DOCUMENT RESUME

ED 053 158 TM 000 667

AUTHOR	Roberts, S. O.; Oppenheim, Don B.
TITLE	The Effect of Special Instruction Upon Test Performance of High School Students in Tennessee.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	RB-66-36
PUB DATE	Jul 66
NOTE	29p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Disadvantaged Youth, Grade 11, Instructional Materials, Mathematics, Negro Students, *Performance, Post Testing, Pretesting, Programed Materials, *Remedial Instruction, Remedial Programs, Scores, Special Programs, *Testing, *Tutorial Programs, Verbal Ability
IDENTIFIERS	Preliminary Scholastic Aptitude Test, PSAT, STEP

ABSTRACT

This study was undertaken to determine whether students who had received less adequate instruction and were less able in scholastic aptitude test performance than the subjects in previous studies might benefit more from special instruction. Specially prepared linear programed materials were developed in the verbal and the mathematical areas. Instruction in either of the two areas was given to eleventh grade students in predominantly Negro high schools. Alternate forms of the PSAT, administered as pre-and posttests, were used to evaluate the effectiveness of the special instruction. While there were statistically significant differences in gain scores between some of the experimental groups and their controls, the outcome of this study was generally negative. The gains in the experimental groups were more than offset by reductions in posttest scores in the control groups. The magnitude of the gains observed in the experimental groups were so small that it does not seem reasonable to expect that similar short-term instruction given on a wide scale would be of significant benefit to disadvantaged students. (Author/AG)

COLLEGE ENTRANCE EXAMINATION BOARD
RESEARCH AND DEVELOPMENT REPORTS

RDR-66-7, No. 1

•

RESEARCH BULLETIN
RB-66-36 JULY 1966

The Effect of Special Instruction Upon Test Performance of High School Students in Tennessee

•

S. O. Roberts
Fisk University

and

Don B. Oppenheim
Educational Testing Service

EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY
BERKELEY, CALIFORNIA

College Entrance Examination Board

RESEARCH AND DEVELOPMENT REPORTS

RDR-66-7, No. 1

THE EFFECT OF SPECIAL INSTRUCTION UPON TEST PERFORMANCE OF HIGH SCHOOL STUDENTS IN TENNESSEE

S. O. ROBERTS, Fisk University

and

DON B. OPPENHEIM, Educational Testing Service

Research Bulletin

RB-66-36

July 1966

EDUCATIONAL TESTING SERVICE

Princeton, New Jersey

The Effect of Special Instruction upon Test Performance of High School Students in Tennessee

Abstract

Previous studies of the effect of coaching on test scores have been essentially negative. The present study was undertaken to determine whether students who had received less adequate instruction and who were less able in scholastic aptitude test performance than the subjects in previous studies might benefit more from special instruction.

Specially prepared linear programed materials, designed to foster an analytic approach to the tasks required on college admission tests and to be used in conjunction with classroom instruction, were developed in the verbal and the mathematical areas. Instruction in either of the two areas was given to eleventh grade students in 18 predominantly Negro high schools. Alternate forms of the PSAT, administered as pre- and posttests, were used to evaluate the effectiveness of the special instruction.

While there were statistically significant differences in gain scores between some of the experimental groups and their controls, the outcome of this study, like the outcome of earlier studies, was generally negative. In fact, the gains in the experimental groups were more than offset by reductions in posttest scores in the control groups. The magnitude of the gains observed in the experimental groups were so small that it does not seem reasonable to expect that similar short-term instruction given on a wide scale would be of significant benefit to disadvantaged students.

The Effect of Special Instruction upon Test Performance of High School Students in Tennessee[1]

Over the past few decades a number of investigators such as Klineberg (1935) and Anastasi (1958) have found, and others as noted by Tumin (1963) have hypothesized, that the educational environment of culturally deprived students is not as stimulating nor as demanding as that of the population at large. Culturally deprived students might, therefore, be identified as a group that is in the early stages of learning where gains due to practice are the greatest. If this is true, some instruction in test-taking techniques might improve their performance on tests such as those typically used for college admissions. Such improvement would lead to social gains as well as to gains made from the instruction itself because the resulting higher test scores would enhance the chances of being accepted for a college education. Thus, if this kind of special instruction in test-taking, or coaching, were successful, and were made available to schools, it could make an important contribution to the remediation of the culturally deprived.

Although findings by those who have explored the effects of coaching on test scores (Dyer, 1953; French, 1955; French & Dear, 1959) have been essentially negative, it is reasonable on theoretical grounds to expect that special instruction might help the culturally deprived. Previous studies to determine the effects of coaching upon performance on the Scholastic Aptitude Test (SAT) have generally been conducted with students who were academically oriented and who had received a rather good level of instruction (Malcolm, 1961). However, in their review of the effects of coaching on aptitude test scores, French and Dear (1959, p.323) commented that students receiving less adequate education might have benefited more from the coaching. In addition, they suggested that

if the subjects of the experiments had been less able and had had less experience with objective tests, greater gains might have resulted.

To explore these hypotheses, a study was undertaken by the College Entrance Examination Board, Educational Testing Service, and Fisk University in 1965 to investigate the effects of special instruction on eleventh-grade students in predominantly Negro high schools in Tennessee. (It was estimated that the pretest scores of these students would be equivalent to about 300 on the SAT, a level of performance appreciably below that found in previous coaching studies.) The study focused on the question of whether specially prepared programed materials, designed to be used in conjunction with classroom instruction, would enable these students to develop better approaches to certain educational tasks and, as a result, to obtain higher scores on college admission tests.

Procedure

The data from 18 Negro high schools in Tennessee were used in this study. (Schools were not selected at random and therefore cannot be considered to be a normative sample. An effort was made, however, to include schools from different geographical regions in the state and to insure representation of both urban and rural institutions.) Six schools -- three urban and three rural -- were chosen to offer verbal instruction. Eight other schools -- four urban and four rural -- were selected to provide instruction in the mathematics area. Four additional schools -- two urban and two rural -- were selected to offer practice in taking tests but to offer no special instruction.

The subjects in the schools where special instruction was given were volunteers who had indicated that they would like to participate in a program

designed to help high school students perform well on the multiple-choice tests that many colleges require for admission and financial assistance (see Appendix B). Each student was assigned randomly to one of two groups: an experimental group which received instruction after pretesting and a control group which was not instructed until after posttesting. (Those assigned to the latter group were promised that they would receive instruction after the posttests.) The instruction for both the experimental and control groups was given in approximately 15 half-hour sessions spaced over a four- to six-week period. To insure a uniform presentation of material with common content as an experimental control, a linear programed instruction format was adopted. (For a complete statement of the rationale of the instruction, see Appendix A.) There were, however, instructors available to answer questions that might be raised about the programed booklets and to supplement the instruction as they felt appropriate.

The subjects in the four schools where no instruction was offered volunteered for and received only practice on the tests. They took the pre- and posttests at the same times as the experimental and control groups, furnishing an independent estimate of the effect of practice in a setting where there was no chance for diffusion of information from instructed groups to those acting as controls. A tabular summary of the experimental design follows:

Experimental Design

Type of Instruction	Number of Schools	Pre- and Post-tests Given	Group	N	Instruction after Pretest	Instruction after Posttest
Verbal	6	PSAT V+M and STEP Read.	Experimental	150	V only	None
			Control	150	None	V only
Mathematics	8	PSAT V+M and STEP Math.	Experimental	150	M only	None
			Control	150	None	M only
None	4	PSAT V+M and STEP Read. and STEP Math.	No Instruction	120	None	None

Alternate forms of the Preliminary Scholastic Aptitude Test (PSAT) Verbal and Mathematical were administered to all subjects immediately before and immediately after the instruction series as pre- and posttests. The PSAT was chosen as a substitute for the SAT chiefly because of the shorter time required for administration. At those schools where instruction was given using verbal items, pre- and posttests of STEP Reading Level 3 were administered. Where the instruction was given using mathematics items, pre- and posttests of STEP Mathematics Level 3 were used. Both STEP Reading Level 3 and STEP Mathematics Level 3 were given as pre- and posttests at schools where there was no instruction. The STEP Reading and Mathematics Tests were to act as lower-level tests in the event that the PSAT proved to be so difficult for these subjects that a large percentage would receive a scaled score of 20, the lowest score which is reported. While the PSAT scores attained were generally at the low end of the 20-80 scale, there was no appreciable piling up of scores at 20 and, as

a consequence, the analysis of data reported here has been confined to performance on the PSAT.

Results

The means and standard deviations of the pre- and posttest scores for those who received verbal instruction, mathematical instruction, and no special instruction are presented in Tables 1, 2, and 3 respectively. Tables 4 and 5 indicate the differences obtained between mean gain scores for the verbal-instructed and mathematics-instructed groups. The mean gain score for a group is defined as the mean posttest score minus the mean pretest score. Subtraction of the mean gain score for a control group from the mean gain score for the corresponding experimental group yields the difference between mean gain scores. The significance of the difference between mean gain scores was tested with ordinary t tests of the difference between means of uncorrelated observations.

As Table 4 shows, the results obtained from testing at schools where verbal instruction was given indicate a significant effect at some schools but not at others, the differences being more striking in the urban group than in the rural group. It should be noted that at schools b and c, where the results are statistically significant, the gain from pretest to posttest for the experimental group is accompanied by a loss in score for the control group. In fact, the loss in score for the control group occurred at five of the six schools represented in Table 1 and will be considered in the discussion.

Table 5 contains information similar to that in Table 4 for the schools where mathematics instruction was given; the scores analyzed are PSAT-M scores.

It should be noted that score gains or losses throughout the table are quite small. None of the differences in mean gain scores are significant except for a small difference in the urban group.

Discussion

The maximum difference observed between the pre- and posttest mean for any experimental group in the study was only 2.47 points for PSAT-V at school b, a gain not significantly different from zero. Conceptually, this maximum gain contains a practice effect, plus a growth effect, plus an equating error, plus a positive sampling error. It is felt that the sum of the practice, growth, and equating effects can be neglected since the corresponding pre- and posttest means of the groups with no instruction, presented in Table 3, are practically identical. The sampling error is presumed to be positive since 2.47 is a maximum difference taken from 14 differences. Hence, 2.47 is a maximum gain for a coached group and is inflated by a positive sampling error. Since the error of measurement of the PSAT is 3 points, the true gain from coaching, if any, must surely be well within the error of measurement of the PSAT and could be expected to obtain admission to college for only a very few borderline cases. Certainly, the magnitude of gains observed are not such as to expect that the introduction of a widespread practice of similar instruction would lead to major social gains for the disadvantaged student.

One should, however, interpret this conclusion with some caution since only seven and one-half hours of instruction were involved. This is less instruction than has been given in other coaching studies or is generally given in commercial coaching schools. Hence, one need not entirely abandon the hypothesis that these

students can be helped by such special instruction, but rather may conclude that if an improvement in test-taking ability of these students can be accomplished with these instructional materials, a more prolonged and extensive interaction between the students and the materials is required.

The preceding discussion implies a comparison between the mean gains of experimental groups with the mean gains at schools where no instruction occurred and at which the mean gain scores were approximately zero. Such a comparison is contrary to the initial plan of the study which was to compare experimental groups with control groups within schools where instruction was given. The change in plan was due to the unexpected observed drop in scores of the control groups.

Normally, one would expect both experimental and control groups to improve on the posttest due to practice and growth effects. However, in this study, the control group experienced a drop in scores on the posttest. Apparently, not receiving instruction had an effect on these students since a comparable drop in scores on the posttest did not occur in schools where the students received practice but no instruction.

The statistical significance that was found when experimental groups were compared with their controls is attributable to the drop in the control group posttest scores. Hence, there is no reason to believe that instruction would have helped if it had been given to all students. To check this, t tests of correlated means were obtained for the gain scores of the total experimental group receiving verbal instruction and the total experimental group receiving mathematical instruction. Neither of these tests, $t = 1.61$ and $t = .49$, was significant at the five-percent level.

Several possible reasons can be offered for the unexpected decline in scores of the control groups. The control groups in schools b, c, i, j, and n had higher pretest

scores than did their respective experimental groups so that regression could account for some of the decline in posttest scores. Alternatively, the control groups were composed of volunteers for instruction who were told that they would have to wait for a second instructional series. Since they did not receive instruction at the same time as some of their peers, they may have been disappointed, which in turn may have resulted in a loss of interest and motivation. Moreover, the apparent loss in motivation, at least for some of the control subjects, might have resulted from something other than simply a lessening of the desire to do well; some may have felt that they were not being treated fairly by having to wait for the instruction that was being given to others. As a result, they may have been determined not to do well on the posttests. A still stronger interpretation is that at least some of those subjects in the control groups did not believe the promises that they would receive instruction after the first series of instruction was completed. As a result, some of those in the control groups might have acted negatively quite deliberately to "beat the power structure."

The preceding conjectures are all after the fact and cannot be tested. Discussions held with students who were instructed and with instructors immediately after the posttests were administered indicated that both groups were enthusiastic about the programed materials used and the instruction sessions in which they participated. Several of the teachers said they thought the materials were valuable. Both students and teachers expressed concern, however, over the lack of time available to work with the materials. Perhaps the same programs used as a basis for a longer training period and made available to the student to take home and work on at his own pace might prove to be more effective than they were in this experimental setting.

The outcome of this study, like those of earlier studies investigating whether coaching can raise aptitude test scores, is essentially negative. The performance of the experimental groups proved to be lower than was expected. Nevertheless, the question of whether one can intervene effectively to supplement the instruction of the culturally deprived high school student persists. Future investigations might concentrate upon the particular learning problems of this population and what techniques might prove to be effective to overcome these problems rather than take the form of additional coaching studies as they have been performed in the past.

References

Anastasi, Anne. Differential psychology: Individual and group differences in behavior. (3rd ed.) New York: Macmillan, 1958.

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

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.

Klineberg, O. Race differences. New York and London: Harper & Brothers, 1935.

Malcolm, D. J. Summary of studies pertaining to the effect of coaching on the performance of students on the Scholastic Aptitude Test. Unpublished manuscript, Educational Testing Service, 1961.

Tumin, M. M. (Ed.) Race and intelligence. New York: Anti-Defamation League of B'nai B'rith, 1963.

Footnote

[1]While many people have contributed to this study, the authors wish to thank William E. Coffman and Thomas F. Donlon for their help throughout most of the phases of the investigation.

Table 1

PSAT-V Scores of the Verbal-Instructed Groups

School	N		Pretest				Posttest			
			Mean		S.D.		Mean		S.D.	
	Exp.	Cont.	Exp.	Cont.	Exp.	Cont.	Exp.	Cont.	Exp.	Cont.
a	43	37	34.77	34.11	8.9	7.0	34.86	34.08	10.1	8.3
b	19	13	27.95	30.77	5.2	6.8	30.42	28.15	6.9	8.4
c	17	10	25.65	27.20	3.4	2.6	26.71	23.10	5.5	2.6
d	16	11	29.75	26.18	6.7	3.6	30.44	28.91	8.9	5.9
e	20	12	27.00	25.92	3.7	4.7	25.80	25.42	3.9	7.0
f	39	28	34.44	31.75	7.7	7.4	35.69	30.61	8.4	8.4
Urban	101	78	33.36	33.71	8.2	7.2	34.35	31.85	9.0	8.5
Rural	53	33	27.40	26.39	4.7	3.7	27.49	25.88	6.2	5.9
Total	154	111	31.31	30.83	7.8	7.0	31.99	30.07	9.0	8.3

Table 2

PSAT-M Scores of the Mathematics-Instructed Groups

School	N		Pretest				Posttest			
			Mean		S.D.		Mean		S.D.	
	Exp.	Cont.	Exp.	Cont.	Exp.	Cont.	Exp.	Cont.	Exp.	Cont.
g	38	25	35.32	32.48	10.1	6.6	35.92	31.28	9.8	6.1
h	18	13	31.94	31.08	6.0	5.8	32.11	33.46	7.6	5.5
i	19	13	30.58	31.77	5.6	5.5	29.42	29.38	4.0	6.5
j	37	22	29.27	30.64	6.5	4.9	29.38	29.50	6.9	4.1
k	19	11	31.89	27.18	7.0	4.5	31.26	29.09	6.8	4.3
l	17	13	40.18	39.46	9.5	10.4	40.41	38.62	10.6	7.5
m	20	12	39.15	31.08	9.5	5.6	40.45	30.42	8.7	6.0
n	20	13	31.40	33.77	5.6	6.7	31.90	32.08	5.4	4.5
Urban	112	72	34.74	32.94	9.7	7.4	35.25	31.92	9.9	6.6
Rural	76	50	31.45	31.10	6.0	6.0	31.17	31.08	6.0	5.5
Total	188	122	33.41	32.19	8.5	6.9	33.60	31.57	8.7	6.1

Table 3

PSAT Scores of the No-Instruction Groups

		PSAT-V				PSAT-M			
		Pretest		Posttest		Pretest		Posttest	
School	N	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
o	28	29.00	6.30	29.07	7.09	32.35	8.72	34.57	8.35
p	31	26.68	4.36	26.35	5.30	29.35	3.77	28.00	4.42
q	35	33.60	6.29	33.91	6.55	35.37	8.42	35.97	7.43
r	19	26.37	4.91	25.84	6.75	30.53	4.18	28.26	4.38
Urban	63	31.56	6.66	31.76	7.16	33.98	8.63	35.35	7.82
Rural	50	26.56	4.53	26.16	5.83	29.80	3.93	28.10	4.36
Total	113	29.35	6.30	29.28	7.15	32.13	7.23	32.14	7.43

Table 4

Gain Scores of the Verbal-Instructed Groups

School	Experimental Mean	Experimental S.D.	Control Mean	Control S.D.	Difference Between Mean Gain Scores (E-C)
a	.09	5.15	- .03	5.55	.12
b	2.47	5.67	-2.62	4.88	5.09**
c	1.06	4.63	-4.10	3.03	5.16**
d	.69	6.45	2.73	5.52	-2.04
e	-1.20	4.45	- .50	5.71	- .70
f	1.25	5.19	-1.14	5.02	2.39*
Urban	.99	5.29	- .86	5.27	1.85*
Rural	.09	5.20	- .51	5.57	.60
Total	.68	5.26	- .76	5.34	1.44*

* Significant at .05 level.

** Significant at .01 level.

Table 5

Gain Scores of the Mathematics-Instructed Groups

School	Experimental Mean	Experimental S.D.	Control Mean	Control S.D.	Difference Between Mean Gain Scores (E-C)
g	.60	5.27	-1.20	4.77	1.80
h	.17	5.77	2.38	5.11	-2.21
i	-1.16	5.18	-2.39	5.35	1.23
j	.11	5.57	-1.14	5.07	1.25
k	- .63	4.89	1.91	4.50	-2.54
l	.23	5.25	- .84	4.82	1.07
m	1.30	7.35	- .66	5.74	1.96
n	.50	3.77	-1.69	5.74	2.19
Urban	.51	5.70	-1.02	4.95	1.53*
Rural	- .28	4.85	- .02	5.48	- .26
Total	.19	5.31	- .62	5.18	.81

* Significant at .05 level.

where mathematics instruction was given; the scores analyzed are PSAT-M scores.

APPENDIX A

APPENDIX A

A GUIDE FOR INSTRUCTORS IN VERBAL TESTING

Tennessee Testing Project

Spring 1965

Introduction

This guide is intended to describe for the instructors in verbal testing the assumptions upon which the development of the materials to be used in this project are based. Basically, these materials consist of "courses" of programed instruction in five areas: four devoted to specific item types and a fifth devoted to test-taking in general. Additional materials consist of practice tasks intended to permit the application of the principles which the programs present.

The decision to use programed materials was based primarily on the need to reduce variability in instruction. Given more time to prepare, a greater use of lectures or instructions could have been planned, and this variability could be controlled, for example, by workshops for the project instructors. In the absence of such preliminary meetings, the use of programs seemed indicated. A related benefit concerns the reproducibility of the instruction: the programs embody a detailed record of our intentions and should be invaluable to anyone who attempts to replicate this study.

In spite of this highly predetermined curriculum, the role of the instructor is crucial. The programs have never been used by any group of students, and there are bound to be difficulties. The instructor must be familiar with the content of the programs and with the goals of the instruction. He must be prepared to help the students to work the programs and to proceed with the practice tasks. Finally, he must institute a period of "normal" instruction in which he undertakes

to summarize the program for the students, to answer questions and clarify concepts, and to develop and sustain motivation. As will be seen below, it will be difficult to achieve these goals within the time allowed.

Organization of the Course

The five programs are:

(1) Sentence completion

(2) Antonyms

(3) Analogies

(4) Reading comprehension

(5) How to take a test

Each program consists of about 150 frames or units. The nature of these frames, their content and the type of response which the student makes, varies a good deal, but in general the student is required to write a word suggested by the context. Because the programs have never been used, we have no reasonable estimate of the time required to work through them. We would hope that each program could be completed in from 45 to 60 minutes. Since there are to be 15 sessions in all, we propose that the instructor devote three sessions to each of the five areas listed. Three sessions will total 90 minutes. If the programs are completed in the 60 minutes which we estimate, there will be 30 minutes during the three sessions which can be devoted to work on the practice materials and/or to talks by the instructor. We would propose that 10 minutes be given to the practice materials and 20 minutes to the discussions by the instructor.

A 30-minute session is all too brief. When it is divided among several activities, these subdivisions are smaller still. Nevertheless, a balanced curriculum requires all three of the major activities proposed here: programed instruction, practice tasks, and instructor's discussions. We propose the following schedule for the three sessions to be spent on a given topic:

I.	Programed Instruction	25 minutes
	Instructor's Summary	5 minutes
II.	Programed Instruction	25 minutes
	Instructor's Summary	5 minutes
III.	Programed Instruction	10 minutes
	Practice Materials	10 minutes
	Instructor's Summary	10 minutes

In general, the students should do slightly more than 40% of each program on each of the first two days, the remaining 15% or so on the third and final day.

The instructor must use his own judgment as to the appropriateness of these suggestions. It is obvious that some time will be needed at the outset to permit the instructor to introduce himself to the class, to allow him to describe the course, and, in general, to organize the group. Further, we cannot anticipate the differences in the speed with which the students can complete the programs or the variety of needs for support and direction. If the instructor will read the programs and if he will read the discussion below, he will probably be able to effect compromises which retain the spirit of this study.

The Rationale for the Instruction

The duration of the instruction has been an important factor in determining the nature and scope of the materials. Seven and one-half hours is very nearly the minimum instruction interval with which one can expect to demonstrate gains. The adoption of this interval necessarily excludes from consideration the correction of certain defects in candidate preparation. For example, a vocabulary deficiency, in the sense of a deficiency in the total numbers of words known, would be without doubt a defect in a candidate's preparation. The correction of this deficiency, however, is certainly not the work of seven and one-half hours of instruction, and the materials which have been developed make little explicit effort to increase vocabulary.

The materials, in general, attempt to focus the attention of the student on those aspects of words which are important in tests. These aspects are basically the relationships which words have among themselves. That is, it is one thing to know the meaning of REQUEST. It is another thing to recognize that it is in many ways the opposite of REJECT. Still another aspect of REQUEST is required to formulate the relationship between "REQUEST and ENTREAT" and to perceive the similarity between this relationship and that of "ADMIRE and IDOLIZE," a similarity which is lacking in the relationship between "REPEAT and PLEAD." Finally, the proper use of REQUEST in context is another aspect of this word. For example, as a replacement for the missing word in the sentence "Our tough sergeant did not just ask us to do K.P., he _____ us to do it," REQUESTED is inappropriate. The three aspects of words which are presented above are directly tested by means of the antonym, analogy, and sentence completion questions. More complex questions involving the meaning of groups of words, rather than individual words, are asked in the reading comprehension passages.

In a sense, the course is based on the assumption that a common defect in student preparation is that students have not learned to organize their vocabulary in certain important ways. That is, the typical verbal test will usually reward a person who has a large vocabulary and who has perceived a certain kind of semantic structure. He can rapidly think of synonyms or antonyms for given words, and he has perceived the broad similarity between certain words, such as that between MASK and ALIAS. The programs and other materials used in this course attempt to heighten the awareness of these relationships among words.

Each type of test question is treated independently, that is, there is a program devoted entirely to this type. In the final phase of the course, some general considerations concerning the taking of multiple-choice tests are covered. It is during this phase that the student receives practice in the mechanics of answer sheet "gridding" (no simple task).

While the programs for the various item types differ somewhat, there is a general similarity among them. Each program is intended to create an analytical approach to the questions discussed. That is, each program attempts to guide the student in thinking about the questions. In the sentence completions program, for example, the student is told to begin by identifying clue words and to attempt to supply his own replacement for the missing word before proceeding to read the optional responses. This sequence of events: identification of key words and suggestion of one's own replacement are characteristics of the performance of the high-scoring student. Typically, the high-scoring student thus comes to the test with a double advantage: a large vocabulary and a solid analytical approach. As stated above, this course cannot increase vocabulary size. It is hoped that it can influence the analytical approach.

This general attempt to induce analysis and a search for a rational basis for an answer is limited to the very basic rules-of-thumb. That is, no attempt has been made to identify classes of wrong answers which are likely to occur in items of a given type, although this could be an approach. For example, in antonym questions one frequently encounters a distracter or wrong answer which is intended to trap the would-be etymologist who knows a little Latin or French. If the key word is "unlawful," one wrong answer might be "lexical," for the Latin word for "law" is "lex" and hence "lexical" might seem to be the equivalent of "lawful." Knowing that items contain such wrong answers, one might explicitly warn students against an etymological approach. These programs do not concern themselves with such sophisticated ideas.

The foregoing discussion is perhaps excessively complete as an example of what is not presented in the course. Nevertheless, because of the important role of the instructor it is necessary that he clearly understand that this course is

intended merely to provide, if possible, some correction for a background which has not stressed those aspects of words which are most relevant to taking verbal tests. In no sense does the course try to help the student to learn the "tricks" of the test-maker. In general, an increase in score that is due to an increase in knowledge of test tricks is probably a false increase in that it does not represent a genuine increase in verbal ability. An increase in score that is due to better organization of one's knowledge of words, on the other hand, is probably very genuine; the effects of the new organization are probably lasting and of use in thinking about problems other than the narrow concerns of doing well on tests. At any rate, these are the assumptions and values which have guided the development of these materials.

Test-Taking Skills

One aspect of test-taking has been treated quite explicitly, however. This concerns the advisability of "guessing." This course attempts to provide the student with a rule-of-thumb to assist him in knowing when to respond and when not to respond; that is, when to mark an answer and when to omit. There are good reasons for presenting this advice, even though it is almost certainly limited to gaining score points on a test. In the first place, the "scoring formula" is too complex and the description thereof in the typical test instructions is too vague for students to develop the correct decision processes on their own. Further, wherever large numbers of the questions are beyond the student, in the sense that he cannot totally determine the answer, he may become discouraged and respond randomly or not at all. This is unfortunate, for multiple-choice questions should always be answered if two of the wrong answers can be identified. The student must continue to attempt the task of excluding wrong answers even when he cannot identify with certainty the correct answer. This is particularly important in reading comprehension questions, where some of the wrong answers

may be easily proven faulty with the help of the information in the passage.

At any rate, given the importance of making responses and of using partial knowledge to eliminate distracters, and given the danger that discouragement and unnecessary fear may lead to random responses or omits, the course explicitly provides a basis for deciding when to respond.

One very important part of test-taking, of course, is confidence. This need not be the confidence of the very bright student who knows that he knows most of the answers; it can be the confidence that comes with knowing some of the answers, however few, and with knowing that trying to figure out the answers can lead to success. Tests can be discouraging simply because they are mysterious. Few students can appreciate that getting 25% of the answers correct may be a better-than-average performance. It is important that the instructor try to instill this confidence; that he reward success, however small; that he approve the work, however much it seems to result in no gain. It must always be remembered that this is an experimental study of the effects of special instruction on test scores. It is by no means certain that there will be any marked effects. Nevertheless, the aspects of words and tests which are emphasized are sound. The instructor can present them with confidence, for they can be of value to the student in other ways even if they fail to improve his test scores. If he communicates this confidence to the students, they will be most likely to benefit from the course.

APPENDIX B

APPENDIX B

IMPORTANT NOTICE

All High School Juniors

Are You Thinking of Going to College?

Several national organizations and universities are working together on a program that is designed to help high school students do well on various tests that are required by many colleges for admission and financial assistance. They would like your cooperation to see if this program can be helpful.

If you have any idea now or in the future that you would like to go to college when you finish high school, it will be in your best interest to volunteer for this program. Since only a limited number can be accommodated at a given time, you should sign up early. The following steps are involved.

1. Check with your counselor or principal if necessary for complete details.
2. The program will include the following:
   a. At an announced time, you will take tests, related to college work, about 2 hours long.
   b. Then from a representative of the project and one of your teachers or counselor, you will receive special instructions for 1.5 hours per week for five weeks. Attendance must be regular if you enter the program.
   c. At the end of 5 weeks, other tests, about 1.5 hours long, will be given to see how well you have done.
   d. Students who could not be included in the first program will then be given the same experiences.
   e. There are no charges or other requirements other than your willingness to attend all sessions of the special program that is being provided.
   f. If you concentrate on this work, it will be of great value not only to you, but to many other high school students like yourself.
   g. SEE YOUR COUNSELOR TODAY.