slopes are approximately 40 percent less than the slopes based on the unadjusted correlations. For 9 of the 10 years, the SAT-M correlation with F-GPA is slightly higher than the SAT-V correlation with F-GPA. Figure 1 indicates that after controlling for institutional variation in VSS studies and internal variation of SAT scores and high school record, the multiple SAT correlation with F-GPA declines by .04 from 1976 to 1985, rather than the .08 indicated in Table 1. Figure 1 also utilizes vertical line segments to indicate bounds of one standard error for each yearly estimate.

Table 6 provides the estimates of the average regression weights for SAT-V and SAT-M that result from the multiple regression of SAT-V, SAT-M, and high school record on F-GPA. Table 6 also provides both the partial SAT correlation with F-GPA controlling for high school record and the incremental validity. Both the partial correlation and the incremental validity are based on the multivariate corrections for restriction of range. The regression weights for both SAT-V and SAT-M indicate a linear decline, with the weights being larger for SAT-V. One reason that SAT-V weights are larger than SAT-M weights is because SAT-M is more highly correlated with high school record than is SAT-V. A decline is present for the partial correlation (slope = -.0059), while for the incremental validity the slope of -.0014 may not be different from zero.

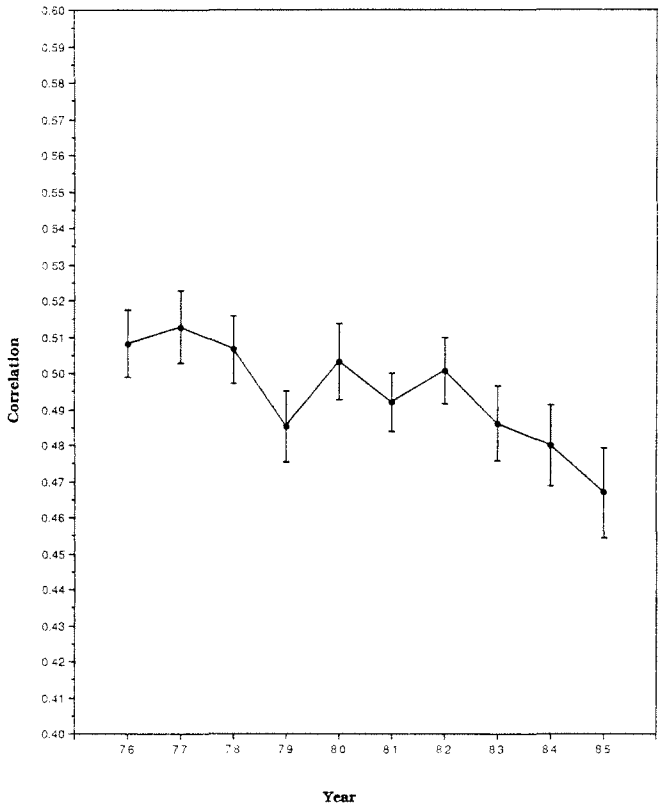

Figure 1. 10-Year Estimates of the Correlation of SAT with F-GPA Corrected for Restriction of Range

For the 222 colleges conducting multiple studies from 1976 to 1985, there was a reduction in the predictive validity of the SAT. The repeated measures estimates of change after correcting for restriction of range, however, produced estimates that show considerably less decline than those in Table 1. The corrected repeated measures estimates of change are half the size of the Table 1 estimates. The finding suggests that a large part of the noted overall decline in correlations can be attributed to increased within-college restriction of range of the predictor variables and to differential college participation in the VSS.

Grouping Variables

The categorization of the colleges found that 81 colleges were public colleges and 141 were private colleges; 143 colleges had freshman enrollments less than 750 students and 79 colleges had enrollments of at least 750; and 119 colleges had pairs of SAT V+M means less than 950 and 108 colleges had pairs of SAT V+M means of at least 950 (five colleges had pairs of means in both SAT V+M categories). The grouping of the data based on type of college control, freshman class size, and average of SAT V+M score also resulted in smaller numbers of colleges providing pairwise estimates for the categories. The smallest number of colleges conducting analyses for a pair of years was two, for both public colleges and less selective colleges in the 1976-1985 pair of years. The median number of colleges for pairs of years was 22 for private colleges, 10 for public colleges, 21 for smaller colleges, 9 for larger colleges, 13 for more selective colleges, and 14 for less selective colleges. As a result of the smaller number of colleges for each pairwise estimate, the standard errors for estimates of both the slope and the yearly correlations are larger. In Appendix B the table for public colleges shows that the sample size is small for several of the pairs of years associated with 1985. As a result, the estimates for public colleges for the year 1985 are problematic.

Table 7 provides estimates of the multiple correlation of SAT scores and high school record with F-GPA, for six college categorizations. The decline in the multiple correlation is largest for public colleges (slope = -.0041). None of the other linear trends are large enough, given their standard errors, to be distinguished from zero. The highest multiple correlations are found for private colleges, smaller colleges, and less selective colleges.

Table 8 provides the yearly estimates of the multiple correlation corrected for restriction of range. Again, public colleges have the largest decline in the correlation with a slope of -.0030. However, this slope like the other five slopes, is not large enough to be judged different from zero.

The estimates within each categorization for the correlation of high school record with F-GPA are found in Table 9. No trends appear for any of the six categories, with the correlations holding relatively steady for the 10 years under study. The estimates of the correlation of high school record with F-GPA after correction for restriction of range are found in Table 10. As with Table 9, no clear trends appear for any category.

Table 11 shows the yearly estimates of the SAT multiple correlation with F-GPA. All categorizations, except more selective colleges, show decline. The slope for private colleges (-.0054) is less than half the slope for public colleges (-.0113). The slope at more selective colleges (-.0045) is less than half the slope at less selective colleges (-.0097). The trends in the multiple correlation for the categories based on freshman class size are not greatly different.

The multiple SAT correlations with F-GPA after correcting for restriction of range are shown in Table 12. While declines are still present, the slopes of the declines are not as steep as in Table 11.

Figure 2 displays the yearly averages for the SAT corrected correlation with F-GPA for public and private colleges. The figure shows that the SAT correlation is higher for private colleges and that the difference between the average correlations has widened from 1976 to 1985. The large drop in the public college line in 1985, in part, is due to the small sample sizes for the pairwise estimates involving 1985. The standard error for this point is twice as large as any other standard error associated with the estimates in the figure. Figure 3 displays the yearly SAT correlations according to the size of the freshman class. Smaller colleges have consistently higher SAT correlations, with the size of these correlations appearing to decline less than for larger colleges. Figure 4 displays the yearly SAT correlations for colleges grouped according to average SAT V+M score. A sharper decline is present for less selective colleges.

Table 6. Estimates of the Average Raw Regression Weights for SAT-V and SAT-M and the Incremental Validity and Partial Correlation after Multivariate Correction for Restriction of Range with Associated Standard Errors and Slopes of Best-Fitting Lines

Year	*SAT-V*	*SAT-M*	*Corrected Partial Correlation*	*Corrected Incremental Validity*
1976	.0014(.0001)	.0011(.0001)	.35(.01)	.072(.005)
1977	.0014(.0001)	.0012(.0001)	.36(.01)	.075(.005)
1978	.0014(.0001)	.0011(.0001)	.35(.01)	.073(.005)
1979	.0013(.0001)	.0010(.0001)	.33(.01)	.068(.005)
1980	.0013(.0001)	.0012(.0001)	.35(.01)	.075(.004)
1981	.0014(.0001)	.0009(.0001)	.33(.01)	.063(.004)
1982	.0013(.0001)	.0010(.0001)	.34(.01)	.068(.004)
1983	.0012(.0001)	.0010(.0001)	.33(.01)	.066(.004)
1984	.0012(.0001)	.0011(.0001)	.32(.02)	.069(.010)
1985	.0011(.0001)	.0008(.0001)	.29(.01)	.056(.005)
Slope	−.000028	−.000025	−.0059	−.0014
S. E. Slope	.000013	.000009	.0017	.0008

Table 7. Yearly Estimates within College Categories of Multiple Correlation of SAT Scores and High School Record with F-GPA along with Associated Standard Errors and Slopes of the Best-Fitting Lines for the Correlations

Year	*Private Colleges*	*Public Colleges*	*Less than 750 Freshmen*	*750+ Freshmen*	*Less than 950 SAT V+M*	*950+ SAT V+M*
1976	.59(.02)	.56(.01)	.61(.02)	.53(.02)	.62(.02)	.54(.02)
1977	.59(.01)	.54(.01)	.59(.01)	.54(.02)	.59(.01)	.55(.02)
1978	.59(.01)	.53(.01)	.60(.01)	.51(.01)	.59(.01)	.54(.01)
1979	.59(.01)	.50(.02)	.59(.02)	.51(.01)	.59(.02)	.52(.02)
1980	.58(.01)	.52(.01)	.58(.01)	.52(.01)	.58(.01)	.53(.01)
1981	.58(.01)	.53(.01)	.59(.01)	.51(.01)	.59(.01)	.53(.02)
1982	.60(.01)	.52(.02)	.60(.01)	.53(.02)	.60(.01)	.53(.01)
1983	.59(.01)	.50(.02)	.58(.01)	.51(.02)	.57(.02)	.54(.01)
1984	.56(.02)	.53(.02)	.58(.01)	.49(.01)	.58(.02)	.52(.02)
1985	.59(.01)	.50(.01)	.59(.01)	.52(.01)	.57(.02)	.54(.01)
Slope	−.0012	−.0041	−.0020	−.0026	−.0034	−.0007
S. E.	.0016	.0016	.0015	.0016	.0023	.0020

Table 8. Yearly Estimates within College Categories of Multiple Correlation of SAT Scores and High School Record with F-GPA after Multivariate Correction for Restriction of Range with Associated Standard Errors and Slopes for the Best-Fitting Lines for the Corrected Correlations

Year	*Private Colleges*	*Public Colleges*	*Less than 750 Freshmen*	*750+ Freshmen*	*Less than 950 SAT V+M*	*950+ SAT V+M*
1976	.65(.02)	.64(.01)	.66(.01)	.61(.02)	.67(.01)	.62(.02)
1977	.65(.01)	.62(.01)	.66(.01)	.61(.01)	.65(.01)	.63(.02)
1978	.65(.01)	.61(.01)	.66(.01)	.58(.02)	.64(.01)	.63(.01)
1979	.64(.01)	.59(.02)	.64(.01)	.58(.01)	.65(.01)	.59(.02)
1980	.64(.01)	.61(.02)	.65(.01)	.60(.02)	.64(.02)	.62(.01)
1981	.65(.01)	.63(.01)	.66(.01)	.61(.01)	.66(.01)	.63(.01)
1982	.66(.01)	.62(.02)	.66(.01)	.61(.02)	.66(.01)	.63(.01)
1983	.65(.01)	.59(.02)	.65(.01)	.59(.02)	.63(.02)	.62(.01)
1984	.63(.01)	.62(.02)	.65(.01)	.59(.02)	.65(.02)	.61(.02)
1985	.66(.01)	.58(.02)	.65(.01)	.60(.02)	.64(.02)	.63(.01)
Slope	−.0003	−.0030	−.0009	−.0012	−.0017	.0000
S. E.	.0014	.0016	.0013	.0017	.0016	.0019

Table 9. Yearly Estimates within College Categories of Correlation of High School Record with F-GPA with Associated Standard Errors and Slopes of the Best-Fitting Lines for the Correlations

Year	*Private Colleges*	*Public Colleges*	*Less than 750 Freshmen*	*750+ Freshmen*	*Less than 950 SAT V+M*	*950+ SAT V+M*
1976	.53(.02)	.50(.01)	.54(.02)	.47(.02)	.55(.02)	.48(.02)
1977	.52(.01)	.47(.02)	.52(.01)	.47(.02)	.53(.02)	.48(.02)
1978	.53(.02)	.46(.01)	.53(.01)	.44(.01)	.52(.02)	.48(.02)
1979	.53(.02)	.44(.02)	.53(.02)	.45(.02)	.53(.01)	.46(.03)
1980	.51(.01)	.46(.01)	.52(.01)	.44(.01)	.52(.02)	.46(.01)
1981	.53(.01)	.47(.01)	.53(.01)	.45(.01)	.54(.01)	.47(.02)
1982	.53(.01)	.47(.02)	.53(.01)	.47(.02)	.54(.01)	.48(.01)
1983	.53(.01)	.44(.02)	.52(.01)	.46(.02)	.51(.02)	.49(.02)
1984	.49(.03)	.47(.02)	.50(.03)	.44(.01)	.53(.02)	.43(.04)
1985	.54(.01)	.47(.01)	.54(.01)	.48(.01)	.53(.02)	.49(.01)
Slope	−.0007	−.0013	−.0016	.0007	−.0010	−.0006
S. E.	.0025	.0017	.0023	.0020	.0023	.0032

Table 10. Yearly Estimates within College Categories of Correlation of High School Record with F-GPA after Multivariate Correction for Restriction of Range with Associated Standard Errors and Slopes of Best-Fitting Lines

Year	*Private Colleges*	*Public Colleges*	*Less than 750 Freshmen*	*750+ Freshmen*	*Less than 950 SAT V+M*	*950+ SAT V+M*
1976	.57(.02)	.58(.01)	.59(.02)	.55(.02)	.58(.01)	.55(.02)
1977	.58(.01)	.55(.01)	.58(.01)	.54(.02)	.58(.01)	.56(.01)
1978	.58(.01)	.54(.02)	.59(.01)	.51(.02)	.56(.01)	.57(.02)
1979	.57(.02)	.53(.02)	.57(.02)	.52(.02)	.58(.01)	.52(.02)
1980	.56(.01)	.54(.02)	.57(.01)	.53(.02)	.56(.01)	.55(.01)
1981	.58(.01)	.58(.01)	.60(.01)	.55(.01)	.60(.01)	.57(.01)
1982	.58(.01)	.57(.02)	.59(.01)	.56(.02)	.59(.01)	.56(.01)
1983	.58(.01)	.53(.02)	.57(.01)	.54(.02)	.56(.02)	.57(.01)
1984	.55(.02)	.57(.02)	.57(.02)	.53(.01)	.59(.02)	.52(.02)
1985	.59(.01)	.56(.02)	.59(.01)	.57(.01)	.58(.02)	.57(.01)
Slope	.0004	.0004	−.0003	.0026	.0012	.0003
S. E.	.0020	.0016	.0019	.0019	.0017	.0027

Table 11. Yearly Estimates within College Categories of Multiple Correlation of SAT Scores with F-GPA with the Associated Standard Errors and the Slopes of the Best-Fitting Lines for the SAT Correlations

Year	*Private Colleges*	*Public Colleges*	*Less than 750 Freshmen*	*750+ Freshmen*	*Less than 950 SAT V+M*	*950+ SAT V+M*
1976	.47(.02)	.42(.02)	.47(.02)	.41(.01)	.50(.02)	.40(.02)
1977	.46(.01)	.40(.01)	.46(.01)	.41(.01)	.46(.01)	.40(.01)
1978	.45(.01)	.41(.01)	.46(.01)	.40(.01)	.45(.01)	.41(.01)
1979	.44(.01)	.37(.02)	.44(.01)	.37(.01)	.45(.01)	.38(.02)
1980	.44(.01)	.38(.01)	.44(.01)	.39(.01)	.44(.02)	.40(.01)
1981	.41(.01)	.36(.01)	.42(.01)	.36(.01)	.42(.01)	.37(.02)
1982	.44(.01)	.35(.02)	.44(.01)	.37(.01)	.44(.01)	.38(.02)
1983	.43(.01)	.36(.02)	.43(.01)	.35(.02)	.43(.02)	.38(.01)
1984	.41(.02)	.37(.02)	.42(.02)	.34(.02)	.40(.02)	.39(.02)
1985	.42(.01)	.27(.05)	.40(.02)	.33(.01)	.39(.02)	.35(.02)
Slope	−.0054	−.0113	−.0064	−.0085	−.0097	−.0045
S. E.	.0015	.0046	.0018	.0017	.0024	.0023

Table 12. Yearly Estimates within College Categories of Multiple Correlation of SAT Scores with F-GPA after Multivariate Correction for Restriction of Range with Associated Standard Errors and Slopes of the Best-Fitting Lines

Year	*Private Colleges*	*Public Colleges*	*Less than 750 Freshmen*	*750+ Freshmen*	*Less than 950 SAT V+M*	*950+ SAT V+M*
1976	.51(.01)	.50(.02)	.52(.01)	.49(.01)	.54(.01)	.49(.01)
1977	.52(.01)	.49(.01)	.53(.01)	.49(.01)	.52(.01)	.49(.02)
1978	.52(.01)	.48(.01)	.52(.01)	.47(.02)	.51(.01)	.50(.01)
1979	.50(.01)	.45(.02)	.50(.01)	.46(.01)	.50(.01)	.46(.01)
1980	.52(.01)	.47(.02)	.51(.01)	.48(.01)	.51(.01)	.49(.01)
1981	.50(.01)	.47(.01)	.51(.01)	.46(.01)	.50(.01)	.48(.01)
1982	.52(.01)	.45(.01)	.52(.01)	.46(.01)	.51(.01)	.49(.01)
1983	.50(.01)	.45(.02)	.51(.01)	.44(.02)	.50(.02)	.47(.01)
1984	.49(.02)	.47(.01)	.50(.01)	.44(.01)	.47(.02)	.48(.02)
1985	.50(.02)	.37(.04)	.48(.02)	.43(.02)	.47(.02)	.45(.02)
Slope	−.0024	−.0090	−.0032	−.0066	−.0058	−.0024
S. E.	.0012	.0029	.0014	.0018	.0021	.0012

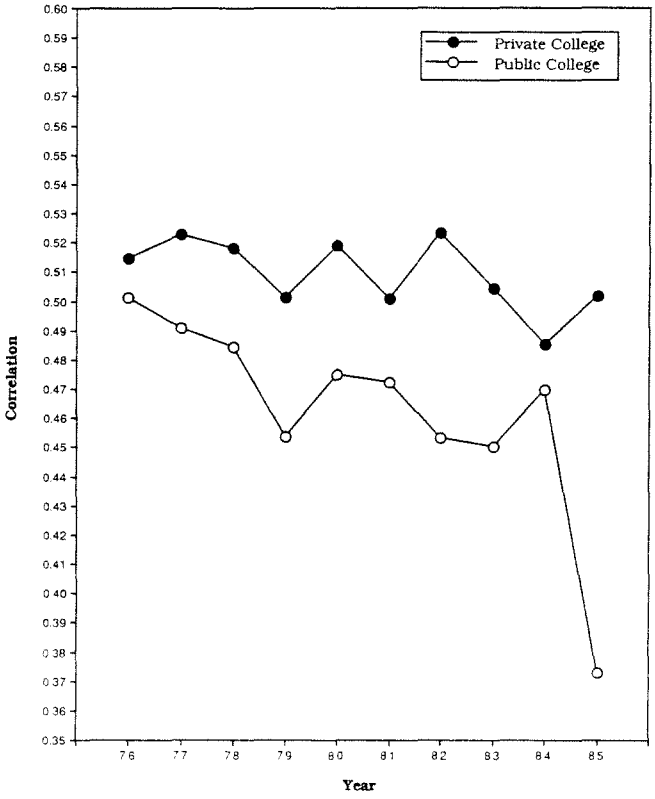

Figure 2. 10-Year Estimates of the Correlation of SAT with F-GPA Corrected for Restriction of Range

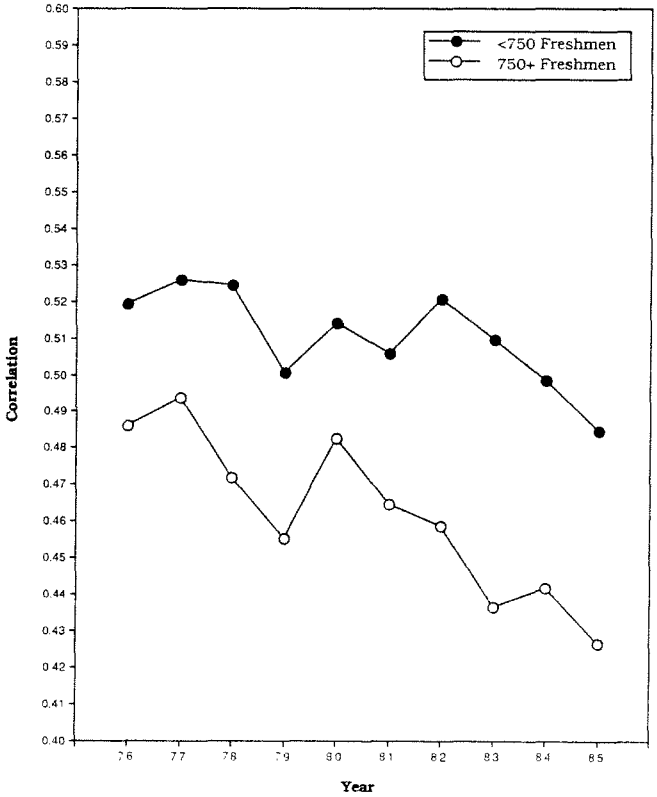

Figure 3. 10-Year Estimates of the Correlation of SAT with F-GPA Corrected for Restriction of Range

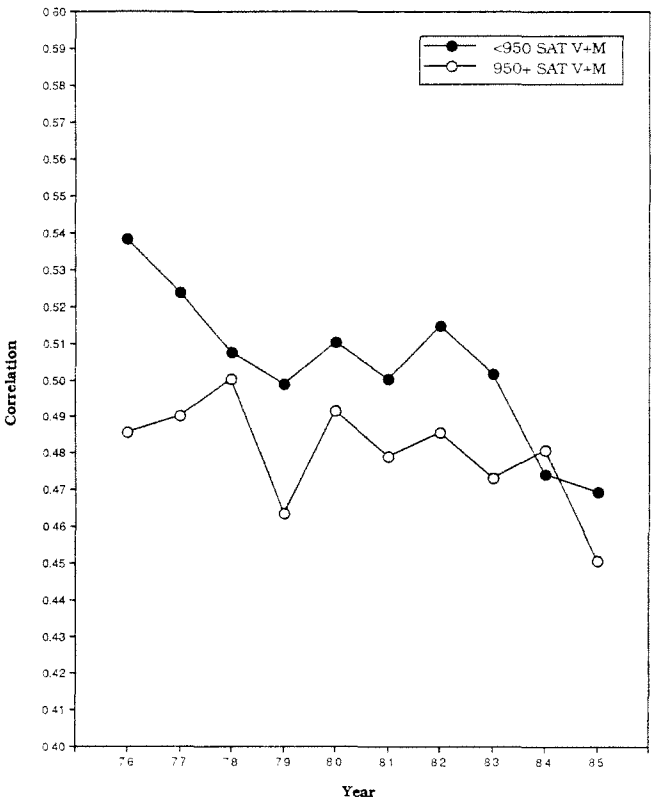

Figure 4. 10-Year Estimates of the Correlation of SAT with F-GPA Corrected for Restriction of Range

The figures show that some decline in predictive validity occurred for all the categorizations of the data. It is clear that the declines varied among these nominal college characteristics, even though the characteristics have no obvious relationship to the factors that might cause a change in predictive validity. This finding suggests that further study of more relevant factors is needed to more fully understand the observed declines.

SUMMARY

The analyses of colleges conducting multiple studies found that the estimates of change in the correlation of SAT scores with F-GPA is smaller than the initial yearly averages indicated. Two factors appear to have exaggerated the amount of change. First, in the period from 1976 to 1985, students within colleges have become a more homogeneous group with respect to SAT scores. A comparison of slopes of the data associated with Tables 4 and 5 indicate that as much as 40 percent of the change in the SAT correlation with F-GPA is due to decreased variability over time of the distribution of SAT scores within colleges conducting validity studies through VSS. Another part of the decline results from differential college participation in VSS. Thus, when the yearly average correlations are computed to account for these two spurious factors, the decline in the predictive validity is most likely a little less than half the decline suggested by the raw VSS yearly correlation averages.

The analyses based on college categorizations indicate that the decline in the correlation of SAT with F-GPA was not limited to a specific type of college. There is some evidence that private colleges, smaller colleges, and more selective colleges had smaller changes in predictive validity than public colleges, larger colleges, and less selective colleges. However, this result is not informative as to why predictive validity changes occurred from 1976 to 1985. Analyses of other characteristics, such as changes in curriculum or grading practices, would more likely help clarify why the relationship between SAT scores and F-GPA has declined.

In conclusion, the repeated measures analyses detected a decline in the predictive validity of SAT scores. The degree to which the decline results from changes in the composition of first-year college students, changes in the SAT, changes in the courses contributing to F-GPA, or changes in the real relationship of SAT scores with F-GPA is a topic for future research. What is certain from these analyses is that the decline is not well characterized by simple comparisons of average correlations based on the total self-selected population of colleges participating in the VSS from one year to another.

REFERENCES

College Board (1985). *National College-Bound Seniors, 1985.* New York: College Entrance Examination Board.

College Board (1976). *National College-Bound Seniors, 1976.* New York: College Entrance Examination Board.

Goldman, R. D., and Widawski, M. H. (1976). A within-subjects technique for comparing college grading standards: Implications in the validity of the evaluation of college achievement. *Educational and Psychological Measurement, 36,* 381-390.

Gulliksen, H. (1950). *Theory of Mental Tests.* New York: Wiley.

Mosteller, F., and Tukey, J. W. (1977). *Data Analysis and Regression.* Reading, Mass. Addison-Wesley.

Ramist, L. (1984). Predictive validity of ATP tests. In T. F. Donlon (Ed.). *The College Board Technical Handbook for the Scholastic Aptitude Test and Achievement Tests* (pp. 141-170). New York: College Entrance Examination Board.

APPENDIX A

Yearly Estimates

Suppose there is some measure of interest (such as a correlation coefficient) that is available for some colleges in some years. Denote it by y_{ij} for college i and year j. For each pair of years, all colleges that have the measure available for both years are used to compute the difference between the two values. Thus for years j and k, if college i has the measure for both years, the difference was computed by:

$$d_{ijk} = y_{ij} - y_{ik}$$

Next such differences for these two years are averaged. If there are n_{jk} colleges with the statistic for the two years, the average is given by

$$d_{jk} = \sum_{i=1}^{n_{jk}} d_{ijk}/n_{jk}$$

Note that

$$d_{kj} = -d_{jk}$$

and let

$$d_{jj} = 0$$

Then for year j, the average deviation effect (a_j) is estimated as follows:

$$a_j = \sum_{k=1}^{10} d_{jk}/10$$

For the actual tables and figures, these deviation effects were added to an overall mean, determined as follows. The mean value of the measure for year j was averaged over all colleges having a value of the statistic for that year. The result was y_j. The unweighted mean of these values was the overall level, namely

$$y = \sum_{j=1}^{10} y_j/10$$

Thus the values $t_j = y + a_j$ describe average yearly estimates for the measure y_{ij}.

Standard Errors of Yearly Estimates

Due to the complexity of the dependencies in the terms used to compute yearly estimates, there does not appear to be any straightforward way to obtain standard errors for the estimates. Consequently, it was decided to use jackknifing to provide approximate standard errors. For an overview of jackknifing, see Mosteller and Tukey (1977), pages 133-136.

The jackknifing proceeded by first computing the vector of estimates based on all colleges, say t_{all}. Next, one college was left out at a time. For each college, this meant leaving out at least two records, since each college included in this analysis must have data for at least two years. For each reduced set of colleges, the corresponding vector of estimates, say $t_{(i)}$ when college i is left out, was obtained. Finally, this set of estimate vectors produced a jackknifed estimate vector $t*$ and an estimated variance-covariance matrix $S*$ for this estimate. From $S*$, standard errors for each element in $t*$ were computed.

More specifically, the jackknifed estimate for year j may be written as

$t_j* = (n)t_{j,all} - (n-1)t_{j(.)}$, where

$t_{j(.)} = \sum_{i=1}^{n} t_{j(i)}/n$ and there are a total of n colleges.

For the estimation method used here, it turns out that $t_j* = t_{j,all}$.

The elements of $S*$ are given by

$$s_{jk}* = [(n-1)/n] \sum_{i=1}^{n} [t_{j(i)} - t_{j(.)}][t_{k(i)} - t_{k(.)}].$$

The jackknife estimate of the standard error for t_j* is

$$s_j* = (s_{jj}*)^{1/2}$$

The standard error for the slope of the linear trend in the yearly estimates is formed by pre- and post-multiplying $S*$ by the vector of linear coefficients $c' = (-9, -7, -5, -3, -1, 1, 3, 5, 7, 9)/165$.

APPENDIX B

Number of Colleges for Each Pairwise Set of Years

All Pairs

	1985	1984	1983	1982	1981	1980	1979	1978	1977	1976
1985										
1984	41									
1983	36	46								
1982	34	34	37							
1981	30	38	39	49						
1980	26	27	38	33	46					
1979	20	30	29	29	38	48				
1978	19	23	23	25	29	32	38			
1977	21	25	25	24	39	36	48	37		
1976	14	22	22	20	23	25	26	25	39	

All Pairs—High School Rank

	1985	1984	1983	1982	1981	1980	1979	1978	1977	1976
1985										
1984	21									
1983	17	25								
1982	17	17	19							
1981	17	21	21	30						
1980	14	15	19	20	24					
1979	10	17	16	17	19	29				
1978	12	14	13	17	16	18	19			
1977	12	16	13	15	21	17	25	15		
1976	9	16	13	11	13	14	14	12	20	

All Pairs—High School GPA

	1985	1984	1983	1982	1981	1980	1979	1978	1977	1976
1985										
1984	19									
1983	18	20								
1982	15	16	17							
1981	12	16	17	19						
1980	12	10	16	13	21					
1979	9	11	12	12	19	18				
1978	7	8	9	7	11	12	15			
1977	9	7	12	9	18	17	22	20		
1976	5	6	9	9	10	9	12	10	17	

Public Colleges

	1985	1984	1983	1982	1981	1980	1979	1978	1977	1976
1985										
1984	12									
1983	9	15								
1982	8	11	15							
1981	7	10	14	17						
1980	5	8	13	11	17					
1979	5	9	9	5	9	15				
1978	5	9	10	8	11	13	13			
1977	8	8	10	6	13	13	14	14		
1976	2	6	7	7	7	9	10	11	14	

Private Colleges

	1985	1984	1983	1982	1981	1980	1979	1978	1977	1976
1985										
1984	29									
1983	27	31								
1982	26	23	22							
1981	23	28	25	32						
1980	21	19	25	22	29					
1979	15	21	20	24	29	33				
1978	14	14	13	17	18	19	25			
1977	13	17	15	18	26	23	34	23		
1976	12	16	15	13	16	16	16	14	25	

750+ Freshmen

	1985	1984	1983	1982	1981	1980	1979	1978	1977	1976
1985										
1984	14									
1983	12	14								
1982	11	11	12							
1981	9	11	14	18						
1980	5	8	12	12	17					
1979	4	7	8	6	11	16				
1978	4	4	6	7	9	12	13			
1977	5	5	6	5	12	14	17	13		
1976	4	6	8	6	7	8	7	8	9	

750 or Fewer Freshmen

	1985	1984	1983	1982	1981	1980	1979	1978	1977	1976
1985										
1984	27									
1983	24	32								
1982	23	23	25							
1981	21	27	25	31						
1980	21	19	26	21	29					
1979	16	23	21	23	27	32				
1978	15	19	17	18	20	20	25			
1977	16	20	19	19	27	22	31	24		
1976	10	16	14	14	16	17	19	17	30	

Average SAT Below 950

	1985	1984	1983	1982	1981	1980	1979	1978	1977	1976
1985										
1984	13									
1983	14	20								
1982	14	19	23							
1981	10	22	21	30						
1980	10	13	20	20	24					
1979	8	14	12	15	22	27				
1978	8	13	12	14	16	15	19			
1977	10	11	13	12	22	22	22	18		
1976	2	6	6	9	11	9	11	11	19	

APPENDIX B *(continued)*

	Average SAT 950+									
	1985	1984	1983	1982	1981	1980	1979	1978	1977	1976
1985										
1984	27									
1983	20	25								
1982	16	12	14							
1981	17	14	15	15						
1980	16	14	16	11	19					
1979	9	14	14	10	12	18				
1978	9	9	9	8	11	13	18			
1977	9	12	10	7	13	12	21	18		
1976	10	13	13	8	9	12	12	13	15	