DOCUMENT RESUME

ED 395 008 TM 025 004

AUTHOR	Cook, Linda L.; And Others
TITLE	Achievement Test Scaling.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RR-88-34
PUB DATE	Jun 88
NOTE	223p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC09 Plus Postage.
DESCRIPTORS	*Achievement Tests; College Entrance Examinations; Comparative Analysis; Correlation; Higher Education; *Mathematics Tests; *Scaling; *Scores; Scoring; Test Results; *Verbal Tests
IDENTIFIERS	*College Board Achievement Tests; Scholastic Aptitude Test

ABSTRACT

Scaling is carried out in an effort to increase the comparability of scores obtained on different tests. This study explored the relationships between College Board Achievement Test scores and potential scaling covariates for various subgroups of the test-taking population with the goal of providing several alternatives to traditionally used scaling procedures. The alternatives vary components such as scaling covariates, scaling samples, and characteristics of the hypothetical reference group. The traditional method of scaling the Achievement Tests uses scores on the Scholastic Aptitude Test (SAT) verbal and mathematics tests as scaling covariates for all 14 achievement tests. To determine additional covariates that might be used, correlations of test scores with 17 potential covariates were examined, and stepwise regressions were carried out. As a result, the Achievement Tests were grouped into five clusters. Two alternative scaling procedures were proposed. The first consists of scaling the tests within the five independent clusters; and the second consists of a two-stage process of scaling within the clusters independently and using the results of the initial scaling to scale the tests again with SAT verbal and mathematics scores as common covariates. (Contains 9 figures, 12 tables, and 3 references.) (SLD)

RESEARCH REPORT

ACHIEVEMENT TEST SCALING

Linda L. Cook

with the collaboration of

William H. Angoff
and
Alicia P. Schmitt

BEST COPY AVAILABLE

Educational Testing Service
Princeton, New Jersey
June 1988

Achievement Test Scaling[1]

by

Linda L. Cook

with the collaboration of

William H. Angoff
and
Alicia P. Schmitt

---

[1]Final Report of Project/Job 201/44, sponsored by the Joint Staff Research and Development Committee.

EXECUTIVE SUMMARY

Although the statistical processes of scaling and test equating are mathematically very similar, the purposes, and end results, of the two procedures are often quite different. This is certainly the case for the Admissions Testing Program (ATP) Achievement Tests.

Test equating is the method that is most often used with the Achievement Tests to maintain form-to-form comparability of test scores. Every time a new form of an Achievement Test is introduced (for example, a new form of the Biology Test), scores on the new form are equated to scores on a previous form of the test. This type of equating can be carried out quite rigorously and assures that scores obtained by a student who may have taken a form of the Biology test in the spring of the year can be compared with confidence to scores obtained by a student who, perhaps, took a different form of the Biology Test in the fall of the year. Form-to-form equating is essential if scores obtained on different forms of the same test are to be compared for college admissions purposes.

Form-to-form equating also plays an important role when scores on the same test are to be evaluated over a long period of time, perhaps in an effort to examine trends in test scores for large populations. Similarly, form-to-form equating is essential if scores on a particular test are to be used for placement purposes. For example, cut scores used to place students into classes at differing levels may be developed and validated using a particular form of a test and then used when students taking alternate forms of the test are tested for placement purposes. If the placement decision is to be effective, scores obtained on the forms of the test must be comparable to scores obtained on the test form, or forms, that were used to initially establish and validate the placement cut score.

Scaling, on the other hand, as it is used for ATP Achievement Tests, is carried out in an effort to increase comparability of scores obtained on different tests, e.g., scores obtained on the Biology Test with those obtained on the Chemistry Test. Hence, if it is necessary for college admissions purposes to compare scores obtained by students taking a Biology Test with those obtained by students taking a Chemistry Test, scaling becomes important.

An issue that has concerned those interested in maintaining both the across test and form-to-form comparability of Achievement Test scores is that it has been very difficult to perform both of those procedures simultaneously. It has been possible to maintain the within test comparability required for placement and some admissions decisions as well as trend analyses by form-to-form equating. However, periodic scalings[1] of the tests, designed to increase across test comparability required for admissions decisions, has interrupted the form-to-form comparability which is necessary for the previously mentioned placement decisions, some admissions decisions, and trend analyses.

Whether or not it is more important to scale the tests to increase across test comparability of scores or to pursue rigorous form-to-form equating is not so much a statistical matter as it is a matter of program priorities and test use. What is a statistical concern is the fact that the procedure used to scale the tests and, hence, to maximize across test comparability (at the expense of form-to-form comparability) has been questioned. It is this latter concern that is the focus of the study summarized in this report.

The purpose of the study described in this report was to explore the relationships between College Board Achievement Test scores and potential scaling covariates for various subgroups of the test taking population. It was hypothesized that such an exploration would lead to the following:

---

[1]Periodic scalings of the ATP Achievement Tests have not been carried out since 1980.

- The selection of additional scaling covariates that might provide improved scaling results for those tests that do not provide scores correlating highly with SAT-V and/or SAT-M scores;
- The review and possible respecification of the sample of students that are used to scale the tests; and
- The review and possible respecification of the hypothetical scaling population.

The goal of the present study was to provide several alternatives to the scaling procedures traditionally used for the Achievement Tests. The alternatives will, by their nature, vary components such as scaling covariates, scaling samples, and characteristics of the hypothetical reference group.

A brief summary of the results of the study follows. First, the results of the investigation of the characteristics of the scaling sample are discussed. Second, the choice of additional scaling covariates and their relationship to clusters of tests are summarized. Third, the results of the examination of the characteristics of the hypothetical reference group are discussed.

Characteristics of the Scaling Sample

The traditional method of scaling the Achievement Tests employs a scaling sample with the following characteristics. The sample is based on high school juniors and seniors. Juniors are selected from those students who take the tests at the May and June administrations and seniors are selected from students who take the tests at the November, December, and January administrations.

Several trends were observed in summarizing the results of a number of grade level analyses done for the study. First, it appears, as has traditionally been the situation, that the tests are taken predominantly by high school seniors at the November, December, and January administrations and by juniors at the May and June administrations. There are, however, some exceptions to this rule.

For one, the Biology Test is taken predominately by sophomores at the June administration. Also, the Literature, Physics, and Mathematics Level II Tests are taken by almost equal numbers of juniors and seniors at the May administration.

One question is, given the fact that the basic test taking patterns (seniors taking the tests in the fall and juniors in the spring) do not seem to have changed greatly over the past several decades, is there a reason for revising the specifications that are presently used to draw the scaling samples?

The results of the analyses conducted for the study indicated that the practice of sampling high school seniors taking the tests in November, December, and January and juniors taking the tests in May and June provides a viable sampling option. A potential alternative to this practice is related to whether to include high school sophomores taking the tests in June. It was decided that an effort should be made to include sophomores is the scaling sample if additional analyses indicate that scores on missing covariates can be estimated adequately and if the regressions show that Achievement Test scores obtained by this group have a similar relationship to the relevant covariates as that displayed by scores obtained by other sampling groups.

Additional Scaling Covariates

The traditional method of scaling the Achievement Tests uses scores on SAT-V and SAT-M as scaling covariates for all 14 tests. In addition, semesters of study (as assessed by Achievement Test Background Questionnaire responses) is used as a covariate for scaling foreign language tests.

Both Braun and Tucker (see Dorans, 1985) have pointed to problems with the traditional scaling method that are related to choice of covariates. Basically, the results of both their studies indicated that when ability has a major impact on test selection (as it does with the Achievement Tests), there is a clear need to rescale the tests. The quality of this scaling will depend, however, on how highly the covariates used to rescale the tests correlate with Achievement Test performance.

In order to determine what additional covariates might be added to the model used to scale the Achievement Tests, correlations of test scores with 17 potential covariates were examined. And, more importantly, stepwise regressions were carried out. The stepwise regressions were used to answer questions such as: If SAT-V and SAT-M scores are used to scale Achievement Test scores, will the scaling improve if semesters of study are added as an additional covariate? In addition, the results of the stepwise regression analyses were compared to determine if several tests could be clustered together, i.e., tests within a cluster show a similar relationship to a common core of covariates and, hence, can all be rescaled using the covariates. It should be noted that, in choosing tests to form the clusters, practical considerations were taken into account, i.e., efforts were made to keep tests together that formed logical clusters such as the science tests, the language tests, etc.

As a result of the regression analyses, the Achievement Tests were organized into five clusters. Cluster 1, referred to as the English Cluster, contains the English Composition Test and the Literature Test. Cluster 2, referred to as the History Cluster, contains the American History and Social Studies Test, and the European History and World Cultures Test. All tests in both the clusters correlate highly with SAT-V scaled scores, reasonably highly with SAT-M scaled scores, and moderately with grades in the respective subject areas. The European History and World Cultures Test scores show the lowest correlation with these covariates. This is particularly true of scores obtained at the December administration of the test.

Cluster 3, the Mathematics Cluster, contains the Mathematics Level I and II Tests. Proposed covariates for these two tests are SAT-M scaled scores and Achievement Test Background Questionnaire responses assessing amount of study in mathematics. High correlations with SAT-M scores and moderate correlations with Background Questionnaire responses are displayed by the Mathematics Level I and Mathematics Level II data.

Proposed covariates for Cluster 4, the Science Cluster, containing the Biology, Chemistry, and Physics Tests, are SAT-V scaled scores, SAT-M scaled scores, and amount of course work in biology and sciences or amount of course work in physical sciences, respectively. The Chemistry and Physics Tests show correlations with the covariates that are slightly more similar to each other than they are to the correlations displayed by the Biology Test scores. Scores obtained on the Biology Test have a tendency to be a little more highly correlated with SAT-V scores and a little less highly correlated with SAT-M scores than scores obtained on the other two tests in the cluster.

Cluster 5 contains three of the five language tests, French, Spanish and Latin. Results of the analyses were not amenable to clustering either the Hebrew Test or the German Test. The additional covariate proposed for the cluster, along with those used by the traditional procedure (SAT-V and SAT-M scaled scores and Achievement Test Background Questionnaire responses assessing amount of study in a language), is grades in foreign languages. All of the tests show a fairly strong correlation with this covariate.

To summarize, the tests were sorted into five scaling clusters based on the results of a stepwise regression analyses, as well as on practical considerations. The results of the clusterings were confirmed by an examination of the correlations of test score with selected covariates. It is suggested that, based on the results of this portion of the study, two alternative scaling procedures be considered. The first procedure consists of scaling the tests within the five independent clusters. No attempt would be made to establish any type of relationship among the scores obtained on tests that are members of different clusters. The second procedure consists of a two-stage process. The first stage consists of scaling the tests within the clusters independently. The second stage would use the results of the initial scaling and scale the tests again, using SAT-V and SAT-M scores as common covariates for all 14 tests.

Hypothetical Scaling Population

The hypothetical reference population that is used in the traditional Achievement Test scaling procedure is assumed to have a scaled score mean of 500 on SAT-V and SAT-M, a scaled score standard deviation of 100 on both these variables, and a correlation between scores obtained on SAT-V and SAT-M of .60. It has been pointed out by Braun (see Dorans, 1985) that these values influence the scaling results. The results of Braun's studies indicated that choice of reference population is of critical importance. Braun found that different reference groups produced different Achievement Test scale alignments.

The analyses conducted for this part of the study indicated that, in general, SAT-V scaled score means for the total groups that take the Achievement Tests at the five administrations range between roughly 500 and 575. SAT-M scaled score means obtained by the same groups range, roughly, between 525 and 675. The total group correlation between scores on these two tests range between .44 and .69 with the majority of the correlation coefficients falling somewhere in the vicinity of .55. Furthermore, examination of the information provided regarding scaled score standard deviations indicates that SAT-V scaled score standard deviations obtained by the total groups taking the 14 tests range between 91 and 117. Scaled score standard deviations on SAT-M obtained for the same groups range between 77 and 111.

It is fairly clear, from the information presented above, that SAT-V and SAT-M scores obtained by the groups taking the 14 tests are quite diverse. Furthermore, it is clear that a hypothetical population with means and standard deviations of 500 and 100, respectively, on SAT-V and SAT-M has little empirical basis (although this population does have historical importance). The correlation between SAT-V and SAT-M scores that is presently used for the hypothetical population (.60) does, however, have some empirical basis.

A number of possibilities exist for alternatives to the way in which the hypothetical reference group is currently specified. Given the diversity of the scores on the covariates obtained by the various groups, an empirical approach might be the most feasible. This approach could be one that involves pooling scores on the various covariates and using the pooled means, standard deviations, etc., as the values to specify for the hypothetical reference group.

One empirical approach is to pool data obtained for the common covariates that are used within a cluster, i.e., five hypothetical reference populations would be established. Next, if a two stage scaling procedure were used, data could be pooled across the covariates common to the 14 tests (SAT-V and SAT-M scores) and used to specify the hypothetical reference group for the second stage of the scaling.

The alternative scaling procedures suggested above represent, to some extent, a compromise between the traditional scaling method and form-to-form equating. Because the procedures involve clustering tests that have similar relationships to a common set of covariates, some of the problems with the traditional scaling method, alluded to by Tucker and Braun, should be reduced. The suggested procedures represent changes in methodology that should lead to greater comparability of scores on tests grouped within clusters, but less comparability on scores obtained on tests that are members of different clusters. The procedures that involve two-stage scaling represent a further compromise between a one-stage procedure and the traditional procedure. Because this procedure involves a second stage scaling that uses common covariates for the 14 tests, scores obtained on tests that are members of different clusters might be considered roughly comparable. However, because of the problem of differential relationship of Achievement Test scores to SAT-V and SAT-M scores, some lack of comparability of the scales established across the clusters will probably be difficult to avoid.

To summarize, based on empirical evidence gathered in this study, alternative scaling procedures have been suggested. These procedures represent a compromise between the current scaling method and form-to-form equating. Which procedure offers maximum comparability of scores either within a cluster or across clusters, while at the same time minimizing scaling problems resulting from the differential relationship of scores to covariates, may be a question that is amenable to future empirical investigation.

ACKNOWLEDGEMENTS

The author and collaborators wish to express their appreciation to the many professionals who provided important contributions to the study and to the content and format of this report. We wish to thank Gretchen Rigol and Robert Cameron of the College Board, New York, William Schrader, ETS consultant, and Daniel Eignor, College Board Statistical Analysis staff member at ETS, for their careful review and insightful comments regarding the content and format of this report. We wish to thank Ledyard Tucker, ETS consultant, and Gary Marco, College Board Statistical Analysis staff member at ETS, for their guidance and direction throughout this project and for their review of our final report.

Finally, we would like to express our appreciation to Nancy Ervin, College Board Statistical Analysis staff member at ETS, for her assistance in performing the data analyses for the study and Georgiana Thurston and Judith Fisher, also College Board Statistical Analysis staff members, for their assistance in the preparation of the tables and text of this report.

Achievement Test Scaling

by
Linda L. Cook

with the collaboration of
William H. Angoff
Alicia P. Schmitt

BACKGROUND

When the objective forms of the College Board Achievement Tests were first introduced in about 1942 for operational use in admissions, the tests were scaled in such a way that the mean for the group choosing to take each of them was set at 500, and the standard deviation at 100. That is to say, the average of each group of candidates taking its test was made to appear equal to the average performance of each of the other groups of candidates taking their tests. Similarly with their standard deviations. As a consequence of this scaling design, the score a candidate received on a test was dependent on, among other things, the ability level of the group of examinees who took the particular test. For example, a candidate would appear more able if he or she took a test typically chosen by less able examinees and less able if he or she chose a test typically taken by high ability examinees. Inasmuch as each candidate in the tested population was (appropriately) regarded as competitive with each of the other candidates in a contest for the same reward, i.e., admission to college, this state of affairs appeared to be--and indeed, was--unfair; any candidate who understood the design of the score scales, and wished to appear to be relatively knowledgeable in his/her field, could adopt the strategy, if they chose to, of selecting the Achievement Test normally taken by the least able candidates. In order to remove this element of unfairness, a scaling system was designed in the middle 1940's to adjust the scales for the several Achievement Tests to reflect the level and dispersion of ability of the candidates taking each test. For example, a test typically taken by a more able group of candidates was made to yield an average scaled score higher than 500, and a test typically taken by a less able group of candidates was made to yield an average scaled score lower than 500.

The operational scale definition adopted to achieve this result was that the candidate of average ability, relative to a hypothetical aggregate of all candidates taking the College Board tests, would earn a score of 500 regardless of the Achievement Test that he or she chose to take; also, that the dispersion of scaled scores for this hypothetical aggregate population would be defined with a standard deviation of 100. Thus, higher-ability Achievement Test groups would automatically have higher means and, correspondingly, lower-ability Achievement Test groups would have lower means. This definition was implemented by defining "general ability" (more specifically, "general academic ability") as measured by the verbal and mathematical Scholastic Aptitude Tests (SAT-V and SAT-M), respectively; and the degree to which the SAT-V and -M scores played a role in this operational definition was a direct function of the relevance of those tests for the particular Achievement Test in question, as measured by the correlation of the SAT scores with the scores on the particular Achievement Test. A further adjustment was later introduced into this system by adding semesters of study to the SAT-V and -M scores for scaling the language tests. This adjustment was intended to account for the fact that some languages were typically studied for longer periods of time than were other languages.

A distinction was made in earlier years, as it is made today, between the efforts to provide optimum calibration or equivalence between one form and another within each of the various Achievement Tests (equating) and the efforts, as just described, to maintain the optimum interrelationships among the scales for the various tests (scaling). The first of these is roughly analogous to the calibration of, say, successively manufactured scales for weighing objects in order to insure that the same object will yield the same reading of weight in pounds, whether the particular weight scales in question were manufactured this past year or whether they were manufactured 20 years ago. Here, we observe, the calibration is carried out as a study of the relation between two instruments that provide measurements in the same domain, namely weight or, in the case of the forms of the individual Achievement Tests--in German, in English Literature, in Physics, etc.--forms that measure the same content domain but which may differ from one another in level of difficulty.

The second of these, scaling, presents some of the same considerations that must be borne in mind when developing height-weight tables, for example. Unlike alternative instruments used for taking measurements in the same domain, as in height or weight or temperature, the relation between measures in different domains, like height vs weight, is seen clearly to vary depending on the nature of the group on which the relationship is based, for example, whether the group consists of males or females; ectomorphs, mesomorphs, or endomorphs; or adolescents, young adults, middle-aged, or elderly people. Clearly, a set of relationships based on one of these types of groups will not necessarily apply to other groups. Similar considerations apply to the Achievement Tests. A set of relationships among the Achievement Tests based on one group of students, say largely male, will not apply to a group of students composed largely of females. Similarly, a set of relationships based on students who have been exposed to one type of curriculum emphasis in a subject-matter area will not apply to students who have been exposed to another type of curriculum emphasis. Yet, it is the intent of the scaling effort to develop one set of interrelationships among the scales for the different Achievement Tests that will apply at least reasonably well to all component subgroups of students. This is a task that is, admittedly, impossible to execute rigorously with respect to all reasonable criteria. Yet, the alternative, to let each test scale be defined by the group who chooses to take that test, is, for reasons already discussed, even less attractive.

Since the introduction of the objective Achievement Tests, it has become clear that they have been useful for two different, and--as it later became evident--somewhat conflicting contexts. One of these was the use of the tests in admission decisions, in which students choose to take different combinations of one, two, or three Achievement Tests at a single sitting. It was for this purpose that the scaling of the Achievement Tests in relation to one another was originally designed. As already indicated, although the methodology used to achieve this between-test comparability fell short of the ideal, it was the best, indeed the only defensible, methodology available. However, since the between-test comparability system designed at one point in time is likely to change with changes in population or changes in secondary school curricular emphasis--just as the height-weight relationship might change with a change in the population or with a change of dietary habits in the population--a far-sighted general plan for scaling, or establishing comparability, should normally provide for a periodic re-examination of the interrelationships among the tests, and an adjustment in the system, if that were found to be indicated.

The second use of the Achievement Tests resided, and still resides, in the context of course placement. In this context, a university administration typically chooses a score level above which a student may be permitted to enter a more advanced level of instruction, and below which the student would be required to take a more elementary, less demanding, or sometimes even a remedial course. In order for the university to maintain the same level of standards from one year to the next, or to be able to change its standards deliberately and with full understanding of the desired consequences, it is essential that the meaning of the "passing," or "admit," score be rigorously maintained from one year to the next. This implies that within each test--Physics, European History and World Cultures, Mathematics Level II, etc.--extreme effort be given to rigorous form-to-form equating. If, however, the system is designed to provide for optimum <u>between-test comparability</u>, to be revised as necessary whenever the existing comparability is found to be out of date, then the <u>within-test equating</u> system is necessarily weakened. The two purposes--within-test equivalence and between-test comparability--are therefore antithetical; they cannot both be served simultaneously and equally well.

On the other hand, it is also true that the purposes of equivalence and comparability do make some of the same demands. Since college applicants may take an Achievement Test at any one of various different administrations, the forms of a test offered at the different administrations, in order to be directly comparable, must be well equated. If, however, the comparability between forms is changed, as it must be to be responsive to changing populations and changing curricula, the score equivalences between the forms of a given test will necessarily be disturbed. The result is that comparisons among individuals offering scores for admission on the same test, but from different administrations, will be affected.

In general, random or systematic statistical errors in equating will have an effect on both course placement decisions and admissions decisions. It will be useful to review the general theory and method of equating scores on alternate forms of the Achievement Tests to demonstrate this.

The conventional way in which scores on these forms are equated is to carry over a subset of the items of which an old form of an Achievement Test is comprised and include the subset in the new form (the form to be equated). Inasmuch as the groups taking the different forms of a test differ, sometimes markedly, from one administration to another, it is necessary to use the information on these common items to make statistical adjustments for the differences between the groups. For ideal equating, the groups taking the two forms should be large in size and only randomly different from one another. Since this is almost never the case in performing equating in the context of the Achievement Tests, it is important that the common items constitute a substantial proportion of the test and that they represent the total test form in all respects--content, difficulty, and discriminating power--in brief, to serve as a parallel miniature of the parent form of the test. Even when these demands are met, it is inevitable that, because of the less-than-perfect methods of equating currently available to us, there will be random drift in the scales for the Achievement Tests, which will, in turn exert a disturbing effect on placement and admissions decisions. But when neither of these conditions are met--when the groups taking the different forms are not randomly equivalent, and, additionally, when common items cannot be drawn in sufficient numbers or in the desired manner, then systematic equating error necessarily results, with a consequent, sometimes appreciable, drift in the Achievement Test scales.

In 1959, Professor Samuel Wilks of Princeton University was engaged to review the work of equating and scaling the SAT and the Achievement Tests. The scope of his review included not only an examination of the particular methods in use at the time, it also included an examination of the system, its philosophy, and its mode of implementation. One of the questions under consideration in his review was that of the relative emphasis to be placed on the efforts to perpetuate the scales of the individual Achievement Tests and provide undisturbed form-to-form score equivalence (equating) versus the emphasis to be placed on the efforts to maintain the appropriate, up-to-date interrelationships among the scales for those tests (scaling). Professor Wilks recommended that the latter (scaling) should be the first order of business. Accordingly, a plan was instituted to rescale the tests each year, and to average the results of the rescaling with the existing between-test or equating relationships, expecting that the differences between the two efforts would in short time diminish to the vanishing point. This plan was put into effect, but the results turned out not as predicted; a review of the work of rescaling in the 1970's revealed that the scaling operation was not moving consistently in one direction, but fluctuated from one rescaling to another, and by sizeable amounts. In the meantime, because of these efforts to develop the correct relationship among tests by periodic rescaling, the existing form-to-form relationships were--as anticipated--disturbed; and this caused some concern in the universities who depended on a constant within-test scale structure for the effective use of the tests in course placement, and in some instances, college admissions.

The model used by Wilks (and used to scale the Achievement Tests until 1980, when all scaling was halted) for defining comparability among the several Achievement Tests specifies a base population for which the comparability applies to be one with a mean of 500 and standard deviation of 100 on both the SAT-V and SAT-M. In addition, it was originally specified, when this comparability was first introduced in the 1940's, that the correlation between SAT-V and SAT-M for that population would be .40, a not unrealistic figure at that time. More recent statistical summaries show that the correlations between Verbal and Math vary between .65 and .70, for the most part. To reflect this change in relationship between SAT-V and -M scores, the correlation between them that was used for the reference population in scaling was raised to .60 in the mid-1970's. Moreover, the means and standard deviations for Verbal and Math are no longer 500 and 100. Verbal means based on a total year's tested group are about 430, and Math means are about 475; the standard deviations are typically higher than 100, about 110 for Verbal and 120 for Math. However, the scale for each Achievement Test had historical significance in that it was based on a mean and standard deviation estimated for the same population for which the Scholastic Aptitude Test Verbal and Math scales were defined, the population for which those means were 500 and standard deviations, 100. In the case of the language Achievement Tests, an additional variable was used, supplementing the Verbal and Math scores, namely, the number of semesters for which candidates typically studied each language.

On the face of it, the foregoing model seems to be an excellent conception for a testing program in which every candidate is asked to take the same core test (the SAT), but in which each candidate is at liberty to take one or more optional tests (the Achievement Tests), chosen from a list available to him or her. However, there are several difficulties with the model which will be described in the following section of this report.

PROBLEM AND PURPOSE

In the early 1980's, Henry Braun and Ledyard Tucker (see Dorans, 1985) conducted studies designed to investigate Achievement Test rescaling (as defined by Wilks, 1959) that used both real and simulated data. These studies were undertaken to gain a better understanding of how operational decisions affected the outcomes of scaling. The effects of changing the definitions of the standard reference group and of changing the definition of the samples used for calculating the scaling equations as well as the effects of various choices of covariates were given particular attention.

One of Braun's studies was conducted using data from the December 1979 Achievement Test administration. In this study, Braun investigated, among other things, the effect of the choice of reference group on the rescaling results. (As mentioned previously, the method used until 1980 employed a hypothetical reference group with a mean of 500 and a standard deviation of 100 on SAT-V and SAT-M and a correlation between the scores on the two tests of .60). The results of Braun's studies indicated that choice of reference group had more than historical importance. Braun found that different reference groups produced different Achievement Test scale alignments. It should be noted that the choice of a suitable reference group has been discussed for many years, but no fully satisfactory way of redefining the standard reference group has been identified to date.

The simulation studies carried out by Tucker and Braun (see Dorans, 1985, for a detailed description of these studies) both employed a simulation model designed by Tucker to approximate the Achievement Test scaling situation. Tucker's model consisted of essentially two inter-related components. One component modeled Achievement and Aptitude Test performance and the second, Achievement Test selection. A maximum of two Achievement Tests could be taken in Tucker's study. The results of Tucker's study demonstrated that (Dorans, 1985, p. 17)

- The set of candidates who take the same pair of Achievement Tests need not have the same means and standard deviations on two properly aligned scales for the two tests. This finding demonstrated that it would be unwise to rescale by attempting to set means and standard deviations equal for all existing subgroups of candidates that took two Achievement Test forms.

- To the extent that developed academic skill plays a significant role in the test selection process, then there will be differential selection biases introduced that affect the rescaling results.

- To the extent that Achievement Tests correlate highly with their respective covariates, the differential selection biases can be overcome in the rescaling process.

As mentioned previously, Braun employed Tucker's simulation model in his study, which was an expansion on Tucker's work. Whereas Tucker limited his simulees (simulated test takers) to a maximum of two Achievement Tests, Braun allowed his simulees to take a maximum of three Achievement Tests. Braun simulated performance on eight hypothetical Achievement Tests, which varied in their relationships to the rescaling covariates, SAT-V and SAT-M scores. Braun also used two reference populations to determine the rescaling equations; the traditionally used hypothetical reference population (with scaled score mean on SAT-V and -M of 500 and standard deviation of 100) and the observed scaled score means and standard deviations of the December 1979 Achievement Test population. Braun's results led him to the following conclusions:

- As was indicated by his empirical studies, the choice of reference population can have a significant impact on rescaling results.

- Similar to Tucker's conclusions, the presently used methodology works well if Achievement Test scores are highly related to SAT-V and SAT-M scores or if the student's choice of Achievement Test is not primarily a function of his/her ability.

This latter point should be expanded upon. Both Tucker and Braun's results indicated that rescaling works well provided that ability has no impact on the test selection process. However, as noted earlier, ability clearly does affect the test selection process. Only if ability did not affect the selection process would rescaling no longer be necessary. When ability has a major impact on test selection, there is a need for rescaling and the quality of the rescaling will depend upon how highly the covariates (SAT-V and SAT-M) used to rescale the test correlate with Achievement Test performance.

To summarize, research results so far indicate that 1) because the test selection process is ability related, rescaling is most probably necessary, and 2) the currently used methodology is influenced by the choice of reference group and by the relationship between Achievement Test scores and the covariates (SAT-V and SAT-M scores) that are presently used to rescale the tests.

The purpose of the study described in this report was to explore the relationships between College Board Achievement Test scores and potential scaling covariates for various subgroups of the test taking population. It was expected that such an exploration would lead to the following:

- The selection of additional scaling covariates that might provide improved scaling results for those tests that d not provide scores correlating highly with SAT-V and/or SAT-M scores;
- The re-specification of the sample of students that are used to scale the test, i.e., such a re-specification might possibly lead to Achievement Test scores that show a higher correlation with selected scaling covariates and;
- The re-specification of the hypothetical scaling population. As Braun and Tucker pointed out, the characteristics of the hypothetical population differentially affect the scales of the tests. A change in specifications for this population might possibly provide scales that are more appropriate for some of the tests.

The final goal of the present study was to provide several alternative scaling procedures for the Achievement Tests. The procedures presented will be based on empirical evidence gathered in the study and will vary components such as scaling covariates, scaling samples and characteristics of the hypothetical reference group.

It was anticipated that all 14 tests would not be amenable to similar treatment; therefore, in presenting results, the tests will be clustered and alternative procedures will be specified for each cluster.

METHODOLOGY

Description of the Tests

The 14 Achievement Tests fall into five general subject areas:

English

- English Composition (two versions: all-multiple choice and multiple-choice with essay)
- Literature

Foreign Languages

- French
- German
- Hebrew
- Latin
- Spanish

History and Social Studies

- American History and Social Studies
- European History and World Cultures

Mathematics

- Mathematics Level I
- Mathematics Level II

Sciences

- Biology
- Chemistry
- Physics

All the Achievement Tests take one hour of testing time, and consist entirely of multiple-choice questions except the English Composition Test with Essay, which consists of a 20 minute essay and 40 minutes of multiple-choice questions. The tests vary in content as well as in the number of multiple choice questions they contain. The reader is referred to Taking the Achievement Tests, for a detailed description of test content. The approximate number of questions contained in each test form is listed in the table below.

Test	Approximate Number of Questions
English Composition with Essay	70 multiple choice items plus one essay
English Composition without Essay	90
Literature	60
French	85
German	80
Hebrew	90
Latin	70
Spanish	85
American History and Social Studies	95
European History and World Cultures	95
Mathematics Level I	50
Mathematics Level II	50
Biology	95
Chemistry	85
Physics	75

Data Collection

The Achievement Test data used to provide a base for this study was obtained from tapes used to select new form samples for operational equating purposes. These tapes are prepared when no less than sixty percent of the total volume of answer sheets from a particular administration have been scanned and scored. All cases available on the tapes were used to form the data base for this study. In no instance did the scaled score mean of the group used for the study differ by more than a few points from the total population of examinees taking the test at a specific administration.

The following Achievement Test administrations were used for the study.

o November	1983, 1984
o December[1]	1983, 1984
o January	1984, 1985
o May[1]	1984, 1985
o June	1984, 1985

Data from the 1983 and 1984, or 1984 and 1985 administrations were collapsed and treated as though they came from the same administration, i.e., all analyses were carried out on a total of five administration groups.

Scholastic Aptitude Test (SAT) scaled scores used for this study were the most recent scores obtained by candidates during the period from March 1983-January 1984 and the period March 1984-January 1985 (depending upon whether candidates took the Achievement Tests with the 1983-84 or 1984-85 cohort).

Analyses

The analyses consisted of obtaining, for each Achievement Test, counts, distributions, summary statistics, and intercorrelations for the following variables:

- o Achievement Test Scaled Score
- o SAT-V Scaled Score
- o SAT-M Scaled Score

---

[1] It should be noted that, at the time data were collected for this study, the small volume tests, European History and World Cultures, German, Hebrew, and Latin, were given only at the December and May administrations. Beginning with testing year 1986-87, these tests are offered only at the December and June administrations.

- Achievement Test Backgroud Questionnaire responses (for the French, German, Latin, Physics, Mathematics Level I and Mathematics Level II tests)
- Student Description Questionnaire (SDQ) responses to questions 5-17, 23, 24, 39, 40, 43.[1]

The analyses were performed for each test by administration by grade level combination.

Several points require noting. First, for this exploratory study, it was decided to analyze a large subset of available variables without particular regard to their suitability for use in operational scaling. The purpose was to assess how much the correlation between Achievement Test scores and the covariates could be increased and to identify the covariates which were most useful for enhancing the correlation. (The reader is referred to Tables 1 and 2 for a detailed description of the variables.) Second, since the data were collected for this study, a new Student Descriptive Questionnaire, as well as additional Achievement Test Background Questionnaires have been introduced in the Admissions Testing Program. The results of the study are limited in that data from those new questionnaires were not available for analyses.

In addition to the counts, summary statistics, and intercorrelations, covariate selection (stepwise regression) analyses were also carried out for each test by administration by grade level combination. The regression procedure used is referred to as the Maximum R^2 Improvement Technique (MAXR) and is available as part of the SAS Institute Statistical Aralyses System (1982). R^2 refers to the squared multiple correlation between any two or more variables.

---

[1] SDQ questions 24, 39, 40 and 43 were used in the regression analyses only. Students fill out the SDQ when they register to take the SAT or Achievement Tests. They have an opportunity to update information collected by this questionnaire whenever they register for another ATP test.

The Maximum R^2 Improvement Technique is considered superior to a conventional stepwise technique and almost as good as all possible regression solutions. Unlike most stepwise procedures, this technique does not settle on a single model. Instead, the procedure tries to find the best one variable model, the best two variable model, and so forth. The MAXR method begins by finding the one variable model producing the highest R^2. Then another variable, the one that yields the greatest increase in R^2, is added. Once the two variable model is obtained, each of the variables in the model is compared to each variable not in the model. For each comparison, MAXR determines if removing one variable and replacing it with the other variable increases R^2. After comparing all possible switches, the one that produces the largest increase in R^2 is made. Comparison begins again, and the process continues until MAXR finds that no switch could increase R^2. The difference between conventional stepwise techniques and the Maximum R^2 Improvement Method is that, in the Maximum R^2 Improvement Method, all switches are evaluated before any particular switch is made. In conventional stepwise procedures the "worst" variable may be removed without considering what adding the "best" remaining variable might accomplish. For the purposes of this report, only the results at the analyses producing the best two-, three-, four-, and five-variable models are reported.

Several points should be made regarding the type of regression analysis carried out for this study. First, the regression procedure could have been controlled in such a way that the covariates historically used to scale the tests (SAT-V, SAT-M, and in some cases, semesters of study) were automatically considered as part of the model. The computer program would then have searched for additional covariates that increased R^2 above that observed for the covariates traditionally used. The researchers chose not to do this for the present study. Instead, they conducted a completely exploratory analysis, resulting in a choice of covariates that maximized R^2 for a particular data set.

Secondly, like all stepwise regression procedures, the MAXR procedure is sample dependent, i.e., ideally, for the most generalizable results, the analyses should be carried out using the population of interest or a representative sample of this population. The current analysis was carried out using all available cases with scores on all variables of interest. This constitutes a very large sample of Achievement Test takers; however, the sample may not be representative of the total population because, as mentioned previously, it was selected when approximately 60% of the answer sheets for a particular administration were scored and because of the restriction that examinees have scores on all variables.

It is important to note that the actual groups used to carry out the correlational, regression, and analyses of summary statistics varied according to the type of analysis done. In order to provide the most representative results, all available cases were used to compute summary statistics for the various test scores. However, because not all students took the same Achievement Tests and because not all students have SAT scores, samples used to compute these summary statistics varied.

The correlational analyses were carried out using all available cases with scores on the two variables that were paired. Thus, sample size varies from pairing to pairing. Finally, as mentioned above, the regression analyses were carried out using only those examinees who had scores on all the variables of interest. Consequently, sample size varies from one regression analysis to another.

Table 1 contains brief descriptions of the variables used as Achievement Test covariates in the correlation and regression analyses. As seen in Table 1, variables 1 and 2 are SAT-V and SAT-M scaled scores respectively. Test booklets containing SAT forms, such as those used in this study, consist of six 30-minute sections: two SAT-V sections, two SAT-M sections, one Test of Standard Written English, and one experimental section. The two SAT-V sections contain a total of 85 five-choice items compose of 25 antonyms, 20 analogies, 15 sentence completions and 5 reading passages each of which is followed by five items based on the passage. The two SAT-M sections contain a total of 60 items composed of 40 five-choice regular mathematics items and 20 four-choice quantitative comparisons items.

Insert Table 1 about here

Variables 3 - 16 and 18 - 21, listed in Table 1, refer to questions on the Student Description Questionnaire (SDQ). The actual texts of the questions used appear in Table 2. Variable 17 is amount of training in a subject as assessed by the Background Questionnaires that were administered with the Achievement Tests. Only the following tests contain a Background Questionnaire: Mathematics Level I and II, Physics, French, German, Spanish and Latin. The Hebrew Test also contains a questionnaire, but the Hebrew Test is in the process of being extensively revised, so data from the Hebrew Test Quesionnaire were not evaluated for the study. Table 3 contains the actual text of the questionnaires.

Insert Tables 2 and 3 about here

Finally, it was necessary to recode responses to the Student Descriptive Questionnaire items as well as responses to the Achievement Test Background Questionnaires prior to analyses. The information given in Table 4 describes how the recoding was carried out.

Insert Table 4 about here

RESULTS

It is important for the reader to keep in mind the purposes of the study as the results of the analyses are reviewed. One purpose of the study was to consider respecification of the hypothetical reference population. SAT scaled scores were collected and summarized by grade level and administration for the groups taking the 14 Achievement Tests. It was expected that examination of these data would lead to confirmation or respecification of the manner in which the hypothetical scaling population is specified.

Achievement Test scaled scores obtained by grade level groups taking the 14 Achievement Tests across the five administrations were also summarized. The data were used to fulfill a second purpose of the study; i.e., examination of the Achievement Test sample used for scaling purposes. It was expected that an examination of these data would lead to a confirmation or respecification of the manner in which samples of Achievement Test examinees are traditionally selected for scaling studies.

Finally, correlational analyses and regression analyses were conducted to examine the relationship between Achievement Test scores and potential scaling covariates. The results of these studies were used to develop clusters of Achievement Test scores that had similar relationships to subsets of the covariates investigated in the study.

The results of the analyses are summarized in Tables 5-12 and Figures 1-9. Table 5 contains summary statistics by grade level and administration for the 14 tests. Summary statistics were computed for all available cases with Achievement Test scores. It should be noted, when reviewing the results found in Table 5, that four of the tests, European History and World Cultures, German, Hebrew and Latin were given only at the December and May Administrations.

Insert Table 5 about here

Table 6 presents summary statistics for SAT-V and SAT-M scaled scores for groups taking the 14 Achievement Tests. As was the case for the information displayed in Table 5, summaries were obtained by grade level for the five administrations. Two points should be noted when reviewing the data presented in Table 6. First, since not all students taking an Achievement Test at a particular administration have SAT scores on their records, the sample sizes reported in Table 6 are somewhat smaller than those reported in Table 5. Second, as was mentioned previously, four of the small volume tests were given only at the December and May Administrations.

Insert Table 6 about here

Table 7 contains correlation coefficients of Achievement Test scaled scores with the variables that were examined in the study as potential scaling covariates. The covariates are designated in the table as variable 1 (Var. 1) through variable 17 (Var. 17). As mentioned previously, titles and detailed descriptions for the variables are found in Tables 1-3. Table 4, as previously mentioned, contains information pertaining to the way in which the variables were recoded prior to conducting the correlation and regression analyses.

The correlations presented in Table 7 were calculated for the total group taking the particular Achievement Test at a particular administration. Grade level information was obtained as part of the study, but is not presented in this report. Again, it should be noted that only the December and May administration summaries report results for all 14 tests. Further, its important to note that the number of candidates used to compute the correlations vary from pairing to pairing. Minimum sample sizes for each test are given in Table 7. Table I, found in the appendix, provides the correlation matrices (with accompaning sample sizes for each correlation coefficient) that are summarized in Table 7.

Insert Table 7 about here

The information presented in Table 8 summarizes the results of the stepwise regressions carried out for the study. The regression analyses were carried out for the 14 tests by grade level and administration date. As mentioned previously, its important to note that the analyses were produced only for those examinees who had scores on all the variables of interest. Table II, (found in the appendix) provides summary statistics and intercorrelation matrices for the various analysis groups. Only those analyses based on sample sizes of 100 cases or more were included in Table 8.

Insert Table 8 about here

The results of the analyses will be presented in the following manner: First, the results of all the analyses conducted for a particular test will be presented for each of the 14 tests. Then the results of the analyses will be compared across the various tests.

English Composition Test: Summary Statistics

Table 5 contains summary statistics by grade level and administration for the English Composition Test. It is important to note that the total group includes some candidates who did not specify grade level.

It is apparent from Table 5 that, as expected, the majority of the test taking population are high school juniors and seniors, with seniors being more prevalent at the November, December and January administrations and juniors representing the major portion of test takers at the May and June administrations. The English Composition Test is not typically taken by high school freshmen. The only administration that contains a sizable number of freshmen candidates is the June Administration. High school sophomores also represent a very small part of the total candidate pool, at least for the November, December and January administrations.

An examination of the Achievement Test scaled score means for the various groups (those for groups with sample sizes less than 25 are not included in Table 5) indicates that high school sophomores have the highest scaled score mean of all grade levels taking the test at the November administration. This mean, however, is based on only 33 candidates and they are apparently very highly self selected. The scaled score means obtained by juniors and seniors taking the test at the November administration are quite similar, with the junior mean being just slightly higher. The highest scaled score mean obtained at the December administration (567) was again obtained by sophomore candidates (N=27). The junior mean for this administration (555) was considerably higher than that obtained by the high school seniors (521).

For the January administration, the highest English Composition Test mean score was obtained by high school juniors, the next highest by high school sophomores, and the lowest by those students reporting their grade level as senior. A similar pattern is found in the mean scores obtained for the English Composition Test given at the May administration. High school juniors obtained the highest mean score (527) and seniors the lowest mean (485) of the three grade levels. Finally, the mean scores obtained by candidates taking the test at the June administration are rank ordered high to low as follows: juniors, sophomores, freshmen, and seniors.

When one considers the achievement levels of the groups (as assessed by mean scores on the English Composition Test) across the five administrations, the following trends become clear. First, seniors produced the lowest English Composition Achievement Test mean scores at all five administrations. The mean scores obtained by seniors for the five administrations are rank ordered high to low as follows: November, December, January; followed by very similar May and June means. The pattern of mean scores for junior candidates is somewhat different from that found for senior candidates, i.e., higher mean scores were obtained by candidates taking the test in December, with January scores being the next highest. Scores obtained by juniors at the November administration are the next highest, followed by those obtained by junior June candidates. The lowest English Composition Test mean scores obtained by juniors are for the May administration. It appears that, in general (with the exception of sophomore scores obtained at the November and December administrations), juniors are the highest scoring group at each of the respective administrations.

As mentioned above, sophomores produced the highest scaled score means for any grade level at both the November and December administrations. But, it needs to be kept in mind, that the sample sizes are quite small for sophomores taking the test at these two administrations. Also, the standard deviation of the scores is quite high for the December sophomore group. It is interesting to note that sophomores produced higher scores than senior candidates at both the May and June administrations. Also interesting is the fact that freshmen taking the test in June obtained a higher scaled score mean than seniors who took the test at this same administration.

Table 6 contains SAT-V and SAT-M summary statistics by grade level for candidates taking the English Composition Test at the five administrations. Again it is important to note that "total" includes those examinees who did not specify grade level.

Considering only the junior and senior SAT-V and -M means for the November administration that are presented in Table 6, it is clear that the junior candidates are more able (as assessed by SAT-V and SAT-M scores) than the high school seniors. It should be noted that the correlation between SAT-V and SAT-M scores is lower for junior candidates (.51) than for senior candidates (.57). SAT-V and SAT-M scaled score means for juniors and seniors taking the English Composition test at the December Administration show similar trends to those exhibited by the November data, i.e., juniors obtained higher mean scores on both SAT-V and SAT-M than senior candidates. One difference between scores obtained at these two administrations can be noted in the correlation between SAT-V and SAT-M scores; this correlation is .55 for both groups.

For the January administration, sample sizes are large enough to consider the mean scores obtained by high school sophomores as well as by juniors and seniors. As previously noted, high school juniors score higher than seniors on both SAT-V and SAT-M. Interestingly, the 26 sophomores whose scores are summarized for the January administration also score higher than the seniors who took the test at this administration. High school sophomores produced the highest correlation between SAT-V and SAT-M scores (.77), and seniors the lowest correlation (.54).

The SAT-V and -M scaled scores summarized for the May administration indicate that juniors taking the test at this administration score higher on both variables than seniors. Sophomores also score higher than seniors on both SAT-V and SAT-M. Sophomore candidates score higher than junior candidates on SAT-M but lower on SAT-V. Scores for the sophomore candidates are based on a small sample (N=38) however, and have fairly large standard deviations (126 and 122, respectively). Correlations between SAT-V and SAT-M scores range from .66 for the sophomore candidates to .53 for the senior candidates.

Perusal of the data obtained for the June group taking the English Composition Test (shown in Table 6) indicates that freshmen (N=36) obtained the highest SAT-V scaled score mean and an SAT-M scaled score mean equal to that of the sophomores. Both freshmen and sophomores obtained higher SAT-V and SAT-M scaled score means than the junior or senior candidates. As was the case for the previous four administrations, juniors appear to be more able (as assessed by SAT-V and SAT-M scores) than seniors. The correlations between SAT-V and SAT-M scores range from a low of .42 for freshmen candidates to a high of .58 for juniors who took the test.

The following trends in both SAT-V and SAT-M scores across the five administrations can be noted. First, considering only junior and senior candidates, scaled score means for SAT-V range from a low of 470 for May senior candidates to a high of 560 for juniors taking the test in December. Scaled score means for SAT-M range from a low of 533 for May senior candidates to a high of 630 for November juniors. Further, it should be noted that the means for these tests vary considerably across grade level and administration and in most instances, they are considerably above a scaled score mean of 500. Secondly, it is important to note the relationship between SAT-V and SAT-M scaled scores, as indicated by the correlation between these scores, varies across grade levels and administrations. These correlations range from a high of .77 for sophomores taking the test in January to a low of .42 for freshmen taking the test in June. In general, most of the correlation coefficients appear to range between .55 and .60.

English Composition Test: Correlational Analyses

Table 7 contains correlation coefficients for Achievement Test scores with 16 of the 17 covariates for the total group taking the English Composition Test at the five administrations. It should be noted that correlations were computed for all examinees taking the test at a particular administration who had scores on both variables of interest. Thus, as mentioned previously, sample sizes vary from one pair of variables to another. Table I, found in the appendix, contains the correlation matrices summarized in Table 7. The matrices found in Table I list the sample size for each cell in the matrix. It should be noted that no correlations of English Composition Test score with variable 17 (Achievement Test Background Questionnaire responses) are given in Table 7. This is because no Background Questionnaire is administered with the English Compostion Test.

Examination of the data given in Table 7 indicates that, for the November administration, correlations of the 16 covariates with English Composition Test scores range from .01 for variable 7 (amount of high school course work in biology and the sciences) to .81 for variable 1 (SAT-V scaled scores). Variable 2 (SAT-M scaled scores) has the second highest correlation with the English Composition Achievement Test score. Other variables that show a correlation with the English Composition Test scores of .30 or greater are variables 3, 10, 11 and 13-16. (See Tables 1 and 2 for a description of the variables.)

An examination of the data given in Table 7 for the group taking the test at the December administration indicates a similar pattern to the one observed for the November correlations. The correlations with Achievement Test score range from .00 for variable 7 to .76 for variable 1. Again, SAT-M scaled scores (variable 2) offer the second highest correlation with English Composition Test scores. Other variables with correlations of .30 or greater are variables 3, 10, 12, 13, and 15. In general, all the correlation coefficients appear to be slightly lower for data collected at the December administration than for those collected in November.

Data obtained at the January administration shows similar trends to those obtained for the November and December administrations. SAT-V and SAT-M scaled scores (variables 1 and 2) correlate most highly with English Composition Test score. However, a number of other variables (variables 3, 10, 13, 15 and 16) show moderate relationships with the English Composition Test scores.

The correlation coefficients observed for the May and June data are similar to those observed for the other administrations. As expected, SAT-V and SAT-M scaled scores are the two variables most highly correlated with English Composition Achievement Test score. For both administrations, variable 7 (amount of high school course work in biology and the sciences) has the lowest correlation with English Composition score, whereas for both administrations (with the exception of SAT-V and SAT-M scores), variable 3 (high school class rank) has the highest correlation with English Composition Test score. Other variables that show a moderate relationship with English Composition Test scores obtained at the two administrations are variables 6, 10, 11, 12, 13, 14 and 15.

English Composition Test: Covariate-Selection Analyses

Results of the stepwise regression analyses carried out for the English Composition Test are summarized in Table 8. It is important to note that the analyses were carried out only for those examinees who had scores on all the variables included in the model. Thus, sample size varies considerably from analysis to analysis. Table II, found in the appendix, contains intercorrelation matrices for the analysis groups. Sample sizes are presented in this table. Analyses were not interpreted for groups with sample sizes of fewer than 100 cases.

Examination of the information presented in Table 8 indicates that, for the sophomore June analysis, the best two variable model is one that includes variables 1 and 2 (SAT-V and SAT-M scaled scores). This model accounts for 62% of the total variance. The best five variable model is one that includes variables 1 and 2 as well as variable 6 (amount of high school course work in foreign languages), variable 16 (number of honors or awards received during high school years), and variable 19 (father's level of education). Use of all 5 variables represents an increase in total variance accounted for of approximately 2.2%.

The regression analyses conducted for high school juniors taking the English Composition Test at the five administrations show some similarity in that the best two variable model for all administrations, with the exception of the December administration, is one that involves SAT-V and SAT-M scores as covariates. For the December data, the best two variable model includes variable 1 (SAT-V scaled scores) and variable 12 (grades received in a foreign language). Using data obtained from seniors taking the English Composition Test in June, the best two variable model was found to be one that includes variable

1 and variable 10 (grades received in English.) The percentage of variance accounted for by the two variable models ranges from a low of 59% for the analysis carried out for seniors taking the test in December to a high of 66% for seniors taking the test in November. The percentage of variance accounted for by the best 5 variable models ranges from a low of 61%, again for December seniors, to a high of 67% for November seniors.

Finally, the analyses carried out for the total groups taking the English Composition Test at the five administrations indicate that, for all groups except those examinees taking the test at the December administration, the best two variable model is one that uses scores on SAT-V and SAT-M. For the December group, the best two variable model appears to be one that uses scores on SAT-V and grades received in foreign languages. It is interesting to note that grades received in foreign languages (variable 12) is included in the best three variable model for all administration groups with the exception of the January group. For this group, the best three variable model includes SAT-V and SAT-M scores and variable 10 (grades received in English). The percentage of variance accounted for by the best five variable models ranges from a low of 61% for the model fit to the data from the December administration to a high of 67% for the model specified for the May group.

Literature Test: Summary Statistics

Summary statistics for the Literature Test by grade level and administration are found in Table 5. As was the case for the English Composition Test, it is important to note that total group includes some candidates who did not specify grade level.

Examination of the data presented in Table 5 indicates that the Literature Test is not typically taken by high school freshmen, and that the number of sophomores taking the test is substantial only for the June administration. The test appears to be taken predominantly by seniors at the November, December and January administrations and juniors at the June administration. The grade level composition of the groups taking the test at the May administration is mostly juniors and seniors.

An examination of the Literature Test scaled score means for all groups containing 25 or more examinees indicates several trends. For the November administration, juniors score higher than seniors; however, it should be noted that junior scores are based on a sample of only 33 examinees, probably very highly self selected and that these scores are quite variable. The data collected for the December administration again shows that juniors score higher than seniors and also that junior scores are considerably more variable than senior scores. The same pattern of scaled score means is noted for the May and June administrations, i.e., juniors score higher on the Literature Test than seniors and their scores are more variable. The greatest discrepancies between junior and senior means (36 points) occur for the June administration. It is interesting to note that the highest scores of any grade level taking the test at the June administration are those obtained by sophomores.

Various trends become clear when the achievement levels of the groups (as assessed by Literature Test scores) are evaluated across the five administrations. High school juniors, taking the test in December, are the most able junior group (as well as the most able of any group with 25 or more candidates) taking the test at any of the five administrations. The least able junior group is the group of students who took the test in May. The remaining junior means can be arranged from high to low as follows: November, June, and January.

The Literature Test scaled score means obtained by those reporting their grade level as senior are much more similar across the five administrations than the junior means. The highest means (430 and 428) were obtained by the groups taking the test at the November and December administrations, respectively. The senior groups taking the test at the May and June administrations also received similar scaled score means (494 and 490, respectively) which were both lower than those received by the fall groups. The senior group taking the test at the January administration exhibited a scaled score mean between those obtained by the fall and spring groups.

Table 6 presents SAT-V and SAT-M scaled score summary statistics for groups taking the Literature Test at the five administrations. Again, only summary statistics based on groups with 25 or more examinees are included in the table. Also, as was noted for the English Composition Test, the total group includes some students who did not specify grade level.

A comparison of SAT-V and SAT-M scaled score means for juniors and seniors taking the test at the December Literature Test administration indicates that juniors score higher on both SAT-V and SAT-M than seniors. One interesting point to note is the similarity between the verbal and math mean scores obtained by the juniors and, also, those obtained by the seniors. Typically, most Achievement Test groups exhibit a sizable difference in these scores, with SAT-M scores being higher than SAT-V scores.

Examination of the data obtained for the January administration indicates that high school juniors scored higher than high school seniors on both SAT-V and SAT-M. This same trend is apparent for both the May and June groups. When examining the June data, it should be noted that the highest scoring group on both SAT-V and SAT-M were the sophomores (N=36) who took the test at this administration.

The relationship between SAT-V and SAT-M scores (as assessed by the correlation coefficient) varies across administrations and grade levels. These correlations range from .49 for juniors taking the test in December to .66 for seniors taking the test in June.

There are several trends in SAT-V and SAT-M scores across the five administrations that are worth noting. Generally, the most able group (as assessed by SAT-V and SAT-M scores) taking the test at all administrations, with the exception of June sophomores, is the junior group. The junior and senior groups taking the test in November, December and June are the highest scoring groups that take the test. The lowest scoring groups, on both SAT-V and SAT-M, are the groups taking the test in May. It should be noted that, as was the case for the groups taking the English Composition Test, the SAT-V and SAT-M scaled score means vary considerably according to grade level and administration and are, in most instances, considerably above a scaled score of 500. It should also be noted that the relationship between SAT-V and SAT-M scores varies with grade level and administration. In general, most of the correlations range between .56 and .64.

One unusual point observed is the magnitude of SAT-V and -M scaled score means when paired for a particular group. As mentioned previously, in several instances, these means are more similar (e.g., December juniors and seniors, June juniors) than typically observed for Achievement Test takers.

Literature Test: Correlational Analyses

Table 7 contains correlation coefficients for the Literature Test score with 16 of the 17 experimental covariates for the total group taking the Literature Test at the five administrations. As mentioned previously, it is important to note that the correlations were computed for all examinees taking the test at a particular administration who had scores on both variables of interest; hence, sample size varies from one correlation coefficient to another. Also, as was the case for the English Composition Test, no Background Questionnaire is administered with the Literature Test, thus, no correlations of Literature Test scores with variable 17 appear in Table 7.

Examination of the data given in Table 7 indicates that, for the November administration, correlations of Literature Test score with the 16 covariates range from a correlation of .84 with variable 1 (SAT-V scaled scores) to (not surprisingly) a correlation of .00 with variable 7 (amount of high school course work in biology and sciences). The variable showing the second highest correlation with Literature Test score (.54) is variable 2 (SAT-M scaled scores). Other variables that show a correlation with the Literature Test score of .30 or greater are variables 2, 3, 10, 13, and 15. (See Tables 1 and 2 for a description of these variables.)

The data given in Table 7 for the December administration of the Literature Test show similar patterns in correlations to those observed for the November data. Again, variable 1 (SAT-V scaled scores) shows the highest correlation with Literature Test score (.82), variable 7 the lowest (-.01), and variable 2 (SAT-M scaled scores) the second highest correlation (.49). Other variables with correlations of .30 or greater with Literature Test scores obtained at the December administration are variables 2, 3 and 15.

Data obtained at the January, May and June Literature Test administrations show similar patterns to those obtained at the two administrations just discussed. In all cases, the variable that correlates most highly with Literature Test scores is SAT-V scaled scores, SAT-M scaled scores provide the second highest correlation and amount of high school course work in biology and sciences (variable 7) shows the lowest correlation.

Literature Test: Covariate-Selection Analyses

Results of the Literature Test regression analyses are summarized in Table 8. As mentioned previously, it is important to note that the analyses were carried out only for those examinees who had scores on all the variables included in the model; thus, sample size varies considerably from one analysis to another. The reader is referred to Table II in the appendix for the intercorrelation matrices and sample sizes used for the analyses. Analyses were not interpreted for groups with sample sizes of less than 100 cases.

The information presented in Table 8 for the Literature Test indicates that, for the junior May analysis, the best two variable model is one that contains variable 1 (SAT-V scaled scores), and variable 4 (amount of course work in English). This model accounts for 72% of the total variance. The best five variable model is one that also includes variable 2 (SAT-M scaled scores), variable 6 (amount of course work in foreign languages) and variable 16 (number of honors or awards received.) Use of all five variables represents an increase, from the two variable model, of 1% in the total variance accounted for.

The analyses for the junior June sample produced slightly different results than those obtained using May data. The best two variable model, which accounts for 69% of the total variance, is one that includes variables 1 and 10 (SAT-V scaled scores and grades in English). The best five variable model, which again represents an increase of 1% in the total variance accounted for, includes variables 1, 6, 8, 10, and 16. These variables are (in the order that they appear in the model) SAT-V scaled scores, course work in foreign languages, course work in physical sciences, grades in English, and number of honors or awards.

The analyses conducted for high school seniors taking the Literature Test at the five administrations shows some similarity in that the best two variable model for the November, December, January and May administrations is one that includes variable 1 (SAT-V scaled scores) and variable 10 (grades in English). For June seniors, the best two variable model appears to be one that employs variable 1 (SAT-V scaled scores) and variable 3 (high school class rank). The percentage of variance accounted for by the two variable models ranges from 68% for the December and May seniors to 71% for seniors taking the test at the other three administrations.

Increasing the number of variables in the model from two to five for the senior level data had little effect on the overall percentage of variance accounted for. The five variable models are not consistent across the different administrations, although they all have in common, variable 1 (SAT-V scaled scores) and variable 10 (grades in English).

The results of the total group analyses for the Literature Test are quite similar to those obtained for the senior data, with the exception that variables 1 and 10 constitute the best two variable model for all administrations. The percentage of variance accounted for by this model ranges from 68% to 70% across the five administrations. As was the case for the senior level data, increasing the number of variables in the model from two to five does not increase the percentage of variance accounted for by much. Again, the only variables that consistently enter into the five variable models, across the five administrations, are variable 1 (SAT-V scaled scores) and variable 10 (grades in English).

<u>American History and Social Studies Test: Summary Statistics</u>

Summary statistics by grade level and administration for the American History and Social Studies Test are found in Table 5. Perusal of the information found in this table indicates that, as was the case for the English Composition Test and the Literature Test, the American History Test is taken predominantly by high school juniors and seniors, with seniors tending to take the test in the fall and juniors tending to take the test in the spring. It is interesting to note that a surprisingly large number of sophomores take the test at the June administration. The only freshmen group containing 25 or more candidates is the one taking the test at the June administration.

Examination of the information presented in Table 5 indicates that high school juniors score higher than seniors at all five administrations. High school sophomores taking the test at the June administration score the same as the juniors who took the test at this administration. And, as was the case for the juniors, the sophomores score higher than the June seniors. It is worth noting that the small number of high school sophomores who took the test at the December administration (N=25) scored higher than either the juniors or seniors who took the test at the same administration.

The achievement level of the groups (as measured by the American History and Social Studies Test) varies considerably across grade levels and administrations. The junior means can be ranked high to low in the following order: June, November, May, December and January. For the senior group, the means, ordered by administration (again high to low), are November, December, January, June and May.

SAT-V and SAT-M summary statistics for students taking the American History and Social Studies Test at the five administrations are found in Table 6. Comparison of the junior and senior means for all five administrations indicates that high school juniors consistently score higher than high school seniors on both SAT-V and SAT-M. Sophomores taking the test in June score considerably higher on both SAT-V and SAT-M than either the juniors or seniors who took the test at this administration.

As expected, the relationship between SAT-V and SAT-M scores, as described by the correlation coefficient, varies by grade level and administration, with a range from .42, for juniors taking the test in December, to .66 for juniors taking the test in January.

Scaled score means for both SAT-V and SAT-M vary considerably across the various grade levels and administrations. SAT-V means range from a low of 480 for seniors who took the test in May to a high of 557 for sophomores taking the test in June. Scaled score means on SAT-M range from a low of 504 obtained by seniors taking the test in June to a high of 603, again for sophomores taking the test in June. In general, the scaled score means on both SAT-V and SAT-M for most grade levels and groups are considerably above 500.

American History and Social Studies Test: Correlational Analyses

Table 7 contains correlation coefficients for the American History and Social Studies Test scores with 16 of the 17 covariates used in the study. The correlations are for the total group taking the test at the five administrations. The reader is again referred to Table I in the appendix for complete correlation matrices and accompanying sample sizes. It should be noted that Table 7 does not contain correlations of American History Test scores with variable 17; this is because no Background Questionnaire is given with the American History and Social Studies Test.

Examination of the data given in Table 7 indicates that, for the November administration, correlations of the 16 covariates with the American History Test score range from a high of .76 for variable 1 (SAT-V scaled scores) to a low of .01 for variable 7 (amount of course work in biology and sciences). Variable 2

(SAT-M scaled scores) has the second highest correlation with American History Test scores. Other variables that show a moderate relationship with the American History and Social Studies Test scores (a correlation of .30 or greater) are variables 3, 10, 13 and 15. These variables are, respectively, high school class rank, grades in English, grades in biology and sciences, and grades in social studies.

Data obtained at the December administration exhibit similar patterns of correlation coefficients. American History Test scores correlate most highly with variable 1 (SAT-V scaled scores) and second highest with variable 2 (SAT-M scaled scores`. Variables 3, 10, and 15 all show correlations greater than .30 with the American History and Social Studies Test scores.

An examination of the correlation coefficients obtained for data from the remaining three administrations shows the same relationship between American History and Social Studies Test scores and the 16 covariates. For each administration, the variable that correlates most highly with American History Achievement Test scores is variable 1 (SAT-V scaled scores), and the second most highly correlated variable is variable 2 (SAT-M scaled scores). It is interesting to note that relationships with variables 6, 11, 12 and 14, which are, respectively, course work in foreign languages, grades in mathematics, grades in foreign languages, and grades in physical sciences, appear to be stronger for data collected at the two spring administrations of the test than for data collected at fall administrations.

American History and Social Studies Test: Covariate-Selection Analyses

Results of the regression analyses carried out for the American History and Social Studies Test are summarized in Table 8. The reader is referred to Table II of the appendix for intercorrelation matrices of the variables used in the analyses.

The regression analyses conducted for the high school sophomores taking the test at the June administration indicate that the best two variable model is one that contains variable 1 (SAT-V scaled scores) and variable 19 (father's level of education). This model accounts for 60% of the total variance. The best five variable model includes variables 1, 4, 5, 15 and 19 which are respectively, SAT-V scaled scores, amount of course work in English, amount of course work in mathematics, grades in social studies, and father's level of education. The five variable model accounts for 65% of the total variance, an increase in 5% over that accounted for by the two variable model.

The analyses carried out for high school juniors who took the test in May and June indicate that the best two variable model for both these groups is one that includes variable 1 (SAT-V scaled scores) and variable 15 (grades received in social studies). The two variable model accounts for 60% of the total variance for the May group and 61% of the total variance for the June group.

The analyses carried out for the seniors indicate, across all five administrations, that the best two variable model is one that includes variables 1 and 15 (SAT-V scaled scores and grades received in social studies, respectively). The percentage of variance accounted for by this model across the five administrations ranges from 56% for the January group to 64% for the June group.

The results of the analyses carried out for the total group were quite similar to those obtained using senior data. The best two variable model for the total group, for all five administrations, is one that contains variable 1 (SAT-V scaled scores) and variable 15 (grades in social studies). The percentage of variance accounted for by this model ranges from 56% for data obtained from the January administration to 63% for data obtained from the May administration. The best three variable model is also consistent across the five administrations. This model includes variables 1 and 15, but also variable 2 (SAT-M scaled scores). Addition of SAT-M scaled scores to the model allows, in most cases, for an additional 1-2% of the total variance to be accounted for. The addition of fourth and fifth variables to the models did not show substantial increases in the percentage of variance accounted for (above that accounted for by the three variable model) for many of the administrations. The three common variables (across the administrations) included in the five variable models were variables 1, 2 and 15 (SAT-V and -M scaled scores and grades in social studies).

Mathematics Level I Test: Summary Statistics

Achievement Test summary statistics for the Mathematics Level I Test are found in Table 5. Similar to tests that have been previously examined, it is apparent that the Mathematics Level I Test is taken predominately by high school juniors and seniors. Sophomores do constitute a reasonable number of the examinees who take the test at the May and June administrations and a sizeable number of freshmen take the test in June.

An examination of the Math Level I Achievement Test scaled score means for the various groups (only those with sample sizes greater than 25 are included in Table 5) indicates that sophomores and juniors score similarly on the November administration of the test, with high school seniors scoring considerably lower at this same administration. For the remaining administrations, sophomores appear to be the highest scoring group and seniors the lowest. Scores obtained by high school juniors are more or less intermediate to those obtained by sophomores and seniors. Scores obtained by freshmen taking the test at the May administration are higher than those obtained by juniors. The opposite effect is observed in the data obtained for the June administration, i.e., freshman taking the test at this administration score lower than high school juniors.

When one considers the achievement levels of the groups (as assessed by mean scores on the Mathematics Level I Test) across the five administrations, the following trends become clear. First, the mean scores obtained by seniors rank consistently, high to low, from the November to the June administration. This consistency in ranking of mean scores is not apparent when one examines the scores obtained by high school juniors. Mean scores obtained by juniors are ranked high to low across the five administrations as follows: November, January, December, June, and May.

Table 6 contains SAT-V and SAT-M summary statistics by grade level for candidates taking the Mathematics Level I Test at the five administrations. Examination of the data presented in Table 6 for the students taking the Mathematics Level I Test indicates that high school juniors score higher than high school seniors on both SAT-V and SAT-M for all five administrations. Considering only high school juniors, SAT-V scaled score means range from a low of 501, for the group taking the test in May, to a high of 534 for the December candidates. Scaled score means on SAT-M obtained by junior candidates range from 554 (for the May group) to 604, obtained by high school juniors who take the test in November. Scaled score means on SAT-V for senior candidates range from 460 (candidates taking the test in January) to 507 for candidates taking the test in November. Senior scaled score means on SAT-M range from 522 for the May group to 564 for the November group.

In general the correlations between SAT-V and SAT-M scores are higher for the junior groups than for the seniors. These correlations range from .53 to .59 for juniors and from .50 to .55 for seniors.

For the May administration, sample sizes are large enough to consider mean scores obtained by sophomores, as well as those obtained by juniors and seniors. It is interesting to note that, for this administration, sophomores score considerably higher on both SAT-V and SAT-M than high school juniors and seniors. In addition, the high correlation between SAT-V and SAT-M scores (.66) obtained by this group is noteworthy.

Sample sizes for groups taking the Mathematics Level I Test at the June administration are large enough to consider mean scores obtained by all four grade levels. Sophomores obtained the highest scores on SAT-V for this administration, but high school freshmen obtained the highest scores on SAT-M. For both freshmen and sophomores, the correlations between SAT-V and SAT-M scores were quite low (.35 for freshmen and .47 for sophomores).

Mathematics Level I Test: Correlational Analyses

Table 7 contains correlation coefficients for Achievement Test scores with 17 covariates for the total group taking the Mathematics Level I Test at the five administrations. It should be noted that a Background Questionnaire is administered with the Mathematics Level I Test; hence, correlations of test scores with variable 17 appear in Table 7. Examination of the data given in Table 7 indicates that, for the November administration, correlations of the 17 covariates with the Mathematics Level I Test scores range from -.03 for variable 7 (amount of high school course work in biology and the sciences) to .84 for variable 2 (SAT-M scaled scores). Variables 1 and 11 (SAT-V scaled scores and grades in mathematics, respectively) correlate the same (.50), and second highest with Math Level I Achievement Test score. Other variables with moderate correlations (.30 or greater) are variables 3, 5, 8, 13, 14 and 17. (See Tables 1 and 2 for a description of these variables.)

An examination of the data given in Table 7 for the December administration indicates a similar correlational pattern to the one observed for November. The correlations with Achievement Test score range from -.04 (variable 7) to .82 for the correlation with variable 2 (SAT-M scaled scores). Other variables showing correlations with Achievement Test score of .30 or greater are variables 1, 3, 5, 8, 11, 13, 14 and 17. In general, most of the correlations exhibited by the December data are slightly lower than those observed for data collected at the November administration.

Results of the analyses carried out using data obtained at the January administration show similar trends to those obtained for the November and December analyses. SAT-M scaled scores (variable 2) and SAT-V scaled scores (variable 1), as well as grades in mathematics (variable 11), correlate most highly with Mathematics Level I scores. However, a number of other variables (variables 3, 5, 8, 13, 14, and 17) show correlations of .30 or greater with Achievement Test score.

The correlation coefficients observed for the May and June data are similar to those observed for the other administrations. As expected, SAT-M scaled scores correlate most highly with scores obtained on the Mathematics Level I test. SAT-V scaled scores and grades in mathematics correlate the second most highly with Achievement Test scores. The correlation of Mathematics Level I scores with variable 17 (Achievement Test Background questionnaire responses) is lower for the May and June data than for data obtained at the fall administrations.

Mathematics Level I Test: Covariate-Selection Analyses

Results of the regression analyses carried out for the Mathematics Level I Test are summarized in Table 8. As mentioned previously, no analyses for groups containing fewer than 100 examinees are tabled. Examination of the information presented in Table 8 indicates that, for the sophomore June analysis, the best two variable model is one that includes variables 2 and 11 (SAT-M scaled scores and grades in mathematics). This model accounts for 63% of the total variance. The best five variable model is one that includes variables 2, 17, 8, 11 and 21. Employment of all five variables results in an increase in the amount of variance accounted for of 2.2%.

The regression analyses conducted for the high school juniors taking the test at the five administrations show some similarity across administrations. The best two variable model for most of the administrations is one that includes variables 2 and 11 (SAT-M scaled scores and grades in mathematics, respectively). The amount of variance accounted for by this model, used with data from the December, May and June administrations, ranges from 71% to 74%.

The best two variable model for juniors taking the test at the November administration appears to be one that includes variable 2 (SAT-M scaled scores) and variable 17 (Achievement Test Background Questionnaire responses). This model accounts for 74% of the variance in the November data. For the January administration, the best two variable model for high school juniors taking the test is one that includes variable 2 and variable 8 (amount of course work in physical sciences).

An examination of the regression analyses carried out for high school seniors taking the Mathematics Level I Test indicates that, in most cases, the best two variable model is one that include variables 2 and 17 (SAT-M scaled scores and Achievement Test Background Questionnaire responses, respectively). The one exception to this pattern is the model fit to data obtained at the June administration. For these data, the best two variable model is one that includes variable 2 and variable 11 (grades in mathematics). It is important to note that the best three variable model for high school seniors taking the test at all five administrations is one that includes variables 2, 17 and 11. The total variance accounted for by this model, applied to data from the five administrations, ranges from 67% to 72%.

The analyses carried out for the total group indicates that the best three variable model for the November - May administrations is one that includes variables 2, 17 and 11. For the total group taking the Mathematics Level I Test at the June administration, the best three variable model is one that includes variables 2, 11 and variable 1 (SAT-V scaled scores). The total variance accounted for by the three variable models applied to data from the five administrations ranges from 67% to 73%.

Mathematics Level II Test: Summary Statistics

Summary statistics for the Mathematics Level II Test by grade level and administration are found in Table 5. Examination of the data presented in Table 5 indicates, as has been the case for the other tests discussed so far, that the Mathematics Level II Test is not typically taken by high school freshmen or sophomores. The only groups for these grade levels containing more than 25 examinees are sophomores taking the test at the January and May administrations and freshmen and sophomores taking the test at the June administration.

An examination of the Mathematics Level II scaled score means summarized in Table 5 indicates several trends. For all five administrations, the highest scaled score means for the Mathematics Level II Test are those obtained by the sophomores who took the test. For the June administration, the highest scaled score mean on the Achievement Test was obtained by high school freshmen and the second highest by sophomores. In general, the highest scores were obtained by the groups taking the test in November and the lowest by the groups who took the test in May.

Table 6 presents SAT-V and SAT-M scaled score summary statistics for groups taking the Mathematics Level II Test at the five administrations. A comparison of SAT-V and SAT-M scaled score means obtained by high school juniors and seniors indicates that juniors score higher than seniors on both tests at all five administrations.

High school sophomores who took the Mathematics Level II Test at the May administration scored higher on SAT-M than either juniors or seniors. This same group scored lower on SAT-V than juniors but higher on this test than seniors. An examination of data obtained at the June administration of the Mathematics Level II Test indicates that high school freshmen taking the test at this administration scored higher on SAT-M than any other grade level. High school freshmen scores on SAT-V were lower than those obtained by sophomores or juniors, but higher than those obtained by seniors taking the test at the June administration.

In general, SAT-V scaled score means obtained by juniors taking the test across the five administrations ranged from 541, for the group taking the test at the November administration, to 576 for the group taking the test at the December administration. Scaled score means on SAT-M for high school juniors ranged from 649 for the group that took the test in May to 675 for the group that took the test in December. Senior SAT-V means ranged from a low of 483 for the group taking the test at the May administration to 557 for the group taking the test at the November administration. Senior scaled score means on SAT-M range from a low of 569 for the group that took the test at the May administration to a high of 656 for the group taking the test in November.

The relationship between SAT-V and SAT-M scores (as assessed by the correlation coefficients) varies across administrations and grade levels. These correlations range from .43 for juniors taking the test at the December administration to .60 for juniors taking the Mathematics Level II Test in November.

Mathematics Level II Test: Correlational Analyses

Table 7 contains correlation coefficients for the Mathematics Level II Test score with 17 covariates for the total group taking the Mathematics Level II Test at the five administrations. It should be noted that a Background Questionnaire is administered with the Mathematics Level II Test; thus, correlations of Achievement Test score with variable 17 appear in Table 7. Examination of the data given in Table 7 indicates that, for the November administration, correlations of the Mathematics Level II Test scores with the 17 covariates range from a correlation of .00 for variable 7 (amount of course work in biology and sciences) to a high of .79 for variable 2 (SAT-M scaled scores). The variable correlating the second highest with Mathematics Level II test score is variable 1 (SAT-V scaled scores). Other variables that show a correlation with Math Level II Achievement Test score of .30 or greater are variables 3, 11, 14 and 17.

The data given in Table 7 for the December administration of the Mathematics Level II Test show similar patterns in correlations to those observed for the November data. Again, variable 2 (SAT-M scaled scores) shows the highest correlation with Mathematics Level II Test scores. Other variables with correlations of .30 or greater with the scores obtained at the December administration are variables 1, 3, 11, 14 and 17.

Data obtained at the January, May and June Mathematics Level II Test administrations show similar patterns to those obtained at the two administraions just described. In all cases, the variable that correlates most highly with Mathematics Level II Test scores is SAT-M scaled scores. SAT-V scaled scores correlate the second highest and variable 11 (grades in mathematics) correlates the third highest.

Mathematics Level II Test: Covariate-Selection Analyses

Results of the Mathematics Level II regression analyses are summarized in Table 8. As has been the case for the regression analyses described so far, analyses are not tabled for groups with sample sizes of fewer than 100.

The information presented in Table 8 for the Mathematics Level II Test indicates that for the sophomore June analysis, the best two variable model is one that includes variables 2 (SAT-M scaled scores) and variable 8 (amount of course work in mathematics.) The best five variable model for this group is one that includes variables 2, 17, 6, 8, and 14. The total variance accounted for by the five variable model is approximately 47%.

The analyses carried out for the December and June junior groups indicate that the best two variable model for both groups is one that includes variable 2 (SAT-M scaled scores) and variable 11 (grades in mathematics). This model accounts for 68% of the total variance for the December group and 62% of the total variance for the June group. The best three variable model for the December and June groups is not consistent across the two administrations. The model specified for the December juniors is one that includes variables 2, 17, and 11 (SAT-M scaled scores, Achievement Test Background Questionnaire responses, and grades in mathematics, respectively). The best three variable model specified for the June group is one that includes variables 2 and 11, but also variable 1 (SAT-V scaled scores).

The analyses conducted for the high school seniors taking the Mathematics Level II Test at the five administrations show some similarity in that the best two variable model specified for all administrations, with the exception of June, is one that includes variables 2 and 17, (SAT-M scaled scores and Achievement Test Background Questionnaire responses). For the June administration, the best two variable model is one that includes variables 2 and 11 (SAT-M scaled scores and grades in mathematics).

The best three variable models specified for seniors taking the test at the five administrations show some similarity in that, for all administrations, variables 2 and 17 are included in the models. Differences occur in models that include variable 1 (SAT-V scaled scores) or variable 11 (grades in mathematics).

The results of the total group analyses for the Mathematics Level II Test are quite similar to those obtained for the senior data. It is interesting to note that, for all administrations, the best four variable model is one that includes variables 1, 2, 11 and 17. The percentage of variance accounted for by this model ranges from a low of 63% for the November and December administration to a high of 73% for the May administration.

Biology Test: Summary Statistics

Summary statistics by grade level and administration for the Biology Test are found in Table 5. Perusal of the information found in this table indicates that, for the fall administrations (November, December, January), the test is taken predominately by high school juniors and seniors, as has been the case for other Achievement Tests discussed so far. The situation is quite different for the spring administrations (May and June). A surprisingly large number of high school freshmen and sophomores take the test at both of these administrations, with freshmen and sophomores constituting the major portion of test takers for the June administration.

Examination of the information presented in Table 5 indicates that high school sophomores that take the test at the fall administrations score higher than either juniors or seniors. For the spring administrations, the trend is reversed and juniors score higher than sophomores. High school freshmen score almost as high as juniors who take the test in May and are the highest scoring group of any grade level to take the test in June. High school seniors are consistently the lowest scoring grade level group to take the test at any of the five administrations. The total group means for the spring administrations are considerably higher than the total group means obtained by the fall groups. This appears to be due mainly to the high scoring freshmen and sophomores who take the test in May and June.

SAT-V and SAT-M summary statistics for students taking the Biology Test at the five administrations are found in Table 6. It is immediately apparent, if one compares the sample sizes for the freshmen and sophomores who take the test in June with the sample sizes reported for the same groups in Table 5, that only a few of the freshmen and sophomores taking the Biology Test in June had available SAT scores.

Comparison of the junior and senior means at all five administrations indicate that high school juniors consistently score higher than high school seniors on both SAT-V and SAT-M. SAT-V scaled score means for high school juniors range from a high of 552 obtained for juniors taking the test in November to a low of 522 for those students taking the test in January. Junior SAT-M means range from a high of 651 for the November group to a low of 579, obtained by the May and June groups. High school senior SAT-V scaled score means range from 519 to 469, with the highest mean obtained by the November group and the lowest by the June group. SAT-M scaled score means obtained by high school seniors range from a high of 584 for the November group to a low of 527 for the June group.

The only freshmen and sophomore groups that contain 25 or more candidates are those students taking the test at the June administration. Both of these groups scored lower on SAT-V than the high school juniors taking the test at this administration; however, freshmen and sophomore SAT-V scores were higher than those obtained by seniors. Students reporting their grade level as freshmen scored higher on SAT-M than any other grade level group taking the Biology Test in June. Those reporting their grade level as sophomores scored lower on SAT-M than freshmen and juniors but higher than seniors taking the test at this administration.

As expected, the relationship between SAT-V and SAT-M scaled scores (as described by the correlation coefficient between these scores) varies by a grade level and administration, with a range from .43 for freshmen taking the test in June to .74 for juniors taking the test in November. In general, the correlation between SAT-V and SAT-M scaled scores for the total group ranges between .56 and .63 across the five administrations.

Biology Test: Correlational Analyses

Table 7 contains correlation coefficients for the Biology Test scores with 16 of the 17 covariates used in the study. The correlations are for the total group taking the test at the five administrations. The reader is referred to Table I of the appendix for complete correlation matrices and accompanying sample sizes. It should be noted that Table 7 does not contain correlations of Biology Test scores with variable 17; this is because no Background Questionnaire is given with the Biology Test.

Examination of the data given in Table 7 indicates that, for the November administration, correlations of the 16 covariates with Biology Test scores range from a high of .75 for variable 1 (SAT-V scaled scores) to a low of .06 for variable 4 (amount of course work in English). Variable 2, (SAT-M scaled scores) has the second highest correlation with Biology Test scores. Other variables that show a moderate relationship with Biology Test scores (a correlation of .30 or greater) are variables 3, 10, 11, 13, 14 and 15. These variables are, respectively, high school class rank, grades in English, grades in mathematics, grades in biology and sciences, grades in physical sciences, and grades in social studies.

Results of the analyses of data obtained at the December administration exhibit similar patterns of correlation coefficients. Biology Test scores correlate most highly with variable 1 (SAT-V scaled scores) and second highest with variable 2 (SAT-M scaled scores). Variables 3, 11, 13, 14 and 15 all show correlations of .30 or greater with Biology Test scores.

An examination of the correlation coefficients obtained for data from the remaining three administrations show the same relationship between Biology Test scores and the 16 covariates. For each administration, the variable that correlates most highly with Achievement Test scores is variable 1 (SAT-V scaled scores) and the second most highly correlating variable is variable 2 (SAT-M scaled scores).

Biology Test: Covariate-Selection Analyses

Results of the regression analyses carried out for the Biology Test are summarized in Table 8. The reader is referred to Table II of the appendix for intercorrelation matrices of the variables used in the analyses. The regression analyses conducted for the high school freshmen taking the test in June indicate that the best two variable model is one that includes variable 18 (highest level of education planned beyond high school) and variable 19 (father's level of education). This model accounts for approximately 40% of the total variance in Biology Test scores. The best five variable model for freshmen taking the test at the June administration accounts for approximately 56 % of the total variance and includes variables 3, 16, 18, 19 and 20 which are, respectively, high school class rank, number of honors or awards received during high school, highest level of education planned beyond high school, father's level of education and mother's level of education.

The analyses carried out for high school sophomores taking the test in June indicates that the best two variable model is one that includes variable 1 (SAT-V scaled scores) and variable 2 (SAT-M scaled scores). This model accounts for 57% of the total variance. The best three variable model for sophomores taking the test in June accounts for approximately 60% of the total variance and is one that includes variables 1 and 2 as well as variable 13 (grades in biology and the sciences).

The analyses carried out for high school juniors who took the test in December and June indicate that the best two variable model, for data collected at both these administrations, is one that includes variable 1 (SAT-V scaled scores) and variable 2 (SAT-M scaled scores). This model accounts for approximately 48% of the total variance for the December group and approximately 55% of the total variance for the June group. The best three variable model for juniors taking the test in December appears to be one that includes variables 1, 2 and 9 (amount of course work in social studies). This model accounts for approximately 52% of the total variance. The best three variable model for juniors taking the test at the June administration is one that includes variables 1, 2, and 13 (grades in biology and sciences). This model accounts for approximately 57% of the total variance.

The analyses carried out for the seniors indicate, across all five administrations, that the best two variable model is one that includes variable 1 and 2. The amount of variance in Biology Test score acccounted for by this model ranges from 55% for seniors taking the test in May to approximately 62% for the groups taking the test in November and June. The best three variable model is fairly consistent across the five administrations with the exception of the model specified for the November seniors. The model specified for this group is one that includes variables 1, 2 and 13. The best three variable model specified for the remaining four administrations is one that specified variables 1, 2 and variable 7 (amount of course work in biology and the sciences). The percentage of variance accounted for by this model across the four administrations ranges from approximately 57% for data obtained at the May administration to 64% for data obtained for seniors taking the test at the June administration.

The results of the analyses carried out for the total groups are fairly consistent across the five administrations. The best two variable model for all five administrations is one that includes variable 1 (SAT-V scaled scores) and variable 2 (SAT-M scaled scores). The percentage of variance accounted for by this model ranges from 56% for data obtained at the June administrion to 64% for data obtained at the May administration.

The best three variable models specified for the total groups vary somewhat from administration to administration. All of the three variable models include variables 1 and 2. The third variable included in the model specified for the November, May and June administrations is variable 13 (grades in biology and sciences). The third variable included in the model specified for the December and January groups is variable 7 (course work in biology and sciences).

Chemistry Test: Summary Statistics

Table 5 contains summary statistics by grade level and administration for the Chemistry Test. As mentioned previously, it is important to note that the total group includes some candidates who did not specify grade level.

It is apparent, from examination of the data present in Table 5, that the Chemistry Test is not typically taken by high school freshmen or sophomores at the November, December or January administrations. A reasonable number of sophomores appeared to have taken the test at the May administration. A similar number of freshmen took the test in June. Finally, similar to the pattern obtained for the Biology Test, high school sophomores consititute a large proportion of the June test takers.

An examination of the Achievment Test scaled score means for the various groups indicates that, as has been observed for all the tests discussed so far, juniors score higher on the Chemistry Test than seniors at all five administrations. High school sophomores taking the test in May also score higher than seniors taking the test at this administration and just slightly lower than the May juniors. The highest scaled score mean obtained by any group taking the test at the June administration was that obtained by students reporting their grade level as freshmen. High school sophomores scored the second highest on the test, followed by juniors and then by high school seniors.

An examination of the total group means across the five administrations indicates that the most able group (as assessed by Chemistry Test scores) is the group taking the test in June. The group with the second highest scaled score mean is the group taking the test in November. The lowest Chemistry scaled score mean was obtained by the total group taking the test at the January administration.

Table 6 contains SAT-V and SAT-M summary statistics by grade level for candidates taking the Chemistry Test at the five administrations. It should be noted, as was the case for the examinees taking the Biology Test, the number of freshmen and sophomores who took the test in May and June with appropriate SAT scores is small compared to the total number who actually took the Chemistry Test at these administrations.

Considering only junior and senior scaled score means presented in Table 6, the same pattern that has been observed for the previous tests discussed is apparent, i.e., juniors score higher than seniors on both SAT-V and SAT-M for all five administrations. SAT-V scaled scores obtained by juniors range from a low of 528 to a high of 555. Those obtained by seniors range from a low of 476 to a high of 545. SAT-M scaled score means obtained by juniors range from 606 to 664 and those obtained by seniors range from 587 to 650.

Sophomores who took the test at the May administration have a higher SAT-M mean than junior or senior candidates but a lower SAT-V mean than juniors taking the test at this administration. The highest SAT-M scaled score mean for the June administration was obtained by freshmen examinees and the second highest by sophomores. High school sophomores taking the test in June obtained the highest SAT-V mean, followed by freshmen, then juniors, and finally, seniors.

An examination of the relationship between SAT-V and SAT-M scores (as assessed by the correlation coefficient) indicates that these correlations range from .41 for juniors taking the test in May to .68 for the small number of sophomores (N=25) also taking the test in May.

Chemistry Test: Correlational Analyses

Table 7 contains correlation coefficients for Achievement Test Score with 16 of the 17 covariates for the total group taking the Chemistry Test at the five administrations. As mentioned previously, it should be noted that sample sizes vary from one pair of variables to another. Table I, found in the appendix, contains the correlation matrices summarized in Table 7. It should also be noted that no correlations of Chemistry Test scores with variable 17 (Achievement Test Background Questionnaire responses) are given in Table 7. This is because no Background Questionnaire is administered with the Chemistry Achievement Test.

Examination of the data given in Table 7 indicates that, for the November administration, correlations of the 16 covariates with Chemistry Test Scores range from .63 for variable 2 (SAT-M scaled scores) to .02 for variable 4 (amount of course work in English). Variable 1 (SAT-V scaled scores) has the second highest correlation with Chemistry Test scores. Other variables that show a correlation with the Chemistry Test scores of .30 or greater are variables 3, 8, 11, 13 and 14. These variables are, respectively, high school class rank, amount of course work in physical sciences, grades in mathematics, grades in biology and sciences, and grades in physical sciences.

An examination of the data given in Table 7 for the group taking the test at the December administration indicates a similar pattern to the one observed for the November correlations. The correlations with Chemistry Achievement Test scores range from .68 with variable 1 (SAT-M scaled scores) to .02 with variable 4 (amount of course work in English). Other variables correlating moderately with Chemistry Test scores are variables 1, 3, 11, 13 and 14.

Data obtained at the remaining three administrations show similar patterns to those observed for the November and December groups. In general, variable 2 (SAT-M scaled scores) shows the highest correlation with Chemistry Test scores and variable 1 (SAT-V scaled scores), the second highest. Variables 3, 8, 11, 13 and 14 show moderate correlations with Chemistry Test scores.

Chemistry Test: Covariate-Selection Analyses

Results of the regression analyses carried out for the Chemistry Test are summarized in Table 8. As mentioned previously, it is important to note that the analyses were carried our for only those examinees who had scores on all the relevant variables. The reader is referred to Table II, found in the appendix, for intercorrelation matrices for all analysis groups. As has been the case for other tests, analyses are not presented for groups containing less than 100 examinees.

Examination of the information presented in Table 8 indicates that, for the sophomore June analysis, the best two variable model is one that includes variable 1 (SAT-V scaled scores) and variable 2 (SAT-M scaled scores). This model accounts for approximately 36% of the total variance. The best five variable model for sophomores taking the test in June appears to be one that includes variables 1, 2, 4, 8 and 11. (See Table 1 for a description of these variables.) The best five variable model fit to June sophomore data accounts for approximately 37% of the total variance.

The best two variable model specified for juniors taking the Chemistry Test at the December and June administrations is not consistent across the two administrations. For the December junior data, the best two variable model accounts for approximately 70% of the total variance and includes variables 2 (SAT-M scaled scores) and variable 10 (grades in English.) The best two variable model for the June group includes variables 1 and 2 and accounts for approximately 73% of the total variance.

The analyses carried out for the high school seniors taking the test indicates that the best two variable model fit to data obtained at each of the five administrations is one that includes variables 1 and 2. The percentage of variance accounted for by this model ranges from 47% for the November group to 58% for the June group.

The best three variable model fit to the senior data is also consistent across the five test administrations. This model includes variables 1 and 2, as well as variable 8 (amount of course work in physical sciences). The percentage of variance accounted for by this model ranges from approximately 50% for data obtained from the November administration to 63% for seniors taking the test in June.

The results of the analyses conducted for the total group were very similar to those obtained for the senior group, i.e., the best two variable model specified is consistent across the five administrations, and is one that contains variables 1 and 2. The best three variable model specified is consistent with the results obtained for the senior analyses with the exception of the model fit to the June total group data. This model includes variables 1, 2, and 14 (grades received in physical sciences).

Physics Test: Summary Statistics

Summary statistics for the Physics Test by grade level and administration are found in Table 5. Examination of the data presented in this table indicates that the test is taken predominately by seniors at the November, December and January administrations and by high school juniors at the June administration. The test appears to be taken in May by similar numbers of juniors and seniors.

The only sizable freshmen group is the group taking the test at the June administration. This administration also has a fairly large number of sophomores taking the test.

An examination of the Physics Test scaled score means for all groups containing 25 or more examinees indicates several trends. In general, as expected, high school juniors score higher than high school seniors for all five administrations. The highest scoring group for the May administration contains those students reporting their grade level as sophomore. For the June administration, the highest scaled score means on the Physics Test were obtained by the high school freshmen who take the test, the next highest by sophomores, followed by juniors and then seniors.

Perusal of the total group data summarized for the five administrations of the test indicates that the highest scaled score mean (616) was obtained by the total group taking the test at the November administration. It should be noted that this mean is very close to that obtained by the June total group (613). The lowest Physics Test mean obtained for a total group is 561, obtained by the January group.

Table 6 presents SAT-V and SAT-M scaled score summary statistics for groups taking the Physics Test at the five administrations. Again, only summary statistics based on groups with 25 or more examinees are presented. For all five administrations, high school juniors score higher than high school seniors on both SAT-V and SAT-M. The range of SAT-V scaled score means for high school juniors is 505 to 562 and for high school seniors the range is 479 to 554. High school junior SAT-M scaled score means range from 648 to 684. SAT-M means obtained by high school seniors range from 594 to 673. High school sophomores taking the test in June obtained higher scaled score means on both SAT-V and

SAT-M than either the juniors or the seniors taking the test at this administration.

The relationship between SAT-V and SAT-M scaled scores varies across the different administrations and grade levels. The correlation between these two scores obtained by high school juniors ranges from .31 for the 34 candidates who took the Physics Test in December to .52 for the January and May groups. The correlation coefficient between SAT-V and SAT-M scores obtained by high school seniors ranges from a low of .44 for May candidates to a high of .50 for June candidates.

Physics Test: Correlational Analyses

Table 7 contains correlation coefficients for the Physics Test score with 17 covariates for the total group taking the test at the five administrations. Unlike the Biology and Chemistry Tests, a Background Questionnaire is administered with the Physics Test; hence, correlations of Physics Test score with variable 17 appear in Table 7.

Examination of the data given in Table 7 indicates that, for the November administration, correlations of Physics Test scores with the 17 covariates range from a correlation of .62 with variable 2 (SAT-M scaled scores) to a correlation of .00 with variable 4 (amount of course work in English). The variable showing the second highest correlation with Physics Test scores (.50) is variable 1 (SAT-V scaled scores). The variable that correlates the third highest with Physics Test scores (.30) is variable 8 (amount of course work in physical sciences).

The results of the analyses given in Table 7 for the December Physics Test administration show similar patterns in correlations to those observed for the November data. Again, variable 2 (SAT-M scaled scores) correlates most highly

(.63) with Physics Test scores. Other variables showing moderate correlations with Physics Test scores are variables 1, 8, 14 and 17, which are respectively, SAT-V scaled scores, amount of course work in physical sciences, grades in physical sciences, and Achievement Test Background Questionnaire responses.

Data obtained at the January, May, and June Physics Test administrations show similar patterns to those observed for the two administrations just discussed. In all cases, the variable that correlates most highly with Physics Test scores is variable 2 (SAT-M scaled scores). Other variables that are somewhat consistent in showing correlations of .30 or greater with Physics Test scores are variables 1, 8, 14, and 17. A few exceptions are worth noting: 1) variable 17 (Achievement Test Background Questionnaire responses) does not correlate as highly with Physics Test scores obtained at the May and June administrations as it does with data obtained from the fall administrations; and 2) variable 3 (high school class rank) and variable 11 (grades in mathematics) show moderate correlations with Physics Test scores for the groups taking the test at the spring administrations.

Physics Test: Covariate-Selection Analyses

Results of the Physics Test regression analyses are presented in Table 8. As mentioned previously, analyses are not presented for groups with sample sizes of fewer than 100 cases. The results of the analyses presented in Table 8 indicate that the best two variable model for juniors taking the test at both the May and June administrations is one that contains variables 1 and 2 (SAT-V and SAT-M scaled scores, respectively). This model accounts for approximately 52% of the total variance for both the May and June analyses. The best three variable model is also consistent for juniors taking the test at the two administrations. This model is one that includes variables 1 and 2, as well as variable 8 (amount of course work in physical sciences.) This model accounts for approximately 54% of the variance in the May data and 53% of the variance for the June analysis.

The analyses conducted for the high school seniors taking the Physics Test at the five administrations show a great deal of consistency across the five administrations. The best two variable model for all five administrations is one that includes variables 1 and 2. This model accounts for a range of 43% to 52% of the total variance in Physics Test score. The best three variable model specified for high school seniors taking the Physics Test is also consistent across the five administrations. This model includes variables 1 and 2, and also variable 17 (Achievement Test Background Questionnaire responses). The best three variable model accounts for a range of 47% to 53% of the total variance in Physics Test score.

The results of the total group analyses for the Physics Test are quite similar to those obtained for the senior data. Variables 1 and 2 constitute the best two variable model for all administrations. The results indicating the best three variable model are not quite consistent with those obtained for the senior data. For the total group analyses, the best three variable model obtained for the fall administrations is one that includes variable 1 (SAT-V scaled scores), variable 2 (SAT-M scaled scores), and variable 17 (Achievement Test Background Questionnaire responses). The best three variable model for data obtained for the spring administrations is one that includes variables 1 and 2, as well as variable 8 (amount of course work in physical sciences).

<u>French Test: Summary Statistics</u>

Table 5 contains summary statistics by grade level and administration for the French Test. Examination of the information presented in Table 5 indicates that the French Test is taken predominately by high school juniors and seniors. High school seniors appear to take the test mostly at the November, December, and January administrations and juniors make up the largest proportion of test takers at the May and June administrations. A considerable number of high school sophomores and a small number of freshmen also take the test at the June administration.

An examination of the Achievement Test scaled score means for the various grade level groups indicates that, for the November administration, high school seniors score slighly higher than juniors; however, for the remaining four administrations, the situation is reversed with juniors being the higher scoring group. High school sophomores (who took the test in May) score lower than juniors and higher than seniors who took the test at that same administration. For the June administration data, scaled score means are arranged high to low by grade level as follows: juniors, freshmen, sophomores, and finally, seniors.

An examination of the total group means for the five administrations indicates that these means range between a high of 555 for the group that took the test at the June administration to a low of 508 for the January group. The means obtained by the November, December, and May total groups are all fairly similar, ranging between 526 and 538.

Table 6 contains SAT-V and SAT-M summary statistics by grade level for candidates taking the French Test at the five administrations. Scaled score means obtained by seniors across the five administration groups range from a low of 529, obtained by the May group, to a high of 558 (obtained at both the November and June administrations). SAT-M means for this group range from a low of 555, again for the May group, to a high of 593 obtained by the November group. Correlations between scores obtained on SAT-V and SAT-M by high school seniors range from .52 to .68.

SAT-V scaled score means obtained by juniors taking the French Test at the four administrations containing groups with 25 or more examinees range from a low of 547 for the June group to a high of 575 for the group taking the test in December. SAT-M scaled score means obtained by this same group range from a low of 581 obtained at the June administration to a high of 613 obtained by those students taking the test in January. The correlation between SAT-V and SAT-M scores obtained by high school juniors ranges between .58 and .62.

It is interesting to note that high school sophomores taking the French Test in June (the only group of sophomores containing 25 or more candidates) have higher scores on both SAT-V and SAT-M than juniors or seniors taking the test at this administration. The correlation between SAT-V and SAT-M scores obtained by sophomores is also quite high (.63).

French Test: Correlational Analyses

Table 7 contains correlation coefficients for French Achievement Test score with 17 covariates for the total group taking the test at the five administrations. As mentioned previously, sample sizes vary from one pair of variables to another. The reader is referred to Table I, found in the appendix, for sample sizes for the correlations summarized in Table 7. It should be noted that a Background Questionnaire is administered with the French Test; thus, correlations with all 17 covariates have been obtained for this test.

Examination of the data given in Table 7 indicates that, for the November administration, correlations with French Test scores range from -.01 for variable 7 (amount of high school course work in biology and sciences) to .58 for variable 1 (SAT-V scaled scores). Variable 17 (Achievement Test Background Questionnaire responses) correlates the second most highly with French Test scores.

An examination of the correlational data given in Table 7 for the December administration indicates a similar pattern to the one observed for the November correlations. Scores on SAT-V correlate most highly with French Test scores and Achievement Test Background Questionnaire responses (variable 17) correlate second highest. The third most highly correlating variable is variable 2 (SAT-M scaled scores). Other variables showing correlations of .30 or greater are variables 6 and 12 (course work in foreign languages and grades in foreign languages, respectively).

Data examined for the January, May, and June administrations show similar trends to those just described. In all cases, the variable correlating the most highly with French Test score is variable 1 (SAT-V scaled scores). Variables that consistently show correlations of .30 or greater with French Test scores are variables 2, 17, 6 and 12. Variable 3 (high school class rank) also shows a moderate correlation with French Test scores for the May and June groups.

French Test: Covariate-Selection Analyses

Results of the regression analyses carried out for the French Test are summarized in Table 8. Examination of the information presented in Table 8 indicates that the best two variable model for juniors taking the French test in May and June is one that includes variable 1 (SAT-V scaled scores) and variable 17 (Achievement Test Background Questionnaire responses). This model accounts for 47% of the total variance for the May group and 41% of the total variance for the June group. A considerable increase in the percentage of variance accounted for by the models fit to the junior data can be accomplished by expanding the models from two to four variables. Examination of the information in Table 8 indicates that the best four variable model for the May and June administrations of the test to this grade level group is one that includes variables 1, 2, 17, and 12, which are SAT-V scaled scores, SAT-M scaled scores, Achievement Test Background Questionnaire responses, and grades in foreign languages, respectively. The four variable model accounts for 53% of the total variance found in the May data and 46% of the total variance in the June data.

The regression analyses carried out for the high school seniors taking the test at the five administrations show some similarity in that the best two variable model is one that consistently employs variables 1 and 17 (SAT-V scaled scores and Achievement Test Background Questionnaire responses). The best four variable model differs somewhat across the five administrations. This model includes variables 1, 2, and 17 for four of the five administrations and either variable 6 or 12 (course work in foreign languages and grades in foreign languages, respectively), as the remaining fourth variable. The best four variable models account for between 40% to 55% of the total variance for the analyses carried out across the five administrations.

Finally, the analyses carried out for the total groups taking the French Test at the five administrations produced results similar to those obtained using the senior data. In all cases, the best four variable model is one that includes some combination of variables 1, 2, 17, 6 and 12. These models account for a range of 40% to 53% of the total variance in Achievement Test score across the five administrations.

Spanish Test: Summary Statistics

Achievement Test summary statistics for the Spanish Test are found in Table 5. An examination of these data indicates that the Spanish Test (similar to the French Test) is taken predominately by high school juniors and seniors, although sophomores do constitute a fairly large proportion of the June test takers.

Similar to the other tests discussed so far, high school juniors obtain higher average scores than high school seniors at all five administrations. Sophomores, taking the test in May, score lower than juniors taking the test at the same administration, but higher than May seniors. The highest scaled score mean obtained by any group at the June administration was that obtained by high school freshmen. The second highest mean for this administration was obtained by juniors, followed by sophomores, and finally, by seniors.

An examinination of the total group means for the Spanish Test given at the five administrations indicates that these means range between a high of 535 for the June group and a low of 485 for students who took the test in January.

Summary statistics for SAT-V and SAT-M scores obtained by examinees taking the Spanish Test at the five administrations are shown in Table 6. As has been the case for previous discussions, summary statistics based on fewer than 25 candidates are not presented.

Perusal of the information found in Table 6 indicates that SAT-V means obtained by senior candidates range from a low of 477 for the May administration group to a high of 515 obtained by both the June and November administration groups. SAT-M means for this group range from a low 528 for the May group to a high of 569 for the November group. The correlations between SAT-V and SAT-M scores obtained by high school seniors taking the Spanish Test range from .57 to .68.

SAT-V scaled score means for juniors taking the Spanish Test range from a low of 514 (for those students taking the Spanish Test in June) to a high of 546 for the December group. SAT-M means for those reporting their grade level as juniors ranged from a low of 563 for June juniors to a high of 600 for December juniors. Correlations between SAT-V and SAT-M scores obtained by juniors range between .58 and .68.

High school sophomores taking the Spanish Test in June received higher SAT-V and SAT-M scores than either the juniors or seniors taking the test at this administration. The correlation between the scores on these two variables obtained by June sophomores was .44.

Spanish Test: Correlational Analyses

Table 7 contains correlation coefficients for Spanish Test score with the 17 proposed covariates for scores obtained by the total group taking the test at the five administrations. It should be noted that, similar to the French Test, a Background Questionnaire is administered with the Spanish Test; thus, correlations with all 17 covariates appear in Table 7.

Review of the information presented in Table 3 indicates that, for the November administration, the variable correlating most highly with Spanish Test score is variable 1 (SAT-V scaled scores). Variable 17 (Achievement Test Background Questionnaire responses) correlates second most highly and variable 2 (SAT-M scaled scores) third most highly. Other variables showing correlations of .30 or greater are variables 6 and 12 (course work in foreign languages and grades in foreign languages, respectively).

December data, presented in Table 7, show a similar pattern of correlation coefficients. In this case, the variable correlating most highly with Spanish Test scores is variable 17 (Achievement Test Background Questionnaire responses). The next highest correlating variable is variable 1, followed by variable 2.

Data for the remaining administrations displayed in Table 7 show similar patterns of correlation coefficients. Achievement Test Background Questionnaire responses (variable 17) shows consistently high correlations with Spanish Test scores, as do variables 1 and 2. It should be noted that, for the January data, the only variable displaying a correlation greater than .30 is Achievement Test Background Questionnaire responses.

Spanish Test: Covariate-Selection Analyses

Results of the regression analyses carried out for the Spanish Test are summarized in Table 8. As mentioned previously, only those analyses conducted on groups with sample sizes of 100 or greater will be discussed. Examination of the information presented in Table 8 for the high school juniors taking the test at the May and June administrations indicates that the best two variable model specified for both these administrations is one that includes variable 1 (SAT-V scaled scores) and variable 17 (Achievment Test Background Questionnaire responses). The best four variable model is also consistent across the two administrations. This model includes variables 1, 2, 17, and 12, which are SAT-V scaled scores, SAT-M scaled scores, Achievement Test Background Questionnaire responses, and grades in foreign languages, respectively. This model accounts for 47% of the total variance in Achievement Test scores for the May administration and 43% of the total variance for the June administration.

The best two variable model specifed for the senior data is not consistent across the five administrations; however, all the models include some combination of variables 1, 17, and 12. The best four variable model specified for seniors is also inconsistent across the five grade levels. The variables that do appear consistently are variable 1 (SAT-V scaled scores), variable 17 (Achievement Test Background Questionnaire responses), and variable 12 (grades in foreign languages). The remaining fourth variable specified is either variable 2 (SAT-M scaled scores), variable 21 (parental income), variable 5 (course work in mathematics), or variable 19 (father's level of education). This model accounts for a range of between 26% to 46% of the total variance in Spanish Test scores.

The results of the analyses carried out for the total group taking the Spanish Test at the five administrations were slightly more consistent than those obtained using only senior data. The best four variable model specified for the total group taking the test at the November, May, and June administrations was one that contains variables 1, 2, 17, and 12. This model accounted for a range of 39% to 42% of the total variance in Spanish Test scores. The best four variable model for the total group taking the test in December and Januarv is one that includes variables 1, 17, 12, and 21 (parental income).

European History and World Cultures Test: Summary Statistics

Summary statistics by grade level and administration for the European Historv and World Cultures Test can be found in Table 5. It should be noted that the European Historv and World Cultures Test was only offered at the December and Mav administrations at the time data were collected for this study. Beginning with testing year 1986-87, the test is offered only at the December and June administrations. Examination of the data presented in Table 5 indicates that the test is taken predominately by high school seniors in December and juniors in May. A considerable number of sophomores also take the test in May.

It is fairly clear, from an examination of the information presented in Table 5, that high school juniors are consistently the highest scoring group that take the test, scoring higher even than the sophomores who take the test in May. It is interesting to note that there is not a large difference between the achievement levels (as assessed by European History Test scores) of the May and December groups. The May group scores somewhat higher than the December group, and this is consistent for both juniors and seniors.

Table 6 presents SAT-V and SAT-M summary statistics for grade level groups taking the European History Test at the two administrations. Again, no statistics based on groups containing fewer than 25 candidates are presented. As has been the case for achievement test groups discussed previously, juniors score higher than seniors on both SAT-V and SAT-M. It is interesing to note that both juniors and seniors taking the test in December appear to be considerably more able (as assessed by SAT-V scores) than their May counterparts. However, SAT-M scores obtained by both juniors and seniors are quite similiar across the two administrations. The correlations between SAT-V and SAT-M scores obtained by juniors and seniors at the two administrations range from a low of .41 obtained by May seniors to a high of .57 obtained by May juniors.

European History and World Cultures Test: Correlational Analyses

Table 7 presents correlations of European History and World Cultures Test scores with 16 of the 17 proposed covariates for the total group taking the test at the December and May administrations. Because no Background Questionnaire is administered with the European History Test, correlations of test scores with variable 17 are not presented in Table 7.

An examination of the data given in Table 7 for both the December and May administrations indicates similar patterns of correlations for data obtained at these two administrations. For both administrations, the variable correlating most highly with European History Test score is variable 1 (SAT-V scaled scores). The variable correlating second most highly, for both administrations, is variable 2 (SAT-M scaled scores). The third most highly correlating variable (across the two administrations) is variable 15 (grades in social studies). Other variables that correlate moderately (correlations of .30 or greater) with

Achievement Test score are variables 10 and 13 (grades in English and grades in biology and sciences).

European History and World Cultures Test: Covariate-Selection Analyses

Table 8 contains the results of the regression analyses carried out for the European History and World Cultures Test. The results of the analyses conducted for the junior May group indicate that the best two variable model is one that includes variables 1 and 2 (SAT-V and SAT-M scaled scores). This model accounts for 51% of the total variance in Achievement Test score. The best three variable model fit to the May junior data is one that includes variables 1, 2, and 15 (grades in social studies). This model accounts for an additional 3% of the total variance.

The best two and three variable models for seniors taking the test at the May and December administrations is not consistent across the two administrations. The best two variable model for the December group is one that includes variables 1 and 2, and this model accounts for 49% of the total variance. The best three variable model for seniors taking the test in December includes variables 1, 2, and 10 (grades in English) and accounts for an additional 1% of the total variance. For senior data obtained at the May administration, the best two variable model is one that includes variables 1 and 7 (amount of course work in biology and sciences). This model accounts for approximately 66% of the total variance. The best three variable model specified for May seniors (which accounts for an additional 1% of the total variance) is one that includes variables 1, 7, and 9 (amount of course work in social studies).

The analyses carried out for the total group taking the test in December and May are somewhat consistent in that the best two variable model specified for both administrations is one that includes variables 1 and 2. The best three variable model specified for the December group (with 49% of the variance accounted for) is one that specifies variables 1, 2, and 10 (grades in English). For the May total group, the best three variables model is one that includes variables 1, 2, and 15 (grades in social studies). This model accounts for approximately 59% of the total variance in European History and World Cultures Test scores.

<u>German Test: Summary Statistics</u>

Summary statistics by grade level and administration for the German Test are found in Table 5. It should be noted that, as was the case for the European History and World Cultures Test prior to testing year 1986-87, the German Test was given only at the December and May administrations. Examination of the information found in Table 5 indicates that the test is taken predominately by seniors in December. A fairly similar number of juniors and seniors take the test in May. This administration also includes a small number of sophomores. The most able group taking the test (as assessed by German Test score) appears to be high school juniors. This is true for both the May and December administrations. Sophomores taking the test in May are more able than seniors, but not as able as juniors. In general, both groups (juniors and seniors) taking the test in December score considerably higher than their May counterparts.

Table 6 contains summary statistics by grade level and administration for candidates taking the German Test. The only summary statistics based on 25 or more candidates are for seniors taking the test in December and May and juniors taking the test in May. Scaled score means obtained for both SAT-V and SAT-M for seniors taking the German Test in December are considerably higher (562 and 598, respectively) than for seniors taking the German Test in May (531 and 576, respectively). The correlation between SAT-V and SAT-M scores for the two administrations are .56 for the December administration and .54 for the May administration. The scaled score means for SAT-V and SAT-M obtained by the juniors taking the test in May are very similar to those obtained by the seniors taking the test in December (565 and 598) but considerably higher than those obtained by seniors taking the test in May. In general, the total group means on both SAT-V and SAT-M are higher for the December group than for the May group.

German Test: Correlational Analyses

Correlation coefficients for German Test score with 17 proposed covariates are shown in Table 7. It should be noted that a Background Questionnaire is administered with the German Test; hence, correlations of Achievement Test score with variable 17 appear in Table 7.

An examination of the data given in Table 7 for both the December and May administrations indicates that the patterns of correlations of the covariates with German Test score are fairly similar for both administrations. The correlations with all of the covariates are the lowest that have been discussed so far. The covariate correlating most highly with German Test scores obtained at the December administration is variable 1 (SAT-V scaled scores); this is followed by variable 17 (Achievement Test Background Questionnaire responses). Variable 6 (course work in foreign languages) is the only other variable with a correlation of .30 or greater with German Test score. It should be noted that variable 2 (SAT-M scaled scores) has a correlation of .29 with German Test scores.

For the May data, the most highly correlating variable is variable 17 (Achievement Test Background Questionnaire responses). This is followed by variable 6 (course work in foreign languages). Variables 1 and 2 (SAT-V and SAT-M scaled scores) show correlations of .29 and .27 with German Test scores.

German Test: Covariate Selection Analyses

Results of the regression analyses carried out for the German Test are summarized in Table 8. Analyses employing the May junior data indicate that the best two variable model is one that contains variable 2 (SAT-M scaled scores) and variable 17 (Achievement Background Questionnaire responses). This model accounts for appoximately 26% of the total variance in German Test score. The best four variable model for this group (which accounts for approximately 31% of the total variance) is one that includes variables 2, 17, and 7 (amount of course work in biology and sciences) and variable 21 (parental income).

The best two variable model for seniors taking the test in December is one that includes variables 1 and 17, and for those taking the test in May, is one that includes variables 2 and 17. The best four variable model for December seniors (accounting for approximatley 30% of the total variance) is one that includes variables 1, 17, and 6 (amount of course work in foreign languages) and 12 (grades in foreign languages). For the May seniors, the best four variable model (which accounts for 43% of the total variance) is one that includes variables 1, 2, 17, and 8 (amount of course work in physical sciences).

The results obtained for the December total group are identical to those obtained for the December seniors. Those obtained for the May total group are slightly different than the results obtained for the May seniors in that the best four variable model is one that includes variables 2, 17, 6, and 8. This model accounts for 30% of the total variance in German Test score for the

December administration and 37% of the total variance for scores obtained by the May total group.

Hebrew Test: Summary Statistics

Table 5 contains summary statistics by grade level and administration for the Hebrew Test. It should be noted that, similar to the German Test and the European History and World Cultures Test prior to testing year 1986-87, the Hebrew Test was offered only at the December and May Achievement Test administrations. Perusal of the information contained in Table 5 indicates that the major administration of the test is the one carried out in December. Seniors constitute the largest proportion of this population. A small number of students also take the test in May and these students are predominately high school juniors. It should be noted that juniors score higher than seniors at both the December and May administrations; also, the scores obtained by the total groups for both administrations are fairly similar.

Summary statistics for SAT-V and SAT-M scores obtained by Hebrew Test takers are found in Table 2. Very few of the summary statistics (those obtained by seniors in December and by the total group in December and May) are based on groups of 25 or more examinees. Both SAT-V and SAT-M total group means are higher for the group taking the test in December (568 and 583, respectively) than for the total group taking the test in May (513 and 553, respectively). The December total group SAT-V and SAT-M means are also higher than those obtained by seniors at that same administration (568 and 583, respectively). Correlations between SAT-V and SAT-M scores range from a low of .48 for the May total group to a high of .69 for the December total group.

Hebrew Test: Correlational Analysis

Table 7 contains correlation coefficients for the Hebrew Test score with 16 of the 17 covariates. Although a Background Questionnaire is administered with the Hebrew Test, it is currently undergoing extensive revision so it will be ignored for these analyses.

Perusal of the data obtained in Table 7 indicates that, for both the December and May administrations, the variable correlating most highly with Hebrew Test score is variable 1 (SAT-V scaled scores). The variable correlating next highest (for data from the December administration) is variable 13 (grades in biology and sciences). This is followed, in order, by variables 12, 2, and 3 (grades in foreign languages, SAT-M scaled scores, and class rank, respectively).

The analyses conducted using data from the May administration indicate that variables 2 and 5 (SAT-M scaled scores and amount of course work in mathematics) tie for the second most highly correlating variables. These variables are followed in order by variables 12, 13, 14, 11, 8, and 10 which are respectively, grades in foreign languages, grades in biology and sciences, grades in physical sciences, grades in mathematics, amount of course work in physical sciences, and grades in English.

Hebrew Test: Covariate-Selection Analyses

Results of the regression analyses conducted for the Hebrew Test are summarized in Table 8. The analyses carried out for December seniors (N=59) and the December total group (N=64) will be discussed in spite of the few number of cases they are based on.

Perusal of the data presented in Table 8 indicates that the best two variable model for the high school seniors taking the Hebrew Test in December is one that includes variable 1 (SAT-V scaled scores) and variable 19 (father's level of education). This model accounts for approximately 35% of the total variance in Hebrew Test score. The best four variable model specified for high school seniors taking the test in December accounts for approximately 52% of the total variance and includes variable 2 (SAT-M scaled scores), variable 17 (responses to the old version of the Achievement Test Background Questionnaire, variable 12 (grades in foreign languages), and variable 19 (father's level of education).

The best two variable model for the total group taking the test in December is one that includes variables 2 and 19. The best four variable model for this group, which accounts for approximately 47% of the total variance in Hebrew Test scores is, as was the case for the senior analyses, one that includes variables 2, 17, 12, and 19.

Latin Test: Summary Statistics

Achievement Test summary statistics by grade level and administration for the Latin Achievement Test are found in Table 5. It should be noted (as is was for the German, Hebrew and European History and World Cultures Test) that prior to testing year 1986-87, the Latin Test was offered at only the December and May Achievement Test administrations.

Examination of the data found in Table 5 for the Latin Test indicates that the test is taken mostly by seniors in December and by high school juniors in May; however, a fairly large number of sophomores, and a sufficient number of freshmen (for the purposes of this discussion), took the test in May. A closer look at the scaled score means presented in Table 5 indicates that juniors taking the test in December obtain an unusually high scaled score mean (630), when compared to that obtained by December seniors (550) or May juniors or seniors (573 and 524, respectively). High school sophomores taking the test in May score higher on average (588) than either the juniors or seniors taking the test. High school freshmen, taking the test in May, achieve a higher mean (567) than the seniors taking the test, but a lower mean than either the sophomores or juniors taking the test. The total group mean obtained for the May group is higher than that obtained in December, reflecting the highly able freshmen and sophomores who take the test at this administration.

SAT-V and SAT-M summary statistics obtained by students taking the Latin Achievement Test are found in Table 6. As mentioned previously, only those statistics based on 25 or more candidates are presented. Scaled score means on both SAT-V and SAT-M obtained by seniors taking the test in December are quite similar to those obtained by the group taking the test in May. For SAT-V, the means are 563 and 561, respectively, and for SAT-M they are, respectively, 594 and 593. Correlations between SAT-V and SAT-M scores obtained by this group are quite high for both administrations, (.55 for the December senior data and .63 for the May senior scores).

SAT-V and SAT-M scaled scores obtained by the juniors taking the test in June are considerably higher than those obtained by the high school seniors. The total group mean for SAT-V and SAT-M scores are higher for the total group taking the test in May than for the total group taking the test in December.

Latin Test: Correlational Analyses

Correlations of the Latin Test score with 17 covariates are summarized in Table 7. An Achievement Test Background Questionnaire is administered with the Latin Test; hence, correlations of Achievement Test scores with variable 17 appear in Table 7.

An examination of the information presented in Table 7 indicates that the patterns of correlation coefficients of the Latin Test score with the 17 covariates are fairly similar for the December and May administrations. The variable correlating most highly with Latin Achievement Test scores, for both administrations, is variable 1 (SAT-V scaled scores). Variable 2 (SAT-M scaled scores) is the variable that correlates the second most highly. The correlation of Achievement Test Background Questionnaire responses (variable 17) with Latin Test scores is dissappointly low for both the December and May data (.29 and .26, respectively). Other variables showing correlations of .30 or greater (examining both the December or May data) are variables 3, 6, 11, 12, 13, and 14. These variables are, respectively, class rank, amount of course work in foreign languages, grades in mathematics, grades in foreign languages, grades in biology and sciences, and grades in physical sciences.

Latin Test: Covariate-Selection Analyses

Results of the regression analyses carried out for the Latin Test are summarized in Table 8. The data presented in Table 8 indicate that the best two variable model for May junior candidates is one that includes variable 1 (SAT-V scaled scores) and variable 12 (grades in foreign languages). This model accounts for approximately 30% of the total variance in Latin Test score. The best four variable model for this group is one that includes variables 1, 2, 17, and 12 (SAT-V scaled scores, SAT-M scaled scores, Achievement Test Background Questionnaire responses, and grades in foreign languages, respectively). Approximately 36% of the total variance in Latin Test score is accounted for by this model.

The best two and four variable models fit to the senior data are not consistent across the two administrations. For the December administration, the best two variable model is specified to be one that includes variables 1 and 2. The best two variable model specified for the May data is one that includes variables 2 and 17.

The best four variable models fit to the senior data collected at the two administrations is one that includes variables 1, 2, and 17 (SAT-V scaled scores, SAT-M scaled scores, and responses to the Achievement Test Background Questionnaire, respectively). The fourth variable included for the December data is variable 12 (grades in foreign languages). The fourth variable included for the May data is variable 3 (class rank). The percentage of variance accounted for by these two models is approximately 42% for the data collected at the December administration and 54% for the May data.

The results of the total group analyses are consistent across the December and May administrations. The best two variable model specified for both administrations is one that contains variables 1 and 2 (SAT-V and SAT-M scaled scores). The best four variable model is one that contains variables 1, 2, 17, and 12 (SAT-V scaled scores, SAT-M scaled scores, Achievement Test Background Questionnaire responses, and grades in foreign languages). This model accounts for approximately 42% of the total variance in the December data and 43% of the total variance in the May data.

<u>Across Test Comparisons: Summary Statistics</u>

Data comparing Achievement Test summary statistics are found in Table 5 and Figure 1. Figure 1 contains plots of total group Achievement Test scaled score means for the 10 Achievement Tests that are given at all five administrations. No plots of Achievement Test means were made for the four tests that are administered only twice a year. The data given in Table 5 will be used to make comparisons for these tests.

Several observations can be made based on an examination of the data presented in Table 5. First, it is generally true, for the 10 tests given at all five administrations, that the majority of students taking the tests at the November, December, and January administrations are high school seniors and those taking the tests at the June administration are high school juniors. The one exception to this rule is the Biology Test, which is taken predominately in June by sophomores. It should be noted that the June Chemistry Test is also taken by a large number of sophomores; however, juniors still make up the majority of the June test taking population.

Considering only those tests administered at all five administrations, the grade level composition of the May administration is not as consistent across tests as the grade level composition observed for the other four administrations. For most of the tests, the majority of May test takers are high school juniors. However, for some tests, almost equal numbers of juniors and seniors participate in the May administration. These tests are Literature, Mathematics Level II, and Physics.

An interesting point to note, when examining the information presented in Table 5 for the tests administered five times a year, is that the largest volume administration for most tests is the December administration. However, the science tests (Biology, Chemistry and Physics) and the American History and Social Studies Test are administered to more candidates in June than in December. This appears mainly to be due to the large number of sophomores that choose to take these tests in June.

The information presented in Table 5 for the Achievement Tests administered only twice a year exhibits similar trends to those observed for the tests administered at all five administrations. Most of the tests administered only twice a year are taken predominately by seniors at the December administration and by juniors at the May administration. It should be noted that high school seniors who took the German Test in May make up a larger proportion of the population than do May seniors taking the remaining tests that are administered only twice annually.

Any comparison of the achievement levels of the various groups (as assessed by scaled scores on the Achievement Tests) must be made with caution. Scores obtained for the different tests may not necessarily be on the same scale and should only be considered roughly comparable. Comparisons within a test, across administrations, can be made with confidence since scores on different forms of the same test which are given at the various administrations have been equated.

An examination of the total group means for the 10 tests given five times annually, which are plotted in Figure 1, indicates that, in general, the most able groups that take the tests are those students who take the tests in June and the least able groups are those who take the tests in January. Reference to Table 5 indicates that the Achievement Tests that are given only twice a year show a similar pattern, i.e., total group means obtained by groups taking the tests in May are higher than those obtained by examinees taking the tests in December.

Insert Figure 1 about here

Considering only the scaled score means obtained on the tests given five times a year, the following observations can be made. A comparison of the scores obtained at the two largest administrations (December and June) shows that the mean scores on the tests rank order high to low in a somewhat similar manner. This ranking is given below:

December Administration

Test	Ach. Test Mean
Mathematics Level II	646
Physics	578
Chemistry	556
Mathematics Level I	532
Literature	528
French	526
English Composition	522
Biology	517
American History	505
Spanish	497

June Administration

Test	Ach. Test Mean
Mathematics Level II	677
Physics	613
Chemistry	585
Biology	562
French	555
American History	555
Mathematics Level I	553
Spanish	535
English Composition	528
Literature	526

Although one must take into account the fact that scaled score means obtained on the different tests are only roughly comparable, it seems reasonable to conclude that the most able groups taking the tests at the December and June administrations are examinees taking the Mathematics Level II, Physics, and Chemistry Tests, with the Mathematics Level II group considerably more able than examinees taking the other two tests. It also seems reasonable to conclude that, generally, those examinees taking the English Composition Test, American History and Social Studies Test, and the Spanish Test are among the least able.

It is interesting to add to this comparison, a comparison of the scaled score means obtained on the tests that are administered only twice a year. For convenience, these are listed below in order from high to low.

December Administration

Test	Ach. Test Mean
Hebrew	625
Latin	555
European History	542
German	542

May Administration

Test	Ach. Test Mean
Hebrew	632
Latin	570
European History	557
German	542

It is clear that the groups taking the Hebrew Test are the most able (as assessed by the Hebrew Test scaled score means). It is also fairly clear that the groups taking the European History and World Cultures Test and the groups taking the German Test are less able than examinees taking the Latin Test.

SAT-V and SAT-M scaled score summary statistics for candidates taking the Achievement Tests are found in Table 6. Figure 2 contains plots of SAT-V scaled score means for the total group taking the 10 Achievement Tests that are given five times annually. Figure 3 displays the same data for SAT-M scaled score means. As was the case for the Achievement Test scaled score means previously discussed, plots were not obtained for the small volume Achievement Tests that are given only twice a year.

Insert Figures 2 and 3 about here

An examination of the information given in Table 6 and Figure 2 indicates that, in general, higher SAT-V scaled score means are obtained by the total groups taking the tests in June and November than those taking the tests in January. In many instances, the November and June SAT-V means for the total group taking a particular test are quite similar. Most of the SAT-V scaled score means obtained by the May groups are similar to, but somewhat higher than, those obtained by the January groups. For most of the tests, the December SAT-V scaled score means obtained by the total groups are lower than those obtained by the November and June groups and higher than those obtained by the January and May groups. Given below is a rank ordering of the 10 tests (from high to low) based on SAT-V scaled score means for the two largest volume administrations (December and June).

December Administration		June Administration	
Test	SAT-V Mean	Test	SAT-V Mean
Mathematics Level II	548	Mathematics Level II	562
French	546	Physics	556
Physics	535	French	548
Chemistry	529	Literature	532
Literature	528	American History	532
English Composition	518	Chemistry	528
Biology	513	Biology	526
Spanish	512	English Composition	521
American History	507	Spanish	514
Mathematics Level I	499	Mathematics Level I	506

Several observations are clear from the orderings given above. First, the groups taking the Mathematics Level II Test are the most able (as assessed by SAT-V scores). However, the discrepancy between SAT-V scores obtained by this group and the next highest scoring group are not nearly as extreme as that observed when the tests were rank ordered by Achievement Test means. Second, it appears that the groups taking the English Composition Test, Spanish Test, and Mathematics Level I Test are the least verbally able; although once again, the large score discrepancies observed between mean Achievement Test scores are not apparent here.

A review of the SAT-V means presented in Table 2 for the total groups taking the tests that are administered only twice a year indicates that, in general, higher SAT-V means are obtained by the groups taking the tests in December than by those taking the tests in May. The total group SAT-V means for tests administered twice annually rank order high to low as follows:

December Administration		May Administration	
Test	SAT-V Mean	Test	SAT-V Mean
Hebrew	571	Latin	578
European History	564	German	548
German	563	European History	542
Latin	563	Hebrew	513

It is interesting to note the similarity among SAT-V means obtained by the European History and World Cultures, German, and Latin Test examinees who took the Achievement Tests at the December administration; also among the SAT-V means obtained by the total groups taking the German and European History and World Cultures Tests in May. The reverse in the ranking of the Hebrew Test means across the two administrations is most likely due to the instability in the data due to small sample sizes.

Perusal of the information given in Table 6 and Figure 3 indicates that, similar to the SAT-V means previously discussed, higher SAT-M scaled score means are obtained by the total groups taking the tests at the November and June administrations. The lowest scaled score means on SAT-M are typically obtained by either the May or January total groups. However, for the language tests given five times annually (Spanish and French), SAT-M scaled score means obtained by the total group who took the test at the May administration are higher than those obtained in either December or January.

The table given below rank orders (from high to low, using total group SAT-M scaled score means) Achievement Test groups who took the tests at the two largest volume administrations (December and June).

December Administration		June Administration	
Test	SAT-M Mean	Test	SAT-M Mean
Physics	644	Mathematics Level II	652
Mathematics Level II	640	Physics	650
Chemistry	624	Chemistry	606
French	570	French	582
English Composition	566	Biology	575
Biology	559	English Composition	572
Spanish	552	Spanish	563
Mathematics Level I	548	American History	562
American History	539	Mathematics Level I	562
Literature	531	Literature	535

The rank ordering of the groups across the two administrations is quite similar. The total groups taking the Physics, Mathematics Level II, Chemistry, and French Tests appear to be the most quantitatively able, and the groups taking the Mathematics Level I, American History and Social Studies, and Literature Tests the least quantitatively able. It is interesting to note that, when the rank ordering obtained above is compared to that obtained using scaled score means on SAT-V, fairly similar conclusions can be reached, i.e., the most able groups taking the Achievement Tests are those that take the Mathematics Level II, French, Physics, and Chemistry Tests. The least able are the groups taking the Mathematics Level I and American History and Social Studies Tests. The only group whose ranking changes considerably, depending upon whether scores on SAT-V or SAT-M are used to rank groups, is the total group taking the Literature Test. This group is, as expected, ranked much higher if SAT-V scaled scores are used for the ranking than if SAT-M scaled scores are used.

A review of the SAT-M scaled score means presented in Table 6 for the total groups taking the tests that are administered only twice a year indicates that, in general, higher SAT-M means are obtained by the total groups taking the tests in December than by those taking the tests in May. The total group SAT-M scaled score means for the groups taking the tests in December and May rank order high to low as follows:

December Administration		May Administration	
Test	SAT-M Mean	Test	SAT-M Mean
German	598	Latin	618
Latin	594	German	588
Hebrew	584	European History	565
European History	559	Hebrew	553

Several points are worth noting when examining the information given in the table above. First, scores on the German, Latin, and to some extent, Hebrew Tests are similar for the groups taking the tests at the December administration. However, scores obtained by the total groups taking the tests at the May administrations are quite diverse, ranging from a high of 618 for Latin Test takers to a low of 553 for the group taking the Hebrew Test. Some similarities in SAT-M mean scores to mean scores collected for SAT-V are worth noting; for example, the group taking the Latin Test in May is very able both verbally and quantitatively. Also, SAT-M scaled score means obtained by Hebrew Test takers are somewhat erratic across the administrations and probably reflect instabilities in the data due to small sample sizes.

Across Test Comparisons: Correlational Analyses

Information summarizing the correlational analyses carried out for the study are found in Table 7 and Figures 4-8. Figure 4 contains plots showing the magnitude of the correlation coefficients of Achievement Test scores with the following covariates: SAT-V scaled scores, SAT-M scaled scores, course grades, Achievement Test Background Questionnaire responses, and amount of course work in a particular subject area. These particular variables were chosen because they contributed most importantly to the regression models that will be discussed in the next section. Information pertaining to tests administered both five times annually and twice annually are included in Figure 4. Two letter symbols have been used in Figure 4 to identify the various Achievement Tests. It should be noted that, since the Achievement Test Background Questionnaire is not administered with all Achievement Tests, only a limited number of points have been plotted for this variable. It should be further noted that Figure 4 contains data pertaining only to the November Achievement

Test administration. Figures 5-8 contain the same information that is displayed in Figure 4 for the remaining four Achievement Test administrations.

Insert Figure 4 about here

One final point needs to be made as introduction to the discussion of the correlational analyses. Two variables that are included in Figures 4-8 vary slightly depending upon which Achievement Test is being considered. These two variables are course grades and amount of course work in a particular subject area. The table given below lists the particular variables used for the correlation coefficients summarized in the figures. The two letter abbreviations given in parentheses are the same as those used in Figures 4-8.

		Course Grades		Amount of Course Work in a Particular Subject Area	
Test	Symbol	Variable Number	Variable Name	Variable Number	Variable Name
Eng. Comp.	(EN)	12	grades in For. Languages	4	c.w. in English
Literature	(LR)	10	grades in English	4	c.w. in English
Amer. Hist.	(AH)	15	grades in Social Studies	9	c.w. in Social Studies
Math I	(M1)	11	grades in Mathematics	5	c.w. in Mathematics
Math II	(M2)	11	grades in Mathematics	5	c.w. in Mathematics
Biology	(BY)	13	grades in Biology	7	c.w. in Biology
Chemistry	(CH)	14	grades in Phys. Sciences	8	c.w. in Phys. Sciences
Physics	(PH)	14	grades in Phys. Sciences	8	c.w. in Phys. Sciences
French	(FR)	12	grades in For. Languages	6	c.w. in For. Languages
Spanish	(SP)	12	grades in For. Languages	6	c.w. in For. Languages
Eur. Hist.	(EH)	15	grades in Social Studies	9	c.w. in Social Studies
German	(GM)	12	grades in For. Languages	6	c.w. in For. Languages
Hebrew	(HB)	12	grades in For. Languages	6	c.w. in For. Languages
Latin	(LT)	12	grades in For. Languages	6	c.w. in For. Languages

Examination of the information provided in Table 7 and Figure 4 for the data collected at the November Achievement Test administration indicates that the correlations of SAT-V scaled scores with Achievement Test scores range from a low of .44 for the Mathematics Level II Test to a high of .84 for the Literature Test. Achievement Tests that have correlations of .70 or greater with this variable are: English Composition, Literature, American History and Social Studies and Biology.

Correlations of Achievement Test scores (for the November administration) with SAT-M scaled scores range from a low of .33 for the Spanish Test to a high of .84 for the Mathematics Level I Test. Only two Achievement Tests have scores that correlate above .70 with this variable; these tests are, as expected, Mathematics Level I and Mathematics Level II.

The third variable to be considered is Achievement Test Background Questionnaire responses. As mentioned previously, only five tests given at the November administration were administered with Background Questionnaires. Correlations with Background Questionnaire responses range from a low of .29 for the Physics Test to a high of .49 for the French and Spanish Tests.

Perusal of the information found in Table 7 and Figure 4 shows that correlations of Achievement Test scores with course grades range from a high of .50 for the Mathematics Level I Test to a low of .29 for the English Composition Test and the Physics Test. In general, the correlations of Achievement Test scores with course grades are considerably lower than the correlations with Achievement Test scores observed for the first three variables discussed.

An examination of the correlations of amount of course work in a particular subject area with Achievement Test scores obtained at the November administration indicates that these correlations range from .09 for the English Composition Test to .38 for the correlation of the amount of course work in foreign languages with scores on the French Test. In general, the correlations of Achievement Test scores with this covariate are the lowest of the correlation coefficients of interest.

Table 7 and Figures 6 and 8 contain information summarizing the correlational analyses carried out for the remaining two administrations at which only 10 Achievement Tests are given (January and June). Examination of the data in these tables and figures indicates that the pattern of correlations discussed for the November data remains relatively unchanged, with a few exceptions, for the data obtained at the January administration of the tests.

Insert Figures 6 and 8 about here

The data collected for the June administration of the tests, however, exhibit a considerable number of discrepancies when compared to the November and January data. Correlations obtained for Achievement Test scores with SAT-V and SAT-M scaled scores are similar to those observed for the November and January administration. Correlations of Achievement Test scores with amount of course work in a particular subject area are somewhat lower, however, than the November and January correlations (ranging from .05 to .31); whereas, those obtained for Achievement Test scores with course grades are somewhat higher than the November and January correlations (ranging from .35 to .51). Correlations of Achievement Test score with Background Questionnaire responses obtained for the June data are the lowest observed across the three administrations (November, January and June).

Correlation coefficients for scores on the 14 Achievement Tests that are given at the December and May administrations with the selected covariates are summarized in Table 7 and Figures 5 and 7. Examination of the data obtained at the December administration (Table 7 and Figure 5) indicates that correlations of Achievement Test scores with SAT-V scaled scores range from .82 for the Literature Test to .37 for the German Test. Achievement Test scores showing correlations with this covariate of .70 or greater are English Composition, Literature, American History and Social Studies, and Biology. It should be noted that the European History and World Cultures Test also shows a high correlation with this variable (.69). Correlations of Achievement Test scores with SAT-M scaled scores range from a low of .29 for the German Test to a high of .82 for the Mathematics Level I Test.

An examination of the correlations of Achievement Test scores with Achievement Test Background Questionnaire responses (Table 7 and Figure 5) shows that the correlations range from .29 for the Latin Test to .44 for the French and Spanish Tests. Correlations of Achievement Test scores with course grades range from a low of .27 for the German Test to a high of .47 for scores obtained on the Mathematics Level I Test.

Insert Figure 5 about here

Finally, perusal of the correlations obtained for Achievement Test scores with amount of course work in a specific subject area indicate that these correlations range from .08 for the English Composition Test to .34 for the French Test.

Comparisons of the correlation coefficients computed using data from the December administration with those obtained using the May data (Table 7 and Figure 5) show considerable similarity between the correlation coefficients computed using scores from these two administrations. In general, the correlations based on scores from the spring administration appear to be slightly higher than those obtained using the fall data. An exception to this generalization are correlations of Achievement Test scores with Background Questionnaire responses, which seem to be slightly higher for all tests given in December except the German Test.

Insert Figure 7 about here

Across Test Comparisons: Covariate-Selection Analyses

The results of the regression analyses that will be compared for the various Achievement Tests are summarized in Table 8 and Figure 9. Figure 9 shows plots of R^2 values obtained for selected regression analyses carried out for the total group taking 13 Achievement tests (R^2 values for the Hebrew Test are not included) at five administrations.

As mentioned earlier in this report, the results of the regression analyses should be interpreted with caution. First, although only total group results (based on large samples) are presented in this section, the results are sample dependent and may not generalize to the total population of Achievement Test takers. Secondly, different conclusions may have been reached if the analyses were not completely exploratory; i.e., if the traditional covariates used to scale Achievement Tests (SAT-V, SAT-M, and, in some cases, semesters of study) were initially specified as variables to be included in the regression analyses.

The 13 tests (all tests except Hebrew) are represented in Figure 9 by the same two character symbols used to plot values of the correlation coefficients found in Figures 4-8. The R^2 values summarize the results of what was judged to be the most promising model for a particular test. Listed below are the models used to provide the R^2 values plotted in Figure 9. Reasons for choosing these models will be presented in the next section of this report. It should be noted that no discussion of the regression analyses conducted for the Hebrew Test will take place here. Reasons for omitting the Hebrew Test from this discussion will also be presented in the next section of this report.

Test	Best Model
English Composition	Three variable model
Literature	Three variable model
American History and Social Studies	Three variable model
Mathematics Level I	Two variable model
Mathematics Level II	Two variable model
Biology	Three variable model
Chemistry	Three variable model
Physics	Three variable model
French	Four variable model
Spanish	Four variable model
European History and World Cultures	Three variable model
German	Four variable model
Latin	Four variable model

An examination of the information presented in Table 8 and Figure 9 indicates that, for the November analyses, the highest R^2 value was obtained for the two variable model fit to the Mathematics Level I data. The smallest R^2 value obtained was for the four variable model fit to the Spanish Test data. It is encouraging that, of the 10 tests administered in December, models can be fit to data obtained for six of these tests that account for at least 60% of the variance in Achievement Test scores. The tests showing the poorest results (in terms of the total variance accounted for) are the foreign language tests and (with the exception of the Biology Test) the science tests.

Insert Figure 9 about here

Results of the analyses carried out for the tests administered in January and June are similar to those obtained using November data, i.e., the tests showing the poorest regression analysis results or the lowest percentage of variance accounted for, are the foreign language tests and the science tests. Still four of the tests administered in January and five of the tests administered in June could be fit by models which accounted for at least 60% of the total variance in Achievement Test scores.

The analyses carried out for the tests given at the December and May administrations indicate that the science tests, the foreign language tests, and the European History and World Cultures Test provide the poorest regression analysis results in terms of total variance accounted for. The German Test appears to provide particularly poor results, showing R^2 values of .30 and .37 for the December and May data, respectively. The R^2 values obtained for the Mathematics Level I and Level II Tests, the American History and Social Studies Test, the Literature Test, and the English Composition Test are all encouragingly high, indicating that it is possible to fit regression models to these data that account for at least 60% of the total variance in Achievement Test scores.

DISCUSSION AND CONCLUSIONS

The purpose of the study described in this report was to explore the relationships between College Board Achievement Test scores and potential scaling covariates for various subgroups of the test taking population. It was expected that such an exploration would lead to the following:

- The selection of additional scaling covariates that might provide improved scaling results for those tests that do not provide scores correlating highly with SAT-V and/or SAT-M scores;
- The respecification of the sample of students that are used to scale the test; and
- The respecification of the hypothetical scaling population.

The goal of the present study was to provide several alternative scaling procedures for the Achievement Tests. The procedures will, by their nature, vary components such as scaling covariates, scaling samples, and characteristics of the hypothetical reference group.

This section of the report will be organized in the following manner. First, the characteristics of the scaling sample will be discussed. Secondly, the choice of additional scaling covariates and their relationship to clusters of tests will be examined. Third, the characteristics of the hypothetical reference group will be discussed.

Characteristics of Scaling Sample

The current method of scaling the Achievement Tests employs a scaling sample based on high school juniors and seniors. The juniors are selected from those students who take the tests at the May and June administrations and the seniors are selected from students who take the tests at the November, December, and January administrations.

Several trends were observed in the section of the report that summarized the results of the grade level analyses. First, it appears that, as has traditionally been the situation, the tests are taken predominately by high school seniors at the November, December, and January administrations and by juniors at the May and June administrations. There are, however, some exceptions to this rule. For one, the Biology Test is taken predominately by sophomores at the June administration. Also, the Literature, Physics, and Mathematics Level II Tests are taken by almost equal numbers of juniors and seniors at the May administration.

Given the fact that the basic test taking patterns (seniors taking the tests in the fall and juniors in the spring) do not seem to have changed greatly over the past several decades, the question is, is there a reason for revising the specifications that are presently used to draw the scaling samples?

One method of discriminating among grade level groups that might provide scores for use in the scaling is to examine the similarity of regression models fit to the data obtained for the various grade level groups taking the tests across the five administrations. Table 9 provides a summary of the variables that constitute the best two-, three-, or four-variable regression models fit to the junior and senior data collected across the five administrations. The decision regarding which regression model is most appropriate for a particular test (the models summarized in Table 9) will be discussed later in this section of the report.

Insert Table 9 about here

Examination of the information provided in Table 9 indicates that, for most of the tests, the greatest consistency among variables specified for a particular regression model is attained for seniors taking the tests in November and December and high school juniors taking the tests in May and June. For example, the information contained in Table 9 for the English Composition Test indicates that the best three variable model fit to the senior data obtained at the November and December administrations of the test is consistent with the best three variable model fit to the scores obtained by juniors taking the test at the May and June administrations.

The consistency observed for the regression models fit to the English Composition Test scores is greater than that observed for some of the other tests; however, it does appear that there is a general trend for models fit to fall senior scores and spring junior scores to be consistent across the November, December, May and June administrations.

An examination of the information presented in Table 9 for junior and senior groups taking the test in January, indicates that, generally, the models specified for the January senior groups are more consistent with those observed for other administrations than are the models specified for the junior January groups.

Values of R^2 are also presented in Table 9 and were considered in making decisions regarding pooling data across grade levels and administrations. These values did not vary substantially for grade level/administration combinations associated with a particular test. The R^2 values did vary considerably (as expected) across the various tests.

Given the above findings, it seems reasonable to conclude that the sampling pattern traditionally used to select Achievement Test examinees for scaling studies is probably still appropriate. An additional question that remains is whether or not scores obtained by high school sophomores at the June administrations of the tests should be included in the scaling sample. There are arguments against including these candidates. For one, not all tests have a large enough sophomore population to be considered as potential candidates for the scaling sample. This inconsistency in sampling by grade level across tests could possibly effect the scaling of the tests in some unknown way. Secondingly, the number of sophomores taking a particular Achievement Test with scores on relevant covariates may be quite small.

One possibility for enlarging the pool of sophomore scores for some tests might be to estimate sophomore scores on the missing covariates. The current scaling method uses a simple linear regression procedure to estimate missing SAT-V and SAT-M scores. This procedure could be used to estimate other missing covariates as well, and might provide a way in which complete data could be obtained for sophomore scores, and thus these scores could be sampled for tests that are taken by a large proportion of sophomores in June. However, once estimates were obtained for scores on missing covariates, regression analyses performed using sophomore scores (actual and estimated) obtained at the June administration of the test would need to be examined carefully to ensure that the relationship of sophomore Achievement Test scores to the salient covariates was consistent with the relationship displayed by other grade levels and tests included in the scaling sample for the cluster.

An examination of the information provided in Table 9 for those tests that are administered only twice a year does not yield data that show a great deal of consistency across December senior and May junior data. The exception to this is the Latin Test data, which are fit consistently with the same four variable model for all grade level and administration combinations.

To summarize, the practice of sampling high school seniors taking the tests in November, December, and January, and juniors taking the tests in May and June seems appropriate. An alternative to this practice is related to the inclusion of high school sophomores taking the tests in June; this alternative would include June sophomores, for some test clusters, if missing covariate scores can be successfully estimated and if the regressions show that Achievement Test scores obtained by this group have a similar relationship to the relevant covariates as that displayed by scores obtained by other sampling groups.

Additional Scaling Covariates

The traditional method of scaling the Achievement Tests uses scores on SAT-V and SAT-M as scaling covariates for all 14 tests. In addition, semesters of study (as assessed by Achievement Test Background Questionnaire responses) is used as a covariate for scaling foreign languages tests.

Both Braun and Tucker (see Dorans, 1985) have pointed to problems with the traditionally used scaling method that are related to choice of covariates. Basically, the results of both their studies indicated that when ability has a major impact on test selection (as it does with the Achievement Tests), there is a clear need to rescale the tests. The quality of this scaling will depend, however, on how highly the covariates used to rescale the tests correlate with Achievement Test performance.

As previously noted, in the sections of this report that discuss the correlational analyses, the various Achievement Tests correlate quite differently with scores on the traditionally used covariates, SAT-V and SAT-M. A number of tests contain scores that are highly related to either scores on SAT-V or SAT-M. Examples of these tests are the English Composition Test, the Literature Test, and the Mathematics Level II Test. On the other hand, some tests provide scores that do not show a very strong relationship to scores obtained on either SAT-V or SAT-M. The foreign language tests show particularly low correlations with scores on these variables.

In order to determine what additional covariates might be added to the model used to scale the Achievement Tests, correlations of test scores with the 17 potential covariates were examined. And, more importantly, stepwise regressions were carried out. The stepwise regressions were used to answer questions such as: if SAT-V and SAT-M scores are used to scale Achievement Test scores, will the scaling improve if semesters of study are added as an additional covariate? In addition, the results of the stepwise regression analyses were compared to determine if several tests could be clustered together, i.e., tests within a cluster show a similar relationship to a common core of covariates and, hence, can all be rescaled using these covariates. It should be noted that, in choosing tests to form the clusters, practical considerations were considered very important, i.e., efforts were made to keep tests together that formed logical clusters such as science, foreign languages, etc.

As noted in a previous section of this report, it should be pointed out that the stepwise regressions carried out for this study were completely exploratory; i.e., the variables shown for the specific models were those that maximized the percentage of variance accounted for in a particular data set. A logical procedure for specifying the regression models could have been used. If this had been the case, variables used for the traditional method of scaling, SAT-V, SAT-M, and in some cases, semesters of study, would have been included in the models prior to the search for additional covariates.

An additional point that was noted previously, and should be noted here, is the sample dependency of the regression analyses that provided the particular models that will be discussed here. Because the samples that were used for this study (although large) may not be representative of the Achievement Test population, the results of the study should be cross validated (using a second sample of Achievement Test takers) to see if the regression models generalize across samples.

The results of the regression analyses that were carried out for the total groups taking the 14 tests are summarized in Table 10. Only the results of the best two-, three-, and four-variable models are presented. First, examine the information presented for the English Composition Test. Examination of the three-variable models specified for this test (across the five administrations) indicates that the models are reasonably consistent and include variable 1 (SAT-V scaled scores), variable 2 (SAT-M scaled scores) and variable 12 (grades in foreign languages).

It is apparent that if the values of R^2 (the percentage of variance in Achievement Test score accounted for by the particular regression model) are compared for the two-, three-, and four-variable models, some gain in R^2 occurs when the models are expanded from two- to three-variable models. However, very little gain accrues when the models are further expanded to include four variables. Hence, one would conclude that probably the best results that can be expected, given the 17 available covariates, are results that can be gained using the best three-variable model.

Insert Table 10 about here

The most logical test to include in a cluster with the English Composition Test is the Literature Test. Examination of the information provided in Table 10 for the Literature Test indicates that the best three-variable model specified across the five administrations is one that consistently includes two variables, variable 1 (SAT-V scaled scores) and variable 10 (grades in English), so there is at least some consistency with the best three-variable model specified for the English Composition Test, although not quite as much consistency as one would hope for. This first grouping of tests, English Composition and Literature, will be referred to as Cluster 1, the English Cluster.

Next, consider the information presented in Table 10 that pertains to the American History and Social Studies Test and the European History and World Cultures Test. A comparison of R^2 values across the best two-, three-, and four-variable models for these two tests indicates that the most appropriate model to select for the two tests is probably a three-variable model. The best three-variable model is specified somewhat consistently across the seven administrations of the two tests (presented in Table 10) and is one that includes variables 1, 2, and 15 (SAT-V and SAT-M scores, and grades in Social Studies, respectively). The second grouping of tests, referred to as the

History Cluster, is one that consists of the American History and Social Studies Test and the European History and World Cultures Test.

The procedure used to form Clusters 1 and 2, the English Cluster and the History Cluster, was followed in the formation of three additional clusters. Cluster 3 is referred to as a Mathematics Cluster and contains the Mathematics Level I and Level II Tests. Cluster 4 is the Science Cluster and contains the Biology, Chemistry, and Physics Tests. The final cluster, Cluster 5, is the Languages Cluster and it contains the French, Spanish, and Latin Tests.

It should be noted that neither the German Test nor the Hebrew Test were placed in clusters. This is because the results of the regression analyses carried out for these tests were inconsistent with those carried out for the other language tests and also because the percentage of variance accounted for by the four-variable model specified for the German data was quite low. A decision to include these tests in the Language Cluster would have to be made on a purely pragmatic basis and comparability of scores on these tests with the remaining language tests would be clearly questionable.

The decision not to include the Hebrew Test in the Language Cluster does not have particularly serious implications. This is because the test is currently being extensively revised and it is anticipated that the revised test will provide data that may be more consistent with that obtained for the other language tests. The decision to exclude the German Test from the Language Cluster does have serious implications and should be considered further. It is important to determine, if possible, what there is about the German Test data that results in such low correlations of test score with potential covariates. It may be possible that if larger samples were used, perhaps data accumulated across five testing years, improved results might be obtained. Both of these factors will be investigated in the second phase of this study.

Table 11 contains a summary of the correlations of Achievement Test scores with the scaling covariates for the proposed clusters. An examination of the information displayed for the tests included in the English Cluster indicates that all tests correlate highly with SAT-V scaled scores, reasonably highly with SAT-M scaled scores and moderately with grades in the respective subject areas. Similar conclusions can be drawn regarding tests comprising the History Cluster. However, it should be noted that the European History and World Cultures Test scores show a lower correlation with the covariates than the American History and Social Studies Test scores. This is particularly true of scores obtained at the December administration of the test.

Insert Table 11 about here

Cluster 3, the Mathematics Cluster, contains the Mathematics Level I and II Tests. Proposed covariates for these two tests are variable 2 (SAT-M scaled scores) and variable 17 (Achievement Test Background Questionnaire responses). High correlations with variable 2 and moderate correlations with variable 17 are displayed by the Mathematics Level I and Mathematics Level II data summarized in Table 11.

Proposed covariates for Cluster 4, the Science Cluster, are variable 1 (SAT-V scaled scores), variable 2 (SAT-M scaled scores), and variables 7 or 8 (amount of course work in biology and sciences or amount of course work in physical sciences, respectively). The Chemistry and Physics Tests show correlations with the covariates that are slightly more similar to each other than they are to the correlations displayed by the Biology Test scores. Scores obtained on the Biology Test have a tendency to be a little more highly correlated with SAT-V scores and a little less highly correlated with SAT-M scores than scores obtained on the other two tests in the cluster.

Cluster 5 contains the three language tests, French, Spanish and Latin. The lower correlations of all of the language tests with variables 1 and 2 (SAT-V and SAT-M scores, respectively) is immediately apparent. The additional covariate (in addition to those used by the current procedure) proposed for this cluster is variable 12 (grades in foreign languages). All of the tests show a fairly strong correlation with this variable. Also, reference to the information provided in Table 10 shows that expansion of the regression model from two to four variables provides a considerable increase in the percentage of variance accounted for by the model.

To summarize, the tests were sorted into five possible scaling clusters based on the results of the stepwise regression analyse. which are summarized in Table 10, as well as on logical considerations. The results of the clusterings are, to some extent, confirmed by an examination of the correlations that are summarized in Table 11. It is suggested that, based on the results of this portion of the study, two alternative scaling procedures be investigated. Alternative One (single stage scaling) would consist of scaling the tests as five independent clusters (using as covariates the variables used to define the clusters). No attempt would be made to establish any type of relationship among the scores obtained on tests that are members of different clusters. Alternative Two (two stage scaling) would consist of a two-stage process. The first stage would consist of scaling the clusters independently, as has been suggested for Alternative One. The second stage would use the results of the cluster scalings and scale the tests again, using SAT-V and SAT-M scores as common covariates for all 14 tests.

Hypothetical Scaling Population

The hypothetical reference population that is used in the current Achievement Test scaling procedure is assumed to have a scaled score mean of 500 on SAT-V and SAT-M, a scaled score standard deviation of 100 on both these variables, and a correlation between scores obtained on SAT-V and SAT-M of .60. As mentioned previously, it was pointed out by Braun (see Dorans, 1985) that these values influence the scaling results. The results of Braun's studies indicated that choice of reference population is of critical importance. Braun found that different reference groups produced different Achievement Test scale alignments.

Examination of the information presented in Table 6 and Figures 2 and 3 indicate that, in general, SAT-V scaled score means for the total groups that take the Achievement Tests at the five administrations range between roughly 500 and 575. SAT-M scaled score means obtained by the same groups range, roughly, between 525 and 675. The total group correlation between scores on these two tests (Table 2) range between .44 and .69 with the majority of the correlation coefficients falling somewhere in the vicinity of .55.

No mention has been made thus far in this report of the magnitude of the scaled score standard deviations obtained on SAT-V and SAT-M by the various groups who take the Achievement Tests. Examination of the information provided in Table 6 indicates that SAT-V scaled score standard deviations obtained by the total groups taking the 14 tests range between 91 and 117. Scaled score standard deviations on SAT-M obtained for the same groups range between 77 and 111.

It is fairly clear, from the information presented above, that SAT-V and SAT-M means obtained by the groups taking the 14 tests are quite diverse. This result strongly emphasizes the need for scaling and for making every effort to improve the quality of scaling operations. It is also apparent that the group of students who are currently taking the Achievement Tests have mean SAT-V and SAT-M scores that differ substantially from the value of 500 for both scores assumed in the definition of the standard reference group. Because substantially all Achievement Test takers take the English Composition Test, a reasonable approximation to the mean score can be obtained by calculating the average score on the basis of data in Table 6. As it happens, the mean for SAT-V is 514 and the mean for SAT-M is 568. An alternative estimate is provided by results reported in National Report: High School Seniors, 1986 published by the College Board. In calculating these means, the SAT scores of each student who took the Achievement Tests in the cohort of 1986 seniors was counted once and only once. For SAT-V, the mean is 577 and for SAT-M, the mean is 576. These results indicate that consideration should be given to redefining the standard reference group so that its statistical characteristics would be more consistent with the characteristics of the group currently taking those tests. Because a change in the definition might result in a serious discontinuity in the score scale for some of the 14 tests, care should be taken in making this change to minimize its effects on the form-to-form comparability of scores.

Table 12 contains summary statistics for scores on the selected covariates obtained by the total groups constituting the five proposed scaling clusters. Perusal of the information in Table 12 indicates that a great deal of diversity of scores on the respective covariates also exists within the five clusters.

Insert Table 12 about here

A number of possibilities exist for alternatives to the way in which the hypothetical reference group is currently specified. Given the diversity of the scores on the covariates obtained by the various groups, an empirical approach might be the most feasible. This approach could be one that involves pooling scores on the various covariates and using the pooled means, standard deviations, etc., as the values to specify for the hypothetical reference group. Pooled covariate data could be used within clusters, if a single stage scaling procedure were used or first, within, and second, across clusters, if a two stage scaling procedure was viewed as desirable.

Suggested Revisions to the Traditional Scaling Methodology

As mentioned previously, the traditionally used Achievement Test scaling method consists of the following components:

- Samples are selected from seniors taking the tests at the November, December, and January administrations and juniors taking the tests at the May and June administrations.
- The hypothetical reference population is specified to have a scaled score mean of 500 and a standard deviation of 100 on both SAT-V and SAT-M and a correlation between scores on these two variables of .60.
- SAT-V and SAT-M scaled scores are used as common covariates for all tests and semesters of study are used as additional unique covariates for the foreign language tests.

As a result of the analyses carried out for this study, some revisions to the traditionally used methodology have been suggested based on empirical evidence.

Sampling Achievement Test Scores

It is suggested that the traditional method of defining the scaling sample (selecting high school seniors who took the tests in November, December, and January, and high school juniors who took the tests in May and June) be augumented by selecting high school sophomores taking the tests in June, for those tests where sophomores comprise a large proportion of June test takers.

Hypothetical Reference Group

It is suggested that the hypothetical reference group be defined empirically. If a single stage scaling approach is to be used, five hypothetical reference groups could be specified. Each group could be specified by pooling data across the covariates on a within cluster basis. If a two stage scaling approach is to be used, six hypothetical reference groups could be specified. First, five groups could be defined as described for the single stage scaling procedure and secondly, the sixth group could be defined by pooling data for the common covariates (SAT-V and SAT-M scores) that are used across the five clusters.

Scaling Models

As alluded to previously in this report, it is suggested that two scaling procedures be evaluated: 1) a single stage scaling procedure, and 2) a two stage scaling procedure. The single stage scaling procedure would involve scaling tests solely on a within cluster basis; i.e., different covariates, different reference populations, and possibly even different sampling procedures would be used for each of the five previously described clusters. The end result would be Achievement Test scales that would be unique to each cluster.

The two stage scaling procedure builds on the single stage procedure; i.e., the scaled scores obtained from the single stage scalings are used as input to the second stage scaling. In this way, an attempt is made to define a single scale that will provide roughly comparable scores across the five clusters.

The purpose of the study described in this report was to explore the relationships between College Board Achievement Test scores and potential scaling covariates. In addition, the characteristics of the various Achievement Test populations were examined to evaluate the traditionally used sampling procedures and the manner in which the reference population is defined.

Although exploratory in nature, several underlying assumptions governed the direction of the study. One assumption was that test taking patterns may have changed over the years both as a result of natural academic changes in the population and as a result of changes in ATP test administration schedules. For this reason, it was thought that the manner in which the scaling sample is selected required review. The results of the scaling study indicate that dramatic changes have not really taken place in the test taking patterns and, with minor revisions, the method traditionally used for selecting scaling samples is most probably still viable.

A second assumption, based on the results of Tucker and Braun's studies (see Dorans, 1985) was that improved scaling results might be obtained if the specifications for the reference population used for the scaling were revised to reflect more closely the characteristics of the current Achievement Test population. Empirical evidence collected in this study indicated that, indeed, the current population (however you wish to define it) has SAT-V and SAT-M summary statistics that are quite different from those traditionally used for the reference scaling population. This is not necessarily bad. The issue is: will improved scaling results be obtained (at least better within cluster comparability of scores) if multiple reference populations, or a single reference population, based on empirical data are used?

Finally, a very important assumption underlying the current study was: given that the traditionally used covariates that are common across the 14 tests produce test scales that are of unknown comparability, definition of separate sets of covariates for tests that appear to have a similar relationship with these covariates, might provide improved scaling results. These results would be in the form of improved comparability of scales for tests included in a cluster.

The regression analyses carried out for this stage of the study did suggest some tentative clusterings of tests. However, a number of weaknesses of the analyses have already been alluded to. For one, a new version of the SDO and also additional Background Questionnaires have been introduced in ATP since data were collected for this study. It is quite likely that the results of the study might have been different if these additional data had been available. Also, scores for the Test of Standard Written English (TSWE) were not available at the time data were collected for this study. It would be interesting to investigate the use of these scores as an additional covariate for tests that show a strong relationship to SAT-V scores. Finally, as mentioned previously, the results of the regression analyses are sample dependent, hence, suggested clusterings should be considered as tentative and subject to cross validation.

In conclusion, it is recommended that further study be carried out. It is suggested that particular care be paid to evaluating the new SDO, Background Questionnaires, and TSWE scores as potential covariates. It is also recommended that the tentative clusters suggested here be cross validated. Furthermore, it would seem important to include, as one of the goals of future studies, a thorough investigation of problems related to the German Test such that ways in which the test might be included in the Language Cluster could be delineated.

REFERENCES

Dorans, N. J. (1985). Achievement Test rescaling: Help or hindrance? SR-85-92. Princeton, NJ: Educational Testing Service.

SAS Institute Statistical Analysis System (1982). SAS user's guide: Statistics 1982 edition. Cary, NC: SAS Institute, Inc.

Wilks, S. S. (1961). Scaling and equating College Board Tests. Princeton, NJ: Educational Testing Service.

Table 1

Descriptions of Variables Used as Covariates in the Correlation and Regression Studies

Variable No.	Description[1]
1.	SAT-Verbal scaled scores
2.	SAT-Mathematical scaled scores
3.	High school class rank
4.	Amount of high school course work in English
5.	Amount of high school course work in Mathematics
6.	Amount of high school course work in foreign languages
7.	Amount of high school course work in Biology and Sciences
8.	Amount of high school course work in Physical Sciences
9.	Amount of high school course work in Social Studies
10.	Latest year and/or midyear grade received in English
11.	Latest year and/or midyear grade received in Mathematics
12.	Latest year and/or midyear grade received in foreign languages
13.	Latest year and/or midyear grade received in Biology and Sciences
14.	Latest year and/or midyear grade received in Physical Sciences
15.	Latest year and/or midyear grade received in Social Studies
16.	Number of honors or awards received during high school years
17.	Amount of training in a subject as assessed by Achievement Test Background Questionnaires (see Figure 3)

Table 1 (continued)

Variable No.	Description[1]
18.	Highest level of education you plan to complete beyond high school
19.	Highest level of education completed by your father or male guardian
20.	Highest level of education completed by your mother or female guardian
21.	Parental income

[1]Tables 2 and 3 contain text of the actual questions responded to by examinees that supplied information for variables 3-17.

Table 2

Selected Questions from Student Descriptive Questionnaire (SDQ)
Used to Collect Covariate Information for Scaling Study

Variable No.	SDQ No.	Text of Question
3	5	What is your most recent high school class rank? (For example, if you are 15th in a class of 100, you are in the second tenth.) If you do not know your rank or rank is not used in your school, give your best estimate. (A) Highest tenth ⎤ (B) Second tenth ⎬ top fifth (C) Second fifth ⎦ (D) Middle fifth (E) Fourth fifth (F) Lowest fifth
4-9	6-11	Questions 6 through 11 ask you to blacken the letter corresponding to the total years of study you expect to complete in certain subject areas. Include in the total only courses you have taken since beginning the ninth grade and those you expect to complete before graduation from high school. Count less than a full year in a subject as a full year. Do not count a repeated year of the same course as an additional year of study. (A) One year or the equivalent (B) Two years or the equivalent (C) Three years or the equivalent (D) Four years or the equivalent (E) More than four years or the equivalent (F) I will not take any courses in the subject area.
4	6	English
5	7	Mathematics
6	8	Foreign Languages
7	9	Biological Sciences (for example, biology, botany, or zoology)
8	10	Physical Sciences (for example, chemistry, physics, or earth science)

Table 2 (continued)

Variable No.	SDQ No.	Text of Question
9	11	Social Studies (for example, history, government, or geography)
10-15	12-17	For each of the subject areas in questions 12 through 17, blacken the latest year-end or midyear grade you received since beginning the ninth grade. For example, if you are a senior and have not taken biology or any other biological science since your sophomore year, indicate that year-end grade. If you are a junior and have completed the first half of the year in an English course, indicate that midyear grade. If you received the grade in an advanced, accelerated, or honors course, also blacken the letter H. (A) Excellent (usually 90-100 or A) (B) Good (usually 80-89 or B) (C) Fair (usually 70-79 or C) (D) Passing (usually 60-69 or D) (F) Failing (usually 59 or below or F) (G) Only "pass-fail" grades were assigned and I received a pass. (H) The grade reported was in an advanced, accelerated, or honors course.
10	12	English
11	13	Mathematics
12	14	Foreign Languages
13	15	Biological Sciences
14	16	Physical Sciences
15	17	Social Studies
16	23	During your high school years how many honors or awards (for example, essay contest, debating tournament, science fair, music, art or theater competition, or membership in a scholastic honors group) have you received? (A) None (B) One or two (C) Three or four (D) Five or six (E) Seven or more

Table 2 (continued)

Variable No.	SDQ No.	Text of Question
18	24	What is the highest level of education you plan to complete beyond high school? (A) A two-year specialized training program (for example, electronics or laboratory technician) (B) A two-year Associate of Arts degree (A.A.) (C) Bachelor's degree (B.A. or B.S.) (D) Master's degree (M.A. or M.S.) (E) Doctor's or other professional degree (such as M.D. or Ph.D.) (F) Other or undecided
19	39	Indicate the highest level of education completed by your father or male guardian. (A) Grade school (B) Some high school (C) High school diploma (D) Buisness or trade school (E) Some college (F) Bachelor's degree (G) Some graduate or professional school (H) Graduate or professional degree
20	40	Using the list in question 39, indicate the highest level of education completed by your mother or female guardian.
21	43	What was the approximate income of your parents before taxes last year? Include taxable and nontaxable income from all sources. (A) Less than $3,000 a year (about $57 a week or less) (B) Between $3,000 and $5,999 a year (from $58 to $114 a week) (C) Between $6,000 and $8,999 a year (from $115 to $173 a week) (D) Between $9,000 and $11,999 a year (from $174 to $230 a week) (E) Between $12,000 and $14,999 a year (from $231 to $288 a week) (F) Between $15,000 and $17,999 a year (from $289 to $346 a week) (G) Between $18,000 and $20,999 a year (from $347 to $403 a week)

Table 2 (continued)

Variable No.	SDQ No.	Text of Question
		(H) Between $21,000 and $23,999 a year
		(I) Between $24,000 and $26,999 a year
		(J) Between $27,000 and $29,999 a year
		(K) Between $30,000 and $34,999 a year
		(L) Between $35,000 and $39,999 a year
		(M) Between $40,000 and $44,999 a year
		(N) Between $45,000 and $49,000 a year
		(O) $50,000 a year or more

Table 3

Achievement Test Background Questionnaires Used to Collect Covariate Information for Scaling Study

---

Mathematics Level I and Level II Tests

To provide information on your training in mathematics, please answer the two questions below. For each question, indicate your answer by blackening one of the specified spaces in the row of nine spaces labeled Q. (You will need to blacken at most TWO of the spaces labeled Q.) The information that you provide is for statistical purposes only and will not influence your score on the test.

Question 1

Which of the following best describes the college preparatory program you have taken or are taking? If you are currently enrolled in an Advanced Placement Mathematics course (Calculus) AB or Calculus BC), blacken space 1, leave spaces 2 through 5 blank, and go on to Question II. If not, please read the statements below and then blacken at most one of spaces 2 through 5 as your answer. If none of the descriptions appropriately describes the mathematics courses you have taken or are taking, leave spaces 1 to 5 blank.

Full-year courses (or their equivalents) elementary algebra (Algebra I), geometry, and intermediate algebra (Algebra II); at least one semester of elementary functions, mathematical analysis, or advanced algebra (Algebra III); and trigonometry as a separate course for at least one semester, with a grade for each mathematics course of C or better. - Blacken space 2.

Full-year (or their equivalents) in elementary algebra (Algebra I), geometry, and intermediate algebra (Algebra II); at least one semester elementary functions, mathematical analysis, or advanced algebra (Algebra III); and trigonometry as part of these courses, with a grade for each mathematics course of C or better. - Blacken space 3.

Table 3 (continued)

Mathematics Level I and Level II Tests (continued)	
Full-year courses (or their equivalents) in elementary algebra (Algebra I), geometry, and intermediate algebra (Algebra II); and at least one semester of elementary functions, mathematical analysis, or advanced algebra (Algebra III), with a grade for each mathematics course of C or better.	- Blacken space 4.
Full-year courses (or their equivalents) in elementary algebra (Algebra I), geometry, and intermediate algebra (Algebra II), with a grade for each mathematics course of C or better.	- Blacken space 5.
<u>Question II</u>	
Which of the following best describes how long it has been since you took a course in <u>plane</u> geometry? Blacken only <u>one</u> of spaces 6 through 9 as your answer to Question II.	
I have never taken a course in plane geometry.	- Blacken space 6.
One school year or less.	- Blacken space 7.
More than one school year but less than two school years.	- Blacken space 8.
More than two school years.	- Blacken space 9.

Table 3 (continued)

Physics Test

To provide information on training in physics, please answer the question below. Indicate your answer by blackening one of the first four spaces labeled Q on your answer sheet. Your response will not influence your test score.

Question 1

How many semesters of instruction have you had in a high school physics course or courses? (There are two semesters in a school year. If you are taking physics in the current semester and the semester is more than half over, count this current semester as a full semester.)

Space 1: One semester

Space 2: Two semesters

Space 3: Three semesters

Space 4: Four semesters

Spaces 5 to 9: (Leave blank)

Table 3 (continued)

French Test[1]

In the group of nine spaces labeled Q, you are to blacken ONE and ONLY ONE space, as described below, to indicate how you obtained your knowledge of French. The information that you provide is for statistical purposes only and will not influence your score on the test.

Question 1

If your knowledge of French does not come primarily from courses taken in grades 9 through 12, blacken space 9 and leave the remaining spaces blank, regardless of how long you studied the subject in school. For example, you are to blacken space 9 if your knowledge of French comes primarily from any of the following sources: study prior to the ninth grade, courses taken at a college, or special study, residence abroad, or living in a home in which French is spoken.

If your knowledge of French does come primarily from courses taken in grades 9 through 12, blacken the space that indicates the level of the French course in which you are currently enrolled. If you are not now enrolled in a French course, blacken the space that indicates the level of the most advanced course in French that you have completed.

Level I:	first or second half	- blacken space 1
Level II:	first half	- blacken space 2
	second half	- blacken space 3
Level III:	first half	- blacken space 4
	second half	- blacken space 5
Level IV:	first half	- blacken space 6
	second half	- blacken space 7
Advanced Placement or course that represents a level of study higher than Level IV: second half		- blacken space 8

If you are in doubt about whether to mark space 9 rather than one of the spaces 1-8, mark space 9.

[1]The same questionnaire (with the appropriate test name) appears in the French, German, Latin and Spanish Tests. The Latin questionnaire differs slightly in that the phrase, "...or living in a home in which [language] is spoken" is eliminated.

Table 4

Specifications for Recoding Responses to Student Descriptive Questionnaire and Achievement Test Background Questionnaires

Student Descriptive Questionnaire			
Question No.	Question	Response	Code
5	High school class rank	F	1
		E	2
		D	3
		C	4
		B	5
		A	6
6-11	Amount of high school course work	F	1
		A	2
		B	3
		C	4
		D	5
		E	6
12-17	Latest year and/or midyear grade in specific subjects	F	1
		D	2
		C&G	3
		B	4
		A	5
		H	If H is marked, advance students code designation by one.
23	Honors or awards	A	1
		B	2
		C	3
		D	4
		E	5
24	Highest level of planned education	A	1
		B	2
		C	3
		D	4
		E	5
		F	(delete)

Table 4 (continued)

Student Descriptive Questionnaire

Question No.	Question	Response	Code
39	Father's level of education	A	1
		B	2
		C	3
		D	4
		E	5
		F	6
		G	7
		H	8
40	Mother's level of education	A	1
		B	2
		C	3
		D	4
		E	5
		F	6
		G	7
		H	8
43	Parental income	A	1
		B	2
		C	3
		D	4
		E	5
		F	6
		G	7
		H	8
		I	9
		J	10
		K	11
		L	12
		M	13
		N	14
		O	15

Table 4 (continued)

Achievement Test Background Questionnaires			
Test	Response		Code
	Question 1	Question 2	
French, Spanish, German and Latin	1	-	Removed from sample
	2	-	3
	3	-	4
	4	-	5
	5	-	6
	6	-	7
	7	-	8
	8	-	9
	9	-	Removed from sample
Math Level I and Math Level II	1	6	5
	1	7	5
	1	8	5
	1	9	5
	2	6	4
	2	7	4
	2	8	4
	2	9	4
	3	6	3
	3	7	3
	3	8	3
	3	9	3
	4	6	2
	4	7	2
	4	8	2
	4	9	2
	5	6	1
	5	7	1
	5	8	1
	5	9	1
Physics	1	-	1
	2	-	2
	3	-	3
	4	-	4

Table 5

Summary Statistics by Grade Level and Administration Date for Fourteen Achievement Tests[1,2]

English Composition Test

Grade Level	November N	November Mean	November S.D.	December N	December Mean	December S.D.	January N	January Mean	January S.D.	May N	May Mean	May S.D.	June N	June Mean	June S.D.
Freshmen													285	513	105
Sophomore	33	549	107	27	567	134	42	523	142	320	519	109	3613	521	102
Junior	301	538	115	2326	555	103	1207	544	111	10473	527	103	80895	531	99
Senior	44092	536	102	145440	521	99	62401	497	98	6159	485	97	5158	486	97
Total	44740	536	102	148455	522	100	64510	498	98	17143	511	103	90198	528	100

Literature Test

Grade Level	November N	November Mean	November S.D.	December N	December Mean	December S.D.	January N	January Mean	January S.D.	May N	May Mean	May S.D.	June N	June Mean	June S.D.
Freshmen															
Sophomore													240	541	105
Junior	32	539	125	132	554	115	105	525	108	889	517	104	5773	530	104
Senior	5447	530	106	15763	528	104	7864	503	104	785	490	103	648	494	110
Total	5536	530	106	16029	528	104	8141	503	105	1734	505	105	6739	526	105

American History and Social Studies Test

Grade Level	November N	November Mean	November S.D.	December N	December Mean	December S.D.	January N	January Mean	January S.D.	May N	May Mean	May S.D.	June N	June Mean	June S.D.
Freshmen													28	472	78
Sophomore				25	604	104				113	548	109	1730	557	92
Junior	36	552	74	192	529	95	198	518	103	2826	539	93	28506	557	99
Senior	8975	512	92	20181	505	91	11552	496	94	886	462	88	819	467	102
Total	9076	512	92	20495	505	91	11929	496	94	3857	521	99	31133	555	100

Mathematics Level I Test

Grade Level	November N	November Mean	November S.D.	December N	December Mean	December S.D.	January N	January Mean	January S.D.	May N	May Mean	May S.D.	June N	June Mean	June S.D.
Freshmen										25	582	99	622	577	109
Sophomore	30	597	121	50	602	113	52	606	108	472	574	92	10269	582	92
Junior	347	597	96	1344	584	91	1238	592	98	8306	542	95	67717	551	93
Senior	30321	548	94	99366	531	84	48787	525	90	4983	513	94	4016	506	96
Total	30937	548	94	101366	532	85	50762	527	92	13943	532	97	82894	553	95

Mathematics Level II Test

Grade Level	November N	November Mean	November S.D.	December N	December Mean	December S.D.	January N	January Mean	January S.D.	May N	May Mean	May S.D.	June N	June Mean	June S.D.
Freshmen													86	737	66
Sophomore							29	751	47	85	700	97	1433	701	84
Junior	110	703	91	337	693	86	426	687	98	2572	657	91	20652	680	82
Senior	12470	665	84	28980	645	88	13953	634	96	2349	565	103	1631	610	102
Total	12688	666	84	29555	646	89	14659	636	97	5049	615	108	23871	677	86

Table 5 (continued)

Summary Statistics by Grade Level and Administration Date for Fourteen Achievement Tests

Biology Test

Grade Level	November N	November Mean	November S.D.	December N	December Mean	December S.D.	January N	January Mean	January S.D.	May N	May Mean	May S.D.	June N	June Mean	June S.D.
Freshmen										139	569	86	15320	577	92
Sophomore	136	590	107	99	568	104	115	577	111	462	545	94	26367	552	96
Junior	225	562	103	383	557	97	291	567	106	916	570	104	7447	570	103
Senior	4847	536	105	14872	515	100	8483	505	105	496	500	105	491	491	114
Total	5258	539	106	15499	517	100	9027	508	106	2022	547	105	49674	562	97

Chemistry Test

Grade Level	November N	November Mean	November S.D.	December N	December Mean	December S.D.	January N	January Mean	January S.D.	May N	May Mean	May S.D.	June N	June Mean	June S.D.
Freshmen													172	605	97
Sophomore										190	566	97	11741	603	91
Junior	124	593	94	234	605	107	348	608	93	1418	568	105	21572	576	99
Senior	6094	570	101	15120	555	100	7914	546	93	404	522	113	424	548	119
Total	6272	571	102	15497	556	100	8442	550	94	2038	559	108	33957	585	98

Physics Test

Grade Level	November N	November Mean	November S.D.	December N	December Mean	December S.D.	January N	January Mean	January S.D.	May N	May Mean	May S.D.	June N	June Mean	June S.D.
Freshmen													73	654	75
Sophomore										27	643	101	670	636	90
Junior	28	651	97	65	629	103	156	642	103	628	605	75	8391	615	91
Senior	3236	614	97	8856	576	96	7590	557	94	708	552	88	743	561	95
Total	3365	616	98	9139	578	96	8012	561	96	1391	579	97	9932	613	93

French Test

Grade Level	November N	November Mean	November S.D.	December N	December Mean	December S.D.	January N	January Mean	January S.D.	May N	May Mean	May S.D.	June N	June Mean	June S.D.
Freshmen													36	554	110
Sophomore										106	508	107	1270	552	104
Junior	55	530	114	260	582	106	246	597	101	1373	554	103	9457	560	98
Senior	3529	539	100	13838	524	96	4529	503	91	544	487	104	689	502	110
Total	3607	538	100	14143	526	96	4814	508	94	2040	533	108	11464	555	100

Spanish Test

Grade Level	November N	November Mean	November S.D.	December N	December Mean	December S.D.	January N	January Mean	January S.D.	May N	May Mean	May S.D.	June N	June Mean	June S.D.
Freshmen													31	542	105
Sophomore										84	489	99	1076	532	95
Junior	43	530	104	162	553	109	172	532	106	1246	537	98	9678	539	91
Senior	4155	525	98	13772	496	93	5476	483	104	632	488	105	741	488	108
Total	4208	525	98	13967	497	94	5694	485	105	1972	519	103	11532	535	93

Table 5 (continued)

Summary Statistics by Grade Level and Administration Date for Fourteen Achievement Tests

European History and World Cultures Test

Grade Level	November N	November Mean	November S.D.	December N	December Mean	December S.D.	January N	January Mean	January S.D.	May N	May Mean	May S.D.	June N	June Mean	June S.D.
Freshmen															
Sophomore										285	550	100			
Junior				128	579	99				558	569	95			
Senior				3461	541	101				166	536	100			
Total				3668	542	102				1037	557	99			

German Test

Grade Level	November N	November Mean	November S.D.	December N	December Mean	December S.D.	January N	January Mean	January S.D.	May N	May Mean	May S.D.	June N	June Mean	June S.D.
Freshmen															
Sophomore										33	561	110			
Junior				70	572	109				331	571	102			
Senior				2947	533	95				226	495	115			
Total				3033	534	96				594	542	114			

Hebrew Test

Grade Level	November N	November Mean	November S.D.	December N	December Mean	December S.D.	January N	January Mean	January S.D.	May N	May Mean	May S.D.	June N	June Mean	June S.D.
Freshmen															
Sophomore															
Junior				34	646	90				43	661	111			
Senior				145	615	96				16	544	104			
Total				193	625	97				60	632	121			

Latin Test

Grade Level	November N	November Mean	November S.D.	December N	December Mean	December S.D.	January N	January Mean	January S.D.	May N	May Mean	May S.D.	June N	June Mean	June S.D.
Freshmen										28	567	79			
Sophomore										274	588	114			
Junior				127	630	113				709	573	107			
Senior				2337	550	104				134	524	120			
Total				2482	555	106				1147	570	111			

[1] Total includes examinees who did not specify grade level.

[2] Summary Statistics were computed for all available cases who had Achievement Test Scores.

Table 6

SAT-Verbal and SAT-Mathematical Summary Statistics by Grade Level[1,2] and Administration Date for Fourteen Achievement Tests

English Composition Test

Grade Level	November						December						January						May						June					
	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}
Freshmen	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	36	537	101	603	99	0.42
Sophomore	--	--	--	--	--	--	--	--	--	--	--	--	26	510	154	578	153	0.77	38	507	122	613	126	0.66	425	533	102	603	98	0.44
Junior	148	536	129	630	101	0.51	764	560	106	602	102	0.55	640	537	120	608	105	0.56	7700	521	108	572	109	0.62	75545	524	98	574	103	0.58
Senior	41773	525	109	592	103	0.57	140877	518	100	566	101	0.55	59720	490	102	553	103	0.54	5820	470	99	533	103	0.53	4814	476	102	539	106	0.54
Total	42136	525	109	592	103	0.57	142183	518	100	566	101	0.55	61062	491	103	554	103	0.54	13691	499	107	555	109	0.60	80997	521	99	572	103	0.58

Literature Test

Grade Level	November						December						January						May						June					
	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}
Freshmen	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Sophomore	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	36	596	99	594	92	0.62
Junior	--	--	--	--	--	--	72	578	108	571	87	0.49	55	539	116	553	102	0.65	603	521	109	528	112	0.64	5264	537	103	537	101	0.59
Senior	5127	536	108	556	100	0.61	15058	528	102	531	99	0.56	7409	511	106	525	101	0.57	731	479	104	507	102	0.56	592	487	109	511	109	0.66
Total	5181	536	108	556	100	0.61	15232	528	103	531	99	0.56	7599	512	106	525	101	0.57	1360	499	109	516	107	0.61	5936	532	105	535	103	0.60

American History and Social Studies Test

Grade Level	November						December						January						May						June					
	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}
Freshmen	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Sophomore	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	117	557	104	603	104	0.56
Junior	--	--	--	--	--	--	65	530	116	581	106	0.42	68	546	114	558	115	0.66	1883	521	105	557	109	0.60	26199	533	98	563	101	0.58
Senior	8466	512	107	562	100	0.56	19340	507	99	539	100	0.53	10959	492	101	529	101	0.53	799	460	104	506	108	0.58	739	463	106	504	110	0.53
Total	8518	512	107	562	100	0.56	19484	507	99	539	100	0.53	11160	492	101	529	101	0.54	2712	503	108	541	111	0.61	27090	532	99	562	102	0.59

Mathematics Level I Test

Grade Level	November						December						January						May						June					
	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}
Freshmen	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	91	511	85	655	80	0.35
Sophomore	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	42	535	112	619	102	0.66	727	524	98	609	87	0.47
Junior	105	515	126	604	100	0.55	478	534	113	593	94	0.57	511	510	116	592	98	0.53	5922	501	104	554	102	0.59	63177	508	95	564	94	0.55
Senior	28542	507	106	564	97	0.55	96017	499	97	548	94	0.53	46588	478	100	539	95	0.52	4689	460	98	522	96	0.50	3716	465	100	523	99	0.52
Total	28822	507	106	565	97	0.55	96970	499	98	548	94	0.53	47644	479	101	540	96	0.52	10768	483	104	540	101	0.57	67897	506	96	562	95	0.55

Mathematics Level II Test

Grade Level	November						December						January						May						June					
	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}
Freshmen	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	48	525	90	688	81	0.53
Sophomore	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	27	546	107	672	89	0.51	215	568	95	678	79	0.46
Junior	63	541	126	666	99	0.60	173	576	115	675	87	0.43	220	553	132	661	89	0.51	1973	560	106	649	83	0.52	19500	568	97	657	77	0.48
Senior	11955	557	110	656	81	0.50	28131	548	103	639	82	0.50	13422	519	111	622	89	0.52	2251	483	96	569	93	0.46	1552	494	104	592	90	0.45
Total	12071	557	111	656	81	0.51	28471	548	104	640	82	0.50	13833	519	112	623	90	0.51	4282	519	109	607	97	0.56	21365	562	100	652	80	0.50

BEST COPY AVAILABLE

Table 6 (continued)

SAT-Verbal and SAT-Mathematical Summary Statistics by Grade Level and Administration Date for Fourteen Achievement Tests

Biology Test

Grade Level	November						December						January						May						June					
	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}
Freshmen	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	327	500	86	584	108	0.43
Sophomore	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	745	514	98	563	101	0.56
Junior	38	552	110	651	119	0.74	122	544	107	594	96	0.49	111	522	125	603	106	0.60	548	534	110	579	108	0.63	6823	533	99	579	101	0.58
Senior	4541	519	111	584	101	0.55	14391	513	101	558	101	0.57	8099	496	104	548	102	0.58	456	472	106	531	104	0.57	444	469	112	527	113	0.60
Total	4624	520	111	584	101	0.56	14621	513	102	559	101	0.57	8331	496	104	548	102	0.58	1038	506	112	558	109	0.63	8369	526	101	575	103	0.58

Chemistry Test

Grade Level	November						December						January						May						June					
	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}
Freshmen	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	35	536	96	645	99	0.54
Sophomore	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	25	519	123	626	93	0.68	703	542	92	627	86	0.43
Junior	43	555	120	664	84	0.47	99	540	122	646	105	0.65	124	553	119	662	83	0.41	987	528	112	617	97	0.55	20155	528	99	606	95	0.54
Senior	5765	545	113	650	85	0.52	14655	529	106	623	91	0.53	7634	514	111	614	94	0.51	375	476	111	587	101	0.46	389	478	129	593	107	0.49
Total	5832	545	113	650	85	0.52	14835	529	106	624	91	0.53	7878	514	111	615	94	0.50	1406	514	114	610	99	0.53	2130	528	100	606	95	0.54

Physics Test

Grade Level	November						December						January						May						June					
	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}
Freshmen	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Sophomore	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	77	567	103	670	88	0.50
Junior	--	--	--	--	--	--	34	507	133	684	65	0.31	80	530	142	660	96	0.52	471	545	113	648	90	0.52	7964	562	98	654	81	0.49
Senior	3050	554	115	673	77	0.45	8551	536	108	643	85	0.48	7266	509	111	624	89	0.47	665	479	110	594	90	0.44	696	495	114	603	93	0.50
Total	3118	553	115	673	77	0.44	8729	535	109	644	85	0.47	7531	509	111	625	89	0.47	1161	506	117	617	95	0.51	8783	556	101	650	84	0.51

French Test

Grade Level	November						December						January						May						June					
	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}
Freshmen	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Sophomore	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	111	559	109	612	95	0.63
Junior	--	--	--	--	--	--	103	575	105	604	100	0.61	121	571	95	613	95	0.62	1114	550	97	590	99	0.61	8946	547	95	581	96	0.58
Senior	3380	558	107	593	94	0.57	13487	545	94	570	93	0.52	4388	530	97	560	95	0.53	505	529	105	555	100	0.63	659	558	107	590	108	0.68
Total	3399	558	107	593	94	0.57	13617	546	94	570	93	0.52	4539	531	97	561	95	0.53	1632	543	100	579	101	0.63	9726	548	96	582	97	0.59

Spanish Test

Grade Level	November						December						January						May						June					
	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}
Freshmen	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Sophomore	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	88	522	90	590	101	0.44
Junior	--	--	--	--	--	--	59	546	114	600	106	0.68	68	534	101	580	99	0.63	916	523	97	573	102	0.60	9084	514	93	563	99	0.58
Senior	3955	515	106	569	102	0.60	13396	512	96	552	98	0.57	5276	494	99	543	100	0.58	603	477	102	528	106	0.63	688	515	109	568	109	0.68
Total	3968	515	106	569	102	0.60	13474	512	96	552	98	0.57	5373	495	99	543	101	0.58	1527	505	102	555	106	0.63	9866	514	94	563	99	0.59

Table 6 (continued)

SAT-Verbal and SAT-Mathematical Summary Statistics by Grade Level and Administration Date for Fourteen Achievement Tests

European History and World Cultures Test

	November						December						January						May						June					
Grade Level	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}
Freshmen							--	--	--	--	--	--							--	--	--	--	--	--						
Sophomore							--	--	--	--	--	--							--	--	--	--	--	--						
Junior							50	578	88	571	105	0.43							321	554	104	567	105	0.57						
Senior							3283	563	103	559	103	0.52							152	518	124	554	104	0.41						
Total							3385	564	103	559	103	0.51							508	542	112	565	105	0.52						

German Test

	November						December						January						May						June					
Grade Level	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}
Freshmen							--	--	--	--	--	--							--	--	--	--	--	--						
Sophomore							--	--	--	--	--	--							--	--	--	--	--	--						
Junior							--	--	--	--	--	--							201	565	95	598	99	0.54						
Senior							2864	562	97	598	93	0.56							215	531	103	576	102	0.54						
Total							2896	563	97	598	94	0.56							424	548	101	588	101	0.55						

Hebrew Test

	November						December						January						May						June					
Grade Level	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}
Freshmen							--	--	--	--	--	--							--	--	--	--	--	--						
Sophomore							--	--	--	--	--	--							--	--	--	--	--	--						
Junior							--	--	--	--	--	--							--	--	--	--	--	--						
Senior							142	568	100	583	105	0.68							--	--	--	--	--	--						
Total							162	571	101	584	102	0.69							31	513	107	553	108	0.48						

Latin Test

	November						December						January						May						June					
Grade Level	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}	N	SAT-V Mean	SAT-V SD	SAT-M Mean	SAT-M SD	r_{vm}
Freshmen							--	--	--	--	--	--							--	--	--	--	--	--						
Sophomore							--	--	--	--	--	--							--	--	--	--	--	--						
Junior							--	--	--	--	--	--							480	582	97	624	91	0.57						
Senior							2286	563	95	594	92	0.55							131	561	106	593	101	0.63						
Total							2315	563	95	594	92	0.55							631	578	99	618	94	0.58						

Table 7

Correlation of Achievement Test Score with Selected Covariates for Five Administrations[1]

November Administration

Test (min. N)[2]		Covariates																
		Var.1	Var.2	Var.3	Var.4	Var.5	Var.6	Var.7	Var.8	Var.9	Var.10	Var.11	Var.12	Var.13	Var.14	Var.15	Var.16	Var.17
Eng. Comp.	(3,718)	0.81	0.56	0.39	0.09	0.15	0.27	0.01	0.21	0.10	0.39	0.31	0.29	0.32	0.32	0.31	0.30	
Literature	(4,392)	0.84	0.54	0.37	0.14	0.13	0.28	0.00	0.19	0.12	0.36	0.25	0.27	0.31	0.28	0.31	0.27	
Amer. Hist.	(7,533)	0.76	0.53	0.35	0.09	0.16	0.24	0.01	0.24	0.18	0.32	0.24	0.21	0.30	0.28	0.41	0.26	
Math I	(25,153)	0.50	0.84	0.45	0.02	0.36	0.13	-0.03	0.38	0.03	0.25	0.50	0.26	0.35	0.43	0.26	0.22	0.44
Math II	(11,417)	0.44	0.79	0.34	0.03	0.15	0.15	0.00	0.26	0.03	0.19	0.39	0.22	0.25	0.36	0.19	0.22	0.32
Biology	(4,336)	0.75	0.61	0.38	0.06	0.20	0.25	0.16	0.28	0.09	0.34	0.35	0.25	0.41	0.39	0.34	0.26	
Chemistry	(5,541)	0.57	0.63	0.34	0.02	0.15	0.18	0.09	0.35	0.05	0.24	0.35	0.23	0.31	0.38	0.24	0.25	
Physics	(2,806)	0.50	0.62	0.24	0.00	0.10	0.16	0.06	0.30	0.00	0.12	0.26	0.10	0.18	0.29	0.15	0.17	0.29
French	(2,891)	0.58	0.45	0.26	0.04	0.04	0.38	-0.01	0.15	0.02	0.22	0.25	0.30	0.21	0.24	0.22	0.18	0.49
Spanish	(3,373)	0.46	0.33	0.26	0.01	0.06	0.33	0.02	0.10	0.05	0.23	0.23	0.33	0.24	0.25	0.22	0.21	0.49

December Administration

Test		Covariates																
		Var.1	Var.2	Var.3	Var.4	Var.5	Var.6	Var.7	Var.8	Var.9	Var.10	Var.11	Var.12	Var.13	Var.14	Var.15	Var.16	Var.17
Eng. Comp.	(121,770)	0.76	0.50	0.36	0.08	0.08	0.27	0.00	0.10	0.08	0.39	0.26	0.31	0.30	0.27	0.31	0.28	
Literature	(12,721)	0.82	0.49	0.36	0.11	0.05	0.22	-0.01	0.09	0.12	0.38	0.22	0.27	0.29	0.23	0.32	0.27	
Amer. Hist.	(17,061)	0.75	0.50	0.34	0.09	0.11	0.18	0.00	0.18	0.17	0.32	0.22	0.21	0.29	0.26	0.39	0.24	
Math I	(83,341)	0.48	0.82	0.42	0.02	0.31	0.10	-0.04	0.30	0.04	0.25	0.47	0.25	0.33	0.40	0.25	0.20	0.42
Math II	(25,714)	0.44	0.78	0.35	0.03	0.16	0.11	-0.02	0.25	0.01	0.21	0.40	0.21	0.26	0.35	0.21	0.23	0.33
Biology	(13,349)	0.72	0.62	0.34	0.08	0.19	0.17	0.15	0.21	0.09	0.29	0.31	0.21	0.36	0.33	0.30	0.23	
Chemistry	(13,728)	0.58	0.68	0.36	0.02	0.15	0.15	0.09	0.31	0.05	0.26	0.37	0.23	0.33	0.38	0.27	0.24	
Physics	(7,614)	0.53	0.63	0.25	0.03	0.10	0.11	0.05	0.30	0.04	0.15	0.26	0.12	0.23	0.32	0.16	0.20	0.37
French	(11,344)	0.51	0.40	0.26	0.02	0.04	0.34	0.00	0.07	0.03	0.19	0.22	0.33	0.21	0.22	0.18	0.18	0.44
Spanish	(11,436)	0.39	0.30	0.23	0.02	0.03	0.29	0.02	0.04	0.03	0.20	0.19	0.31	0.20	0.20	0.19	0.19	0.44
Europ. Hist.	(3,010)	0.69	0.46	0.25	0.02	0.08	0.18	-0.04	0.12	0.14	0.21	0.16	0.15	0.21	0.19	0.31	0.19	
German	(2,519)	0.37	0.29	0.17	0.02	0.03	0.33	0.02	0.04	-0.02	0.15	0.17	0.27	0.11	0.16	0.11	0.17	0.35
Hebrew	(118)	0.42	0.34	0.33	0.16	0.16	0.20	0.23	0.18	0.17	0.29	0.16	0.35	0.41	0.27	0.28	0.25	
Latin	(1,948)	0.53	0.48	0.33	0.02	0.06	0.29	-0.06	0.05	0.01	0.24	0.31	0.36	0.26	0.30	0.23	0.25	0.29

January Administration

Test		Covariates																
		Var.1	Var.2	Var.3	Var.4	Var.5	Var.6	Var.7	Var.8	Var.9	Var.10	Var.11	Var.12	Var.13	Var.14	Var.15	Var.16	Var.17
Eng. Comp.	(53,877)	0.79	0.54	0.38	0.10	0.12	0.21	0.00	0.15	0.10	0.40	0.28	0.28	0.32	0.29	0.31	0.30	
Literature	(6,499)	0.85	0.52	0.35	0.13	0.08	0.21	0.00	0.15	0.13	0.40	0.21	0.26	0.32	0.25	0.34	0.30	
Amer. Hist.	(9,800)	0.73	0.50	0.31	0.10	0.15	0.18	-0.03	0.20	0.18	0.32	0.21	0.20	0.30	0.27	0.38	0.23	
Math I	(42,275)	0.45	0.81	0.40	0.03	0.35	0.10	-0.03	0.33	0.06	0.24	0.46	0.23	0.31	0.38	0.24	0.20	0.43
Math II	(12,424)	0.47	0.81	0.38	0.04	0.22	0.14	0.00	0.28	0.02	0.23	0.42	0.23	0.29	0.38	0.22	0.24	0.34
Biology	(7,628)	0.72	0.62	0.36	0.05	0.18	0.18	0.18	0.24	0.10	0.29	0.33	0.22	0.38	0.34	0.30	0.24	
Chemistry	(7,231)	0.55	0.64	0.33	0.04	0.17	0.16	0.10	0.35	0.06	0.25	0.35	0.21	0.31	0.37	0.26	0.23	
Physics	(6,568)	0.54	0.63	0.27	0.05	0.15	0.10	0.09	0.32	0.08	0.16	0.26	0.11	0.23	0.30	0.17	0.20	0.36
French	(3,927)	0.44	0.37	0.24	0.02	0.05	0.35	0.05	0.10	0.00	0.16	0.21	0.27	0.19	0.21	0.14	0.17	0.42
Spanish	(4,725)	0.23	0.17	0.21	-0.01	0.01	0.24	0.04	0.04	0.03	0.17	0.14	0.28	0.14	0.17	0.15	0.16	0.41

Table 7 (continued)

Correlation of Achievement Test Score with Selected Covariates for Five Administrations

May Administration

Test		Covariates																
		Var.1	Var.2	Var.3	Var.4	Var.5	Var.6	Var.7	Var.8	Var.9	Var.10	Var.11	Var.12	Var.13	Var.14	Var.15	Var.16	Var.17
Eng. Comp.	(12,977)	0.81	0.62	0.39	0.09	0.13	0.32	-0.02	0.14	0.08	0.41	0.33	0.36	0.35	0.34	0.34	0.27	
Literature	(1,256)	0.83	0.55	0.37	0.10	0.10	0.30	0.02	0.16	0.14	0.39	0.25	0.27	0.35	0.28	0.34	0.26	
Amer. Hist.	(2,836)	0.77	0.60	0.46	0.07	0.19	0.33	0.01	0.22	0.16	0.42	0.36	0.34	0.42	0.40	0.51	0.29	
Math I	(10,448)	0.51	0.83	0.43	0.03	0.33	0.20	-0.02	0.31	0.04	0.30	0.51	0.32	0.37	0.44	0.31	0.22	0.35
Math II	(4,043)	0.53	0.83	0.33	0.03	0.21	0.31	-0.01	0.26	-0.04	0.22	0.48	0.25	0.30	0.42	0.23	0.23	0.30
Biology	(1,400)	0.74	0.70	0.39	0.11	0.22	0.25	0.12	0.27	0.14	0.32	0.38	0.28	0.46	0.38	0.37	0.22	
Chemistry	(1,604)	0.60	0.70	0.36	0.07	0.17	0.15	0.05	0.35	0.10	0.27	0.41	0.23	0.35	0.47	0.33	0.24	
Physics	(1,077)	0.58	0.67	0.32	0.10	0.20	0.19	0.12	0.30	0.07	0.29	0.39	0.18	0.29	0.42	0.24	0.25	0.22
French	(1,491)	0.54	0.48	0.27	0.01	0.05	0.40	0.03	0.09	-0.02	0.24	0.27	0.39	0.23	0.26	0.21	0.18	0.51
Spanish	(1,446)	0.45	0.40	0.30	0.03	0.11	0.34	-0.02	0.06	0.02	0.25	0.30	0.40	0.27	0.28	0.24	0.19	0.45
Europ. Hist.	(508)	0.75	0.51	0.29	0.12	0.04	0.23	0.00	0.16	0.16	0.35	0.22	0.20	0.33	0.26	0.42	0.18	
German	(424)	0.29	0.27	0.08	0.02	-0.01	0.42	-0.01	-0.07	-0.01	0.10	0.09	0.20	0.03	0.03	0.00	0.05	0.51
Hebrew	(40)	0.50	0.46	0.41	0.14	0.46	0.14	0.25	0.33	0.27	0.32	0.38	0.43	0.21	0.40	0.26	0.28	
Latin	(738)	0.56	0.53	0.37	0.02	0.06	0.29	-0.07	0.12	-0.08	0.24	0.34	0.39	0.29	0.34	0.28	0.26	0.26

June Administration

Test		Covariates																
		Var.1	Var.2	Var.3	Var.4	Var.5	Var.6	Var.7	Var.8	Var.9	Var.10	Var.11	Var.12	Var.13	Var.14	Var.15	Var.16	Var.17
Eng. Comp.	(68,542)	0.80	0.59	0.42	0.08	0.11	0.27	0.00	0.13	0.08	0.41	0.34	0.35	0.36	0.35	0.35	0.29	
Literature	(5,078)	0.84	0.55	0.42	0.13	0.09	0.27	-0.01	0.09	0.14	0.42	0.27	0.29	0.34	0.30	0.37	0.26	
Amer. Hist.	(23,146)	0.77	0.59	0.45	0.08	0.17	0.24	0.01	0.21	0.14	0.39	0.36	0.33	0.41	0.39	0.46	0.29	
Math I	(61,096)	0.51	0.84	0.46	0.03	0.27	0.13	-0.02	0.29	0.05	0.30	0.51	0.31	0.40	0.45	0.31	0.23	0.31
Math II	(19,242)	0.46	0.79	0.37	0.04	0.13	0.14	-0.02	0.24	0.04	0.21	0.44	0.26	0.30	0.40	0.24	0.20	0.19
Biology	(18,884)	0.67	0.64	0.41	0.05	0.17	0.17	0.05	0.17	0.09	0.31	0.40	0.28	0.44	0.40	0.33	0.16	
Chemistry	(23,971)	0.59	0.67	0.40	0.02	0.14	0.14	0.03	0.22	0.08	0.28	0.40	0.25	0.38	0.43	0.29	0.21	
Physics	(7,613)	0.57	0.68	0.30	0.02	0.11	0.12	-0.02	0.23	0.04	0.17	0.33	0.16	0.27	0.36	0.18	0.16	0.04
French	(8,101)	0.55	0.48	0.32	0.04	0.05	0.30	0.00	0.06	0.02	0.24	0.28	0.38	0.27	0.25	0.22	0.20	0.41
Spanish	(8,676)	0.51	0.42	0.32	0.04	0.09	0.31	0.01	0.08	0.07	0.27	0.27	0.39	0.28	0.27	0.26	0.20	0.40

1 Correlation coefficients computed for variables 18-21, which appear in table 8, are not presented in table 7.

2 Correlation coefficients were computed for all available cases with scores on the two variables of interest. Correlation matrices, which include sample size for each respective pairing, maybe found in Table I of the appendix.

Table 8

Summary of Stepwise Regression Procedure[1]
English Composition Test

Sophomore

	November					December					January					May					June			
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
																				2	Var.1 Var.2	0.7897	0.6237	-
																				3	Var.1 Var.2 Var.19	0.7986	0.6377	0.0140
																				4	Var.1 Var.2 Var.16 Var.19	0.8018	0.6429	0.0052
																				5	Var.1 Var.2 Var.6 Var.16 Var.19	0.8037	0.6459	0.0030

Junior

	November					December					January					May					June			
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.8630	0.7448	-	2	Var.1 Var.6	0.8027	0.6444	-	2	Var.1 Var.2	0.8169	0.6674	-	2	Var.1 Var.2	0.8173	0.6679	-	2	Var.1 Var.2	0.8048	0.6477	-
3	Var.1 Var.2 Var.6	0.8708	0.7583	0.0135	3	Var.1 Var.2 Var.6	0.8123	0.6598	0.0154	3	Var.1 Var.2 Var.21	0.8216	0.6751	0.0077	3	Var.1 Var.2 Var.12	0.8239	0.6788	0.0109	3	Var.1 Var.2 Var.12	0.8112	0.6580	0.0103
4	Var.1 Var.2 Var.5 Var.6	0.8748	0.7652	0.0069	4	Var.1 Var.2 Var.6 Var.20	0.8179	0.6690	0.0092	4	Var.1 Var.2 Var.7 Var.21	0.8255	0.6815	0.0064	4	Var.1 Var.2 Var.6 Var.10	0.8271	0.6841	0.0053	4	Var.1 Var.2 Var.6 Var.10	0.8137	0.6621	0.0041
5	Var.1 Var.2 Var.5 Var.6 Var.19	0.8772	0.7695	0.0043	5	Var.1 Var.2 Var.6 Var.10 Var.20	0.8202	0.6728	0.0038	5	Var.1 Var.2 Var.7 Var.10 Var.21	0.8284	0.6863	0.0048	5	Var.1 Var.2 Var.6 Var.8 Var.10	0.8287	0.6867	0.0026	5	Var.1 Var.2 Var.6 Var.8 Var.10	0.8154	0.6649	0.0028

Summary of Stepwise Regression Procedure
English Composition Test

(continued)

Senior

Step No.	November Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	December Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	January Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	May Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	June Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.8132	0.6613	-	2	Var.1 Var.12	0.7667	0.5879	-	2	Var.1 Var.2	0.7966	0.6346	-	2	Var.1 Var.2	0.7838	0.6143	-	2	Var.1 Var.10	0.7892	0.6229	-
3	Var.1 Var.2 Var.12	0.8173	0.6680	0.0067	3	Var.1 Var.2 Var.12	0.7714	0.5951	0.0072	3	Var.1 Var.2 Var.10	0.8032	0.6452	0.0106	3	Var.1 Var.2 Var.10	0.7950	0.6321	0.0178	3	Var.1 Var.2 Var.10	0.7976	0.6362	0.0133
4	Var.1 Var.2 Var.6 Var.10	0.8198	0.6721	0.0041	4	Var.1 Var.2 Var.6 Var.10	0.7765	0.6029	0.0078	4	Var.1 Var.2 Var.10 Var.12	0.8055	0.6489	0.0037	4	Var.1 Var.2 Var.10 Var.12	0.7989	0.6382	0.0061	4	Var.1 Var.2 Var.10 Var.12	0.8003	0.6405	0.0043
5	Var.1 Var.2 Var.6 Var.10 Var.16	0.8212	0.6743	0.0022	5	Var.1 Var.2 Var.6 Var.8 Var.10	0.7788	0.6065	0.0036	5	Var.1 Var.2 Var.8 Var.10 Var.12	0.8070	0.6513	0.0024	5	Var.1 Var.2 Var.10 Var.12 Var.16	0.8007	0.6412	0.0030	5	Var.1 Var.2 Var.10 Var.12 Var.16	0.8016	0.6425	0.0020

Total Group

Step No.	November Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	December Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	January Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	May Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	June Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.8133	0.6615	-	2	Var.1 Var.12	0.7671	0.5885	-	2	Var.1 Var.2	0.7979	0.6366	-	2	Var.1 Var.2	0.8169	0.6673	-	2	Var.1 Var.2	0.8062	0.6500	-
3	Var.1 Var.2 Var.12	0.8174	0.6682	0.0067	3	Var.1 Var.2 Var.12	0.7718	0.5957	0.0072	3	Var.1 Var.2 Var.10	0.8043	0.6469	0.0103	3	Var.1 Var.2 Var.12	0.8244	0.6797	0.0124	3	Var.1 Var.2 Var.12	0.8127	0.6605	0.0105
4	Var.1 Var.2 Var.6 Var.10	0.8200	0.6724	0.0042	4	Var.1 Var.2 Var.6 Var.10	0.7769	0.6036	0.0079	4	Var.1 Var.2 Var.10 Var.12	0.8066	0.6506	0.0037	4	Var.1 Var.2 Var.6 Var.10	0.8272	0.6843	0.0046	4	Var.1 Var.2 Var.6 Var.10	0.8152	0.6645	0.0040
5	Var.1 Var.2 Var.6 Var.10 Var.16	0.8213	0.6745	0.0021	5	Var.1 Var.2 Var.6 Var.8 Var.10	0.7792	0.6072	0.0036	5	Var.1 Var.2 Var.8 Var.10 Var.12	0.8080	0.6529	0.0023	5	Var.1 Var.2 Var.6 Var.10 Var.12	0.8290	0.6872	0.0029	5	Var.1 Var.2 Var.6 Var.8 Var.10	0.8168	0.6672	0.0027

Table 8 (continued)

Summary of Stepwise Regression Procedure[1]
Literature Test

Junior

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
															2	Var.1 Var.4	0.8475	0.7182	-	2	Var.1 Var.10	0.8307	0.6900	-
															3	Var.1 Var.2 Var.4	0.8504	0.7231	0.0049	3	Var.1 Var.10 Var.16	0.8322	0.6925	0.0025
															4	Var.1 Var.2 Var.4 Var.16	0.8522	0.7262	0.0031	4	Var.1 Var.6 Var.10 Var.16	0.8335	0.6947	0.0022
															5	Var.1 Var.2 Var.4 Var.6 Var.16	0.8531	0.7278	0.0016	5	Var.1 Var.6 Var.8 Var.10 Var.16	0.8344	0.6962	0.0015

Senior

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.10	0.8415	0.7082	-	2	Var.1 Var.10	0.8240	0.6790	-	2	Var.1 Var.10	0.8413	0.7078	-	2	Var.1 Var.10	0.8270	0.6839	-	2	Var.1 Var.3	0.8455	0.7149	-
3	Var.1 Var.10 Var.19	0.8425	0.7098	0.0016	3	Var.1 Var.3 Var.10	0.8252	0.6809	0.0019	3	Var.1 Var.2 Var.10	0.8422	0.7093	0.0015	3	Var.1 Var.6 Var.10	0.8292	0.6875	0.0036	3	Var.1 Var.3 Var.20	0.8478	0.7187	0.0038
4	Var.1 Var.3 Var.10 Var.19	0.8430	0.7107	0.0009	4	Var.1 Var.3 Var.8 Var.10	0.8260	0.6823	0.0014	4	Var.1 Var.2 Var.10 Var.16	0.8428	0.7103	0.0010	4	Var.1 Var.6 Var.10 Var.16	0.8308	0.6903	0.0028	4	Var.1 Var.3 Var.10 Var.20	0.8490	0.7208	0.0021
5	Var.1 Var.3 Var.8 Var.10 Var.19	0.8434	0.7114	0.0007	5	Var.1 Var.3 Var.8 Var.10 Var.19	0.8264	0.6830	0.0007	5	Var.1 Var.2 Var.10 Var.16 Var.19	0.8432	0.7110	0.0007	5	Var.1 Var.6 Var.10 Var.15 Var.16	0.8316	0.6915	0.0012	5	Var.1 Var.3 Var.10 Var.16 Var.20	0.8498	0.7221	0.0013

Total Group

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.10	0.8414	0.7080	-	2	Var.1 Var.10	0.8243	0.6795	-	2	Var.1 Var.10	0.8416	0.7083	-	2	Var.1 Var.10	0.8389	0.7037	-	2	Var.1 Var.10	0.8341	0.6958	-
3	Var.1 Var.10 Var.19	0.8423	0.7095	0.0015	3	Var.1 Var.3 Var.10	0.8253	0.6812	0.0017	3	Var.1 Var.2 Var.10	0.8424	0.7097	0.0014	3	Var.1 Var.10 Var.16	0.8408	0.7070	0.0033	3	Var.1 Var.3 Var.10	0.8357	0.6984	0.0026
4	Var.1 Var.3 Var.10 Var.19	0.8428	0.7103	0.0008	4	Var.1 Var.3 Var.8 Var.10	0.8262	0.6826	0.0014	4	Var.1 Var.2 Var.10 Var.16	0.8430	0.7107	0.0010	4	Var.1 Var.6 Var.10 Var.16	0.8421	0.7091	0.0021	4	Var.1 Var.3 Var.8 Var.10	0.8367	0.7000	0.0016
5	Var.1 Var.3 Var.8 Var.10 Var.19	0.8433	0.7111	0.0008	5	Var.1 Var.3 Var.8 Var.10 Var.19	0.8266	0.6833	0.0007	5	Var.1 Var.2 Var.10 Var.16 Var.19	0.8434	0.7114	0.0007	5	Var.1 Var.4 Var.6 Var.10 Var.16	0.8431	0.7109	0.0018	5	Var.1 Var.3 Var.8 Var.10 Var.16	0.8375	0.7014	0.0014

Table 8 (continued)

Summary of Stepwise Regression Procedure[1]
American History and Social Studies Test

Sophomore

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
																				2	Var.1 Var.19	0.7767	0.6033	
																				3	Var.1 Var.5 Var.19	0.7921	0.6275	0.0242
																				4	Var.1 Var.5 Var.12 Var.19	0.8023	0.6437	0.0162
																				5	Var.1 Var.4 Var.5 Var.15 Var.19	0.8092	0.6548	0.0111

Junior

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
															2	Var.1 Var.15	0.7761	0.6023	-	2	Var.1 Var.15	0.7788	0.6066	-
															3	Var.1 Var.2 Var.15	0.7901	0.6243	0.0220	3	Var.1 Var.2 Var.15	0.7922	0.6276	0.0210
															4	Var.1 Var.2 Var.9 Var.15	0.7939	0.6303	0.0060	4	Var.1 Var.2 Var.9 Var.15	0.7930	0.6288	0.0012
															5	Var.1 Var.2 Var.9 Var.15 Var.16	0.7950	0.6321	0.0018	5	Var.1 Var.2 Var.9 Var.15 Var.18	0.7935	0.6297	0.0009

Summary of Stepwise Regression Procedure
American History and Social Studies Test
(continued)

Senior

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.15	0.7750	0.6007	-	2	Var.1 Var.15	0.7642	0.5840	-	2	Var.1 Var.15	0.7462	0.5568	-	2	Var.1 Var.15	0.7663	0.5872	-	2	Var.1 Var.15	0.8021	0.6433	-
3	Var.1 Var.2 Var.15	0.7819	0.6114	0.0107	3	Var.1 Var.2 Var.15	0.7738	0.5988	0.0148	3	Var.1 Var.2 Var.15	0.7535	0.5678	0.0110	3	Var.1 Var.8 Var.15	0.7776	0.6047	0.0175	3	Var.1 Var.2 Var.15	0.8119	0.6592	0.0159
4	Var.1 Var.2 Var.10 Var.15	0.7833	0.6135	0.0021	4	Var.1 Var.2 Var.9 Var.15	0.7755	0.6014	0.0026	4	Var.1 Var.2 Var.8 Var.15	0.7549	0.5699	0.0021	4	Var.1 Var.2 Var.8 Var.15	0.7811	0.6101	0.0054	4	Var.1 Var.2 Var.14 Var.15	0.8144	0.6632	0.0040
5	Var.1 Var.2 Var.9 Var.10 Var.15	0.7842	0.6149	0.0014	5	Var.1 Var.2 Var.9 Var.12 Var.15	0.7761	0.6024	0.0010	5	Var.1 Var.2 Var.8 Var.9 Var.15	0.7560	0.5715	0.0016	5	Var.1 Var.2 Var.8 Var.13 Var.15	0.7832	0.6134	0.0033	5	Var.1 Var.2 Var.9 Var.14 Var.15	0.8160	0.6658	0.0026

Total Group

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.15	0.7746	0.6000	-	2	Var.1 Var.15	0.7644	0.5843	-	2	Var.1 Var.15	0.7460	0.5565	-	2	Var.1 Var.15	0.7965	0.6344	-	2	Var.1 Var.15	0.7828	0.6127	-
3	Var.1 Var.2 Var.15	0.7815	0.6108	0.0108	3	Var.1 Var.2 Var.15	0.7740	0.5990	0.0147	3	Var.1 Var.2 Var.15	0.7534	0.5676	0.0111	3	Var.1 Var.2 Var.15	0.8088	0.6542	0.0198	3	Var.1 Var.2 Var.15	0.7960	0.6336	0.0209
4	Var.1 Var.2 Var.10 Var.15	0.7828	0.6128	0.0020	4	Var.1 Var.2 Var.9 Var.15	0.7757	0.6017	0.0027	4	Var.1 Var.2 Var.8 Var.15	0.7548	0.5697	0.0021	4	Var.1 Var.2 Var.9 Var.15	0.8125	0.6601	0.0059	4	Var.1 Var.2 Var.9 Var.15	0.7967	0.6348	0.0012
5	Var.1 Var.2 Var.9 Var.10 Var.15	0.7837	0.6142	0.0014	5	Var.1 Var.2 Var.9 Var.10 Var.15	0.7764	0.6028	0.0011	5	Var.1 Var.2 Var.8 Var.10 Var.15	0.7558	0.5713	0.0016	5	Var.1 Var.2 Var.6 Var.9 Var.15	0.8152	0.6646	0.0045	5	Var.1 Var.2 Var.9 Var.15 Var.18	0.7974	0.6358	0.0010

Table 8 (continued)

Summary of Stepwise Regression Procedure[1]
Mathematics Level I Test

Sophomore

	November				December				January				May				June			
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
																				2	Var.2 Var.11	0.7949	0.6318	
																				3	Var.2 Var.11 Var.19	0.7992	0.6387	0.0069
																				4	Var.2 Var.17 Var.11 Var.21	0.8036	0.6458	0.0071
																				5	Var.2 Var.17 Var.8 Var.11 Var.21	0.8084	0.6535	0.0077

Junior

November					December					January				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.2 Var.17	0.8610	0.7414	-	2	Var.2 Var.11	0.8593	0.7384	-	2	Var.2 Var.8	0.8461	0.7159	-
3	Var.2 Var.17 Var.7	0.8671	0.7519	0.0105	3	Var.2 Var.17 Var.11	0.8631	0.7449	0.0065	3	Var.2 Var.17 Var.8	0.8537	0.7288	0.0129
4	Var.2 Var.17 Var.7 Var.19	0.8708	0.7583	0.0064	4	Var.2 Var.17 Var.11 Var.20	0.8668	0.7513	0.0064	4	Var.2 Var.17 Var.8 Var.12	0.8588	0.7376	0.0088
5	Var.2 Var.17 Var.7 Var.12 Var.15	0.8748	0.7653	0.0070	5	Var.2 Var.17 Var.11 Var.19 Var.20	0.8690	0.7551	0.0038	5	Var.2 Var.17 Var.8 Var.12 Var.19	0.8627	0.7442	0.0066

May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.2 Var.11	0.8521	0.7260	-	2	Var.2 Var.11	0.8449	0.7138	-
3	Var.2 Var.17 Var.11	0.8569	0.7343	0.0083	3	Var.1 Var.2 Var.11	0.8485	0.7200	0.0062
4	Var.1 Var.2 Var.17 Var.11	0.8594	0.7386	0.0043	4	Var.1 Var.2 Var.17 Var.11	0.8522	0.7262	0.0062
5	Var.1 Var.2 Var.17 Var.3 Var.11	0.8608	0.7410	0.0024	5	Var.1 Var.2 Var.17 Var.3 Var.11	0.855[illegible]	0.7281	0.0019

Summary of Stepwise Regression Procedure
Mathematics Level I Test
(continued)

Senior

Step No.	November Var. Ent.	November Mult. R	November R-Sq.	November Inc. R-Sq.	Step No.	December Var. Ent.	December Mult. R	December R-Sq.	December Inc. R-Sq.	Step No.	January Var. Ent.	January Mult. R	January R-Sq.	January Inc. R-Sq.	Step No.	May Var. Ent.	May Mult. R	May R-Sq.	May Inc. R-Sq.	Step No.	June Var. Ent.	June Mult. R	June R-Sq.	June Inc. R-Sq.
2	Var.2 Var.17	0.8456	0.7151	-	2	Var.2 Var.17	0.8269	0.6837	-	2	Var.2 Var.17	0.8142	0.6630	-	2	Var.2 Var.17	0.8146	0.6636	-	2	Var.2 Var.11	0.8306	0.6899	-
3	Var.2 Var.17 Var.11	0.8514	0.7249	0.0098	3	Var.2 Var.17 Var.11	0.8318	0.6919	0.0082	3	Var.2 Var.17 Var.11	0.8197	0.6719	0.0089	3	Var.2 Var.17 Var.11	0.8226	0.6766	0.0130	3	Var.2 Var.17 Var.11	0.8346	0.6966	0.0067
4	Var.2 Var.17 Var.8 Var.11	0.8536	0.7286	0.0037	4	Var.1 Var.2 Var.17 Var.11	0.8347	0.6967	0.0048	4	Var.1 Var.2 Var.17 Var.11	0.8223	0.6762	0.0043	4	Var.2 Var.17 Var.5 Var.11	0.8263	0.6827	0.0061	4	Var.2 Var.17 Var.5 Var.11	0.8368	0.7002	0.0036
5	Var.1 Var.2 Var.17 Var.8 Var.11	0.8554	0.7317	0.0031	5	Var.1 Var.2 Var.17 Var.8 Var.11	0.8360	0.6989	0.0022	5	Var.1 Var.2 Var.17 Var.5 Var.11	0.8243	0.6795	0.0033	5	Var.2 Var.17 Var.5 Var.8 Var.11	0.8278	0.6853	0.0026	5	Var.2 Var.17 Var.3 Var.5 Var.11	0.8384	0.7029	0.0027

Total Group

Step No.	November Var. Ent.	November Mult. R	November R-Sq.	November Inc. R-Sq.	Step No.	December Var. Ent.	December Mult. R	December R-Sq.	December Inc. R-Sq.	Step No.	January Var. Ent.	January Mult. R	January R-Sq.	January Inc. R-Sq.	Step No.	May Var. Ent.	May Mult. R	May R-Sq.	May Inc. R-Sq.	Step No.	June Var. Ent.	June Mult. R	June R-Sq.	June Inc. R-Sq.
2	Var.2 Var.17	0.8458	0.7154	-	2	Var.2 Var.17	0.8272	0.6842	-	2	Var.2 Var.17	0.8155	0.6650	-	2	Var.2 Var.11	0.8401	0.7057	-	2	Var.2 Var.11	0.8465	0.7165	-
3	Var.2 Var.17 Var.11	0.8516	0.7252	0.0098	3	Var.2 Var.17 Var.11	0.8321	0.6924	0.0082	3	Var.2 Var.17 Var.11	0.8210	0.6741	0.0091	3	Var.2 Var.17 Var.11	0.8455	0.7148	0.0091	3	Var.1 Var.2 Var.11	0.8498	0.7222	0.0057
4	Var.2 Var.17 Var.8 Var.11	0.8538	0.7289	0.0037	4	Var.1 Var.2 Var.17 Var.11	0.8350	0.6972	0.0048	4	Var.1 Var.2 Var.17 Var.11	0.8235	0.6781	0.0040	4	Var.2 Var.17 Var.5 Var.11	0.8482	0.7195	0.0047	4	Var.1 Var.2 Var.17 Var.11	0.8530	0.7276	0.0054
5	Var.1 Var.2 Var.17 Var.8 Var.11	0.8555	0.7319	0.0030	5	Var.1 Var.2 Var.17 Var.8 Var.11	0.8363	0.6994	0.0022	5	Var.1 Var.2 Var.17 Var.5 Var.11	0.8255	0.6815	0.0034	5	Var.1 Var.2 Var.17 Var.5 Var.11	0.8507	0.7237	0.0042	5	Var.1 Var.2 Var.17 Var.3 Var.11	0.8541	0.7295	0.0019

Table 8 (continued)

Summary of Stepwise Regression Procedure[1]
Mathematics Level II Test

Sophomore

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
																				2	Var.2 Var.8	0.6448	0.4158	-
																				3	Var.2 Var.17 Var.8	0.6614	0.4375	0.0217
																				4	Var.2 Var.17 Var.8 Var.14	0.6749	0.4555	0.0180
																				5	Var.2 Var.17 Var.6 Var.8 Var.14	0.6828	0.4662	0.0107

Junior

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
					2	Var.2 Var.11	0.8258	0.6819	-	2	Var.2 Var.17	0.8205	0.6733	-	2	Var.2 Var.11	0.8007	0.6412	-	2	Var.2 Var.11	0.7880	0.6210	-
					3	Var.2 Var.17 Var.11	0.8465	0.7166	0.0347	3	Var.2 Var.17 Var.6	0.8249	0.6805	0.0072	3	Var.2 Var.17 Var.11	0.8076	0.6522	0.0110	3	Var.1 Var.2 Var.11	0.7930	0.6289	0.0079
					4	Var.2 Var.17 Var.11 Var.20	0.8545	0.7301	0.0135	4	Var.2 Var.17 Var.10 Var.16	0.8284	0.6863	0.0058	4	Var.2 Var.17 Var.11 Var.14	0.8107	0.6573	0.0051	4	Var.1 Var.2 Var.17 Var.11	0.7955	0.6328	0.0039
					5	Var.2 Var.17 Var.11 Var.12 Var.16	0.8605	0.7404	0.0103	5	Var.2 Var.17 Var.6 Var.10 Var.16	0.3326	0.6932	0.0069	5	Var.1 Var.2 Var.17 Var.11 Var.14	0.8133	0.6614	0.0041	5	Var.1 Var.2 Var.17 Var.11 Var.14	0.7966	0.6346	0.0018

Summary of Stepwise Regression Procedure
Mathematics Level II Test
(continued)

Senior

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
	November					December					January					May				June			
2	Var.2 Var.17	0.7875	0.5202	-	2	Var.2 Var.17	0.7847	0.6158	-	2	Var.2 Var.17	0.8070	0.6512	-	2	Var.2 Var.17	0.8191	0.6710	-	Var.2 Var.11	0.7899	0.6240	-
3	Var.2 Var.17 Var.11	0.7930	0.6289	0.0087	3	Var.2 Var.17 Var.11	0.7901	0.6242	0.0084	3	Var.1 Var.2 Var.17	0.8114	0.6583	0.0071	3	Var.1 Var.2 Var.17	0.8230	0.6773	0.0063	Var.2 Var.17 Var.11	0.7949	0.6318	0.0078
4	Var.1 Var.2 Var.17 Var.11	0.7950	0.6320	0.0031	4	Var.1 Var.2 Var.17 Var.11	0.7929	0.6287	0.0045	4	Var.1 Var.2 Var.17 Var.11	0.8150	0.6642	0.0059	4	Var.1 Var.2 Var.17 Var.11	0.8258	0.6820	0.0047	Var.2 Var.17 Var.11 Var.14	0.7977	0.6364	0.0046
5	Var.1 Var.2 Var.17 Var.11 Var.14	0.7959	0.6334	0.0014	5	Var.1 Var.2 Var.17 Var.8 Var.11	0.7942	0.6307	0.0020	5	Var.1 Var.2 Var.17 Var.8 Var.11	0.8161	0.6660	0.0018	5	Var.1 Var.2 Var.17 Var.8 Var.11	0.8270	0.6839	0.0019	Var.2 Var.17 Var.5 Var.11 Var.14	0.8004	0.6406	0.0042

Total Group

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
	November					December					January					May				June			
2	Var.2 Var.17	0.7880	0.6209	-	2	Var.2 Var.17	0.7850	0.6162	-	2	Var.2 Var.17	0.8075	0.6520	-	2	Var.2 Var.11	0.8485	0.7200	-	Var.2 Var.11	0.7987	0.6380	-
3	Var.2 Var.17 Var.11	0.7934	0.6295	0.0086	3	Var.2 Var.17 Var.11	0.7904	0.6247	0.0085	3	Var.1 Var.2 Var.17	0.8117	0.6589	0.0069	3	Var.1 Var.2 Var.11	0.8521	0.7260	0.0060	Var.1 Var.2 Var.11	0.8032	0.6451	0.0071
4	Var.1 Var.2 Var.17 Var.11	0.7954	0.6326	0.0031	4	Var.1 Var.2 Var.17 Var.11	0.7933	0.6293	0.0046	4	Var.1 Var.2 Var.17 Var.11	0.8153	0.6647	0.0058	4	Var.1 Var.2 Var.17 Var.11	0.8555	0.7319	0.0059	Var.1 Var.2 Var.17 Var.11	0.8052	0.6483	0.0032
5	Var.1 Var.2 Var.17 Var.11 Var 14	0.7962	0.6340	0.0014	5	Var.1 Var.2 Var.17 Var.8 Var.11	0.7945	0.6313	0.0020	5	Var.1 Var.2 Var.17 Var.8 Var.11	0.8164	0.6665	0.0018	5	Var.1 Var.2 Var.17 Var.6 Var.11	0.8569	0.7342	0.0023	Var.1 Var.2 Var.17 Var.11 Var.14	0.8064	0.6503	0.0020

Table 8 (continued)

Summary of Stepwise Regression Procedure[1]
Biology Test

Freshmen

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
																				2	Var.18 Var.19	0.6297	0.3965	
																				3	Var.3 Var.18 Var.19	0.6904	0.4766	0.0801
																				4	Var.3 Var.18 Var.19 Var.20	0.7258	0.5268	0.0502
																				5	Var.3 Var.16 Var.18 Var.19 Var.20	0.7517	0.5651	0.0383

Sophomore

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
																				2	Var.1 Var.2	0.7566	0.5724	
																				3	Var.1 Var.2 Var.13	0.7677	0.5893	0.0169
																				4	Var.1 Var.2 Var.13 Var.16	0.7725	0.5967	0.0074
																				5	Var.1 Var.2 Var.13 Var.16 Var.19	0.7769	0.6035	..0068

Junior

November					December					January					May					June				
Step No	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
					2	Var.1 Var.2	0.6901	0.4763	-	2	Var.1 Var.21	0.8392	0.7043	-	2	Var.1 Var.2	0.8070	0.6513	-	2	Var.1 Var.2	0.7430	0.5520	-
					3	Var.1 Var.2 Var 9	0.7179	0.5154	0.0391	3	Var.1 Var.20 Var.21	0.8463	0.7163	0.0120	3	Var.1 Var.2 Var.13	0.8179	0.6690	0.0177	3	Var.1 Var.2 Var.13	0.7552	0.5704	0.0184
					4	Var.1 Var.2 Var 9 Var 11	0.7414	0.5496	0.0342	4	Var.1 Var.3 Var.20 Var.21	0.8539	0.7291	0.0128	4	Var.1 Var.2 Var.7 Var.13	0.8223	0.6761	0.0071	4	Var.1 Var.2 Var.7 Var.13	0.7577	0.5741	0.0037
					5	Var.1 Var 2 Var.6 Var.9 Var.11	0.7601	0.5777	0.0281	5	Var.1 Var.3 Var.10 Var.20 Var.21	0.8606	0.7407	0.0116	5	Var.1 Var.2 Var.7 Var.8 Var.13	0.8250	0.6806	0.0045	5	Var.1 Var.2 Var.7 Var.13 Var.14	0.7593	0.5765	0.0024

Summary of Stepwise Regression Procedure
Biology Test

(continued)

Senior

November

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7854	0.6168	-
3	Var.1 Var.2 Var.13	0.7929	0.6287	0.0119
4	Var.1 Var.2 Var.7 Var.13	0.7983	0.6373	0.0086
5	Var.1 Var.2 Var.7 Var.13 Var.14	0.7998	0.6397	0.0024

December

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7622	0.5810	-
3	Var.1 Var.2 Var.7	0.7742	0.5994	0.0184
4	Var.1 Var.2 Var.7 Var.13	0.7772	0.6040	0.0046
5	Var.1 Var.2 Var.7 Var.12 Var.13	0.7779	0.6052	0.0012

January

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7545	0.5693	-
3	Var.1 Var.2 Var.7	0.7685	0.5906	0.0213
4	Var.1 Var.2 Var.7 Var.13	0.7740	0.5990	0.0084
5	Var.1 Var.2 Var.7 Var.10 Var.13	0.7749	0.6005	0.0015

May

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7426	0.5514	-
3	Var.1 Var.2 Var.7	0.7577	0.5741	0.0227
4	Var.1 Var.2 Var.7 Var.13	0.7652	0.5855	0.0114
5	Var.1 Var.2 Var.7 Var.10 Var.13	0.7703	0.5934	0.0079

June

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7894	0.6231	-
3	Var.1 Var.2 Var.7	0.7986	0.6377	0.0146
4	Var.1 Var.2 Var.7 Var.11	0.8029	0.6447	0.0070
5	Var.1 Var.2 Var.7 Var.11 Var.20	0.8050	0.6480	0.0033

Total Group

November

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7859	0.6176	-
3	Var.1 Var.2 Var.13	0.7933	0.6293	0.0117
4	Var.1 Var.2 Var.7 Var.13	0.7986	0.6377	0.0084
5	Var.1 Var.2 Var.7 Var.13 Var.14	0.8001	0.6401	0.0024

December

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7614	0.5797	-
3	Var.1 Var.2 Var.7	0.7734	0.5982	0.0185
4	Var.1 Var.2 Var.7 Var.13	0.7764	0.6028	0.0046
5	Var.1 Var.2 Var.7 Var.12 Var.13	0.7772	0.6040	0.0012

January

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7564	0.5722	-
3	Var.1 Var.2 Var.7	0.7704	0.5935	0.0213
4	Var.1 Var.2 Var.7 Var.13	0.7757	0.6017	0.0082
5	Var.1 Var.2 Var.7 Var.10 Var.13	0.7768	0.6034	0.0017

May

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7981	0.6369	-
3	Var.1 Var.2 Var.13	0.8073	0.6518	0.0149
4	Var.1 Var.2 Var.7 Var.13	0.8142	0.6629	0.0111
5	Var.1 Var.2 Var.7 Var.8 Var.13	0.8165	0.6667	0.0038

June

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7483	0.5600	-
3	Var.1 Var.2 Var.13	0.7606	0.5785	0.0185
4	Var.1 Var.2 Var.7 Var.13	0.7628	0.5819	0.0034
5	Var.1 Var.2 Var.7 Var.13 Var.14	0.7647	0.5847	0.0028

Table 8 (continued)

Summary of Stepwise Regression Procedure[1]
Chemistry Test

Sophomore

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
																				2	Var.1 Var.2	0.6017	0.3620	-
																				3	Var.1 Var.2 Var.8	0.6120	0.3746	0.0126
																				4	Var.1 Var.2 Var.4 Var.8	0.6239	0.3892	0.0146
																				5	Var.1 Var.2 Var.4 Var.8 Var.11	0.6272	0.3934	0.0042

Junior

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
					2	Var.2 Var.10	0.8356	0.6982	-	2	Var.2 Var.5	0.6969	0.4856	-	2	Var.1 Var.2	0.7469	0.5579	-	2	Var.1 Var.2	0.7288	0.5312	-
					3	Var.1 Var.2 Var.13	0.8455	0.7149	0.0167	3	Var.2 Var.5 Var.21	0.7420	0.5505	0.0649	3	Var.[illegible] Var.[illegible] Var.8	0.7622	0.5810	0.0231	3	Var.1 Var.2 Var.14	0.7403	0.5480	0.0168
					4	Var.1 Var.2 Var.7 Var.13	0.8529	0.7275	0.0126	4	Var.2 Var.5 Var.12 Var.21	0.7712	0.5947	0.0442	4	Var.1 Var.2 Var.8 Var.14	0.7713	0.5949	0.0139	4	Var.1 Var.2 Var.8 Var.14	0.7463	0.5570	0.0090
					5	Var.1 Var.2 Var.7 Var.9 Var.13	0.8570	0.7344	0.0069	5	Var.2 Var.4 Var.5 Var.12 Var.21	0.8081	0.6531	0.0584	5	Var.1 Var.2 Var.8 Var.10 Var.14	0.7736	0.5984	0.0035	5	Var.1 Var.2 Var.8 Var.13 Var.14	0.7474	0.5586	0.0016

Summary of Stepwise Regression Procedure
Chemistry Test

(continued)

Senior

November

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.6869	0.4718	-
3	Var.1 Var.2 Var.8	0.7063	0.4988	0.0270
4	Var.1 Var.2 Var.8 Var.14	0.7165	0.5134	0.0146
5	Var.1 Var.2 Var.8 Var.14 Var.18	0.7189	0.5168	0.0034

December

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7270	0.5285	-
3	Var.1 Var.2 Var.8	0.7432	0.5523	0.0238
4	Var.1 Var.2 Var.3 Var.14	0.7499	0.5624	0.0101
5	Var.1 Var.2 Var.7 Var.8 Var.14	0.7517	0.5651	0.0027

January

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.6934	0.4808	-
3	Var.1 Var.2 Var.8	0.7178	0.5153	0.0345
4	Var.1 Var.2 Var.8 Var.14	0.7243	0.5246	0.0093
5	Var.1 Var.2 Var.7 Var.8 Var.14	0.7265	0.5278	0.0032

May

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7153	0.5116	-
3	Var.1 Var.2 Var.8	0.7469	0.5578	0.0462
4	Var.1 Var.2 Var.8 Var.18	0.7556	0.5710	0.0132
5	Var.1 Var.2 Var.8 Var.14 Var.18	0.7617	0.5802	0.0092

June

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7639	0.5836	-
3	Var.1 Var.2 Var.8	0.7892	0.6228	0.0392
4	Var.1 Var.2 Var.4 Var.8	0.7968	0.6349	0.0121
5	Var.1 Var.2 Var.4 Var.8 Var.13	0.8034	0.6454	0.0105

Total Group

November

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.6866	0.4714	-
3	Var.1 Var.2 Var.8	0.7054	0.4976	0.0262
4	Var.1 Var.2 Var.8 Var.14	0.7158	0.5123	0.0147
5	Var.1 Var.2 Var.8 Var.14 Var.18	0.7183	0.5159	0.0036

December

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7272	0.5288	-
3	Var.1 Var.2 Var.8	0.7434	0.5526	0.0238
4	Var.1 Var.2 Var.8 Var.14	0.7498	0.5622	0.0096
5	Var.1 Var.2 Var.7 Var.8 Var.14	0.7519	0.5653	0.0031

January

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.6946	0.4824	-
3	Var.1 Var.2 Var.8	0.7185	0.5162	0.0338
4	Var.1 Var.2 Var.8 Var.14	0.7246	0.5250	0.0088
5	Var.1 Var.2 Var.7 Var.8 Var.14	0.7270	0.5285	0.0035

May

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7467	0.5575	-
3	Var.1 Var.2 Var.8	0.7650	0.5853	0.0278
4	Var.1 Var.2 Var.8 Var.14	0.7740	0.5991	0.0138
5	Var.1 Var.2 Var.8 Var.10 Var.14	0.7767	0.6033	0.0042

June

Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.7276	0.5294	-
3	Var.1 Var.2 Var.14	0.7388	0.5458	0.0164
4	Var.1 Var.2 Var.8 Var.14	0.7450	0.5550	0.0092
5	Var.1 Var.2 Var.8 Var.13 Var.14	0.7462	0.5568	0.0018

Table 8 (continued)

Summary of Stepwise Regression Procedure[1]
Physics Test

Junior

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
															2	Var.1 Var.2	0.7217	0.5209	-	2	Var.1 Var.2	0.7197	0.5179	-
															3	Var.1 Var.2 Var.8	0.7337	0.5383	0.0174	3	Var.1 Var.2 Var.8	0.7274	0.5291	0.0112
															4	Var.1 Var.2 Var.12 Var.14	0.7483	0.5599	0.0216	4	Var.1 Var.2 Var.8 Var.14	0.7303	0.5333	0.0042
															5	Var.1 Var.2 Var.17 Var.12 Var.14	0.7543	0.5689	0.0090	5	Var.1 Var.2 Var.8 Var.10 Var.14	0.7357	0.5412	0.0079

Senior

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.6583	0.4334	-	2	Var.1 Var.2	0.6943	0.4820	-	2	Var.1 Var.2	0.6983	0.4876	-	2	Var.1 Var.2	0.6522	0.4254	-	2	Var.1 Var.2	0.7218	0.5210	-
3	Var.1 Var.2 Var.17	0.6839	0.4677	0.0343	3	Var.1 Var.2 Var.17	0.7308	0.5340	0.0520	3	Var.1 Var.2 Var.17	0.7285	0.5307	0.0431	3	Var.1 Var.2 Var.17	0.6872	0.4723	0.0469	3	Var.1 Var.2 Var.17	0.7348	0.5399	0.0189
4	Var.1 Var.2 Var.17 Var.8	0.6946	0.4824	0.0147	4	Var.1 Var.2 Var.17 Var.8	0.7373	0.5436	0.0096	4	Var.1 Var.2 Var.17 Var.8	0.7348	0.5399	0.0092	4	Var.1 Var.2 Var.17 Var.8	0.6998	0.4897	0.0174	4	Var.1 Var.2 Var.17 Var.15	0.7393	0.5466	0.0067
5	Var.1 Var.2 Var.17 Var.8 Var.18	0.6975	0.4865	0.0041	5	Var.1 Var.2 Var.17 Var.8 Var.10	0.7393	0.5465	0.0029	5	Var.1 Var.2 Var.17 Var.8 Var.10	0.7372	0.5435	0.0036	5	Var.1 Var.2 Var.17 Var.8 Var.14	0.7068	0.4995	0.0098	5	Var.1 Var.2 Var.17 Var.14 Var.15	0.7448	0.5548	0.0082

Total Group

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.6586	0.4338	-	2	Var.1 Var.2	0.6923	0.4793	-	2	Var.1 Var.2	0.6986	0.4881	-	2	Var.1 Var.2	0.7118	0.5066	-	2	Var.1 Var.2	0.7279	0.5299	-
3	Var.1 Var.2 Var.17	0.6843	0.4683	0.0345	3	Var.1 Var.2 Var.17	0.7301	0.5330	0.0537	3	Var.1 Var.2 Var.17	0.7305	0.5337	0.0456	3	Var.1 Var.2 Var.8	0.7342	0.5390	0.0324	3	Var.1 Var.2 Var.8	0.7357	0.5412	0.0113
4	Var.1 Var.2 Var.17 Var.8	0.6952	0.4833	0.0150	4	Var.1 Var.2 Var.17 Var.8	0.7367	0.5428	0.0098	4	Var.1 Var.2 Var.17 Var.8	0.7372	0.5434	0.0097	4	Var.1 Var.2 Var.17 Var.8	0.7428	0.5518	0.0128	4	Var.1 Var.2 Var.8 Var.14	0.7384	0.5453	0.0041
5	Var.1 Var.2 Var.17 Var.8 Var.18	0.6979	0.4871	0.0038	5	Var.1 Var.2 Var.17 Var.8 Var.10	0.7388	0.5458	0.0030	5	Var.1 Var.2 Var.17 Var.8 Var.10	0.7398	0.5473	0.0039	5	Var.1 Var.2 Var.17 Var.8 Var.14	0.7508	0.5637	0.0119	5	Var.1 Var.2 Var.8 Var.10 Var.14	0.7436	0.5530	0.0077

Table 8 (continued)

Summary of Stepwise Regression Procedure[1]
French Test

Sophomore

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
																				2	Var.2 Var.17	0.7246	0.5251	
																				3	Var.2 Var.17 Var.14	0.7695	0.5921	0.0670
																				4	Var.2 Var.17 Var.4 Var.14	0.8104	0.6567	0.0646
																				5	Var.2 Var.17 Var.4 Var.14 Var.21	0.8392	0.7042	0.0475

Junior

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
					2	Var.17 Var.12	0.5212	0.2717	-	2	Var.1 Var.17	0.7309	0.5342	-	2	Var.1 Var.17	0.6827	0.4661	-	2	Var.1 Var.17	0.6375	0.4064	-
					3	Var.17 Var.12 Var.18	0.5649	0.3191	0.0474	3	Var.1 Var.17 Var.9	0.7523	0.5660	0.0318	3	Var.1 Var.17 Var.12	0.7177	0.5151	0.0490	3	Var.1 Var.17 Var.12	0.6642	0.4411	0.0347
					4	Var.17 Var.12 Var.15 Var.18	0.5885	0.3463	0.0272	4	Var.1 Var.17 Var.9 Var.21	0.7701	0.5930	0.0270	4	Var.1 Var.2 Var.17 Var.12	0.7275	0.5292	0.0141	4	Var.1 Var.2 Var.17 Var.12	0.6790	0.4611	0.0200
					5	Var.17 Var.12 Var.13 Var.15 Var.18	0.6102	0.3723	0.0260	5	Var.1 Var.17 Var.9 Var.19 Var.21	0.7836	0.6140	0.0210	5	Var.1 Var.2 Var.17 Var.12 Var.15	0.7363	0.5422	0.0130	5	Var.1 Var.2 Var.17 Var.6 Var.12	0.6842	0.4681	0.0070

Summary of Stepwise Regression Procedure
French Test
(continued)

Senior

Step No.	November Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	December Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	January Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	May Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	June Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.17	0.6920	0.4789	-	2	Var.1 Var.17	0.6259	0.3918	-	2	Var.1 Var.17	0.6048	0.3658	-	2	Var.1 Var.17	0.7087	0.5022	-	2	Var.1 Var.17	0.7009	0.4912	-
3	Var.1 Var.2 Var.17	0.7035	0.4949	0.0160	3	Var.1 Var.17 Var.12	0.6455	0.4167	0.0249	3	Var.1 Var.17 Var.12	0.6219	0.3867	0.0209	3	Var.1 Var.17 Var.16	0.7249	0.5255	0.0233	3	Var.1 Var.17 Var.12	0.7206	0.5192	0.0280
4	Var.1 Var.2 Var.17 Var.6	0.7108	0.5053	0.0104	4	Var.1 Var.2 Var.17 Var.12	0.6568	0.4314	0.0147	4	Var.1 Var.2 Var.17 Var.12	0.6328	0.4004	0.0137	4	Var.1 Var.17 Var.6 Var.16	0.7411	0.5493	0.0238	4	Var.1 Var.17 Var.6 Var.12	0.7320	0.5358	0.0166
5	Var.1 Var.2 Var.17 Var.6 Var.12	0.7163	0.5131	0.0078	5	Var.1 Var.2 Var.17 Var.6 Var.12	0.6648	0.4420	0.0106	5	Var.1 Var.2 Var.17 Var.6 Var.12	0.6433	0.4138	0.0134	5	Var.1 Var.17 Var.6 Var.12 Var.16	0.7466	0.5574	0.0081	5	Var.1 Var.17 Var.6 Var.9 Var.12	0.7383	0.5451	0.0093

Total Group

Step No.	November Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	December Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	January Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	May Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	June Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.17	0.6893	0.4751	-	2	Var.1 Var.17	0.6240	0.3894	-	2	Var.1 Var.17	0.6077	0.3693	-	2	Var.1 Var.17	0.6899	0.4760	-	2	Var.1 Var.17	0.6351	0.4034	-
3	Var.1 Var.2 Var.17	0.7020	0.4928	0.0177	3	Var.1 Var.17 Var.12	0.6437	0.4144	0.0250	3	Var.1 Var.2 Var.17	0.6232	0.3884	0.0191	3	Var.1 Var.17 Var.6	0.7126	0.5078	0.0318	3	Var.1 Var.17 Var.12	0.6611	0.4370	0.0336
4	Var.1 Var.2 Var.17 Var.6	0.7094	0.5033	0.0105	4	Var.1 Var.2 Var.17 Var.12	0.6553	0.4294	0.0150	4	Var.1 Var.2 Var.17 Var.6	0.6354	0.4037	0.0153	4	Var.1 Var.17 Var.6 Var.12	0.7292	0.5318	0.0240	4	Var.1 Var.2 Var.17 Var.12	0.6745	0.4549	0.0179
5	Var.1 Var.2 Var.17 Var.6 Var.12	0.7151	0.5113	0.0080	5	Var.1 Var.2 Var.17 Var.6 Var.12	0.6635	0.4402	0.0108	5	Var.1 Var.2 Var.17 Var.6 Var.12	0.6441	0.4149	0.0112	5	Var.1 Var.2 Var.17 Var.6 Var.12	0.7421	0.5507	0.0189	5	Var.1 Var.2 Var.17 Var.6 Var.12	0.6824	0.4657	0.0108

Table 8 (continued)

Summary of Stepwise Regression Procedure[1]
Spanish Test

Sophomore

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
																				2	Var.2 Var.12	0.6597	0.4352	
																				3	Var.2 Var.12 Var.13	0.7614	0.5797	0.1445
																				4	Var.2 Var.17 Var.12 Var.13	0.7915	0.6265	0.0468
																				5	Var.2 Var.17 Var.12 Var.13 Var.16	0.8058	0.6493	0.0228

Junior

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.17 Var.15	0.9885	0.9772	-	2	Var.1 Var.17	0.7453	0.5554	-	2	Var.7 Var.16	0.5212	0.2716	-	2	Var.1 Var.17	0.6325	0.4001	-	2	Var.1 Var.17	0.6061	0.3673	-
3	Var.17 Var.7 Var.15	0.9970	0.9940	0.0168	3	Var.1 Var.17 Var.4	0.7792	0.6072	0.0518	3	Var.7 Var.12 Var.16	0.6041	0.3649	0.0933	3	Var.1 Var.2 Var.17	0.6628	0.4393	0.0392	3	Var.1 Var.17 Var.12	0.6448	0.4158	0.0485
4	Var.17 Var.7 Var.9 Var.15	0.9999	0.9998	0.0058	4	Var.1 Var.17 Var.4 Var.5	0.8018	0.6429	0.0357	4	Var.7 Var.12 Var.15 Var.16	0.6763	0.4574	0.0925	4	Var.1 Var.2 Var.17 Var.12	0.6837	0.4674	0.0281	4	Var.1 Var.2 Var.17 Var.12	0.6538	0.4274	0.0116
5	Var.17 Var.3 Var.7 Var.9 Var.15	1.0000	1.0000	0.0002	5	Var.1 Var.17 Var.4 Var.12 Var.15	0.8204	0.6730	0.0301	5	Var.1 Var.7 Var.12 Var.15 Var.16	0.6910	0.4775	0.0201	5	Var.1 Var.2 Var.17 Var.10 Var.12	0.6871	0.4721	0.0047	5	Var.1 Var.2 Var.17 Var.6 Var.12	0.6581	0.4331	0.0057

Summary of Stepwise Regression Procedure
Spanish Test
(continued)

Senior

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.17	0.6036	0.3643	-	2	Var.1 Var.17	0.5517	0.3044	-	2	Var.17 Var.12	0.4732	0.2239	-	2	Var.17 Var.12	0.5283	0.2791		2	Var.1 Var.17	0.6522	0.4253	-
3	Var.1 Var.17 Var.12	0.6171	0.3808	0.0165	3	Var.1 Var.17 Var.12	0.5743	0.3298	0.0254	3	Var.1 Var.17 Var.12	0.4907	0.2408	0.0169	3	Var.1 Var.17 Var.12	0.5524	0.3051	0.0260	3	Var.1 Var.17 Var.12	0.6681	0.4464	0.0211
4	Var.1 Var.2 Var.17 Var.12	0.6204	0.3849	0.0041	4	Var.1 Var.17 Var.12 Var.21	0.5804	0.3369	0.0071	4	Var.1 Var.17 Var.12 Var.21	0.5093	0.2594	0.0186	4	Var.1 Var.17 Var.12 Var.19	0.5777	0.3337	0.0286	4	Var.1 Var.17 Var.5 Var.12	0.6725	0.4522	0.0058
5	Var.1 Var.2 Var.17 Var.12 Var.21	0.6241	0.3895	0.0046	5	Var.1 Var.2 Var.17 Var.12 Var.21	0.5855	0.3428	0.0059	5	Var.1 Var.17 Var.12 Var.16 Var.21	0.5163	0.2666	0.0072	5	Var.1 Var.17 Var.8 Var.12 Var.21	0.5863	0.3437	0.0100	5	Var.1 Var.17 Var.5 Var.12 Var.14	0.6771	0.4584	0.0062

Total Group

November					December					January					May					June				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.17	0.6037	0.3644	-	2	Var.1 Var.17	0.5523	0.3050	-	2	Var.17 Var.12	0.4703	0.2212	-	2	Var.1 Var.17	0.5785	0.3347	-	2	Var.1 Var.17	0.6053	0.3664	-
3	Var.1 Var.17 Var.12	0.6173	0.3811	0.0167	3	Var.1 Var.17 Var.12	0.5749	0.3305	0.0255	3	Var.1 Var.17 Var.12	0.4890	0.2391	0.0179	3	Var.1 Var.17 Var.12	0.6111	0.3735	0.0388	3	Var.1 Var.17 Var.12	0.6426	0.4129	0.0465
4	Var.1 Var.2 Var.17 Var.12	0.6206	0.3852	0.0041	4	Var.1 Var.17 Var.12 Var.21	0.5811	0.3377	0.0072	4	Var.1 Var.17 Var.12 Var.21	0.5070	0.2570	0.0179	4	Var.1 Var.2 Var.17 Var.12	0.6232	0.3884	0.0149	4	Var.1 Var.2 Var.17 Var.12	0.6507	0.4234	0.0105
5	Var.1 Var.2 Var.17 Var.12 Var.21	0.6244	0.3899	0.0047	5	Var.1 Var.2 Var.17 Var.12 Var.21	0.5863	0.3437	0.0060	5	Var.1 Var.17 Var.12 Var.16 Var.21	0.5129	0.2631	0.0061	5	Var.1 Var.2 Var.17 Var.6 Var.12	0.6327	0.4003	0.0119	5	Var.1 Var.2 Var.17 Var.6 Var.12	0.6567	0.4312	0.0078

Table 8 (continued)

Summary of Stepwise Regression Procedure[1]
European History and World Cultures Test

Junior

December: Step No.	December: Var. Ent.	December: Mult. R	December: R-Sq.	December: Inc. R-Sq.	May: Step No.	May: Var. Ent.	May: Mult. R	May: R-Sq.	May: Inc. R-Sq.
					2	Var.1 Var.2	0.7174	0.5147	-
					3	Var.1 Var.2 Var.15	0.7363	0.5422	0.0275
					4	Var.1 Var.2 Var.11 Var.15	0.7453	0.5554	0.0132
					5	Var.1 Var.[illegible] Var.[illegible] Var.11 Var.15	0.7517	0.5651	0.0097

Senior

December: Step No.	December: Var. Ent.	December: Mult. R	December: R-Sq.	December: Inc. R-Sq.	May: Step No.	May: Var. Ent.	May: Mult. R	May: R-Sq.	May: Inc. R-Sq.
2	Var.1 Var.2	0.7014	0.4919	-	2	Var.1 Var.7	0.8119	0.6592	-
3	Var.1 Var.2 Var.10	0.7048	0.4968	0.0049	3	Var.1 Var.7 Var.9	0.8210	0.6741	0.0149
4	Var.1 Var.2 Var.10 Var.15	0.7124	0.5075	0.0107	4	Var.1 Var.2 Var.7 Var.9	0.8279	0.6854	0.0113
5	Var.1 Var.2 Var.10 Var.13 Var.15	0.7141	0.5099	0.0024	5	Var.1 Var.2 Var.7 Var.9 Var.16	0.8300	0.6889	0.0035

Total Group

December: Step No.	December: Var. Ent.	December: Mult. R	December: R-Sq.	December: Inc. R-Sq.	May: Step No.	May: Var. Ent.	May: Mult. R	May: R-Sq.	May: Inc. R-Sq.
2	Var.1 Var.2	0.7000	0.4900	-	2	Var.1 Var.2	0.7564	0.5722	-
3	Var.1 Var.2 Var.10	0.7035	0.4949	0.0049	3	Var.1 Var.2 Var.15	0.7681	0.5900	0.0178
4	Var.1 Var.2 Var.10 Var.15	0.7108	0.5053	0.0104	4	Var.1 Var.2 Var.9 Var.15	0.7734	0.5982	0.0082
5	Var.1 Var.2 Var.9 Var.10 Var.15	0.7125	0.5077	0.0024	5	Var.1 Var.2 Var.9 Var.11 Var.15	0.7765	0.6030	0.0048

Table 8 (continued)

Summary of Stepwise Regression Procedure[1]
German Test

Junior

December Step No.	December Var. Ent.	December Mult. R	December R-Sq.	December Inc. R-Sq.	May Step No.	May Var. Ent.	May Mult. R	May R-Sq.	May Inc. R-Sq.
					2	Var.2 Var.17	0.5120	0.2621	-
					3	Var.2 Var.17 Var.7	0.5383	0.2898	0.0277
					4	Var.2 Var.17 Var.7 Var.21	0.5536	0.3065	0.0167
					5	Var.2 Var.17 Var.3 Var.7 Var.14	0.5767	0.3326	0.0261

Senior

December Step No.	December Var. Ent.	December Mult. R	December R-Sq.	December Inc. R-Sq.	May Step No.	May Var. Ent.	May Mult. R	May R-Sq.	May Inc. R-Sq.
2	Var.1 Var.17	0.4953	0.2453	-	2	Var.2 Var.17	0.6177	0.3815	-
3	Var.1 Var.17 Var.6	0.5274	0.2782	0.0329	3	Var.2 Var.17 Var.9	0.6406	0.4104	0.0289
4	Var.1 Var.17 Var.6 Var.12	0.5490	0.3014	0.0232	4	Var.1 Var.2 Var.17 Var.8	0.6589	0.4341	0.0237
5	Var.1 Var.17 Var.6 Var.9 Var.12	0.5561	0.3092	0.0078	5	Var.1 Var.2 Var.17 Var.8 Var.9	0.6765	0.4576	0.0235

Total Group

December Step No.	December Var. Ent.	December Mult. R	December R-Sq.	December Inc. R-Sq.	May Step No.	May Var. Ent.	May Mult. R	May R-Sq.	May Inc. R-Sq.
2	Var.1 Var.17	0.4946	0.2446	-	2	Var.2 Var.17	0.5667	0.3211	-
3	Var.1 Var.17 Var.6	0.5258	0.2765	0.0319	3	Var.2 Var.17 Var.6	0.5946	0.3536	0.0325
4	Var.1 Var.17 Var.6 Var.12	0.5472	0.2994	0.0229	4	Var.2 Var.17 Var.6 Var.8	0.6070	0.3684	0.0148
5	Var.1 Var.17 Var.6 Var.9 Var.12	0.5545	0.3075	0.0081	5	Var.1 Var.2 Var.17 Var.6 Var.8	0.6159	0.3793	0.0109

Table 8 (continued)

Summary of Stepwise Regression Procedure[1]
Hebrew Test

Senior										Total Group									
December					May					December					May				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.19	0.5899	0.3480	-						2	Var.2 Var.19	0.5540	0.3069	-					
3	Var.1 Var.16 Var.19	0.6487	0.4208	0.0728						3	Var.2 Var.16 Var.19	0.6314	0.3987	0.0918					
4	Var.2 Var.17 Var.12 Var.19	0.7217	0.5208	0.1000						4	Var.2 Var.17 Var.12 Var.19	0.6876	0.4728	0.0741					
5	Var.2 Var.17 Var.12 Var.16 Var.19	0.7488	0.5607	0.0399						5	Var.2 Var.17 Var.12 Var.16 Var.19	0.7231	0.5229	0.0501					

Table 8 (continued)

Summary of Stepwise Regression Procedure[1]
Latin Test

Junior									
December					May				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
					2	Var.1 Var.12	0.5451	0.2971	-
					3	Var.1 Var.2 Var.12	0.5752	0.3309	0.0338
					4	Var.1 Var.2 Var.17 Var.12	0.5991	0.3589	0.0280
					5	Var.1 Var.2 Var.17 Var.9 Var.12	0.6120	0.3745	0.0156

Senior									
December					May				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.5992	0.3590	-	2	Var.2 Var.17	0.6831	0.4666	-
3	Var.1 Var.2 Var.17	0.6291	0.3958	0.0368	3	Var.2 Var.17 Var.3	0.7374	0.5438	0.0772
4	Var.1 Var.2 Var.17 Var.12	0.6492	0.4214	0.0256	4	Var.1 Var.2 Var.17 Var.3	0.7634	0.5828	0.0390
5	Var.1 Var.2 Var.17 Var.6 Var.12	0.6572	0.4319	0.0105	5	Var.1 Var.2 Var.17 Var.3 Var.6	0.7743	0.5996	0.0168

Total Group									
December					May				
Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.	Step No.	Var. Ent.	Mult. R	R-Sq.	Inc. R-Sq.
2	Var.1 Var.2	0.5955	0.3546	-	2	Var.1 Var.2	0.6012	0.3615	-
3	Var.1 Var.2 Var.17	0.6251	0.3907	0.0361	3	Var.1 Var.2 Var.17	0.6369	0.4057	0.0442
4	Var.1 Var.2 Var.17 Var.12	0.6452	0.4163	0.0256	4	Var.1 Var.2 Var.17 Var.12	0.6556	0.4298	0.0241
5	Var.1 Var.2 Var.17 Var.6 Var.12	0.6543	0.4281	0.0118	5	Var.1 Var.2 Var.17 Var.6 Var.12	0.6637	0.4405	0.0107

[1] Regression analyses were carried out using only those candidates who had scores on all covariates chosen for the model. Analyses were not carried out for groups with N < 100. Table II of the appendix contains intercorrelation matrices for the various analysis groups.

Table 9

Summary of Best Two-, Three-, or Four-Variable Regression Models Specified by Grade Level for Five Achievement Test Administrations[1]

Test	Grade Level	November	December	January	May	June
		Administration/Variables				
Eng. Comp.	Junior	1,2,6	1,2,6	1,2,21	1,2,12	1,2,12
	R^2	.76	.66	.68	.68	.66
	Senior	1,2,12	1,2,12	1,2,10	1,2,10	1,2,10
	R^2	.67	.60	.65	.63	.64
Literature	Junior				1,2,4	1,2,4
	R^2				.72	.69
	Senior	1,10,9	1,3,10	1,2,10	1,6,10	1,3,20
	R^2	.71	.68	.71	.69	.72
Amer. Hist.	Junior				1,2,15	1,2,15
	R^2				.62	.63
	Senior	1,2,15	1,2,15	1,2,15	1,8,15	1,2,15
	R^2	.61	.60	.57	.61	.66
Math I	Junior	2,17	2,11	2,8	2,11	2,11
	R^2	.74	.74	.72	.73	.71
	Senior	2,17	2,17	2,17	2,17	2,11
	R^2	.72	.68	.66	.66	.69
Math II	Junior		2,11	2,17	2,11	2,11
	R^2		.68	.67	.64	.62
	Senior	2,17	2,17	2,17	2,17	2,11
	R^2	.62	.62	.65	.67	.62
Biology	Junior		1,2,9	1,20,21	1,2,13	1,2,13
	R^2		.52	.72	.67	.57
	Senior	1,2,13	1,2,7	1,2,7	1,2,7	1,2,7
	R^2	.64	.60	.54	.57	.64
Chemistry	Junior		1,2,13	2,5,21	1,2,8	1,2,14
	R^2		.71	.55	.58	.55
	Senior	1,2,8	1,2,8	1,2,8	1,2,8	1,2,8
	R^2	.50	.55	.52	.56	.62
Physics	Junior				1,2,8	1,2,8
	R^2				.54	.53
	Senior	1,2,17	1,2,17	1,2,17	1,2,17	1,2,17
	R^2	.47	.53	.53	.47	.54
French	Junior		17,12,15,18	1,17,9,21	1,2,17,12	1,2,17,12
	R^2		.35	.59	.53	.46
	Senior	1,2,17,6	1,2,17,12	1,2,17,16	1,17,6,16	1,17,6,12
	R^2	.51	.43	.40	.55	.54
Spanish	Junior	17,7,9,15	1,17,4,5	7,12,15,16	1,2,17,12	1,2,17,12
	R^2				.47	.43
	Senior	1,2,17,12	1,17,12,21	1,17,12,21	1,2,17,12	1,2,17,12
	R^2	.38	.34	.26	.33	.45
Eur. Hist.	Junior	-		-	1,2,15	-
	R^2				.54	
	Senior	-	1,2,10	-	1,2,15	-
	R^2		.50		.67	
German	Junior	-		-	1,2,17,8	-
	R^2				.31	
	Senior	-	1,17,6,12	-	2,17,7,21	-
	R^2		.30		.43	
Hebrew	Junior	-	-	-		-
	R^2					
	Senior	-	-	-		-
	R^2					
Latin	Junior	-	-	-	1,2,17,12	-
	R^2				.36	
	Senior	-	1,2,17,12	-	1,2,17,12	-
	R^2		.42		.58	

1

Table 10

Best Two, Three or Four Variable Regression Models for Total Groups Taking the Achievement Tests[1]

English Composition Test

	Administration				
	November	December	January	May	June
2 Var. Model	Var. 1,2	Var. 1,12	Var. 1,2	Var. 1,2	Var. 1,2
R^2	.66	.59	.64	.67	.65
3 Var. Model	Var. 1,2,12	Var. 1,2,12	Var. 1,2.10	Var. 1,2,12	Var. 1,2,12
R^2	.67	.60	.65	.68	.66
4 Var. Model	Var. 1,2,6,10	Var. 1,2,6,10	Var. 1,2,10,12	Var. 1,2,6,10	Var. 1,2,6,10
R^2	.67	.60	.65	.68	.66
N	37,188	121,770	53,877	12,977	68,547

Literature Test

	Administration				
	November	December	January	May	June
2 Var. Model	Var. 1,10	Var. 1,10	Var. 1,10	Var. 1,10	Var. 1,10
R^2	.71	.68	.71	.70	.70
3 Var. Model	Var. 1,10,19	Var. 1,3,10	Var. 1,2,10	Var. 1,10,16	Var. 1,3,10
R^2	.71	.68	.71	.71	.70
4 Var. Model	Var. 1,3,10,19	Var. 1,2,10,16	Var. 1,2,10,16	Var. 1,2,10,16	Var. 1,3,8,10
R^2	.71	.68	.71	.71	.70
N	4,392	12,721	6,499	1,256	5,078

American History and Social Studies Test

	Administration				
	November	December	January	May	June
2 Var. Model	Var. 1,15	Var. 1,15	Var. 1,15	Var. 1,15	Var. 1,15
R^2	.60	.58	.56	.63	.61
3 Var. Model	Var. 1,2,15	Var. 1,2,15	Var. 1,2,15	Var. 1,2,15	Var. 1,2,15
R^2	.61	.60	.57	.65	.63
4 Var. Model	Var. 1,2,10,15	Var. 1,2.9,15	Var. 1,2,8,15	Var. 1,2,9,15	Var. 1,2,9,15
R^2	.61	.60	.57	.66	.64
N	7,533	17,061	9,800	2,836	23,146

Table 10 (continued)

Best Two, Three or Four Variable Regression Models for Total Groups Taking the Achievement Tests[1]

Mathematics Level I Test					
	Administration				
	November	December	January	May	June
2 Var. Model	Var. 2,17	Var. 2,17	Var. 2,17	Var. 2,17	Var. 2,17
R^2	.72	.68	.67	.71	.72
3 Var. Model	Var. 2,17,11	Var. 7,17,11	Var. 2,17,11	Var. 2,17,11	Var. 1,2,11
R^2	.73	.69	.67	.71	.72
4 Var. Model	Var. 2,17,8,11	Var. 1,2,17,11	Var. 1,2,17,11	Var. 2,17,5,11	Var. 2,17,5,11
R^2	.73	.70	.68	.72	.73
N	25,153	83,341	42,275	10,448	61,096

Mathematics Level II Test					
	Administration				
	November	December	January	May	June
2 Var. Model	Var. 1,17	Var. 2,17	Var. 2,17	Var. 2,11	Var. 2,11
R^2	.63	.62	.66	.73	.65
3 Var. Model	Var. 2,17,11	Var. 2,17,11	Var. 1,2,17	Var. 1,2,17	Var. 1,2,11
R^2	.63	.62	.66	.73	.65
4 Var. Model	Var. 1,2,17,11	Var. 1,2,17,11	Var. 1,2,17,11	Var. 1,1,17,11	Var. 1,2,17,11
R^2	.63	.63	.66	.73	.65
N	11,332	25,690	12,674	4,033	18,755

Biology Test					
	Administration				
	November	December	January	May	June
2 Var. Model	Var. 1,2	Var. 1,2	Var. 1,2	Var. 1,2	Var. 1,2
R^2	.62	.58	.57	.64	.56
3 Var. Model	Var. 1,2,13	Var. 1,2,7	Var. 1,2,7	Var. 1,2,13	Var. 1,2,13
R^2	.63	.60	.59	.65	.58
4 Var. Model	Var. 1,2,7,13	Var. 1,2,7,13	Var. 1,2,7,13	Var. 1,2,7,13	Var. 1,2,7,13
R^2	.64	.60	.60	.66	.58
N	4,336	13,349	7,628	1,400	18,884

Table 10 (continued)

Best Two, Three or Four Variable Regression Models for Total Groups Taking the Achievement Tests[1]

Chemistry Test

	Administration				
	November	December	January	May	June
2 Var. Model	Var. 1,2	Var. 1,2	Var. 1,2	Var. 1,2	Var. 1,2
R^2	.47	.53	.48	.56	.53
3 Var. Model	Var. 1,7,8	Var. 1,7,8	Var. 1,2,8	Var. 1,2,8	Var. 1,2,8
R^2	.50	.55	.52	.59	.55
4 Var. Model	Var. 1,2,8,14	Var. 1,2,8,14	Var. 1,2,8,14	Var. 1,2,8,14	Var. 1,2,8,14
R^2	.51	.56	.53	.60	.56
N	5,773	14,112	7,412	1,635	24,384

Physics Test

	Administration				
	November	December	January	May	June
2 Var. Model	Var. 1,2	Var. 1,2	Var. 1,2	Var. 1,2	Var. 1,2
R^2	.43	.48	.49	.51	.53
3 Var. Model	Var. 1,2,17	Var. 1,2,17	Var. 1,2,17	Var. 1,2,8	Var. 1,2,8
R^2	.47	.53	.53	.54	.54
4 Var. Model	Var. 1,2,17,8	Var. 1,2,17,8	Var. 1,2,17,8	Var. 1,2,17,8	Var. 1,2,8,14
R^2	.48	.54	.54	.55	.54
N	3,040	8,118	6,867	1,114	8,065

French Test

	Administration				
	November	December	January	May	June
2 Var. Model	Var. 1,17	Var. 1,17	Var. 1,17	Var. 1,17	Var. 1,17
R^2	.48	.39	.37	.48	.40
3 Var. Model	Var. 1,2,17	Var. 1,17,12	Var. 1,2,17	Var. 1,17,6	Var. 1,17,12
R^2	.49	.41	.39	.51	.44
4 Var. Model	Var. 1,2,17,6	Var. 1,2,17,12	Var. 1,2,17,6	Var. 1,17,6,12	Var. 1,2,17,12
R^2	.50	.43	.40	.53	.45
N	2,891	11,344	3,927	1,491	8,101

Table 10 (continued)

Best Two, Three or Four Variable Regression Models for Total Groups Taking the Achievement Tests[1]

Spanish Test

	Administration				
	November	December	January	May	June
2 Var. Model	Var. 1,17	Var. 1,17	Var. 17,12	Var. 1,17	Var. 1,17
R^2	.36	.31	.22	.33	.37
3 Var. Model	Var. 1,17,12	Var. 1,17,12	Var. 1,17,12	Var. 1,17,12	Var. 1,17,12
R^2	.38	.34	.24	.37	.41
4 Var. Model	Var. 1,2,17,12	Var. 1,17,12,21	Var. 1,17,12,21	Var. 1,2,17,12	Var. 1,2,17,12
R^2	.39	.34	.26	.39	.42
N	3,373	11,436	4,725	1,446	8,676

European History and World Cultures Test

	Administration				
	November	December	January	May	June
2 Var. Model		Var. 1,2		Var. 1,2	
R^2		.49		.57	
3 Var. Model		Var. 1,2,10		Var. 1,2,15	
R^2		.49		.59	
4 Var. Model		Var. 1,2,10,15		Var. 1,2,9,15	
R^2		.51		.60	
N		3,010		697	

German Test

	Administration				
	November	December	January	May	June
2 Var. Model		Var. 1,17		Var. 2,17	
R^2		.24		.32	
3 Var. Model		Var. 1,17,6		Var. 2,17,6	
R^2		.28		.35	
4 Var. Model		Var. 1,17,6,12		Var. 2,17,6,8	
R^2		.30		.37	
N		2,519		461	

Table 10 (continued)

Best Two, Three or Four Variable Regression Models for Total Groups Taking the Achievement Tests[1]

Hebrew					
	Administration				
	November	December	January	May	June
2 Var. Model		Var. 2,19			
R^2		.31			
3 Var. Model		Var. 2,16,19			
R^2		.40			
4 Var. Model		Var. 2,17,12,19			
R^2		.47			
N		118			

Latin Test					
	Administration				
	November	December	January	May	June
2 Var. Model		Var. 1,2		Var. 1,2	
R^2		.35		.36	
3 Var. Model		Var. 1,2,17		Var. 1,2,17	
R^2		.39		.41	
4 Var. Model		Var. 1,2,17,12		Var. 1,2,17,12	
R^2		.42		.43	
N		1,948		738	

[1]Variable names are found in Table 1.

Table 11

Correlation of Scaling Covariates with Achievement Test Score for Proposed Scaling Clusters

	Administrations				
	Nov.	Dec.	Jan.	May	June
Cluster One (English)					
English Composition Test					
Var 1	.81	.76	.79	.81	.80
Var 2	.56	.50	.54	.62	.59
Var 12	.29	.31	.28	.36	.35
Literature Test					
Var 1	.84	.82	.85	.83	.84
Var 2	.54	.49	.52	.55	.55
Var 10	.36	.38	.40	.39	.42
Cluster Two (History)					
American History and Social Studies Test					
Var 1	.76	.75	.73	.77	.77
Var 2	.53	.50	.50	.60	.59
Var 15	.41	.39	.38	.51	.46
European History and World Cultures Test					
Var 1		.69		.75	
Var 2		.46		.51	
Var 15		.31		.42	

Table 11 (continued)

Correlation of Scaling Covariates with Achievement Test Score for Proposed Scaling Clusters

Cluster Three (Mathematics)

	Administrations				
	Nov.	Dec.	Jan.	May	June
Mathematics Level I Test					
Var 2	.84	.82	.81	.83	.84
Var 17	.44	.42	.43	.35	.31
Mathematics Level II Test					
Var 2	.79	.78	.81	.83	.79
Var 17	.32	.33	.34	.30	.19

Cluster Four (Science)

	Administrations				
	Nov.	Dec.	Jan.	May	June
Biology Test					
Var 1	.75	.72	.72	.74	.67
Var 2	.61	.62	.62	.70	.64
Var 7	.28	.21	.24	.27	.17
Chemistry Test					
Var 1	.57	.58	.55	.60	.59
Var 2	.63	.68	.64	.70	.67
Var 8	.35	.31	.35	.35	.22
Physics Test					
Var 1	.50	.53	.54	.58	.57
Var 2	.62	.63	.63	.67	.68
Var 8	.30	.30	.32	.30	.23

Table 11 (continued)

Correlation of Scaling Covariates with Achievement Test Score for Proposed Scaling Clusters

Cluster Five (Languages)

	Administrations				
	Nov.	Dec.	Jan.	May	June
French Test					
Var 1	.58	.51	.44	.54	.55
Var 2	.45	.40	.37	.48	.48
Var 17	.49	.44	.42	.51	.41
Var 12	.30	.33	.27	.39	.38
Spanish Test					
Var 1	.46	.39	.23	.45	.51
Var 2	.33	.30	.17	.40	.42
Var 17	.49	.44	.41	.45	.40
Var 12	.33	.31	.28	.40	.39
Latin Test					
Var 1		.53		.56	
Var 2		.48		.53	
Var 17		.29		.26	
Var 12		.36		.39	

Table 12

Summary Statistics for Covariates Selected for Proposed Scaling Clusters[1]

Cluster 1

Test	November						December						January						May						June					
	SAT-V		SAT-M		Course Grades		SAT-V		SAT-M		Course Grades		SAT-V		SAT-M		Course Grades		SAT-V		SAT-M		Course Grades		SAT-V		SAT-M		Course Grades	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Eng. Comp.	525	108	591	103	4.68	.82	518	100	566	101	4.81	.85	491	103	554	103	4.52	.88	499	107	555	109	4.55	.92	521	99	572	103	4.61	.90
Lit.	536	108	556	100	5.01	.81	528	103	531	99	4.88	.84	512	106	525	101	4.83	.85	499	109	516	107	4.68	.89	532	105	535	103	4.79	.89

Cluster 2

Test	November						December						January						May						June					
	SAT-V		SAT-M		Course Grades		SAT-V		SAT-M		Course Grades		SAT-V		SAT-M		Course Grades		SAT-V		SAT-M		Course Grades		SAT-V		SAT-M		Course Grades	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Amer. Hist.	512	107	562	100	4.98	.74	507	99	539	99	4.88	.77	492	101	529	101	4.80	.76	503	108	541	111	4.84	.89	532	99	562	102	4.94	.84
Eur. Hist.	-	-	-	-	-	-	564	103	559	103	5.00	.75	-	-	-	-	-	-	542	112	565	105	5.10	.83	-	-	-	-	-	-

Cluster 3

Test	November				December				January				May				June			
	SAT-M		Ach Test BQ		SAT-M		Ach Test BQ		SAT-M		Ach Test BQ		SAT-M		Ach Test BQ		SAT-M		Ach Test BQ	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Math Level I	565	97	3.19	1.48	548	94	3.24	1.39	540	96	3.28	1.37	540	101	2.68	1.38	562	95	2.57	1.31
Math Level II	656	81	4.45	1.01	640	82	4.41	1.01	623	90	4.26	1.08	607	97	3.66	1.13	652	80	3.50	1.02

Table 12 (continued)

Summary Statistics for Covariates Selected for Proposed Scaling Clusters

Cluster 4

Test	November SAT-V Mean	SD	SAT-M Mean	SD	Course Work in Subject Mean	SD	December SAT-V Mean	SD	SAT-M Mean	SD	Course Work in Subject Mean	SD	January SAT-V Mean	SD	SAT-M Mean	SD	Course Work in Subject Mean	SD
Biology	520	111	584	101	2 88	84	513	102	559	101	2 97	81	496	104	548	102	2 94	.83
Chemistry	545	113	650	85	3 89	85	529	106	624	91	3 93	82	514	111	615	94	3.90	87
Physics	553	115	673	77	3 98	89	535	109	644	85	3 98	89	509	112	625	89	3 85	87

Test	May SAT-V Mean	SD	SAT-M Mean	SD	Course Work in Subject Mean	SD	June SAT-V Mean	SD	SAT-M Mean	SD	Course Work in Subject Mean	SD
Biology	516	112	558	109	2 92	88	524	111	575	103	2.70	.87
Chemistry	514	114	610	99	3 79	91	528	100	606	95	3.68	.80
Physics	516	117	627	95	3.76	.88	556	101	650	84	3.82	83

Cluster 5

Test	November SAT-V Mean	SD	SAT-M Mean	SD	Ach Test BQ Mean	SD	Course Grades Mean	SD	December SAT-V Mean	SD	SAT-M Mean	SD	Ach Test BQ Mean	SD	Course Grades Mean	SD	January SAT-V Mean	SD	SAT-M Mean	SD	Ach Test BQ Mean	SD	Course Grades Mean	SD
French	558	107	593	94	7 25	1 40	4 98	72	546	94	570	93	7 28	1 37	4 86	79	531	97	561	95	7 04	1 45	4 86	78
Spanish	515	106	569	102	6 92	1 45	4 93	72	512	99	552	98	6 95	1.46	4.87	.77	495	99	543	101	6 78	1 56	4.84	.77
Latin	-	-	-	-	-	-	-	-	563	95	594	92	6 67	1 45	5 01	75	-	-	-	-	-	-	-	-

Test	May SAT-V Mean	SD	SAT-M Mean	SD	Ach Test BQ Mean	SD	Course Grades Mean	SD	June SAT-V Mean	SD	SAT-M Mean	SD	Ach Test BQ Mean	SD	Course Grades Mean	SD
French	543	102	579	101	6 79	1 52	4 84	86	548	96	582	97	6.98	1.33	4 87	.83
Spanish	505	102	555	109	6 45	1 52	4 80	.84	514	94	563	99	6.68	1 36	4 89	.81
Latin	578	97	618	94	6 34	1 54	4 95	.84	-	-	-	-	-	-	-	-

[1] Minimum sample sizes are given in Table 7

Figure 1

Achievement Test Scaled Score Means for the Total Groups
Taking the Tests at the Five Administrations

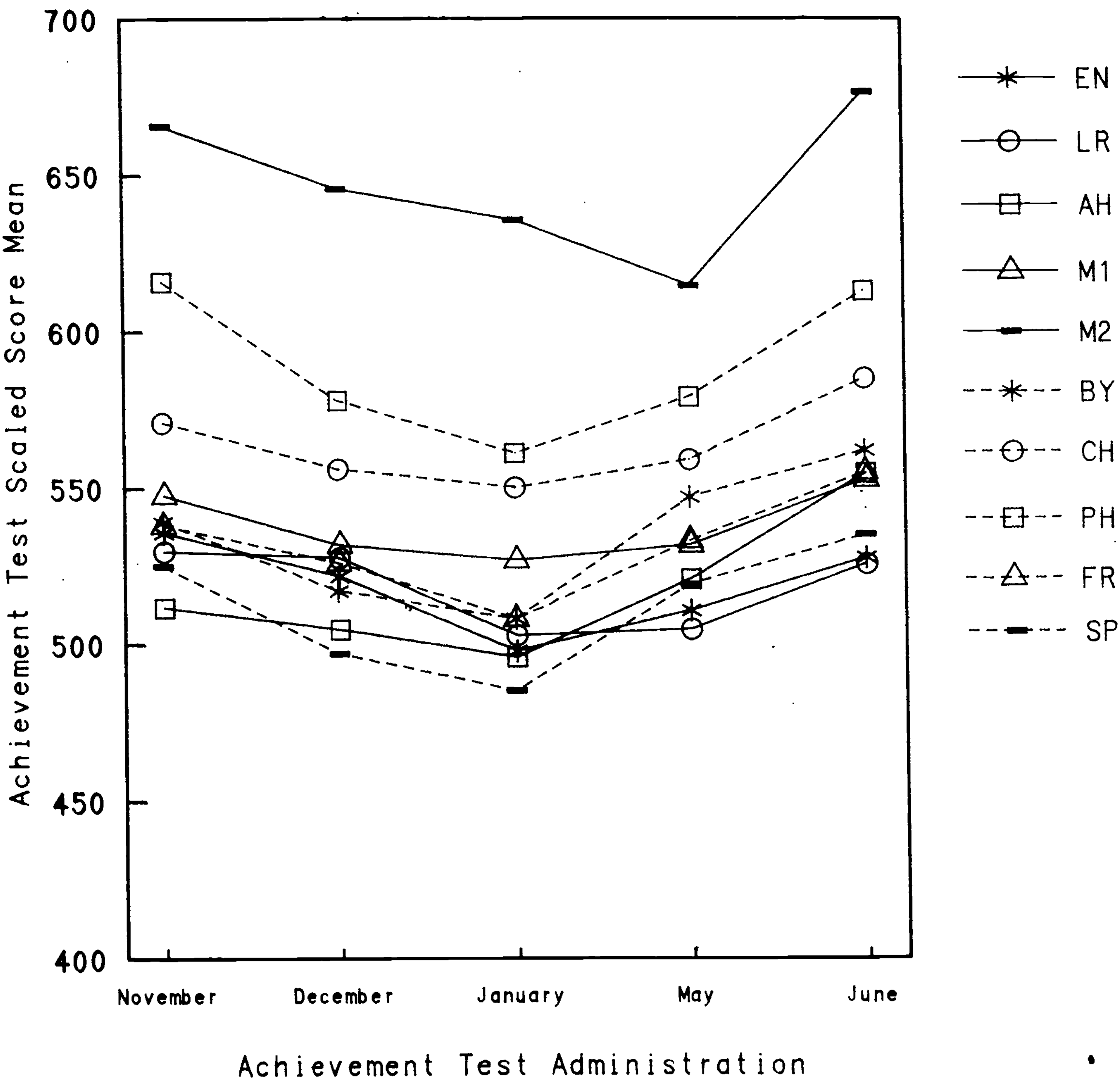

Figure 2

SAT-V Scaled Score Means for the Total Groups Taking the Achievement Tests at the Five Administrations

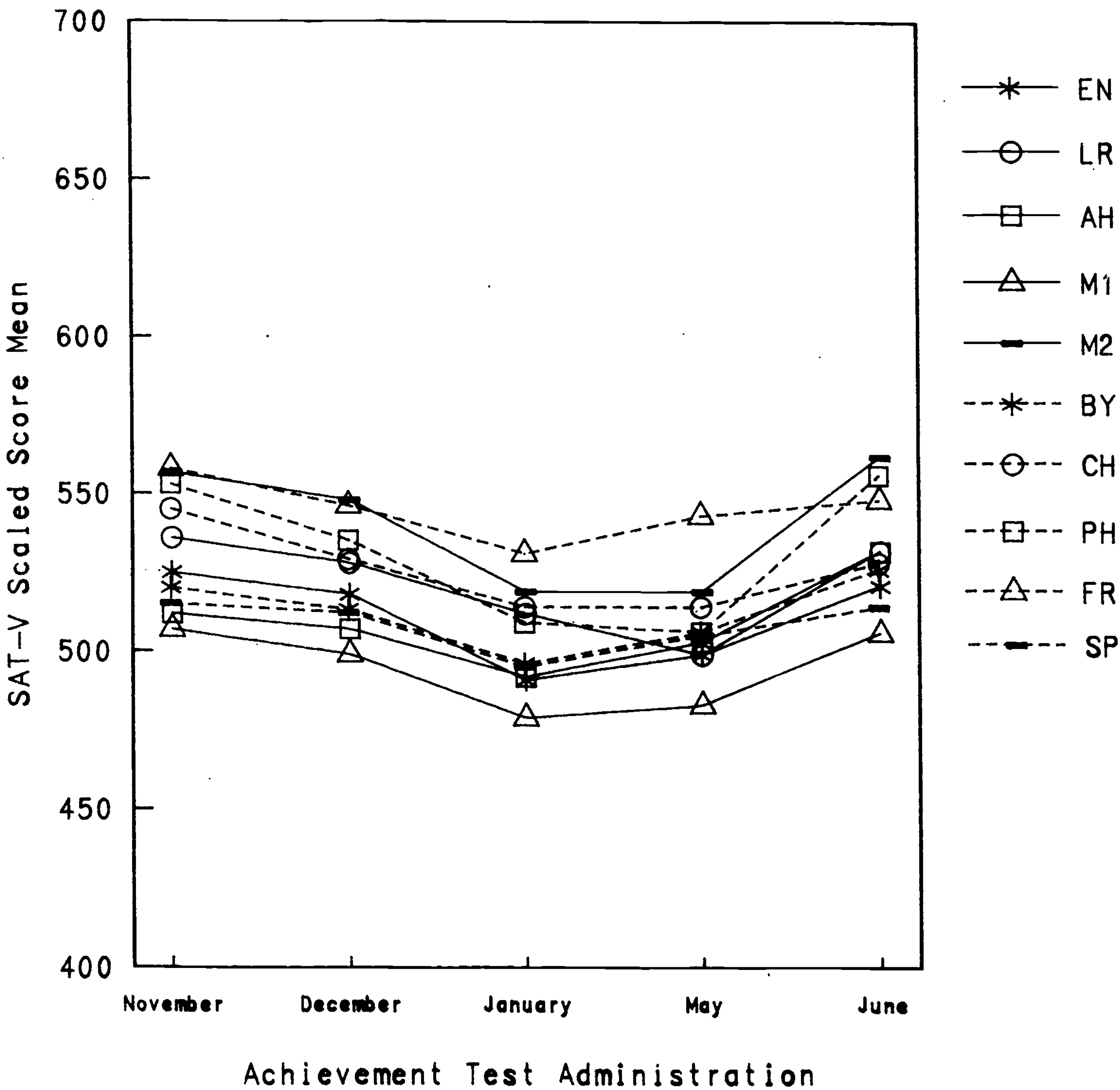

Figure 3

SAT-M Scaled Score Means for the Total Groups Taking the Tests at the Five Administrations

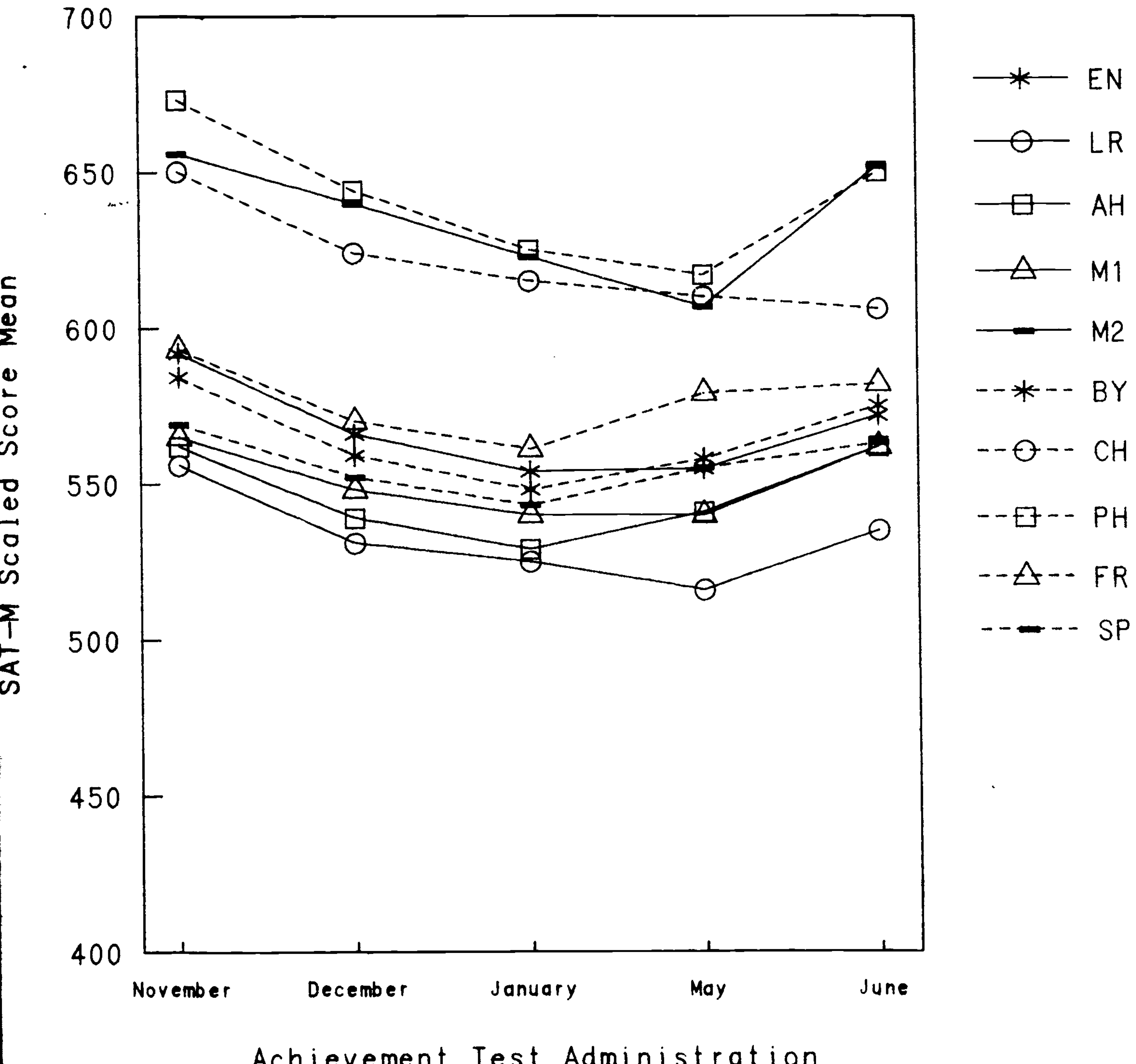

Figure 4

Correlations of Achievement Test Scores with Selected Covariates for November Achievement Test Administration

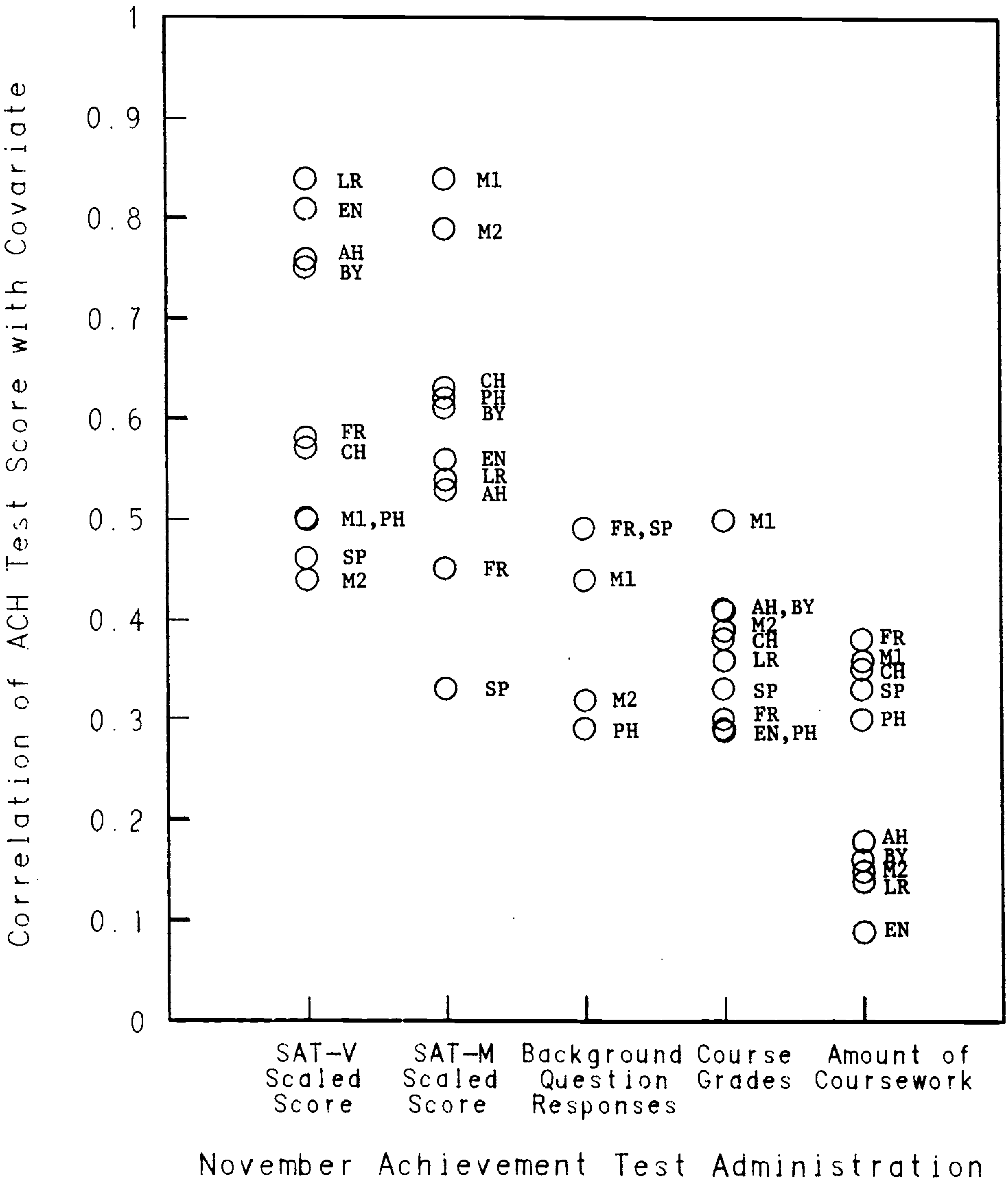

Figure 5

Correlations of Achievement Test Scores with Selected Covariates for December Achievement Test Administration

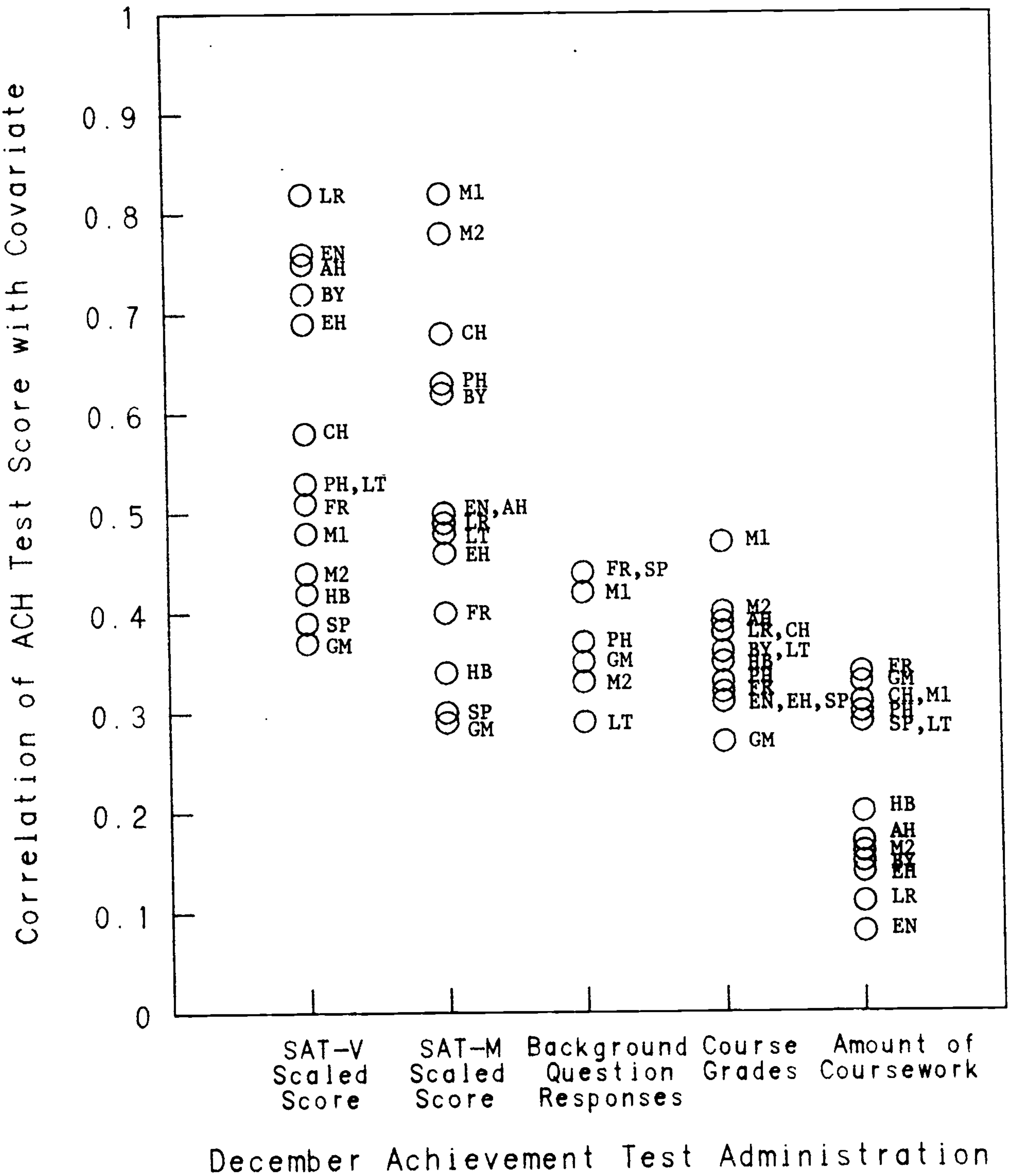

Figure 6

Correlations of Achievement Test Scores with Selected Covariates for January Achievement Test Administrations

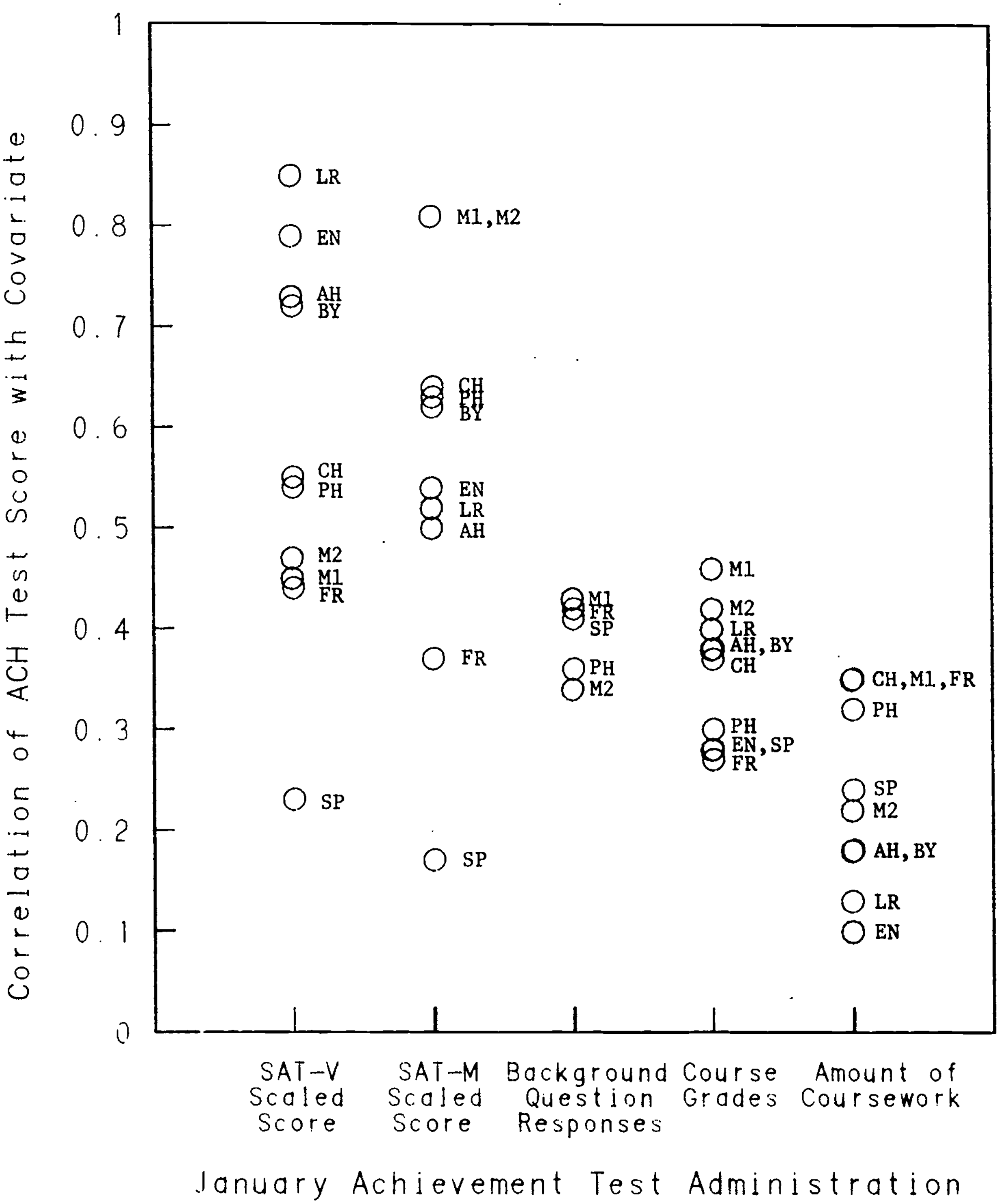

Figure 7

Correlations of Achievement Test Scores with Selected Covariates for May Achievement Test Administration

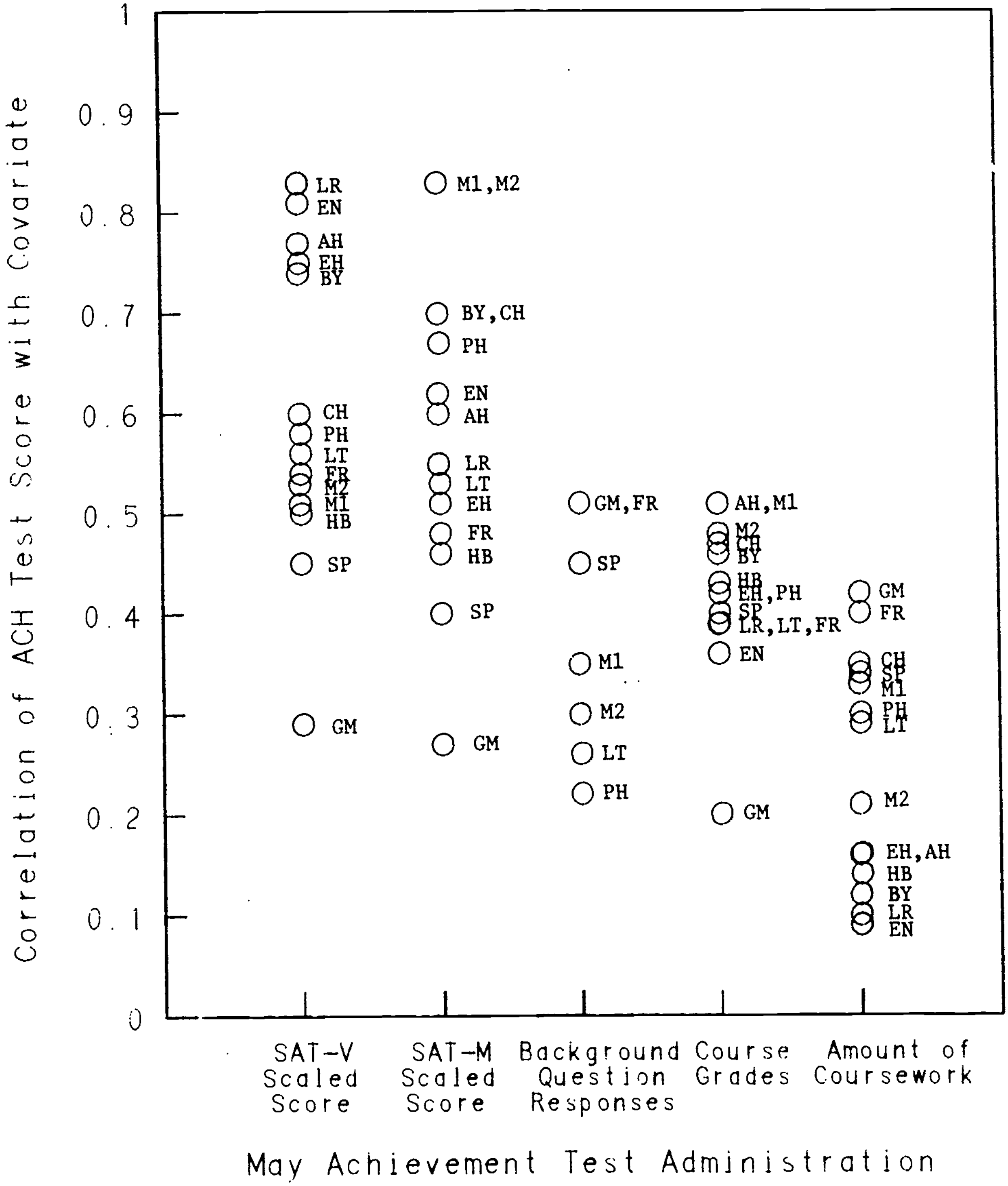

Figure 8

Correlations of Achievement Test Scores with Selected Covariates for June Achievement Test Administration

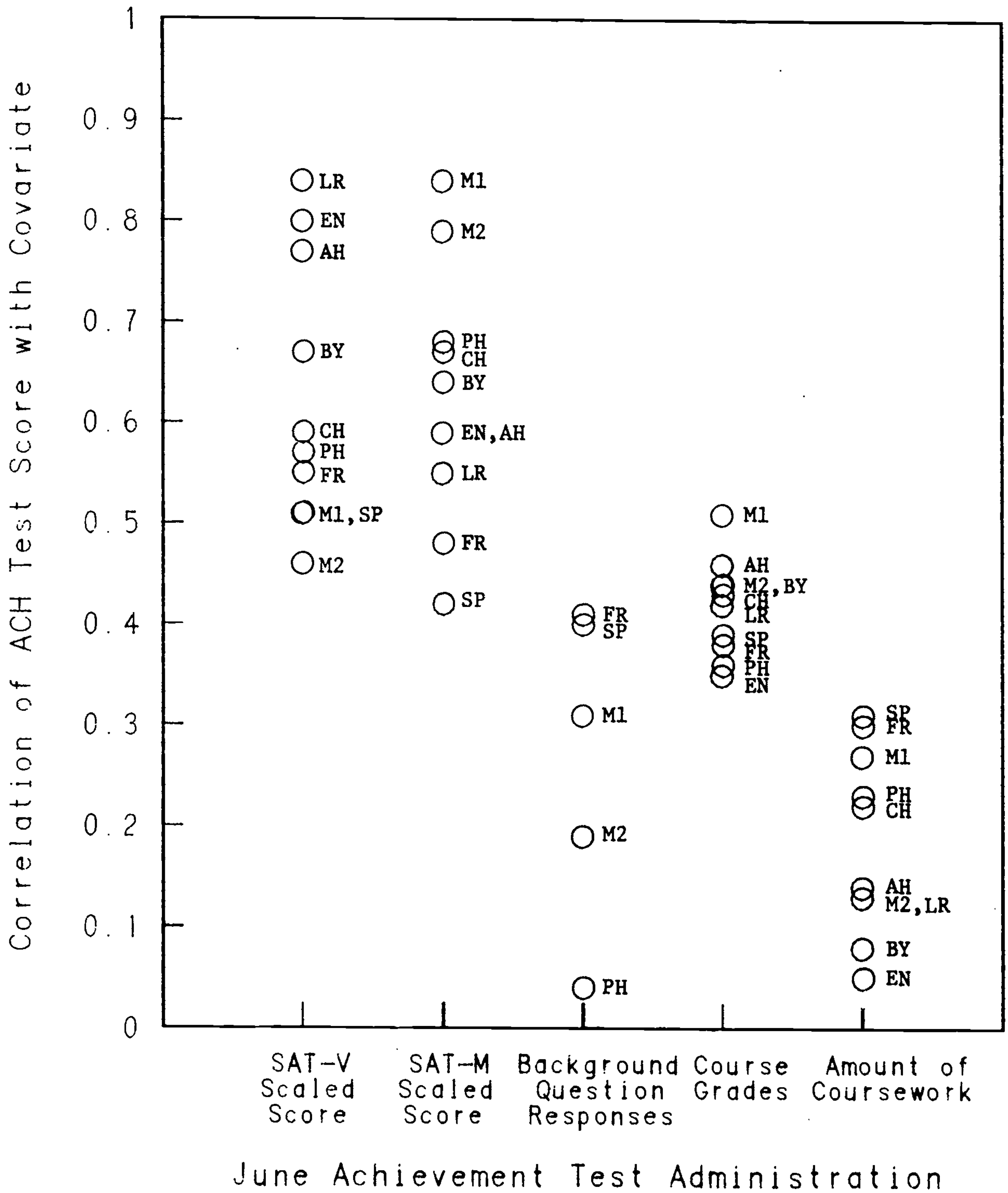

Figure 9

R^2 Values for Best Two-, Three-, and Four-Variable Models
for Fourteen Achievement Tests