To the Graduate Council:

I am submitting herewith a dissertation written by Linda Williams Keefauver entitled "The Effects of a Program of Coaching on Scholastic Aptitude Test Scores of High School Seniors Pretested as Juniors." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Education, with a major in Educational Psychology and Guidance.

Robert L. Williams

Robert L. Williams
Major Professor

We have read this dissertation
and recommend its acceptance:

Charles L. Thompson

Richard D. Bray

Timothy Kindall

Accepted for the Council:

Hilton A. Smith

Vice Chancellor
Graduate Studies and Research

THE EFFECTS OF A PROGRAM OF COACHING ON SCHOLASTIC APTITUDE TEST SCORES OF HIGH SCHOOL SENIORS PRETESTED AS JUNIORS

A Dissertation

Presented for the

Doctor of Education

Degree

The University of Tennessee, Knoxville

Linda Williams Keefauver

August 1976

1288329

ABSTRACT

This study examined the effects of a program of coaching on Scholastic Aptitude Test Scores. The coaching program developed for the study dealt with both verbal and mathematical sections of the SAT and entailed 17, 50-minute sessions of coaching using Mathematics Workbook for College Entrance Examinations (Brownstein, 1975) and Verbal Workbook for College Entrance Examinations (Weiner, 1975).

Fifty-two seniors at The Webb School of Knoxville served as subjects. These 52 students, 32 males and 20 females, consisted of all the members of the 1976 class of 87 who took the SAT in April 1975 and again in December 1975. All subjects were very similar in respect to background (mostly professional parents), ability (I.Q. scores over 115) and the kinds of teaching (college prep) they had received.

Three group conditions—(1) exposure to the coaching program, (2) some preparation on their own (1-10 hours), and (3) no preparation—constituted the independent variable. The dependent variables were Math, Verbal, Reading Comprehension, and Vocabulary scores on the SAT, each considered independently.

The 16 students in the coached group specified their particular needs to be met through the coaching program and were allowed to arrange and perform practice exercises

individually. Seven practice tests were administered to the experimental group, 14 hours were spent "in class" and the experimental subjects accumulated "outside work" times varying from 0 to 36 hours.

The following conclusions were derived from the study:

1. There were no differences among groups on the four dependent variables.
2. Those who needed high SAT scores to get into the college of their choice did not do significantly better than those who did not.
3. Those who expected to improve did not do significantly better than those who did not expect to improve.
4. Those taking a senior year math course did not obtain significantly higher math scores than those not in senior math.
5. Volunteering for the coaching program had no significant effect.
6. Males in the coached group improved their math scores significantly more than coached females, but there were no other statistically significant differences between males and females.
7. Low-scoring students posttested significantly higher than those with acceptable pretest scores on Vocabulary but not on Math, Verbal or Reading.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	1
	Problem Area	1
	College Board Testing Programs	2
	Limitations of the SAT	3
	Advantages of Using SAT Scores	4
	Research Background	6
	Growth, Practice and Sources of Error	13
	Statement of the Problem	18
II.	METHOD	21
	Subjects and Setting	21
	Instruments	22
	Procedure	23
	Practical Hints	24
	Day-by-Day Description of the Coaching Sessions	26
	Student Journals	32
III.	RESULTS	34
	Effects of Independent Variable on Dependent Variables	34
	Effects on Scores of Students Who Needed High Scores Versus Those Who Did Not	35
	Effects of Expectation of Improvement	35
	Effects of Volunteering for the Coaching Program Versus Being Placed in It	38

CHAPTER PAGE

Effects of Taking Senior Math on Math Scores. 39

Effects of Sex on Coaching. 39

Score Changes on Pretest Scores Below 450 Versus Pretest Scores Above 450 41

Comparison of All Subjects' Posttest Scores with Pretest Scores 41

Large Score Increases 45

IV. DISCUSSION. 48

Conclusions 55

BIBLIOGRAPHY 56

APPENDICES 61

Appendix A 62

Appendix B 64

Appendix C 67

Appendix D 69

Appendix E 71

Appendix F 73

Appendix G 75

VITA 77

LIST OF TABLES

TABLE		PAGE
1.	Effects of Independent Variable on Dependent Variables.	34
2.	Pre and Post Scores of Treatment Groups	35
3.	Effects on Scores of Students Who Needed High Scores Versus Those Who Did Not.	36
4.	Effects of Expectation of Improvement	37
5.	Effects of Volunteering for the Coaching Program Versus Being Placed in It.	38
6.	Effects of Taking Senior Math on Math Scores . .	39
7.	Effects of Coaching on Students not Taking Senior Math.	40
8.	Effects of Sex on Coaching	42
9.	Score Changes on Pretest Scores Below 450 Versus Pretest Above 450.	43
10.	Comparison of All Subjects' Posttest Scores with Pretest Scores	44
11.	Pretest to Posttest Score Differences on SAT-Verbal	45
12.	Pretest to Posttest Score Differences on SAT-Math	46
13.	Pretest to Posttest Score Differences on SAT-Vocabulary	46
14.	Pretest to Posttest Score Differences on SAT-Reading.	47

CHAPTER I

INTRODUCTION

Problem Area

High school students attempting to enter prestigious colleges and universities encounter a major hurdle in Scholastic Aptitude Tests. While the College Board states that approximately 65% of all students who take the SAT a second time find that their scores increase, the increase is minimal and shouldn't affect anyone's standing or chances for admission to college (CEEB, 1971). In spite of this official statement, thousands of copies of "Preparing for SAT's" books have been and are being sold to concerned students. Previous research in the area of preparing for these tests has been conflictive. Dyer (1953), French (1955), Dear (CEEB, 1965) and Pallone (1961) found statistically significant differences between coached and uncoached groups with differing levels of practical significance, while Whitla (1962), Frankel (1960) and Lass (1958) found no statistical differences. The present study investigated the question of whether coaching with a smaller sample of students and using different materials than those previously employed would facilitate performance on the SAT.

College Board Testing Programs

Many colleges today require their applicants to submit the results of College Entrance Examination Board (CEEB) tests to be considered for admission. The usual criteria under consideration in the evaluation procedure are high school grades, rank in class, recommendations from high school counselors, and these standardized test scores. In addition to the Scholastic Aptitude Test (SAT), colleges often require achievement tests in various subject matter fields. These achievement tests are also administered by the Admissions Testing Program (ATP), a division of the Educational Testing Service (ETS) of the College Board, and are designed to measure specific knowledge acquired through study. The SAT however, as its name indicates, is an aptitude test, and its purpose is to measure academic skills acquired over a period of many years in order to identify candidates who have the ability to do college work. The SAT is divided into Verbal and Mathematical subtests, and the Verbal subtest results are further subdivided into Reading Comprehension and Vocabulary scores. In addition, beginning in 1974, an experimental test of "standard written English" was included in the score printouts but not considered as an evaluative part of the SAT. The purpose of these aptitude tests is to measure a student's general reasoning ability with both verbal and mathematical symbols.

In college-preparatory secondary schools, especially where there is competition to gain admission to highly selective colleges, students feel a great deal of pressure to excel on the SAT. Evidence of this is the large number of high school juniors taking the test in order to practice for their senior year testing. This trend began in the early 1950's. Duggan (1959) reported that while there was 250% increase in seniors taking the SAT between 1953 and 1958, there was a 600% increase in juniors taking the test during the same period. Many independent secondary schools today require their juniors to take the SAT for guidance purposes.

Limitations of the SAT

Since the test is designed to give an admissions committee an indication of academic potential, it should be useful as a predictor of success at any given college. Research studies point out varying degrees of high positive correlation between SAT scores and achievement in specific fields of study or particular colleges (Dicter, 1937; Lukee, 1951; Mozur, 1968; Olsen, 1957; Pierson, 1960). Most use course grades as indices of achievement, and even though that correlation is often high, it is questionable whether A-students are really what colleges want as their alumni. Professional accomplishments and later-life performance many times are a function of neither high test scores nor good

grades. Persistence, originality and organizational ability are other factors which need to be taken into consideration in the selection process, as is the ability to pursue an interest independently. Students who have actively engaged in hobbies are more likely to become outstanding scholars, artists, or researchers in later life (DeMott, 1974). DeMott further declares that, "admissions offices continue to read major differences of talent and aptitude in minor statistically insignificant differences of SAT scores and continue to weigh academic-skills differences more heavily than real-world talent accomplishments" (p. 68).

The assumption underlying the design of the SAT is that aptitude is a relatively constant condition in a person. It fails to take into account the process aspect of academic aptitude, the importance of interaction between students and their environment and resulting development of motivation which is probably the most crucial determiner of success. Neither does it take into account personality factors, which are certain to have an effect on achievement, nor individual learning styles which would function in day-to-day class activities.

Advantages of Using SAT Scores

Despite these negative factors, the use of SAT scores has had a beneficial effect in giving minority students

access to quality higher education (Hoy, 1967). In the past, only those with wealth and a prep school background could hope to gain entrance to our country's more selective colleges. Students from rural areas and educationally backward schools only attended state supported colleges close to home. The SAT, by establishing a national standard, allowed admissions officers to choose exceptional students who had overcome their deprived backgrounds. In the Ivy League colleges the undergraduate body of the early part of this century was a homogeneous group with respect to socioeconomic background and a heterogeneous one with respect to intellectual ability. Now the reverse is true. Undergraduates in these colleges come from widely varying socioeconomic backgrounds and possess a generally high intellectual ability.

Secondary schools are so divergent in curriculum, teacher pay, per pupil expenditure, and grading systems that the tests do give a common meeting ground for applicants. For these reasons, submitting SAT scores will most likely continue to be a part of the college entrance experience for those applying to highly selective schools. As long as this requirement must be met, students will continue to search for ways to improve their scores. Since 1957 when the first "How to Do Well on SATs" book was published, millions of copies of practice books have been sold to concerned students and various methods of practice and coaching have been used in schools.

Research Background

The official word from the College Board concerning results of research studies on coaching is that

> despite variable factors from one study to another, the net result across all studies is that score gains directly attributable to coaching amount, on the average, to fewer than 10 points—a difference of such small magnitude on the SAT score scale of 200 to 800 that it is unreasonable to expect it to affect college admissions decisions. The magnitude of the gains resulting from coaching vary slightly, but they are always small regardless of the coaching method used or the differences in the students coached. (CEEB, 1968, p. 4)

The booklet from which this quote was taken presents seven studies dating from 1953 to 1962, four of which were conducted by the Educational Testing Service for the College Board, and three of which were carried out by independent researchers.

In 1953, Dyer conducted the first study on the effects of SAT coaching. He was then an associate director of the College Board and is now a vice-president of Educational Testing Service. His study drew, as subjects, boys of above average ability in two independent preparatory schools. Seniors at both schools took the SAT in September and another edition of the test in March. In the time between, the boys in one school were coached for an average of six hours on verbal material and an average of four hours on mathematical material by teachers. This time included outside preparation of each exercise. Students at the other school were not

coached in any way. A comparison of scores showed that the verbal coaching had essentially no effect. For those students who were not taking any mathematics courses, the math coaching seemed to have been a refresher because their average SAT-M score was raised 29 points. Students enrolled in mathematics classes showed no significant gain as a result of coaching. Although Dyer's study was statistically well-designed (the number of years each student had been enrolled in his school before September 1951, the number and level of foreign language and mathematics courses taken during senior year, and the student's initial SAT-V and SAT-M scores earned in September were all used as covariates), there was an obvious time limitation. It is surprising that those not enrolled in math classes, but coached, gained significantly on the SAT-M scale, after only four hours of mathematical preparation.

A more detailed study was carried out under the sponsorship of the College Board in 1955 by French, a research psychologist at Educational Testing Service. In his larger study he looked into the effect of coaching with unspecified sample test items. His sample included public school students, both boys and girls (French, 1955). Most of his subjects were planning for college, but a few, in one coaching program, were taking technical courses. In a smaller study, which was extended from the larger one, he

investigated coaching with the same questions that would appear in the edition of the SAT the students would take after coaching.

For both the larger and smaller studies, three schools were used. One school provided the control group in both studies. One provided 126 students (coached four and one-half hours on verbal materials), 44 of whom undertook a four hour additional verbal study. The other school provided 188 college preparatory and technical students who were coached with both verbal and mathematical materials, for an unspecified amount of time. All students took the SAT as a pretest in September and as a posttest in March. It was found that the scores of the coached groups had increased more than scores of the uncoached students in the first study. While statistical significance was found, French contends that it was of no practical importance. The average score difference of 18 points was compared with the 32-point standard error of measurement inherent in the test itself. Thus, French concludes that

> specific coaching for the SAT does raise the test scores slightly, on the average, but so slightly that the same amount of study on regular school assignments is likely to help just as much on the test and much more toward the student's understanding of the subject. (p. 7)

It is interesting to note that at both schools there was evidence that girls benefited more than boys from coaching; there was no consistent difference between scores of those

taking mathematics courses during their senior year and those not taking such courses; when ability differences were discounted, the coached college prep candidates made no greater gains than did the technical students; and the coached students worked more slowly, attempting fewer items on the posttest.

French's second study, his "identical item" research, was done with the same groups as his larger study but involved only the questions which were used in coaching and then presented in identical form on the test itself (1955). It was found that identical item coaching produced an average increase of 47 points higher on the verbal score and 15 points higher on the mathematical score over regular coaching scores. Thus, there was a 65-point overall increase on SAT-V which the College Board interpreted by stating that "if it were possible to coach with as many as one-tenth of the items that would be used on a future form of the SAT—a virtual impossibility—the effect on the scores of students coached with these items would be negligible" (CEEB, 1968, p. 18). There is an obvious lack of rationale for this explanation.

Another investigation in 1956 and 1957 (CEEB, 1968) conducted by Dear, another researcher for ETS, produced statistically significant gains in mathematical scores (26 points) which were promptly negated because they were less than the SAT standard error of measurement. This study was

particularly well-executed in that 36 hours of intensive coaching was done in small tutorial groups. Tutors were instructed to use whatever methods seemed best for the individual student, and they met midway in the experiment for a one-day workshop to exchange ideas. Dear concludes that "the gains are probably not of sufficient magnitude to give coached students an unfair advantage in college admissions" (CEEB, 1968, p. 21).

A rather decisive finding by Frankel (1960) is reported in CEEB's Effects of Coaching on Scholastic Aptitude Test Scores (1968). Though this study dealt with the effects of growth and practice, as well as coaching on SAT scores, only the coaching results were reported by CEEB. The 45 experimental subjects had received 30 hours of drill in commercial coaching programs. There was no significant gain in either the verbal or the mathematical average score of the coached group over the matched control group. The complete study (Frankel, 1960) reveals that the effect of growth was studied by comparing the SAT scores of 74 seniors who had not previously taken the test. Fifty-one took the December, 1958, SAT and 23 the January, 1959, tests. No statistically significant difference in scores for these two groups was found. It is not surprising that a difference of one month did not reveal any effects of growth. The degree to which practice affected scores was examined by comparing senior test scores

with junior SAT experience and those without such experience. On the average, students with junior year SAT experience scored 18 points higher on their SAT-V scores and 30 points higher on their SAT-M scores than those without test experience. The greatest gain in all groups occurred in the lower score range.

Another type of coaching was carried out in 1959 and 1960 with public school seniors in vocabulary, reading skills, concept mastery and verbal problem solving (Whitla, 1962). Although Whitla names his study "Effect of Tutoring on Scholastic Aptitude Test Scores" the tutoring consisted of the program offered by the Reading Institute of Boston, a school organized to help students develop study habits and reading skills, not to raise SAT scores. No significant differences were found between those who had attended the five two-hour sessions of the Reading Institute and those who had not. Whitla justifies this study by stating that "the Institute seeks to affect just that set of skills which the Scholastic Aptitude Test is intended to measure" (p. 33) and the College Board (CEEB, 1968) quotes Whitla as concluding his article with

> What we have examined is whether training beyond that offered in suburban Boston high schools of both high and low socioeconomic communities contributes to scoring on the Scholastic Aptitude Test. We have found that it does not. (CEEB, 1968, p. 25)

This direct quote is not in the article (Whitla, 1962) but the same information is conveyed. The actual purpose and goals of the Institute were not taken into account.

The CEEB publication Effects of Coaching on Scholastic Aptitude Scores (1968) reports the results of an informal unpublished study done in 1958 by Abraham Lass, principal of the school. Three groups of students in the school's senior class were considered in this study. All had taken the SAT in the spring of their junior year. One group of 38 had taken commercial coaching programs outside the school; the second group of 21 was given an orientation program for the SAT by the school to alert students to the way in which the SAT is administered and the kinds of questions to expect; and the third group of 82 served as controls, who had no preparation for the test. No differences were found after senior year testing.

As an alternative to the previous coaching studies, Nathaniel Pallone attempted to study the effects of short- and long-term developmental reading courses upon SAT verbal scores (Pallone, 1961). Emphasis was placed on reading comprehension with individualized instruction. The short-term program (six-weeks, 90-minute daily summer instruction) yielded a median 109 point gain while the long-term program (September, 1959, to March, 1960, meeting daily for 50 minutes) yielded a 121.9 point median gain. In both cases, those who

had initially scored below the national SAT-V median of 500 showed greater improvement than those whose scores were nearer the upper end of the scale. This outstanding increase in verbal scores is in direct contrast to Dear's study (CEEB, 1968) which concluded that "the verbal section of the test showed no susceptibility to coaching efforts," and to Frankel's (1960) which also emphasized the futility of trying to raise verbal scores.

Growth, Practice and Sources of Error

Levine and Angoff (1958) would attribute part of Pallone's increase in verbal scores to the combined effect of practice (firsthand experience with the test) and growth (change in ability over time resulting from causes other than practice). Most any coaching results can be discounted when no control group scores are utilized, by attributing increases in scores to these growth or practice factors, or to the test's standard error of measurement.

Levine and Angoff (1958) stated that the combined effect of growth and practice is a 30 point gain for eight months on the verbal test. Duggan (1959) has suggested that for students who are or will be studying mathematics in the senior year, a total of about 50 points should be added to the junior SAT-M score to estimate the senior score. Different forms of the same test are given each year. There are no junior forms and different senior forms.

Levine and Angoff (1958) also determined that the effect of SAT practice was an average increase of 10 points on both math and verbal tests. However, another study (Frankel, 1960) found that practice added 18 points to SAT-V and almost 30 points to SAT-M. Pearson (1948) suggests a 20 point verbal practice effect and a growth factor of 3-1/2 points per month.

The College Board does not officially make statements about growth and practice effects. In a booklet to students (CEEB, 1971) the following explanation is made:

> If you are a junior, you may be interested in estimating the scores you would receive if you repeated a test in your senior year. Approximately 65 percent of all students who take the SAT a second time find that their scores increase, and about 35 percent find that their scores decrease. Both increases and decreases are usually small, and only a small percentage of students obtain significantly higher or lower scores when they repeat a test in their senior year. The average score change is a small increase, but changes as small as the average should not affect anyone's standing or chances for admission to college. Therefore, it is best to consider your junior-year scores as rough approximations of your probable senior-year scores. (p. 14)

This statement is in direct contradiction to studies investigating the growth and practice effects. In a booklet designed for high school and college counselors, however, these topics are dealt with in a nonstatistical manner (CEEB, 1974):

> A reason for score changes may be a real increase in the student's ability. Students may learn between the time they first take the test and the time they repeat it. We have found that a

> student's growth in verbal ability, as measured by the SAT, is relatively independent of the curriculum he is exposed to in his senior year. But his growth in mathematical ability, as shown by the test, is larger if he continues to take math as a senior than if he does not. Another possible reason for score changes is the student's familiarity with the test's format and procedures when he repeats it. However, because of the widespread use of such tests today, fewer students gain for this reason than did years ago, and their gains are smaller. (p. 20)

The SAT's standard error of measurement is taken into consideration in most studies presented, with some variation in actual error size. French (1955) reported the test's standard error as 32 points and Dear (1958) reported it as 34. Other researchers who have mentioned it report it the same as the current CEEB estimate of 30 points (CEEB, 1974). "In 2 out of 3 cases, a reported SAT-Verbal score will fall within a band, the limits of which are 30 points above to 30 points below the score representing the student's true ability" (p. 2). Another section of the same source (CEEB, 1974), in explaining why a student's scores change when he/she repeats the SAT, states:

> The most powerful influence is the imprecision inherent in any test, which is indicated by its standard error of measurement. Since two tests are involved in score changes, there are two sources of error. Thus, score changes can be as large as two standard errors of measurement and still not indicate any difference in the student's actual ability—and whether the score change is an increase or a decrease is completely a matter of chance. (p. 20)

Thus, anytime a score increase is less than 60 points the cause could be attributed to the test's standard error of measurement.

In addition, Frankel (1960) mentions a 10 point equating error, defined as the variation due to differences in forms of the same test. The CEEB Guidebook (CEEB, 1974) explains the equating procedure but does not mention it as a source of error.

Based on the studies cited, it appears that investigations of growth and practice effects and experimental coaching programs have yielded confusing empirical results. Despite the College Board's admonitions against coaching, these results leave the coaching question open to more research, none of which has been attempted with the SAT since 1962.

While several studies have examined the effects of coaching on SAT scores, few were done by independent, objective experimenters. Most were conducted by College Board personnel who have an interest in protecting the validity of the test. Even when statistical significance was achieved, it was negated as being below the test's standard error of measurement. Statistical differences were obtained by Dyer (1953) with only a maximum of 10 hours coaching; by French (1955) with 8-1/2 maximum hours spent in drill; by Dear in 1957 (CEEB, 1965) with 36 hours; and by Pallone (1961) after

both a six-week course and a year-long course in developmental reading. Whitla (1962) after 10 hours, Frankel (1960) after 30 hours, and Lass (1958) after an unspecified amount of time, showed no statistical differences.

Only Pallone (1961) achieved a median and mean increase in test scores over the standard error of measurement. However, verbal scores alone were evaluated. His methodology was based on emphasizing exercises in college-level reading manuals, supplemented by intensive practice in using various Science Research Associates reading materials, films, and tachistoscopic devises. The course was counted as a half-unit of English credit and grades were given. There was no control group for this study and all subjects were in a college preparatory school. While pretest scores were generally high, it was found that the greatest gains were in the lower score ranges.

Frankel (1960) also demonstrated that greater gains could be expected from low-scoring students, and Dyer (1953) suggested that perhaps more overall gains would have taken place had the students not been so bright. A study developed from these suggestions investigated the effects of special preparation for Preliminary Scholastic Aptitude Test (PSAT) scores with educationally disadvantaged students (Roberts & Oppenheim, 1966). Specially prepared linear programmed materials were developed in verbal and mathematical areas,

and instruction in one or the other of the two areas was given to eleventh grade students in 18 predominantly black high schools. Alternate forms of the PSAT, administered as pre and posttests, were used to evaluate the program's effectiveness. Some statistically significant differences were found in gain scores between some of the experimental groups and their controls, but the outcome was generally negative. The conclusion was that the magnitude of the gains observed in the experimental groups was so small that the experimenters didn't expect similar short-term instruction to be beneficial to disadvantaged students.

Statement of the Problem

Despite these findings, this researcher would have preferred to use disadvantaged students as subjects. Availability, however, required the use of a college preparatory sample. The coaching program developed for this study dealt with both verbal and mathematical sections of the SAT, entailed seventeen 50-minute sessions of coaching, and used a different set of materials than those in previous studies. The 16 students specified their particular needs to be met through the coaching program and were dealt with individually. None of the other studies have dealt with a group of this size. Most have had impersonal drill work with large numbers of students, and no consideration has been made of test anxiety.

SAT-Verbal subscores for Reading Comprehension and Vocabulary were reported for the first time in 1974-75, and no research has been reported on the correlation between these two subsections. It is suggested in the Guide for High Schools and Colleges (CEEB, 1974) that the subscores are expected to be highly correlated. A correlation between the subscores obtained from this study was computed.

The purpose of this study was to evaluate empirically the influence of coaching for the Scholastic Aptitude Test on test results of 16 Webb School seniors. Three group conditions: (1) exposure to a coaching program, (2) some preparation on their own (1-10 hours), and (c) no preparation, constituted the independent variable. The dependent variables were Math, Verbal, Reading Comprehension, and Vocabulary scores on the SAT, each considered independently. This investigation sought to answer the primary question:

1. Does coaching affect Math, Verbal, Vocabulary, and Reading scores on the SAT?

and the following secondary questions:

2. Did those who need high SAT scores to get into the college of their choice do significantly better than those who did not?
3. Did those who expected to improve do so significantly more than those who did not expect to improve?

4. Did those taking a senior year math course obtain a significantly higher Math score than those not taking a senior year math course?
5. Did those who volunteered for the coaching program do better than those placed in it?
6. Did the coaching have a more significant effect on girls or boys?
7. Did students who initially scored low make greater gains than those who initially scored high?

CHAPTER II

METHOD

Subjects and Setting

Fifty-two seniors at The Webb School of Knoxville served as subjects for this study. These 52 students, 32 males and 20 females, consisted of all the members of the 1976 class of 87 who took the SAT in April 1975 and again in December 1975. Sixteen of the subjects underwent the coaching program. Five of these experimental subjects volunteered for coaching. The other 12 were placed in it because of lack of strong commitment to any of the other enrichment period offerings. The remaining 36 subjects were questioned shortly after the December testing to determine if they had prepared for the test independently. If they had spent from one to ten hours in preparation, they were placed in the practiced control group. Eleven students were included in this group. Those who had spent less than one hour in preparation were placed in the no-practice control group. Twenty-five students comprised the no-practice control subjects. All subjects were very similar in respect to background (mostly professional parents), ability (I.Q. scores over 115) and the kinds of teaching (college prep) they had received.

The Webb School of Knoxville is a coeducational independent day school which was founded by Robert Webb in 1955. Its basic objective is the graduation of students prepared for college admission and achievement. During the school's history, over 99% of its graduates have entered college; and over 93% have successfully completed a four-year degree program, with 45% going on to graduate school. The school is located on a 63-acre campus, 10 miles west of downtown Knoxville.

The school's rotating schedule of classes allows two enrichment periods per week, the last 50 minutes of the school day on Monday and Wednesday. Students are encouraged to choose a noncredit, no-grade class each quarter based on interest. The classes are led by faculty with interest in the specific area. "College Board Preparation" was offered as a course and was taught by this experimenter. Other offerings were football, dance troupe, international cooking, cheerleading, volleyball, soccer, photography, chorus, newspaper, yearbook, and supervised study. All subjects were involved in some form of enrichment.

Instruments

Subjects were pre and posttested using the College Board's Scholastic Aptitude Test. Each experimental subject bought copies of both the Mathematics Workbook (Brownstein, 1975) at $2.75 and Verbal Workbook for College Entrance

Examinations (Weiner, 1975) at $2.25 to review for the posttest. The mathematics workbook offers intensive practice in basic math, algebra, geometry, number series, figure analogies, data interpretation and data sufficiency. It also contains sample Scholastic Aptitude Tests with correct answer sheets and analysis of possible errors for many of the questions. The verbal workbook is identically structured and involves vocabulary building and usage, sentence completion, word relationships and reading comprehension.

Procedure

The coaching program began in September and continued through the first week in December for 50 minutes, two days a week. Seven practice tests were administered to the experimental group, 14 hours were spent "in class" and the experimental subjects accumulated "outside work" times varying from 0 to 36 hours. Practice exercises which were done both in and out of class consisted of verbal and mathematical items presented in the same form as items which appear in the test itself. The practice exercises in mathematics were approximately twice as long as those in verbal skills.

The general plan was to allow each student to select the exercises he/she most needed drill in and work on them in or out of class, after which the instructor and students

went over the material and analyzed the reasoning processes necessary to obtain correct responses. Attention was given to appropriate processes.

The students scored their own work, using the workbook key, and discussed questions on which they had failed. It is difficult to estimate how much of the 14 hour time spent in the coaching sessions was spent on verbal material and how much on math because of the individualization that was done.

Each mathematics exercise consisted of 25 or 50 multiple choice problems. See Appendix A for exact coverage. The 80 shorter verbal exercises were appropriately divided among the types of questions that appear in the test (Appendix B).

Practical Hints

Practical hints for test taking were utilized in general discussions during the coaching sessions, most of which were extracted from different "getting ready for college" books (Fine, 1960; Hoy, 1967; Lass & Wilson, 1965; Morgan & Deese, 1957; Wilson & Buckner, 1961) and then reassembled for topic relevance.

The students were taught to identify the relevant cues in test directions and to answer questions they were sure of first. Then, if time permitted, they were to return to those questions previously omitted and answer any for which they

could identify one or more of the choices as wrong. The CEEB Student Handbook (CEEB, 1971) states:

> If you have some knowledge of a question and can eliminate one or more answer choices, guessing among the remaining choices can usually be to your advantage. Since the scores on the SAT will be based on the number of questions you answer correctly minus a fraction of the number you answer incorrectly, random or haphazard guessing probably won't affect your scores significantly. (p. 5)

It was explained to the students that for questions with five answer choices the raw score is calculated by subtracting 1/4 of a point for each wrong answer from the number of right answers. For questions with four answer choices, 1/3 of a point is subtracted for each wrong answer. The raw score is then converted to a College Board scaled score (from 200-800) so it can be compared with students who have taken other editions of the SAT.

Students were encouraged to survey the test before starting to work in order to divide their work time wisely and to indicate answers in exactly the way specified in the directions. They were encouraged to scrutinize qualifiers carefully and to beware of ambiguities that are true sometimes and not true other times.

Eliminating obviously incorrect choices before choosing a more correct choice was encouraged, as was a final rereading of the test if time permitted. Students were told that if they felt strongly an answer should be changed, they

should change it, but if they wavered, it would be best to leave the initial answer.

It was brought to the subjects' attention that a testing expert (Chan, 1975) suggests a quick, up-down motion, which she says is adequate for the sensitive grading machines, despite the instructions that say to "fill in the boxes carefully and completely." The present researcher felt unsure enough about this, however, to caution against sloppy marking.

Day-by-Day Description of the Coaching Sessions

1. Wednesday, September 3, 1975

Rapport was established. The nature of the SAT in general and college admissions procedures were discussed. The experimenter asked for suggestions from the group on how to organize class time and what to do to make the course beneficial to them. It was suggested to practice skimming questions to save time, to chart practice scores in order to observe improvements, to take timed tests under pressure, and to analyze different types of items. The group agreed to do exercises out of class and to keep records of what they did and how much time they spent. Timed tests were to be done in class with the experimenter pacing the students by calling out ten minute segments of time passed. The students desired then to go back and discuss individual items. As

homework, the students were encouraged to work through the Math "Fundamentals" section on their own and to record results in the Exercise Progress Chart in their workbooks (Appendix C). Deep breathing exercises were demonstrated by the experimenter and practiced by the group to insure relaxation and easing of tension.

2. Monday, September 8, 1975

Deep breathing exercises were practiced. The group discussed the timetable taken from the Verbal Workbook (Appendix D) to see types of verbal items and the proportion of each type on the tests. The experimenter demonstrated how to chart progress in front and back of the Verbal book. They were to check off exercises as they were done, state the date, and indicate the time it took to do each exercise. It was explained that in the back of the Verbal Workbook on "note" pages, each student was to indicate the date of each class attendance, his/her in-class activities for that day, and how helpful it was personally. Folders were distributed in which to put notes and math answer sheets. The Self-Study Manual for Test Anxiety (Richardson, 1973) was explained and offered to whoever needed it. No one used the material. Anxiety didn't seem to be a major concern with this group. The first 30-minute Verbal Aptitude Test was given, but no time remained in which to discuss it. It was suggested that in the future the tests be given the first 30 minutes of the class period to allow time for discussion of the items.

3. Wednesday, September 10, 1975

Deep breathing exercises were done. The test taken Monday was analyzed with the largest problem area being analogies (_____is to _____ as _____ is to _____). The key seemed to be in establishing the functional relationship of the first two words. For the Reading Comprehension section, where a paragraph of text is followed by a series of questions about it, it was suggested that the students read the questions before reading the text and then pull out the answers. Students were encouraged to do practice exercises at home for the verbal sections they found most difficult. The second Verbal Aptitude Test was given the last 30 minutes.

4. Monday, September 15, 1975

Deep breathing exercises were discontinued. The group was bored with them. The majority of students took the first regular math test, but a few chose to do another verbal test, which was allowed. The class chose to discuss the last verbal test rather than the math test, and there was some dispute about techniques for doing Reading Comprehension. Several students could not do well by reading the questions first, so it was suggested that everyone develop his or her own style by doing the practice exercises in various ways to see which worked best. Most students preferred to skim the questions and then read the full test.

Wednesday, September 17, 1975

Faculty meeting—no enrichment.

5. Monday, September 22, 1975

Pacing was done on the second math test by estimating how many problems needed to be done every five minutes and during the testing, every five minutes of time was called off. The students were encouraged to speed up their answering and then compare their test score with the previous one. A vote was taken which was more helpful—taking tests, doing exercises or having discussions. The majority chose taking tests, and a request was made to do a verbal test orally next time.

6. Wednesday, September 24, 1975

One-half of the third Verbal Aptitude Test was done orally by the group. The experimenter had worked through the test, checked answers, looked up words, and analyzed thought processes before class. The students were told there would be an exercise progress check at the next meeting.

Monday, September 29, 1975

Special program for the entire school—no enrichment.

7. Wednesday, October 1, 1975

Progress sheets were checked and students were encouraged to do more out of class exercises. General hints, described earlier, were discussed. The remainder of class time was spent on individual exercises, e.g., Appendix E.

8. Monday, October 6, 1975

Students worked through individual exercises while the experimenter discussed problem areas with individual students.

9. Wednesday, October 8, 1975

The second half of the third Verbal Aptitude Test was analyzed orally. Emphasis was placed on concentration in the Reading Comprehension section.

10. Monday, October 13, 1975

The third Mathematics Aptitude Test was given with ten minute pacing and then discussed.

Wednesday, October 15, 1975

Faculty meeting—no enrichment.

Monday and Wednesday, October 20 and 22, 1975

Academic classes meeting—change in schedule for the entire week so the students could spend an entire afternoon seeing Macbeth at the University of Tennessee.

11. Monday, October 27, 1975

Registration for the December testing was completed in class. There was a discussion on general hints, and the remainder of class time was spent with exercises of the individual student's choosing.

12. Wednesday, October 29, 1975

The fourth math test was given and discussed. Students were encouraged to make shrewd guesses when they did not know the answers and then to go back to check what percentage

of their guesses was correct. The experimenter realized how ill-prepared she was for math coaching and enlisted the assistance of Webb's Math Department Chairperson for the next two sessions.

13. Monday, November 3, 1975

The Math Department Chairperson worked out math exercise problems with the group.

14. Wednesday, November 5, 1975

The Math Department Chairperson continued working out math exercise problems with the group. Student reaction was favorable.

Monday, November 10, 1975

Special program for the entire school—no enrichment.

Wednesday, November 12, 1975

Faculty meeting—no enrichment.

15. Monday, November 17, 1975

Students chose the practice test of their choice to take during the first 30 minutes. Note was taken of who was having difficulty finishing on time. Individual conferences concerned with where these students were bogging down and how to avoid it were had. The rest of the group did individual exercises.

16. Wednesday, November 19, 1975

Students again chose a practice test to take during the first 30 minutes. The remainder of time was spent getting folders in order to turn in next session.

17. Monday, December 1, 1975

Questions concerning this class were answered on the front of each student's folder, the content of which is dealt with in Student Journals (p. 25). Last minute suggestions were to relax the night before testing and to get up early the test morning, to go into the test with a positive attitude, to take a watch for pacing, a pillow to sit on, some chocolate to eat, to underline key words, and to eliminate obviously wrong answers.

Student Journals

Each student kept a journal of in-class activities (what he/she did each period and his or her reaction to it) in the back of the Verbal Workbook. Each was given a folder at the beginning of the coaching sessions in which to put math answer sheets in order to keep a record of work done (Appendix F). Verbal work was recorded on the charts included in the Verbal Workbook and at the end of the course, placed in the folder (Appendix G). Along with the "date taken" information, each student recorded the approximate amount of time involved with each exercise. The journal sheet also was included in the folder on the last class session.

On the front of his or her folder, each student (a) calculated the total amount of time spent in preparation for

the SAT (50 minutes for each class attended plus whatever time was spent outside of class on exercises); (b) described how he/she happened to be in the class; (c) indicated what aspect of the program had been the most helpful to him or her personally; (d) indicated whether he/she expected to improve (defined loosely as at least 40 or 50 points); (e) indicated whether he/she needed a high score (defined loosely as 600+) to get into the college of his/her choice.

The purpose of the journal was to help the students keep a record of their own progress and to help the experimenter identify possible explanations for scores obtained by various students.

CHAPTER III

RESULTS

Effects of Independent Variable on Dependent Variables

An analysis of covariance was used in analyzing differences on posttest scores across the three treatment groups (coached, practiced and not practiced). This technique served to reveal the effect of coaching on the December 1975 test by holding constant the April 1975 test scores, which served as covariates.

Table 1 indicates that there were no significant differences between groups on any section of the SAT. Neither coaching nor practice had any effect on any of the groups' mean posttest scores. Table 2 presents pre and post scores of all three groups in Math, Verbal, Vocabulary and Reading.

Table 1

Effects of Independent Variable on Dependent Variables

Factor	Dependent Variable Posttest	Covariate Pretest	F	Significance Level
Treatment	Math	Math	1.63	.205
Treatment	Verbal	Verbal	.25	.999
Treatment	Vocabulary	Vocabulary	.029	.999
Treatment	Reading	Reading	.044	.999

Table 2

Pre and Post Scores of Treatment Groups

Groups	Math		Verbal		Vocabulary		Reading	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Coached	53.88	53.65	48.06	52.25	50.13	51.19	45.88	53.13
Practiced Controls	60.00	60.45	56.09	57.82	55.09	55.64	57.64	60.55
No-practice Controls	57.12	59.12	49.88	52.88	51.64	52.08	48.16	54.04

Effects on Scores of Students Who Needed High Scores Versus Those Who Did Not

To determine whether students needing high scores achieved greater gain, the experimenter compared the gain scores of students indicating a need for 600+ scores to get into the college of their choice with those not indicating a need for 600+ scores. Results, as shown in Table 3, indicate that there was no significant difference on any gain scores between those who needed high scores and those who did not.

Effects of Expectation of Improvement

In order to examine whether a student's expectations of improvement affected his/her posttest scores, a t-test was again used. Table 4 shows that there were no significant differences between the scores of those who expected to improve and those who did not expect to improve.

Table 3

Effects on Scores of Students Who Needed High Scores Versus Those Who Did Not

Variable	Number of Cases	Mean Gain Scores	Standard Deviation	T Value	2-Tail Probability
Gain Math					
Need	33	0.7273	5.082	0.48	0.633
No-need	19	1.4211	4.891		
Gain Verbal					
Need	33	3.4242	4.294	-0.52	0.610
No-need	19	2.5263	6.842		
Gain Reading					
Need	33	1.0000	5.635	-0.50	0.617
No-need	19	0.0526	7.906		
Gain Vocabulary					
Need	33	6.0000	7.710	-0.38	0.706
No-need	19	5.1053	8.956		

Math and Verbal scores, reported by CEEB on a 200-800 scale, were converted to the 20-80 Reading and Vocabulary scale for this analysis. Thus gain scores are based on two digits rather than three.

Table 4

Effects of Expectation of Improvement

Variable	Number of Cases	Mean Gain Scores	Standard Deviation	T Value	2-Tail Probability
Gain—Math					
Expected gain	48	0.9375	4.987	0.22	0.831
Expected no gain	4	1.5000	5.568		
Gain—Verbal					
Expected gain	48	3.3542	5.269	-1.22	0.229
Expected no gain	4	0.0	5.657		
Gain—Reading					
Expected gain	48	1.0000	6.338	-1.34	0.186
Expected no gain	4	-3.5000	7.937		
Gain—Verbal					
Expected gain	48	5.9583	8.303	-0.88	0.385
Expected no gain	4	2.2500	4.646		

Math and Verbal scores, reported by CEEB on a 200-800 scale, were converted to the 20-80 Reading and Vocabulary scale for this analysis. Thus gain scores are based on two digits rather than three.

Effects of Volunteering for the Coaching Program Versus Being Placed in It

Differences between the experimental subjects who had volunteered for or had been placed into the coaching program were also analyzed with a t-test. Table 5 indicates that there were no significant differences in gain scores for those who volunteered or were placed in this experimental group.

Table 5

Effects of Volunteering for the Coaching Program Versus Being Placed in It

Variable	Number of Cases	Mean Gain Scores	Standard Deviation	T Value	2-Tail Probability
Gain—Math					
Volunteered	5	-3.0000	3.240	1.68	0.116
Placed	11	1.0000	4.817		
Gain—Verbal					
Volunteered	5	3.6000	3.975	0.31	0.761
Placed	11	4.4545	5.502		
Gain—Reading					
Volunteered	5	0.6000	3.130	0.28	0.781
Placed	11	1.2727	4.819		
Gain—Vocabulary					
Volunteered	5	6.6000	5.899	0.24	0.813
Placed	11	7.5455	7.764		

Math and Verbal scores, reported by CEEB on a 200-800 scale, were converted to the 20-80 Reading and Vocabulary scale for this analysis. Thus gain scores are based on two digits rather than three.

Effects of Taking Senior Math on Math Scores

To determine if senior math increased SAT-Math scores, a t-test was performed on SAT-Math gain scores for subjects taking math and those not taking math. The difference between these groups was not significant (Table 6).

Table 6

Effects of Taking Senior Math on Math Scores

Variable	Number of Cases	Mean Gain Scores	Standard Deviation	T Value	2-Tail Probability
Gain—Math					
Math	35	1.5714	5.271	-1.23	0.223
No Math	17	-0.2353	4.191		

Math and Verbal scores, reported by CEEB on a 200-800 scale, were converted to the 20-80 Reading and Vocabulary scale for this analysis. Thus gain scores are based on two digits rather than three.

A t-test was also performed on Math gain scores for those subjects not taking math but coached and those neither taking math nor coached. Table 7 indicates no significant difference between these two groups.

Effects of Sex on Coaching

To answer the question of whether the coaching program had more of an effect on male or female students, a t-test

Table 7

Effects of Coaching on Students not Taking Senior Math

Variable	Number of Cases	Mean	Standard Deviation	T Value	2-Tail Probability
Gain—Math					
No math—coached	11	0.0909	4.230	-0.42	0.678
No math—not coached	6	-0.8333	4.446		

Math and Verbal scores, reported by CEEB on a 200-800 scale, were converted to the 20-80 Reading and Vocabulary scale for this analysis. Thus gain scores are based on two digits rather than three.

examined gain scores on all four dependent variables, SAT-M, SAT-V, Vocabulary subtest and Reading subtest, with comparison made on the variable of sex. Table 8 indicates that males had higher means on gain scores in each case, and the difference is statistically significant on Math gain scores.

Score Changes on Pretest Scores Below 450 Versus Pretest Scores Above 450

To determine if low-scoring students gain more than higher-scoring students, a t-test compared gains made by all those students with pretest scores on or below 450 (45 on the subtests) with gains made by all those with pretest scores above 450 (Verbal and Math) or 45 (Vocabulary and Reading). Results indicate there is a significant difference only for the Vocabulary subtest, with those having pretest scores on or below 45 showing greater gains (Table 9).

Comparison of All Subjects' Posttest Scores with Pretest Scores

The lack of significance found in the preceding results led to an analysis of all subjects' pre and posttest scores examined together to determine if posttest means were significantly higher than pretest means. A matched pairs t-test indicated that the Verbal section and the Vocabulary subsection showed significant overall increases. The Math section and the Reading subsection did not (Table 10).

Table 8

Effects of Sex on Coaching

Variable	Number of Cases	Mean Gain Scores	Standard Deviation	T Value	2-Tail Probability
Gain—Math					
Males	9	2.3333	3.317	-3.17	0.007
Females	7	-3.5714	4.158		
Gain—Verbal					
Males	9	6.0000	5.500	-1.77	0.098
Females	7	1.8571	3.132		
Gain—Reading					
Males	9	2.4444	4.667	-1.53	0.147
Females	7	-0.7143	3.147		
Gain—Vocabulary					
Males	9	9.4444	7.796	-1.47	0.165
Females	7	4.4286	5.159		

Math and Verbal scores, reported by CEEB on a 200-800 scale, were converted to the 20-80 Reading and Vocabulary scale for this analysis. Thus gain scores are based on two digits rather than three.

Table 9

Score Changes on Pretest Scores Below 450
Versus Pretest Above 450

Variable	Number of Cases	Mean Gain Scores	Standard Deviation	T Value	2-Tail Probability
Gain—Math					
450 or below	6	4.5000	4.461	-1.89	0.065
Above 450	46	0.5217	4.898		
Gain—Verbal					
450 or below	10	5.3000	6.945	-1.47	0.147
Above 450	42	2.5714	4.809		
Gain—Reading					
450 or below	10	3.8000	7.099	-1.74	0.088
Above 450	42	-0.0952	6.199		
Gain—Vocabulary					
450 or below	18	9.2778	8.223	-2.44	0.018
Above 450	34	3.7647	7.484		

Math and Verbal scores, reported by CEEB on a 200-800 scale, were converted to the 20-80 Reading and Vocabulary scale for this analysis. Thus gain scores are based on two digits rather than three.

Table 10

Comparison of All Subjects' Posttest Scores with Pretest Scores

Variable	Number of Cases	Mean	Standard Deviation	T Value	2-Tail Probability
Math—Pre		56.7308	9.107		
Math—Post	52	57.7115	9.123	-4.42	0.161
Verbal—Pre		50.6346	7.944		
Verbal—Post	52	53.7308	8.559	-4.20	0.001
Reading—Pre		51.9038	9.304		
Reading—Post	52	52.5577	8.110	0.73	0.471
Vocab—Pre		49.4615	9.403		
Vocab—Post	52	55.1346	9.565	-5.04	0.001

Large Score Increases

In order to examine whether the coaching program increased the probability of individuals making large increases in scores, the percentage of students in each group scoring above the 35 point expected gain, but no higher than a 100 point gain, and those making a 100+ Verbal gain is given in Table 11. A comparable table was compiled for the SAT-Math scores (Table 12), the SAT-Vocabulary scores (Table 13), and the SAT-Reading scores (Table 14). The Vocabulary and Reading score increases are on a two-digit scale. There was negligible difference among the three groups in regard to likelihood of individual large score increases.

Table 11

Pretest to Posttest Score Differences on SAT-Verbal

	35-99 points			100 or more points		
	E	PC	NPC	E	PC	NPC
Increase	6 (38%)	3 (30%)	10 (40%)	2 (13%)	---	2 (8%)
Decrease	---	---	5 (20%)	---	---	---

E = Experimental group; PC = Practiced controls; NPC = No practice controls.

Table 12

Pretest to Posttest Score Differences
on SAT-MATH

	35-99 points			100 or more points		
	E	PC	NPC	E	PC	NPC
Increase	3 (19%)	2 (18%)	6 (24%)	---	---	2 (8%)
Decrease	6 (38%)	2 (18%)	4 (16%)	---	---	---

E = Experimental group; PC = Practiced controls;
NPC = No Practice controls.

Table 13

Pretest to Posttest Score Differences
on SAT-VOCABULARY

	3-9 points			10 or more points		
	E	PC	NPC	E	PC	NPC
Increase	4 (25%)	2 (18%)	3 (12%)	6 (38%)	3 (30%)	9 (36%)
Decrease	1 (6%)	2 (18%)	3 (12%)	---	---	---

E = Experimental group; PC = Practiced controls;
NPC = No practice controls.

Table 14

Pretest to Posttest Score Differences on SAT-Reading

	3-9 points			10 or more points		
	E	PC	NPC	E	PC	NPC
Increase	2 (13%)	1 (9%)	5 (20%)	1 (6%)	1 (9%)	2 (8%)
Decrease	3 (19%)	3 (30%)	4 (16%)	---	---	3 (12%)

E = Experimental group; PC = Practiced controls; NPC = No practice controls.

CHAPTER IV

DISCUSSION

This study was designed to examine the effects of a program of coaching on SAT scores. Concurrent with this examination, the effects of practicing on one's own was determined.

The coaching program was presented to 24 students, 16 of whom served as subjects for the study. Four of the others were required to take a Latin Achievement Test on December 6 and had to postpone taking their second SAT until another day. Two students were juniors. Another had no pretest scores and another had February, instead of April, pretesting.

The purpose of the coaching program was to enable students to approach the SAT in a more systematic fashion. While it is evident that some students reach their ceiling level, after which little gain can be made, the normal expectation in gains from an April to December testing is about 35 points. Not only were no significant differences found between groups on any SAT section in this study, but when all 52 subjects were analyzed as a whole, the only increase greater than 35 points was on the Vocabulary posttest where the increase was a mean of 6 points (60 when considered in terms of a three-digit scale). The total 52 subject mean

increase for Math was 9.98 points; for Verbal, 30.96; and for Reading, 1.83 (10.83 when considered in terms of 35 point increases). Only the overall gains on Verbal and Vocabulary from pre to posttest were considered statistically significant.

Expectation and motivation (defined respectively as whether one thought he/she would improve and whether one needed high scores for the college of choice) had no significant effect on scores. Neither did the motivation implied in volunteering for the coaching program have an effect.

In contrast to Dyer's (1953) finding and in support of French's (1955) finding, taking math in the senior year had no effect. In contrast to French's (1955) evidence that girls benefited more than boys from coaching, there were no significant differences between male or female students regarding the effects of coaching except for Math scores which were significantly higher for males. It may be that males in the coaching program identified with the male reviewer who was used two days for math coaching.

The only statistical significance found in terms of the questions listed at the end of Chapter I were:

1. Males in the coached group improved their Math scores significantly more than coached females.

2. Prescores below 45 on Vocabulary were improved significantly more than Vocabulary prescores above 45.

Another way of examining scores was to determine the frequency of large increases. Even though the coaching program did not have a significant effect on the average scores of the experimental group, it was of concern to determine if the program increased the probability of individuals making 35+ increases in scores. It appears not to have increased that probability.

Despite the general lack of improvement from first to second testing, there were several large increases in all three groups. In an attempt to determine what factors influenced the individual large score changes, an evaluation was done on score increases or decreases in excess of 100 points (10 on the subtests). Each student who obtained such scores was interviewed.

No coached students experienced decreases of more than 100 (10). Six coached students had increases of 100+ (10+) and two of the six increased 100+ in more than one area. All but one attributed the increases to something he/she did in the coaching sessions. The one exception attributed his 110 Verbal increase and concomitant 10 point Vocabulary subscore increase to "studying vocabulary books all summer—learning lots of new words." The other five subjects, all of whom had Vocabulary increases, attributed their score increases to vocabulary review and to doing a lot of vocabulary exercises during the coaching sessions. They did indeed do more

vocabulary exercises than any other type, as shown by the record of work in their folders. Two of these subjects also mentioned their senior English class as instrumental in score increases, and when asked "in what way?", one responded with "the vocabulary review we've been doing" and the other answered "all that reading." Both of these students' 10+ increases were in vocabulary alone. Two of the experimental subjects contended that familiarity with the testing procedure was the key, and one said that in knowing the process, he was more confident in his abilities. There were no Math score increases of 100 or more points for the experimental subjects.

The practiced controls also sustained no decreases of more than 100 (10) points and no SAT-M increases of more than 100 points. Four members of this group, however, improved 10+ points on Verbal subtests (three on Vocabulary and one on Reading). The student with the Reading increment attributed it to "studying vocabulary and reading a lot this year." Three of the others in this group attributed their increased vocabulary scores to the vocabulary studies done in their senior English classes. Some of the words they had learned just this year were on the December test. One of these two students also attributed his vocabulary increase to extra reading, and the other also mentioned that her April scores were much lower than she had expected. The other student

said she had also studied the vocabulary section of a "Preparing for SATS" book.

Of the no-practice control subjects, three students had decreases of more than 100 (10) points. One of these students whose reading score dropped said he "just didn't try—scores in the other sections went down some too." When asked why he didn't try this time, he replied that he was already accepted at a college which didn't require SAT scores so how he did on the test didn't matter. The other two students with such significant decreases also had significant increases. One student whose vocabulary subscore came up 25 points and whose reading subscore went down 12 points remarked that the test this time didn't have as many hard analogies as the April testing, that the vocabulary study in English used many of the same words which were on the December test and that the reading passages were more difficult and less interesting the second time. While this student's opinion was not based on systematic objective data, it is true that a great deal of variability exists among different forms of the same test. A student's strengths could be tapped with one form and his/her weakness with another form. Any test is a sample of a student's performance and thus any score is subject to sampling error. It is possible that many individual score changes are due to this type of error. The other student who had a Vocabulary increase and Reading

decrease said he attributed it to luck. He "had no idea." Two other students with 100+ increases, one of which was on SAT-M, responded with, "no idea unless it was just having taken the test once before." Two additional students in this group felt that their first scores were unrealistically low and the December score increases weren't really caused by anything. They were just the true scores. Three students mentioned their English class vocabulary drill as causes for both Vocabulary and Reading increases. One of these students added, "it's hard to tell while you're taking the test how you're doing."

One of the no-practice controls attributed reading on his own during the summer to his Vocabulary increase, and another attributed his Vocabulary increase to a good night's sleep the night before.

In contrast to the member of this group who had a 10+ decrease because he didn't need the scores at all, is another student who also didn't need the scores but who increased 100+ (10+) in everything except SAT-M. He replied that he was so much more relaxed knowing nothing crucial was involved in the testing and knowing his first scores were pretty good anyway so he was "free" to improve.

Caution should be exercised in generalizing from the results of this study. Several factors prevent conclusive indication that coaching is ineffective in increasing SAT

scores. Because subjects' pretest scores were higher than the national mean, a general regression toward the mean might have occurred. Perhaps if the experimental group had not been coached, those students might have scored much lower than those in the control groups. This is not likely, however, since all the subjects resemble each other. Because of the superior nature of The Webb School, there was little that could be done in a coaching program which had not already been done in regular classes. Also, it cannot be determined if the coaching methods used were the most appropriate for the type of students used in the study. Perhaps the coaching activities were too unrelated to the skills and principles reflected on the tests. To design a coaching program to be consistent with the tests we need to identify more closely what the tests actually measure.

Student motivation may be the key factor in the success or lack of success in coaching. The experimental subjects in this study did not appear to be especially motivated during the coaching sessions. Thus for significant improvement to occur there seems to be a need for room for improvement, some evidence that the overall academic program has been deficient, more item analysis in order to know specifically what the tests measure, and students who want to improve their scores. Future investigations which combine the latter consideration might well generate beneficial effects from coaching.

Conclusions

1. There were no differences among groups on the four dependent variables.

2. Those who needed high SAT scores to get into the college of their choice did not do significantly better than those who did not.

3. Those who expected to improve did not do significantly better than those who did not expect to improve.

4. Those taking a senior year math course did not obtain significantly higher math scores than those not in senior math.

5. Volunteering for the coaching program had no significant effect.

6. Males in the coached group improved their math scores significantly more than coached females, but there were no statistically significant differences on Verbal, Reading or Vocabulary scores, with comparison made on the variable of sex.

7. Low-scoring students (below 450 or 45 on pretesting) posttested significantly higher than those with pretest scores below 450 (45) on Vocabulary but not on Math, Verbal or Reading.

BIBLIOGRAPHY

BIBLIOGRAPHY

Angoff, W. H. The College Board admissions testing program. Princeton, New Jersey: Educational Testing Service, 1971. (ERIC Document Reproduction Service No. ED 013 371).

Brownstein, S. C. Barron's mathematics workbook for college entrance examinations. New York, New York: Barron's Educational Series, Inc., 1975.

C.E.E.B. A statement by the College Board trustees on test coaching. College Board Review, 1958, 38, 3.

C.E.E.B. Effects of coaching on scholastic aptitude test scores. (College Entrance Examination Board Research Report). Princeton, New Jersey: Educational Testing Service, 1968.

C.E.E.B. Guide for high schools and colleges (booklet for secondary schools and colleges). Princeton, New Jersey: College Entrance Examination Board, 1974.

C.E.E.B. Your College Board ATP report (booklet for students). Princeton, New Jersey: Educational Testing Service, 1971.

Chan, J. The art of test taking. McCalls, April 1975, 38.

DeMott, B. Beyond the S.A.T.s: A whole person catalog that works? Saturday Review World, 1974, 1, 68070.

Dickter, M. R. Scholastic aptitude test scores and college grades in mathematics and science. Unpublished doctoral dissertation, University of Pennsylvania, 1937.

Duggan, J. M. Puzzles and powers in junior S.A.T. scores. College Board Review, 1959, 37, 37.

Dyer, H. S. Does coaching help? College Board Review, 1953, 19, 331-335.

Fine, B. How to be accepted by the college of your choice. Great Neck, New York: Channel Press, 1960.

Frankel, E. Effects of growth, practice and coaching on S.A.T. scores. Personnel and Guidance Journal, 1960, 38, 713-719.

French, J. W. An answer to test coaching. College Board Review, 1955, 27, 5-7.

French, J. W., & Dear, R. E. Effect of coaching on an aptitude test. Educational and Psychological Measurement, 1959, 19, 319-330.

Furst, E. J. Question of abuses in the use of aptitude and achievement tests. Theory into Practice, 1963, 2, 199-204.

Goslin, D. A. Criticism of standardized tests and testing. Science, 1968, 159, 851-855.

Goslin, D. A. What's wrong with tests and testing—part II. College Board Review, 1968, 66, 33-37.

Hoy, J. C. Choosing a college. New York: Dell, 1967.

Jacobs, P. I. Large score changes on the scholastic aptitude tests. Personnel and Guidance Journal, 1966, 45, 150-156.

Karas, S., Relles, F., Barbara S., & Watkins, R. Change in scholastic aptitude test scores of candidates tested in the junior and senior years 1960-63. (College Entrance Examination Board Research and Development Report). Princeton, New Jersey: Educational Testing Service, 1963.

Lass, A., & Wilson, E. The college student's handbook. New York: David White Company, 1965.

Levine, R. S., & Angoff, W. H. The effect of practice and growth on scores on the scholastic aptitude test. (C.E.E.B. Statistical Report). Princeton, New Jersey: Educational Testing Service, 1958.

Lukee, A. G. Scholastic aptitude and scholastic achievement at a college of education. Unpublished doctoral dissertation, University of Northern Colorado, 1951.

Millman, J., Bishop, C., & Elul, R. An analysis of test wiseness. Educational and Psychological Measurement, 1965, 25, 707-727.

Moll, R. W. Optional S.A.T.—two years of testing the tradition of tests at Bowdoin. College Board Review, 1973, 86, pp. 10-13, 39-41.

Morgan, C. T., & Deese, J. How to study. New York: McGraw-Hill, 1957.

Moore, J. C., Schutz, R. E., & Baker, R. L. The application of a self-instructional technique to develop a test-taking strategy. American Education Research Journal, 1966, 3, 3-17.

Mozur, J. L. Validity of scholastic aptitude scores as predictors of achievement. Unpublished doctoral dissertation, Case Western Reserve University, 1968.

Oakland, T., & Wailert, E. The effects of test-wiseness materials on standardized test performance of preschool disadvantaged children. Austin, Texas: The University of Texas, 1971. (ERIC Document Reproduction Service No. ED 047 050.

Olsen, M. Summary of the main findings on the validity of the College Entrance Examination Board tests of developed ability as predictors of college grades. (C.E.E.B. Statistical Report). Princeton, New Jersey: Educational Testing Service, 1957.

Pallone, N. J. Effects of short- and long-term developmental reading courses upon S.A.T. verbal scores. Personnel and Guidance Journal, 1961, 39, 654-657.

Pearson, R. Effects of growth and retesting on S.A.T. verbal scores. College Board Review, 1948, 1, pp. 57; 68-71.

Pierson, G. N. The effectiveness of high school rank and S.A.T. scores in predicting academic success of freshman students. Unpublished doctoral dissertation, University of Denver, 1960.

Richardson, F. C. A self-study manual for students on coping with test-taking anxiety. Austin, Texas: The University of Texas, 1973. (ERIC Document Reproduction Service No. ED 084 835).

Roberts, S. O., & Oppenheim, D. B. The effect of spacial instruction upon test performance of high school students in Tennessee. (College Entrance Examination Board Research and Development Report). Princeton, New Jersey: Educational Testing Service, 1966.

Stalmaker, R. C., & Stalmaker, J. M. Effect on a candidate's score or repeating the scholastic aptitude test of College Entrance Examination Board. Journal of Educational and Psychological Measurement, 1946, 6, 495-503.

Wedeen, S. U. Uses and misuses of aptitude tests. Clearing House, 1960, 35, 11-12.

Weiner, M. Barron's verbal aptitude workbook for college entrance examination. New York: Barron's Educational Series, Inc., 1975.

White, J., & Yanker, G. College board crash courses: Psyching out the tests? Or psyching up the takers? Saturday Review of Education, 1972, 55, 71-72.

Whitla, D. K. Effect of tutoring on scholastic aptitude test scores. Personnel and Guidance Journal, 1962, 41, 32-37.

Wilson, E. S., & Buckner, C. A. College ahead! New York: Harcourt, Brace and Co., 1961.

APPENDICES

APPENDIX A

MATH CONTENTS

MATH CONTENTS

1. FUNDAMENTALS

UNIT A Addition	1	UNIT F Decimals	5
UNIT B Subtraction	2	UNIT G Per Cent	6
UNIT C Multiplication	2	Practice Test on Fundamentals	7
UNIT D Division	3	ANSWERS	8
UNIT E Fractions	4		

2. VERBAL PROBLEMS

UNIT A Fractions	11	UNIT I Mixture	63
UNIT B Algebraic Operations	19	UNIT J Geometry	69
UNIT C Using Algebra	25	UNIT K Data Sufficiency	77
UNIT D Per Cent	31	UNIT L Interpreting Data	83
UNIT E Average	39	UNIT M Quantitative Comparisons	95
UNIT F Ratio and Proportion	45	UNIT N Number Series	99
UNIT G Motion	51	UNIT O Figure Analogies	101
UNIT H Work	57	ANSWERS AT THE END OF EACH UNIT	

3. APTITUDE TESTS

TEST I	107	TEST VII	143
TEST II	113	TEST VIII	149
TEST III	119	TEST IX	155
TEST IV	125	TEST X	161
TEST V	131	ANSWERS AT THE END OF EACH TEST	
TEST VI	137		

REFERENCE TABLES 167

TABLE OF SQUARES & SQUARE ROOTS 169

APPENDIX B

VERBAL CONTENTS

VERBAL CONTENTS

1. VOCABULARY

EXERCISE I	1	EXERCISE XIV	8
EXERCISE II	2	EXERCISE XV	8
EXERCISE III	2	EXERCISE XVI	9
EXERCISE IV	3	EXERCISE XVII	9
EXERCISE V	3	EXERCISE XVIII	10
EXERCISE VI	4	EXERCISE XIX	10
EXERCISE VII	4	EXERCISE XX	11
EXERCISE VIII	5	EXERCISE XXI	12
EXERCISE IX	5	EXERCISE XXII	12
EXERCISE X	6	EXERCISE XXIII	13
EXERCISE XI	6	EXERCISE XXIV	13
EXERCISE XII	7	EXERCISE XXV	14
EXERCISE XIII	7	ANSWERS	14

2. SENTENCE COMPLETION

EXERCISE I	17	EXERCISE XI	25
EXERCISE II	18	EXERCISE XII	25
EXERCISE III	19	EXERCISE XIII	26
EXERCISE IV	19	EXERCISE XIV	27
EXERCISE V	20	EXERCISE XV	28
EXERCISE VI	21	EXERCISE XVI	29
EXERCISE VII	22	EXERCISE XVII	30
EXERCISE VIII	22	EXERCISE XVIII	30
EXERCISE IX	23	EXERCISE XIX	31
EXERCISE X	24	EXERCISE XX	32
		ANSWERS	33

3. WORD RELATIONSHIPS

EXERCISE I	34	EXERCISE VI	37
EXERCISE II	35	EXERCISE VII	38
EXERCISE III	35	EXERCISE VIII	38
EXERCISE IV	36	EXERCISE IX	39
EXERCISE V	36	EXERCISE X	39

EXERCISE XI 40
EXERCISE XII 41
EXERCISE XIII 41
EXERCISE XIV 42
EXERCISE XV 42
EXERCISE XVI 43
EXERCISE XVII 44
EXERCISE XVIII 44
EXERCISE XIX 45
EXERCISE XX 45
ANSWERS 46

4. READING

EXERCISE I 48
EXERCISE II 51
EXERCISE III 54
EXERCISE IV 57
EXERCISE V 60
EXERCISE VI 62
EXERCISE VII 65
EXERCISE VIII 68
EXERCISE IX 71
EXERCISE X 74
EXERCISE XI 77
EXERCISE XII 80
EXERCISE XIII 83
EXERCISE XIV 85
EXERCISE XV 88
ANSWERS 90

5. TEN VERBAL APTITUDE TESTS

EXERCISE I 92
EXERCISE II 94
EXERCISE III 97
EXERCISE IV 99
EXERCISE V 102
EXERCISE VI 104
EXERCISE VII 106
EXERCISE VIII 109
EXERCISE IX 111
EXERCISE X 113
ANSWERS 116

APPENDIX C

EXERCISE PROGRESS CHART FOR MATH FUNDAMENTALS

EXERCISE PROGRESS CHART FOR MATH

FUNDAMENTALS

EXERCISE PROGRESS CHART

TO THE STUDENT: *When this chart is completed you will have a clear picture of the extent of your preparation. It will point out possible areas in which you need additional drill.*

Fundamentals	*Page*	*Date*	*Time Begun*	*Time Completed*	*Number Completed*	*Number Not Completed*	*Number Correct*
Addition	1						/20
Subtraction	2						/20
Multiplication	2						/20
Division	3						/20
Fractions	4						/40
Decimals	5						/20
Per Cent	6						/20
Practice Test	7						/25

APPENDIX D

TIMETABLE FOR THE SAT

TIMETABLE FOR THE SAT

Total Time: 3 Hours

9:00 to 9:30	Part 1 **Verbal Aptitude Test**	10 Vocabulary Questions 10 Word Relationship Questions 10 Sentence Completion Questions 10 Reading Comprehension Questions
9:30 to 10:00	Part 2 **Verbal Aptitude Test**	10 Vocabulary Questions 10 Word Relationship Questions 10 Sentence Completion Questions 10 Reading Comprehension Questions
10:00 to 10:30	Part 3 **Test of Standard Written English**	50 Grammar and Usage Questions
10:30 to 11:00	Part 4 **Mathematics Aptitude Test**	25 or 40 Questions
11:00 to 11:05	RECESS	
11:05 to 11:35	Part 5 **Mathematics Aptitude Test**	25 or 40 Questions
11:35 to 12:05	Part 6 Verbal Aptitude Test (40 Questions) or Mathematics Aptitude Test (25 or 40 Questions)	

APPENDIX E

SAMPLE VERBAL EXERCISE

SAMPLE VERBAL EXERCISE

Select best relationship.

16. epaulet : shoulder :: 1 medal : chest 2 knapsack : back 3 sash : window 4 sword : scabbard 5 decoration : uniform 16. ____
17. lungs : blood :: 1 heart : circulation 2 arteries : veins 3 carburetor : car 4 glands : secretions 5 carburetor : gasoline 17. ____
18. adamant : unyielding :: 1 diamond : glittering 2 incoherent : sensible 3 invincible : loquacious 4 warning : admonitory 5 hard : nails 18. ____
19. eulogy : castigation :: 1 cemetery : church 2 praise : denounce 3 praise : worship 4 worship : church 5 curse : bless 19. ____
20. marsupial : kangaroo :: 1 Africa : Australia 2 crustacean : lobster 3 egg : pouch 4 larva : egg 5 fly : hop 20. ____
21. scissors : sever :: 1 bandages : suture 2 millstone : grinds 3 tinkers : dam 4 dam : water 5 cutters : stern 21. ____
22. meat : decayed :: 1 bread : stale 2 fish : live 3 odor : fetid 4 food : defective 5 quality : inferior 22. ____
23. ascetic : acrimonious :: 1 monk : shrew 2 bitter : harmonious 3 acid : alkali 4 acid : base 5 sweet : shrewish 23. ____
24. mutton : sheep :: 1 calf : cow 2 goal : filly 3 venison : deer 4 doe : deer 5 elephant : cow 24. ____
25. mine : me :: 1 his : it 2 his : her 3 his : its 4 his : him 5 his :it's 25. ____
26. sheep : wool :: 1 fodder : animal 2 otter : fur 3 plant : cotton 4 animal : vegetable 5 stupid : good 26. ____
27. fox : cunning :: 1 fields : court 2 beaver : insouciant 3 beaver : industrious 4 vixen : cute 5 hunt : hounds 27. ____
28. novelty : tradition :: 1 revolution : custom 2 sport : heredity 3 innovation : intrusion 4 new : custom 5 false : trying 28. ____
29. perjury : oath :: 1 false : witness 2 embezzlement : trust 3 automobile : defective 4 court : testimony 5 jury : faith 29. ____
30. bushel : potatoes :: 1 basket : fruit 2 carrot : diamond 3 wood : cord 4 points : diamond 5 bricks : mortar 30. ____

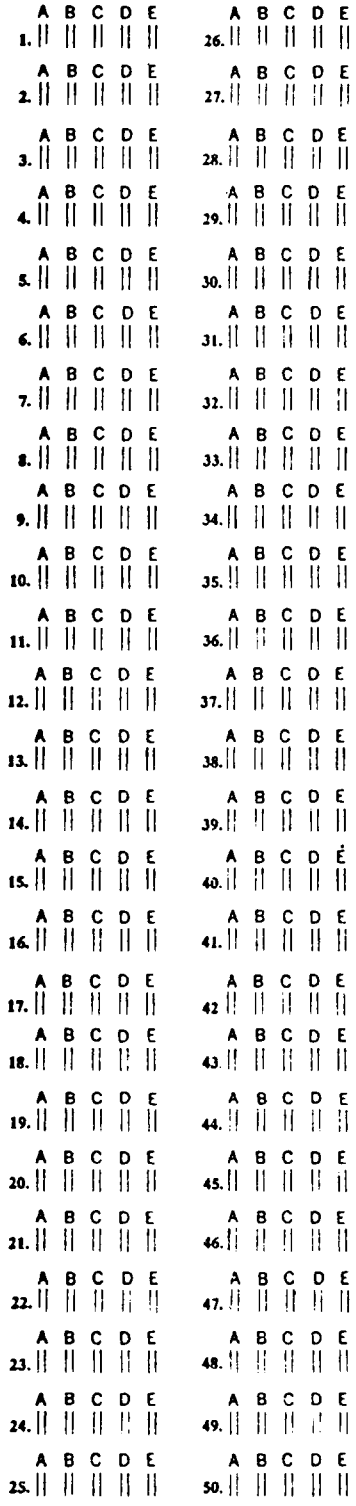
Answers on page 46

APPENDIX F

SAMPLE MATH ANSWER SHEET PLACED IN STUDENTS' FOLDERS FOR RECORD KEEPING

SAMPLE MATH ANSWER SHEET PLACED IN STUDENTS'

FOLDERS FOR RECORD KEEPING



Fold on the dotted line for the answers

Date _____

Time Begun _____

Time Completed _____

Number Completed _____

Number Not Completed _____

Number Correct _____

APPENDIX G

SAMPLE VERBAL CHARTS PLACED IN STUDENTS' FOLDERS FOR RECORD KEEPING

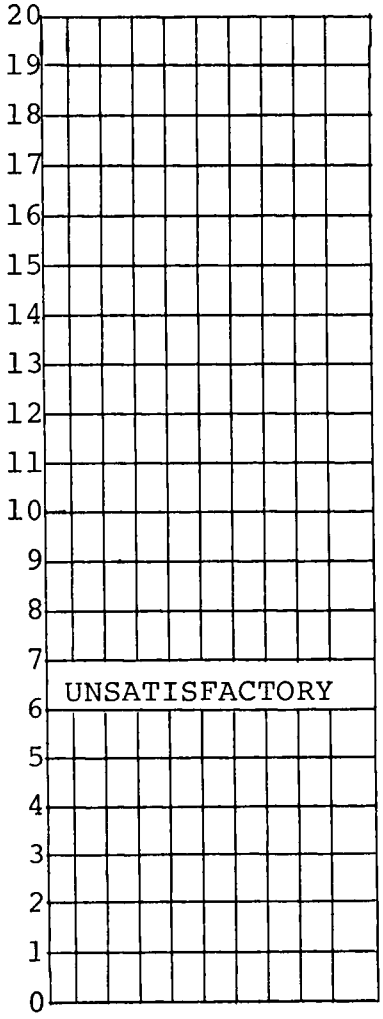
SAMPLE VERBAL CHARTS PLACED IN STUDENTS'

FOLDERS FOR RECORD KEEPING

CHART YOUR PROGRESS

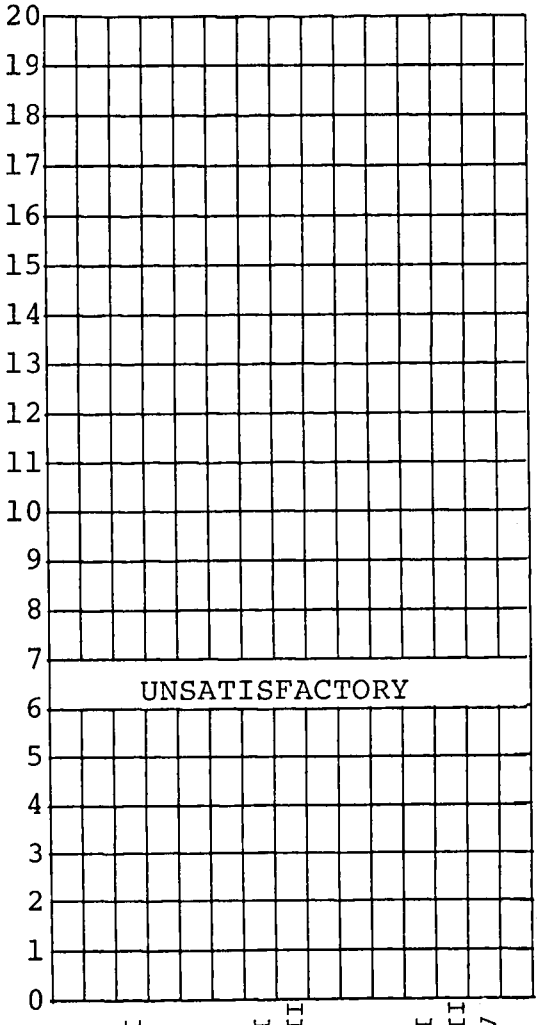
Find out: How good you are
Whether you are improving
Areas of strength and weakness

I
VOCABULARY
(Synonyms)



1._____ 6._____
2._____ 7._____
3._____ 8._____
4._____ 9._____
5._____ 10._____

II
VOCABULARY
(Antonyms)



11.____ 16._____ 21._____
12.____ 17._____ 22._____
13.____ 18._____ 23._____
14.____ 19._____ 24._____
15.____ 20._____ 25._____

VITA

Linda Williams Keefauver was born in Greeneville, Tennessee, on March 17, 1946. She was graduated from Greeneville High School in 1964. In 1968 she graduated cum laude from East Tennessee State University with a Bachelor of Science degree in Elementary Education. There she served as president of Kappa Delta Pi and was twice named to Who's Who in American Colleges and Universities. While working two years as a public elementary school teacher, she received a Master of Science degree in Educational Psychology in 1970, from The University of Tennessee, Knoxville. She then worked as a guidance counselor in two private high schools and entered the doctoral program in Educational Psychology and Guidance at UTK during the winter of 1974. She was awarded a departmental teaching assistantship for the 1975-76 academic year. During her doctoral work she was initiated into membership by Pi Lambda Theta. She completed the requirements for the Doctor of Education degree with a major in Educational Psychology and Guidance in May, 1976, and was graduated in August, 1976.