DOCUMENT RESUME

ED 386 472 TM 023 975

AUTHOR Powers, Donald E.; Wilson, Susan T.
TITLE Passage Dependence of the New SAT Reading Comprehension Questions. College Board Report No. 93-3.
INSTITUTION College Board, New York, NY.; College Entrance Examination Board, New York, N.Y.
REPORT NO ETS-RR-93-60
PUB DATE 93
NOTE 23p.
AVAILABLE FROM College Board Publications, Box 886, New York, NY 10101-0886 ($12).
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *College Entrance Examinations; Construct Validity; High Schools; *High School Students; Questionnaires; Reading Comprehension; *Reading Tests; Test Construction; *Test Items; *Verbal Ability
IDENTIFIERS *Passage Dependency; *Scholastic Aptitude Test; Strategy Choice

ABSTRACT

It has been reasonably well established that test takers can sometimes answer correctly some reading comprehension passages without reading the passages on which the questions are based. This issue was studied with the new Scholastic Aptitude Test (SAT) in a study designed to determine the strategies by which examinees are able to achieve better-than-chance performances without reading the passages. Sets of sample reading comprehension questions were administered, without passages, to 350 verbally-able students in 8 secondary schools across the country. After completing the task, students were asked to complete a questionnaire describing the strategies they had used. The most often cited strategies involved choosing answers on the basis of consistency with other questions and reconstructing the main theme of a missing passage from all the questions and answers in a set. These strategies were also more likely to result in successful performance on individual test items than any of the many other possible approaches. Implications for the construct validity of the new SAT are discussed. One figure and 14 tables present analysis results. (Contains 25 references.) (Author/SLD)

College Board Report No. 93-3

Passage Dependence of the New SAT® Reading Comprehension Questions

DONALD E. POWERS and SUSAN T. WILSON

The College Board
Educational Excellence for *All* Students

College Board Report No. 93-3
ETS RR No. 93-60

Passage Dependence of the New SAT® Reading Comprehension Questions

DONALD E. POWERS and SUSAN T. WILSON

College Entrance Examination Board, New York, 1993

Acknowledgments

Many people made significant contributions at various stages of this project. At the early stages of proposal development, Warren Willingham provided useful advice on the design of the study. Barry Druesne, Diane Rdesinski, and Cathy Wendler helped identify appropriate files from which we were able to contact students. Diane, along with Pete Smith, assisted in matching study records with SAT files. Sydell Carlton, Ed Curley, Patricia Klag, Gerry May, and Eric Wimmers helped identify possible strategies that students might use and provided advice about other aspects of the study. Irv Katz discussed think-aloud protocol analysis with us. Gloria Weiss was instrumental in identifying schools to participate in the study and in advising us on enlisting their help. Nancy Burton, Fred Schuppan, and Nancy Wright provided item analysis data for the reading questions we used, as well as other information about SAT takers. Inge Novatkoski set up our files and conducted early data analyses. Drew Bowker, Dave Freund, Laura Jenkins, and Pat O'Reilly conducted the data analyses, and Ken Wilson provided the final analyses of the data. Corrine Cohen helped code and edit the data. Ruth Yoder provided administrative assistance throughout the project and prepared the manuscript. Nancy Burton and Lawrence Stricker provided helpful comments on an earlier draft. The final word of thanks goes to the anonymous teachers and students at the eight schools at which the study was conducted.

Donald E. Powers is a senior research scientist at ETS.
Susan T. Wilson is a senior research associate at ETS.

Additional copies of this report may be obtained from College Board Publications, Box 886, New York, New York 10101-0886. The price is $12.

Printed in the United States of America.

Contents

Abstract

It has been reasonably well established that test takers can, to varying degrees, answer some reading comprehension questions correctly without reading the passages on which the questions are based. This is true even for carefully constructed measures such as the College Board's SAT. As a result, the use of reading scores, including those from the SAT, as valid indicators of reading comprehension has been challenged.

The major aim of this study was to determine the strategies by which examinees are able to achieve better-than-chance performances without reading the passages. The focus of the research was a sample of reading comprehension questions similar to those that will be used in the revised SAT, to be introduced in 1994. Sets of reading comprehension questions were administered, without the passages, to a sample of verbally able students in eight secondary schools across the country. After attempting the task, these students were asked to complete a questionnaire describing the strategies they had used. The most often cited strategies involved choosing answers on the basis of consistency with other questions and reconstructing the main theme of a missing passage from all the questions and answers in a set. These strategies were also more likely to result in successful performance on individual test items than were any of the many other possible (and less construct-relevant) strategies. Implications are discussed with regard to the construct validity of the new SAT reading comprehension questions and the advice that should be given to prospective SAT takers.

Introduction

A primary objective of test makers is to ensure that test performance is not unduly influenced by factors outside the proper focus of measurement. This intent is formalized in the *Standards for Educational and Psychological Testing* (AERA, APA, and NCME 1985). Irrelevancies such as test takers' level of familiarity with testing procedures or their use of particular test-taking strategies ought to play only a small role in determining success on any examination. Strategies that raise test scores but bear little if any relationship to what the test was designed to measure may diminish the predictive power of a test or dilute the meaning of scores derived from it.

For tests of reading comprehension, one strategy that examinees may use (and one over which test administrators have little if any control) is to answer test questions without reading the passages on which they are based. Unfortunately, little information exists on the prevalence of this approach, but the evidence that is available suggests that adult readers do not employ this strategy often (Farr, Pritchard, and Smitten 1990). Nonetheless, there has been a recent revival of interest in the extent to which success in answering reading comprehension questions depends on having read the passages with which they are associated. Katz, Lautenschlager, Blackburn, and Harris (1990) found that college students can perform at better-than-chance levels on SAT reading questions even when they do not have access to the reading passages. This finding is, of course, not a new discovery. It has long been acknowledged that reading comprehension questions vary in the degree to which they depend on the passages (Preston 1962), and this phenomenon has frequently been studied (Preston 1964; Pyrczak 1972, 1974, 1975; Tuinman 1973-74; Weaver and Bickley 1967). Nor is this finding surprising. As Conlan (1990) argued, it would be highly unusual, given current conceptions of reading comprehension, if examinees were not able to extract some information from the test questions themselves.

The renewed interest in passage dependence is timely in light of the planned revisions to the current SAT. Beginning in 1994 the verbal reasoning portion of the SAT will place greater emphasis on critical reading, and vocabulary will be measured in context rather than with discrete antonym items. Fewer reading passages will be used (four instead of six), but the passages will be longer. The net result is that the proportion of reading questions on the SAT will increase by nearly 60 percent.

The principal objective of the present study was to determine the particular ways in which test takers may glean useful information solely from the kinds of reading comprehension questions that will be used in the revised SAT. A second aim was to assess the construct relevance of these strategies, i.e., the extent to which they entail legitimate verbal reasoning abilities versus testwiseness or other such skills that are less germane to measuring reading comprehension and readiness for postsecondary study. The final purpose was to determine the degree to which the kinds of reading questions to be used in the revised SAT can be answered correctly without the passages and to gain insight into the question characteristics that may contribute to their susceptibility to this strategy.

Method

Checklist Development

A checklist of possible strategies was developed to determine how test takers might approach tests when reading passages are not available. Several activities were under-

taken to gain insight into this process. First, two experienced test developers were asked to answer sets of reading questions without the passages. Because these particular developers were familiar with most of the currently available SAT passages, reading passages from disclosed forms of the Law School Admission Test (with which they were not familiar) were used. A number of useful insights resulted from this process. The most interesting involved the use of one strategy in particular. By carefully reading an entire set of questions, these specialists were able to reconstruct, to a considerable degree, the passage on which the questions were based. In effect, for these highly literate people, the task became an extended cloze test requiring sophisticated verbal reasoning.

One of the authors also conducted individual think-aloud sessions with six local high school students as they attempted to answer sets of SAT questions without access to reading passages. These encounters generated additional hypotheses about how students might approach this unconventional assignment.

Next, a sample of approximately 300 students was surveyed and asked to answer a set of 9 to 12 reading questions without the passages. Every third student received a different set of questions. Students were selected from SAT registrant files so that they would represent a range of academic ability (about two-thirds reported being A students and one-third C students). A total of 33 students, nearly all A students, responded to our request to attempt the questions and then to indicate for each one the strategies they had employed. It should be noted that the participants in these pilot activities were not prompted in any way about how to approach the task. The objective at this stage was to uncover any strategies that might be employed naturally in attempting the exercise. Finally, a number of test preparation books were examined for any suggestions that might relate to the study. Selected literature on testwiseness (e.g., Millman, Bishop, and Ebel 1965) was consulted for additional clues.

The data from these activities were used as the basis for the categories in the final checklist. Possible strategies were classified as reasoning strategies, use of personal knowledge or experience, strategies for vocabulary questions, guessing, and use of features of answer choices. The checklist was pilot tested on a small number of local students, reviewed by several Educational Testing Service (ETS) staff members, and revised before it was used in the study proper.

Selection of Reading Passages

Six reading passages and associated questions were selected from among those pretested in earlier trials of the revised SAT. The nature of the passages and the number of questions associated with each one were as follows:

1. A passage of approximately 900 words on language, in which the author, a Japanese American, recounts an experience he had just after the United States entered the Second World War (12 questions).
2. A 500-word passage adapted from an excerpt of a memoir written by Elizabeth Bishop about the poet Marianne Moore (6 questions).
3. An 800-word passage about Clarence Darrow and the Communist trial of 1920 (9 questions).
4. A 600-word passage that presents a theory about the nature of the object that exploded above Tunguska in 1908 (9 questions).
5. Two passages totaling about 800 words that present two views of the architectural design of cities. One discusses planned, medium-sized cities; the other offers a critique of modern cities (13 questions).
6. A 500-word passage excerpted from a book of literary criticism analyzing the work of Richard Wright (1908–1960) (5 questions).

Sample Selection

The objective was to recruit students who would be most likely to perform well on the task at hand (i.e., answering reading questions without reading the passages) and be able to indicate the strategies involved in their performance. A list was compiled of secondary schools that had volunteered (but were not selected) to participate in the earlier field trials of the revised SAT. From this list, a number of schools having above-average mean SAT-Verbal scores were identified and invited to participate in the study. To further ensure that the sample included a preponderance of verbally able students, school personnel targeted Advanced Placement or honors English classes. Eight schools were eventually selected. These schools were located in California, Colorado, Florida, Maryland, New Jersey, New York, and West Virginia.

Data Collection

Three different test forms were assembled, each consisting of 18 questions from the first and second, third and fourth, or fifth and sixth passages described above. The number of questions was deliberately limited, in order to maintain students' motivation to do this unconventional task, for which they had no particular reason to exert any

This booklet contains two sections. The first section consists of two groups of reading comprehension questions without the passages on which the questions are based. We would like you to answer these questions to the best of your ability without first reading through their accompanying passages. Feel free to use any type of strategy that you need to solve these problems.

The only instructions that we ask you to follow in this section are:

1) *Indicate the time you started answering the questions and the time you finished in the spaces provided for each group of questions.*

2) *Circle your answer to each question on the test form.*

3) *In addition, for each question, please "cross out" (X) any choice that you are able to rule out. (Do this when you are able to "narrow down" the answers, rather than guess among all the answers.)*

Here's an example of what this might look like on your test form:

> *The "unfinished work" referred to in line 50 is the*
>
> *(A) battle of Gettysburg*
> *(B) defense of freedom*
> *(C) establishment of a government*
> *(D) dedication of the battlefield*
> *(E) honoring of the fallen soldiers*

4) *Please answer each question. You should definitely guess when you are uncertain about the answer to a question. Unlike the real SAT, it is to your benefit to guess at an answer in this study. No points are taken off your score for wrong answers.*

FIGURE 1. Directions to study participants.

extraordinary effort. The only concrete incentive was the opportunity for one student from each school to win a $100 prize on the basis of test performance. Specifically, students at each school were told that those earning test scores of at least 30 percent would be eligible for the prize. This score, just above a chance performance of 20 percent, was thought to be within the reach of every student and therefore potentially motivating.

Classroom teachers administered the tests during regular classroom periods in the spring of 1992. The three test forms were alternated within each classroom so that every third student received a different form. That is, students 1, 4, 7, and so on got Form A; students 2, 5, 8, and so on got Form B; and students 3, 6, 9, and so on got Form C. The checklist of possible strategies was included with each test form. Instructions to students are shown in Figure 1.

Students were asked to indicate the times at which they started and finished answering each group of questions in order to evaluate the efficiency of answering questions without passages as well as the effectiveness of the strategy used. Students' actual SAT scores and a variety of other background data were retrieved from SAT files.

Results

Description of the Sample

Records were available from SAT files for 271 of the 350 students (all high school juniors) who participated in the study. It is likely that those students whose records could not be located had not taken the SAT. Table 1 compares the study sample with all 1992 college-bound seniors who took the SAT on selected variables. The study sample was, as expected, very able. The test scores of the study sample were on average higher by about one standard deviation than were the scores of SAT takers in general. In addition, the study sample reported higher grade-point averages and class ranks than did all college-bound seniors. For example, nearly half (47 percent) of the sample were in the top 10 percent of their classes, compared with 21 percent of all college-bound seniors. A majority (66 percent) of the sample indicated that they had taken honors English courses, compared with about 32 percent in the college-bound senior population (College Board 1992).

Descriptive Statistics for Test Scores

Table 2 shows the mean score and mean percentage of items correct for each of the six sets of reading questions when administered without the reading passages. For each set, performance was better than would be expected from random guessing on the five-choice questions (i.e.,

TABLE 1

Comparison of Study Sample with All College-Bound Seniors Who Took the SAT in 1991–92

		Study Sample (N=350)	*All 1991–92 College-Bound Seniors (N=1,030,000)*
Sex (% female)		54	52
Honors English taken (%)		66	32
High school rank (%)			
Top tenth		47	21
Second tenth		27	22
Second fifth		17	28
Third fifth		9	25
Fourth fifth		0	4
Fifth fifth		0	1
High school GPA (%)			
A+		22	5
A		23	12
A–		20	14
B		35	52
C		<1	17
D, E, or F		0	0
SAT scores			
SAT-V	*M*	542	423
	SD	88	112
SAT-M	*M*	606	476
	SD	102	123
TSWE	*M*	52	42
	SD	7	11

Source: College Board, *College Bound Seniors: 1992 Profile of SAT and Achievement Test Takers* (Princeton, N.J.: Educational Testing Service, 1992). *Note:* Statistics for study sample are based on SAT records for 271 of 350 participants.

20 percent). However, for three of the sets the percentage of correct responses (26–29 percent) was not much better than would be expected by chance. Only for the smallest set of questions was performance (59 percent correct) commensurate with that typically observed when reading passages are available. As might be expected, given both the nature of the task and the relatively small number of items, the estimated reliability of each test form was low. Internal consistency estimates (coefficient alpha) were .44, .47, and .19 for the three forms, respectively. The internal consistency also varied for question sets within forms: .35 and .36 for those in Form A, .43 and .32 for Form B, and .09 and .37 for Form C. In comparison, the reliability of a 25-item pretest of the new reading comprehension questions was estimated to be .78 when administered with passages (Lawrence 1992).

With respect to time required to complete the task, examinees reported taking about 13 minutes on average to answer the 18 questions in each form. The mean times for each test form, as determined from examinee reports, were 12.3 minutes (SD = 3.1), 13.6 minutes (SD = 3.9), and 13.8 minutes (SD = 4.7), respectively. Thus, on average, examinees devoted about 40 to 50 seconds per question. It is interesting to note that according to calculations made by Katz et al. (1990), the average time allotted for each reading comprehension question on the SAT is about 65 seconds, which includes the time required to read the passages.

Although overall performance on each set of questions is of interest, it does not tell the complete story. Table 3 compares the performance of the students in our study on each test question (without the passages) with that of students who participated in an earlier large-scale pretesting of the same questions (with the passages). For most test items, the with-passage statistics are based on the performance of nearly 10,000 examinees who participated in tryouts of the new SAT items from fall 1990 through spring 1991. These pretest examinees represented all SAT takers reasonably well with respect to demographics but were slightly less able than the general population of SAT takers. They were, therefore, considerably less able, as defined by SAT performance, than the students who participated in the current study.

TABLE 2

Mean Test Scores for Reading Passage Sets

Form/Passage	*Number of Items*	*Mean Number Correct*	SD	*Mean Percentage Correct*
A. Language	12	3.1	1.8	26
A. Marianne Moore	6	1.7	1.3	29
B. Clarence Darrow	9	3.4	1.7	38
B. Tunguska	9	3.3	1.7	37
C. Architecture	13	3.7	1.6	28
C. Richard Wright	5	2.9	1.2	59

Note: A total of 135 students took Form A, 107 took Form B, and 108 took Form C.

Because of the differences between samples, Table 3 only approximates the differential difficulty of each question when administered with and without the passages. Questions are ranked according to their difficulty when the passage was available. The standard error of the difference between percentages correct based on the sample sizes is .048 or less. Thus, a difference of about 12 percent can be used to gauge the significance of differences at the .01 level and about 9 percent can be used at the .05 level. As is apparent from Table 3, the questions were, with few exceptions, substantially easier for the more typical sample of test takers who had access to the reading passages than for the more able students in our study who did not have access. If the two samples had been more nearly equal in ability, the differences in performance would probably have been even greater.

Item-Level Comparisons

Clearly, according to Table 3, restricting access to reading passages did not depress performance equally for each test item. The ordering of items by difficulty with and without the passages shows a less-than-perfect correspondence. Items that were difficult when passages were available tended to be difficult when the passages were not available. The difficulty of some items, e.g., question 2 from the passage on Language, question 7 from the Clarence Darrow passage, and question 5 from the Richard Wright passage, was apparently affected very little by removing access to the passages on which they were based. Others were influenced dramatically. For instance, although question 11 from the passage on Language was answered correctly by 58 percent of test takers who had the passage, it was answered accurately at only a chance level (20 percent) when the passage was not available. The most striking change in difficulty occurred for question 1 from the Architecture selection. Although this question was very easy (85 percent correct) when the passage was accessible, it was answered correctly at only a chance level without the passage. Question 2 from the same passage was quite easy (63 percent correct) with the passage but was answered at a less-than-chance level (6 percent) when the passage was unavailable.

Aside from the content of the items, which is considered below, there are some readily available explanations for their differential difficulty under the two testing conditions. Table 4 is informative in this regard. It displays

TABLE 3

Performance on Reading Comprehension Questions Given with and without Reading Passages

Form/Passage	Question	Percentage Correct with the Passage	Percentage Correct without the Passage	Difference
A. Language	10	67	29	38
	11	58	20	38
	7	54	35	19
	9	54	24	30
	12	51	40	11
	2	48	54	−6
	6	48	25	23
	5	46	17	29
	1	42	12	30
	8	28	24	4
	3	23	16	7
	4	20	15	5
	Median	48	24	21
A. Marianne Moore	1	85	57	28
	5	70	42	28
	4	48	19	29
	3	45	18	27
	6	39	24	15
	2	36	16	20
	Median	47	22	28
B. Clarence Darrow	5	76	53	23
	4	73	38	35
	3	62	51	11
	7	59	69	−10
	2	58	6	52
	9	56	52	4
	8	55	37	18
	1	33	24	9
	6	25	12	13
	Median	58	38	13
B. Tunguska	3	73	43	30
	4	68	38	30
	2	67	55	12
	5	54	47	7
	6	51	26	25
	8	45	30	15
	1	41	23	18
	7	39	41	−2
	9	32	33	−1
	Median	51	38	15
C. Architecture	1	85	20	65
	6	65	56	9
	2	63	6	57
	3	62	38	24
	9	59	41	18
	4	56	30	26
	11	46	19	27
	7	45	30	15
	8	43	11	32
	13	39	21	18
	12	37	38	−1
	10	29	22	7
	5	23	37	−14
	Median	46	30	18
C. Richard Wright	5	80	74	6
	4	77	62	15
	2	74	51	23
	1	69	55	14
	3	59	52	7
	Median	74	55	14

Note: The Language, Marianne Moore, Tunguska, and Richard Wright item statistics are each based on more than 9,500 students tested in 1990–91. The Clarence Darrow statistics are based on approximately 2,500 to 3,500 students tested in 1990–91. The Architecture passage item statistics are also based on more than 9,500 examinees, except for question 5, which is based on 115 students.
Differences of approximately 9 percent are significant beyond the .05 level, two-tailed; differences of approximately 12 percent are significant at the .01 level.

in descending order of difficulty for each question the without-passage statistics on the percentage of examinees who answered each question correctly, eliminated the correct answer, chose selected incorrect options (the most and least popular), and eliminated selected incorrect options (the most and least popular).

From Table 4 it appears that some items, e.g., questions 1 and 5 from the Language passage, question 2 from the Clarence Darrow passage, and questions 2 and 8 from the Architecture passage, were difficult in large part because examinees were much more likely to eliminate the correct choice than to select it. The most extreme case is question 2 from the Architecture passage: the correct answer was chosen by only 6 percent of the sample but eliminated by 55 percent.

Other items appeared to be difficult because particular incorrect options were attractive. An example of this is question 2 from the passage on Marianne Moore, which was answered correctly by 16 percent of the sample. A majority (57 percent) selected the same incorrect answer.

Some items appeared to be answerable without the passage partly because particular incorrect options were eliminated as plausible answers. Examples are question 9 from the Architecture passage (one choice was eliminated by 66 percent of the sample) and question 3 from the Clarence Darrow text.

Table 5 compares examinee responses, with and without the passages, to several of the questions mentioned above. The particular questions selected for discussion here were chosen because they represent the most dramatic examples of different response patterns. Question 1 from the Language selection appears to have been more difficult without the passage because the correct answer was eliminated about three times as often as it was chosen (36 percent versus 12 percent) and because two incorrect options (D and E) were more attractive to examinees who did not have the passage than to those who did. Question 2 from the Clarence Darrow passage had a similar outcome, with about five times as many without-passage examinees eliminating the correct answer as selecting it (32 percent versus 6 percent). Each of

TABLE 4

Descriptive Statistics for 54 Reading Comprehension Questions Administered without Reading Passages

		Correct Answer		Incorrect Options			
				Percent Choosing		Percent Eliminating	
Form/Passage	Question	Percent Choosing	Percent Eliminating	Least Popular	Most Popular	Least Popular	Most Popular
A. Language	2	54	13	4	32	20	38
	12	40	8	5	40	16	22
	7	35	19	8	24	8	37
	10	29	25	0	33	10	38
	6	25	14	15	23	16	22
	8	24	10	13	28	12	21
	9	24	17	8	46	17	35
	11	20	16	0	45	6	36
	5	17	43	5	29	13	37
	3	16	22	4	48	10	37
	4	15	18	4	44	13	24
	1	12	36	9	36	15	35
A. Marianne Moore	1	57	4	6	21	12	36
	5	42	9	7	30	10	31
	6	24	14	1	43	12	37
	4	19	13	2	29	11	19
	3	18	19	14	27	17	28
	2	16	11	6	57	10	27
B. Clarence Darrow	7	69	6	5	11	12	49
	5	53	8	4	19	18	31
	9	52	14	2	30	14	35
	3	51	16	6	19	26	50
	4	38	12	4	36	17	39
	8	37	20	9	28	21	26
	1	24	31	5	36	23	54
	6	12	28	13	35	14	38
	2	6	32	5	32	10	43
B. Tunguska	2	55	8	8	18	18	31
	5	47	10	1	19	20	37
	3	43	11	2	32	12	36
	7	41	12	10	25	15	29
	4	38	13	1	39	11	27
	9	33	15	9	25	20	26
	8	30	20	4	31	17	36
	6	26	24	7	32	12	31
	1	23	21	3	37	17	41
C. Architecture	6	56	7	1	18	19	43
	9	41	14	3	35	18	66
	3	38	17	6	30	25	51
	12	38	9	6	24	12	25
	5	37	11	8	20	15	25
	4	30	16	10	35	8	28
	7	30	19	4	36	19	36
	10	22	18	8	26	15	27
	13	21	24	7	36	7	24
	1	20	47	3	44	13	32
	11	19	23	5	34	14	38
	8	11	46	6	41	8	48
	2	6	55	4	44	12	45
C. Richard Wright	5	74	6	3	14	14	44
	4	62	6	7	12	26	37
	1	55	13	1	17	22	49
	3	52	12	6	20	17	36
	2	51	13	3	29	17	46

three incorrect choices (B, C, and D) was chosen more frequently without the passage than with it.

For question 2 from the Architecture passage, a majority (55 percent) of examinees eliminated the correct choice when they did not have the passage, and only 6 percent selected it. Besides tending to eliminate the correct answer for this item, examinees without the passage found two incorrect choices (A and E) to be much more plausible than did examinees who had access to the passage.

Question 2 about Marianne Moore appears to have been answered correctly somewhat less often than ex-

TABLE 5

Comparison of Responses for Selected Questions Given with and without Passages

Passage/Question	*Test Item and Options*	*Percent Choosing with Passage*	*Without Passage*	
			Percent Choosing	*Percent Eliminating*
Language/Question 1	The phrase "older and deeper functions" (line 6) refers to the			
	(A) grammatical structure of a language	12	9	35
	*(B) expression of emotions through sound	42	12	36
	(C) transmission of information	16	10	28
	(D) statement of cultural values	8	36	15
	(E) original meanings of words	18	34	20
	Omitted	3	—	—
Clarence Darrow/Question 2	Which of the following best captures the meaning of the word "consideration" in line 17?			
	*(A) leniency	58	6	32
	(B) contemplation	7	30	14
	(C) due respect	18	29	15
	(D) reasoned judgment	10	32	10
	(E) legal rights	2	5	43
	Omitted	6	—	—
Architecture/Question 2	In passage A, the reference to "next month's issue of an architectural periodical" (lines 22–23) serves to			
	(A) show that the plans for the garden cities are well thought of in professional journals	13	44	12
	(B) indicate that what seems like a random process is actually an ordered process	10	12	44
	*(C) suggest that some people lack their own firm ideals of beauty	63	6	55
	(D) imply that only those who are knowledgeable about a subject should offer their opinions	6	4	45
	(E) emphasize the importance of what the experts say	6	35	18
	Omitted	3	—	—
Marianne Moore/Question 2	The major purpose of the passage is to			
	*(A) describe the events that led to a milestone in the author's life	36	16	11
	(B) reveal the character of a college librarian	5	14	27
	(C) relate the significant events of the author's college years	4	6	15
	(D) analyze the impact of Marianne Moore's poetry on the author	38	57	10
	(E) show the unexpected surprises that can happen in an ordinary life	16	7	22
	Omitted	2	—	—
Clarence Darrow/Question 3	By "They can afford it if you members of the jury can" (line 22), Darrow means that			
	(A) no harm will come to the defendants if they are convicted in this case	8	13	44
	(B) the jurors will be severely criticized by the press if they convict the defendants	13	19	26
	(C) the defendants are indifferent about the outcome of the trial	5	6	50
	(D) the verdict of the jury has financial implications for all of the people involved in the trial	6	11	44
	*(E) a verdict of guilty would be a potential threat to everyone's rights	62	51	16
	Omitted	6	—	—

*Correct choice.

pected by chance (16 percent) by no-passage test takers, mainly because of the attractiveness of option D. Question 3 from the Clarence Darrow selection remained relatively easy even without the passage because each of the incorrect options was eliminated relatively frequently. These examples illus ate that the effect of restricting access to reading passages is far from uniform across test items and that differential effects on item difficulty appear to result from several different processes.

Descriptive Statistics on Test-Taking Strategies

The extent to which study participants reported using various strategies to answer questions without the passages is presented in Table 6. For most of the strategies, participants were asked to indicate the proportion of questions for which they had used a strategy (all or nearly

all, about 75 percent, about 50 percent, about 25 percent, or few or no questions). Table 6 presents frequency of use, combining the first two and the last two categories. Generally, each possible strategy was used frequently by at least some students and infrequently by others.

Each strategy considered to involve reasoning of some kind was used by a majority or near majority of students for a preponderance of questions. For instance, the strategy of choosing an answer because it seemed to be consistent with something stated in the other questions was reportedly used by 65 percent of examinees for 75 percent or more of the questions. In addition, a slight majority (55 percent) of the sample said that they, like the experienced test developers consulted early in the study, had tried to reconstruct the theme or main idea of the missing passage by reading all the questions and answers.

Personal knowledge or experience was invoked relatively infrequently. About 20 percent of examinees stated that they had used their personal knowledge for 75 percent or more of the questions for each set of questions. There were some differences among test forms in the use of personal knowledge, with students least likely to call on personal knowledge for the questions on the Architecture passage and most likely to summon it for questions on the Richard Wright selection.

Generally, few students maintained that they recognized the passages or knew their sources. Equally few were familiar with the author of the passage. The exception was examinees' greater familiarity with Richard Wright and his works. Specifically, 30 percent said that they recognized the Richard Wright passage or knew where it came from, and 37 percent said that they recognized the author.

Of the possible strategies for answering vocabulary questions, the most frequently used was to choose an answer because its meaning seemed to fit best with the general theme or context suggested by the questions. Knowing that the vocabulary word or answer could have more than one meaning was also helpful. Knowledge of prefixes, suffixes, or word roots or of the most general or common meaning of a word was used less often.

With respect to strategies involving guessing, random and patterned guessing strategies were utilized relatively infrequently. Two other strategies that involved guessing among choices that could not be ruled out and using vague hunches or intuition were applied by a near majority of the sample.

We speculated that certain features of answer choices might be used either to select the correct answer or to eliminate incorrect alternatives. Table 7 shows, in descending order of prevalence, students' tendencies to use these characteristics. None of the features listed was regarded in the same way by all examinees. Some students used a particular strategy to choose an answer, and others used the same strategy to eliminate options. Some features, however, were used much more frequently in one way than in the other. The clearest tendency was to choose the more carefully worded or qualified alternative or to rule out a less qualified one. About three-quarters of test takers said that they used this approach. Students also reported choosing the more concrete options and eliminating the more abstract ones, and they preferred the specific alternatives over the general ones. Some features of answer choices, such as their length, the degree to which they were regarded as ambiguous, and whether or not they seemed obvious, were used somewhat less often than other features. More important, these features were reportedly used about as often to choose as to eliminate alternatives.

The extent to which students actually eliminated answer choices from consideration was also calculated. On average, students ruled out approximately one (from .9 to 1.3) choice per question on each set of questions. The mean number of choices that were eliminated differed among questions, ranging from .7 to 1.8. Eliminating choices appears to have been a consistently employed strategy. The correlation between the number of choices ruled out for the first set of questions within a test form correlated highly with the number eliminated for the second set—.75, .80, and .86, respectively, for the three forms.

In order to determine which strategies were used for particular questions, examinees were asked to indicate for some sets of questions (those based on the Marianne Moore passage, the Tunguska passage, and the Richard Wright passage) the single most helpful strategy of all those listed on the checklist. Table 8 shows the two most frequently mentioned strategies for each question and the percentage of test takers who listed each one. There was no consensus regarding the helpfulness of any particular strategy for any given item. Nonetheless, students referred to some strategies more often than others. In fact, of the 45 different possibilities, only 10 were mentioned most or second most frequently for any item. Choosing an answer because it seemed to be consistent with something stated in other questions was regarded by a plurality of test takers as the most helpful strategy for 6 of the 20 items and as the second most helpful for 7 others. Attempting to reconstruct the theme or main idea of the missing passage from the other questions and answers was the next most often reported strategy. The use of personal knowledge about a topic was mentioned as being most helpful for four items from the Richard Wright passage.

TABLE 6

Frequency of Use of Various Test-Taking Strategies

Strategy	Percentage Using Strategy for Few or No Questions or about 25% of Questions	Percentage Using Strategy for All or Nearly All Questions or about 75% of Questions	Percentage Answering Yes
REASONING STRATEGIES			
R1 Tried to determine the meaning of a word, or phrase, or the way in which it was used, from the other questions in the set	15	54	
R2 Assumed, guessed, or knew the answers to some questions and then, on the basis of these answers, reasoned what the answer to a later question would have to be (or what it could not be)	22	48	
R3 Chose an answer because it seemed to be consistent with something stated in the other questions	13	65	
R4 Ruled out an answer because it seemed to contradict something in the other questions	8	53	
R5 Chose an answer because it resembled something in the question: I associated a word, phrase, or idea in the question with something in the answer I chose	29	37	
R6 Tried to reconstruct the theme or main idea of the missing passage by reading all the questions and answers			55
PERSONAL KNOWLEDGE OR EXPERIENCE			
P1 Used my personal knowledge about the topic that I learned either inside or outside of school			
First passage	55	18	
Second passage	59	22	
P2 I recognized a passage or knew where it came from			
First passage			4
Second passage			15
P3 I recognized the author of the passage and was somewhat familiar with his/her opinions, etc.			
First passage			3
Second passage			14
STRATEGIES FOR VOCABULARY			
Ruled out an answer because:			
S1 I didn't think the vocabulary word could have that meaning	27	44	
S2 Everyone would know the meaning of that answer, so I thought it was too obvious	63	13	
Chose an answer because:			
S3 I knew that the vocabulary word or answer could have more than one meaning	19	55	
S4 Its meaning seemed to fit best with the general theme or context suggested by the questions	11	69	
S5 It was the most general or common meaning of the vocabulary word	44	23	
S6 I knew about the vocabulary word's root, prefix, or suffix	53	24	
GUESSING			
G1 I guessed randomly among all of the choices	77	9	
G2 I guessed among two or more choices I couldn't rule out	24	48	
G3 I used vague hunches or intuition, but I can't say exactly how	29	45	
G4 I tended to guess a particular choice (e.g., A or C)	85	8	

Note: The use of strategies R3, P1, P2, P3, and S5 varied significantly ($p<.05$) by test form. Percentages are based on all 350 students who completed the questionnaire.

TABLE 7

Frequency of Use of Test-Taking Strategies Involving Features of Answer Choices

Features of Answer Choices (in Relation to Other Choices)	*Percentage Using Feature to Choose Answers*	*Rule Out Answers*
More carefully worded/qualified	76	18
Less qualified	13	72
More concrete	69	22
More abstract	30	61
More general	37	61
More specific	68	33
Too extreme	19	68
Too neutral	24	60
Too similar to other choices	23	65
At the center of all the other choices, something in common with all of them	61	25
More definite/absolute	64	29
Less definite/more relative	32	59
In tune with current thinking	63	17
Outdated, old-fashioned	21	55
More positive/less critical in tone or mood	60	26
More negative/more critical in tone or mood	28	56
Simpler	39	51
More complex	53	40
Most obvious	47	47
Least obvious	40	48
More common/normal	47	34
More uncommon/unusual	38	47
Most ambiguous	29	48
Least ambiguous	40	33
Longer	35	27
Shorter	31	31

Note: Percentages are based on all 350 questionnaire respondents.

It is also interesting to note which strategies were not mentioned as being helpful. Of the 26 possible features of answer choices, only one—selecting an answer because of its positive or less critical tone—was mentioned by enough students to be among the two most helpful strategies. This strategy was mentioned in conjunction with the first question from the Marianne Moore passage, for which the correct response was that Marianne Moore's poetry was (choice D) "inspiring and well crafted."

Study participants were also asked to indicate any other strategies they used that we had not anticipated. A total of 19 percent of respondents said that they had used some strategy that was not specified. Most of these, however, were merely restatements of strategies that were on the checklist. By far the most frequent restatement pertained to reconstructing the theme or main idea of the missing passage by reading all the questions and answers. Students also offered a number of alternatives to the terms we had used—"radical and opinionated" for "definite/absolute," "politically correct" for "in tune with current thinking," and "makes a personal assault on the author or specific people" for "negative/critical in tone." Thus, although students may have used different terminology, they did not appear to use any strategies that we had not anticipated.

Relationships between Strategy Use and Without-Passage Performance

Frequency of use of each strategy (5 = strategy used for all or nearly all questions, 4 = used for about 75 percent of questions, 3 = used for about 50 percent of questions, 2 = used for about 25 percent of questions, 1 = used for few or no questions) was correlated with test performance. Only eight of the many strategies listed were related significantly to performance on any of the six sets of questions administered without the passages. These strategies are enumerated in Table 9. However, none of the strategies was consistently related to performance across the different sets of questions. Given the relatively large number of correlations generated, we are not inclined to make any extensive interpretations of those that were found to be significant. It does appear, however, that random guessing was not an effective strategy, as might be anticipated. For two of the sets, students who guessed randomly with some regularity performed worse than those who did not (–.30 and –.32). Nor was choosing a general or common meaning of a word a good approach for these two sets of questions. Correlations between performance and the use of this strategy were –.29 and –.36.

Correlations between students' use of features of answer choices—either to choose or to rule out alternatives—and their performance on each of the six sets of questions were also computed. Each feature whose use correlated significantly ($p<.05$) with performance on any of the six question sets is listed in Table 10 (for those used to choose answers) and Table 11 (for those used to rule out answers). When used to select answers, only 11 of the 26 possible features correlated significantly with performance on any question set. When used to rule out alternatives, only 7 of the 26 features correlated significantly with test performance on any set. Furthermore, any significant relationship between feature use and test performance was not consistent across question sets. Thus there does not appear to be any systematic relationship between the use of answer-choice features and test performance without passages.

It is possible that general tendencies to use particular strategies do not relate consistently to overall performance because specific strategies may be required for certain items. Strategies that work well for one item may not

TABLE 8

Strategies Most Frequently Listed as Most Helpful for Individual Test Questions

Form/Passage	*Question*	*Most Frequent Strategy*	*Percentage Listing*	*Second Most Frequent Strategy*	*Percentage Listing*
A. Marianne Moore	1	G1	17	R3 F5	15 15
	2	R3	19	R6	17
	3	R4	13	G3	10
	4	R3	17	R6	15
	5	G1	12	R6 R3	11 11
	6	R6	16	R3	14
B. Tunguska	1	R3	17	R1	13
	2	S4	30	R3	11
	3	R3	15	R6	13
	4	S4	31	R3	11
	5	G1	20	G3	15
	6	R3	20	G3	9
	7	G3	14	R6	14
	8	R6	15	R3	14
	9	R3	15	G3	12
C. Richard Wright	1	R6	25	P1	13
	2	P1	17	R3	17
	3	P1	17	S4	11
	4	P1	21	G2	15
	5	P1	20	R6	9

Key:

F5 = More positive/less critical in tone or mood
G1 = Guessing randomly among all the choices
G2 = Guessing among two or more choices unable to rule out
G3 = Using vague hunches or intuition
P1 = Using personal knowledge about the topic
R1 = Trying to determine the meaning of a word, or phrase, or the way in which it was used, from the other questions in the set
R3 = Choosing an answer because it seemed to be consistent with something stated in other questions
R4 = Ruling out an answer because it seemed to contradict something in the other questions
R6 = Trying to reconstruct the theme or main idea of the missing passage by reading all the questions and answers
S4 = Meaning seemed to fit best with the general theme or context suggested by the questions

be effective overall. Therefore, we attempted to relate test performance to strategy use more precisely by correlating performance on specific items with the use of strategies that were regarded as being the most helpful ones for individual items. One section of the questionnaire asked examinees to list for a subset of items the particular strategy that was most helpful for each item. The analysis entailed the following. For each item we specified the strategies that were mentioned as most helpful by about 10 percent or more of the examinees. Then, for each item, performance (wrong = 0, right = 1) was regressed on the use (0 or 1) of each of the strategies in the set. For the 20 test items involved, most of the sets included from two to four strategies, although one had five and three had only one. Because the dependent variable was dichotomous, logistic regression was used. Table 12 contains a summary of this analysis. For 6 of the 20 items at least one strategy in the set contributed significantly ($p<.05$) to predicting performance on the item. These significant predictors are listed in Table 12 along with the regression weights, standard errors of the weights, and the appropriate significance test (Wald statistic). The far right column shows the statistic exp(B), which is e, the base of the natural logarithms, raised to the power of B, the beta weight. This statistic can be interpreted as the factor by which the odds of correctly answering the question increased when examinees used the strategy listed. For the first and the last questions listed, the estimated chance of a correct answer

TABLE 9

Significant Correlations between Use of Strategies and Performance on Question Sets

	Question Set					
Strategy	*Language*	*Marianne Moore*	*Clarence Darrow*	*Tunguska*	*Architecture*	*Richard Wright*
R1	.02	−.08	.14	.02	.05	.22*
R2	.18*	.09	.14	.06	−.10	−.05
R5	−.11	−.17*	−.13	−.14	.03	−.22*
S1	−.03	−.01	.04	−.05	.23*	.06
S2	−.22**	−.02	−.07	.06	−.14	−.06
S5	−.00	.05	−.29**	−.36***	.04	−.08
S6	−.04	−.08	−.18	−.10	.05	−.21*
G1	−.10	.00	−.30**	−.32***	−.16	−.10

Key:

R1 = Tried to determine the meaning of a word, or phrase, or the way in which it was used, from the other questions in the set

R2 = Assumed, guessed, or knew the answers to some questions and then, on the basis of these answers, reasoned what the answer to a later question would have to be (or what it could not be)

R5 = Chose an answer because it resembled something in the question. Associated a word, phrase, or idea in the question with something in the answer chosen

S1 = Didn't think the vocabulary word could have that meaning

S2 = Everyone would know the meaning of that answer, so thought it was too obvious

S5 = It was the most general or common meaning of the vocabulary word

S6 = Knew about the vocabulary word's root, prefix, or suffix

G1 = Guessed randomly among all the choices

*$p<.05$, **$p<.01$, ***$p<.001$.

TABLE 10

Significant Correlations between Test Performance and Use of Answer Choice Features to Choose Answers

	Question Set					
Feature	*Language*	*Marianne Moore*	*Clarence Darrow*	*Tunguska*	*Architecture*	*Richard Wright*
More general	−.25**	−.05	−.01	−.14	.07	−.02
More specific	.15	.12	.09	.21*	−.03	.06
More positive/less critical in tone or mood	.02	.01	−.20*	−.03	−.04	.00
Most obvious	.00	−.12	−.20*	−.21*	−.13	−.08
Least obvious	.03	−.12	.09	.13	.03	.19*
More carefully worded/qualified	.08	−.01	.11	.11	.15	.21*
Shorter	−.10	.02	−.20*	.05	−.06	−.19
More concrete	.11	.17*	.09	.01	.07	.05
Too neutral	−.16	−.03	−.08	−.23*	−.04	.03
More uncommon/ unusual	−.03	−.12	−.04	.19*	.00	−.03
Outdated, old-fashioned	−.13	−.18*	−.08	−.01	−.03	−.02

*$p<.05$, **$p<.01$.

TABLE 11

Significant Correlations between Test Performance and Use of Answer Choice Features to Rule Out Answers

	Question Set					
Feature	*Language*	*Marianne Moore*	*Clarence Darrow*	*Tunguska*	*Architecture*	*Richard Wright*
Simpler	−.06	−.00	−.11	.15	.08	.30**
More negative/more critical in tone or mood	.08	.13	−.27**	−.06	.04	.05
Least obvious	.06	−.01	−.21*	−.16	−.08	−.14
More abstract	.14	−.19*	.02	.00	−.01	−.03
Too similar to other choices	−.02	.04	.14	.20*	.01	.13
Most ambiguous	.18*	.11	.03	−.01	.11	−.05
In tune with current thinking	−.06	−.14	.04	−.07	−.00	−.21*

*$p<.05$, **$p<.01$.

increased substantially when examinees chose the alternative that was more positive/less critical in tone than the other options for the first item, and used personal knowledge about the topic for the last item. For each of the other items, the odds shifted less dramatically with the use of particular strategies but still changed by a factor of more than two to nearly six for various items. For three of the six items listed, attempting to reconstruct the theme or main idea of the missing passage was linked to better performance. Table 13 gives the percentages of examinees who answered these items correctly according to whether or not they used these strategies. The different rates of success are consistent with the results of the logistic regressions. The product moment correlations of strategy use with item performance are also given for those readers who may find this statistic easier to interpret than the logistic regression statistics.

Relationships between Without-Passage Scores and SAT Scores

Table 14 displays, for each of the six sets of questions, the correlations between performance without passages and SAT scores, Test of Standard Written English (TSWE) scores, and high school grades (average and rank in class). The patterns of correlations are quite different for each question set. Performance on three of the sets (Language, Clarence Darrow, and Tunguska) correlated significantly with SAT-Verbal total scores, with verbal subscores, and with TSWE scores. SAT-Mathematical scores did not correlate with performance on any of the six sets. On two of the three sets for which correlations with the test scores were noted (Clarence Darrow and Tunguska), performance also correlated significantly with high school grades. The lower correlation of the Architecture passage with all measures is probably largely a function of the lower reliability of this set of questions (coefficient alpha = .09) compared with the estimated reliability of the other sets (alpha = .33 to .43). These differential patterns have implications for interpreting performance. Particularly noteworthy is the finding that, at least for some sets, those who can glean information from the test questions (without the passages) tend to be those who get good grades.

Student Recommendations

As a final question, study participants were asked if they would ever recommend answering SAT questions without reading the passages. A minority (about 16 percent) said that they would. When asked to indicate the circumstances under which they would follow this strategy, about half said that the approach might help primarily if time were running short. Some students suggested variants of the tactic, such as skimming a passage or reading the first and last sentences in each paragraph. Still others hinted that they would first read the questions and then selectively read the passage, searching for the answers. Several others contended that questions referring to specific lines in a passage could be answered by consulting only those lines instead of reading the entire passage. The second most often mentioned circumstance involved the use of personal knowledge of the subject of a passage.

Finally, students were given an opportunity to provide any other comments of their choosing, and many took the opportunity to do so. Many study participants opined that answering questions without the passages is probably not a major factor in test takers' performance: "I don't really know of many people who do answer the

TABLE 12

Logistic Regression of Performance on Most Helpful Strategies for Selected Questions

Explanatory Variable	*Beta* (B)	*Logistic Regression Statistics* *Standard Error*	*Wald Statistic*	*Exp*(B)
Marianne Moore/Question 1				
F5 More positive/less critical in tone or mood	2.71	1.06	6.5*	15.0
R6 Trying to reconstruct the theme or main idea of the missing passage by reading all the questions and answers	1.70	0.81	4.4*	5.5
Tunguska/Question 1				
R1 Trying to determine the meaning of a word, or phrase, or the way in which it was used, from the other questions in the set	1.58	0.66	5.7*	4.8
Tunguska/Question 2				
S4 Meaning seemed to fit best with the general theme or context suggested by the questions	0.95	0.49	3.8*	2.6
Tunguska/Question 7				
R6 Trying to reconstruct the theme or main idea of the missing passage by reading all the questions and answers	1.72	0.73	5.6*	5.6
Richard Wright/Question 1				
R6 Trying to reconstruct the theme or main idea of the missing passage by reading all the questions and answers	1.51	0.61	6.2*	4.5
Richard Wright/Question 4				
P1 Using personal knowledge about the topic	2.44	1.06	5.3*	11.5

*$p<.05$.

TABLE 13

Percentage of Test Takers Who Answered Selected Questions Correctly, by Use of Selected Strategies

Explanatory Variable	*Did Not Use Strategy*	*Used Strategy*	r
Marianne Moore/Question 1			
F5 More positive/less critical in tone or mood	53	94	.28
R6 Trying to reconstruct the theme or main idea of the missing passage by reading all the questions and answers	55	85	.23
Tunguska/Question 1			
R1 Trying to determine the meaning of a word, or phrase, or the way in which it was used, from the other questions in the set	20	50	.21
Tunguska/Question 2			
S4 Meaning seemed to fit best with the general theme or context suggested by the questions	51	69	.24
Tunguska/Question 7			
R6 Trying to reconstruct the theme or main idea of the missing passage by reading all the questions and answers	37	73	.18
Richard Wright/Question 1			
R6 Trying to reconstruct the theme or main idea of the missing passage by reading all the questions and answers	48	81	.29
Richard Wright/Question 4			
P1 Using personal knowledge about the topic	56	94	.30

TABLE 14

Correlations between SAT Scores and Performance on Questions without Passages

Form/Passage	Total	SAT-Verbal Reading	Vocabulary	TSWE	SAT-Mathematical	High School Grades Average	High School Grades Rank
A. Language	.48**	.46**	.40**	.37**	.14	.09	.11
A. Marianne Moore	.09	.17	.02	.06	−.14	−.08	−.10
B. Clarence Darrow	.37**	.37**	.32**	.34**	.14	.28*	.19
B. Tunguska	.32**	.34**	.29**	.26*	.19	.23*	.34**
C. Architecture	.17	.19	.12	.13	.06	.01	−.05
C. Richard Wright	.06	.07	.04	.02	−.16	.14	.13

Note: Correlations between the SAT and TSWE and questions in Forms A, B, and C are based on 95, 82, and 91 students, respectively, for whom SAT scores were available. Correlations involving high school grade-point averages are based on 86, 75, and 81 cases, respectively, for the three test forms. Correlations involving high school ranks are based on 76, 65, and 74 cases, respectively.
*$p<.05$, two-tailed; **$p<.01$, two-tailed.

questions without reading the passage first. So I don't think the testing service has much to be worried about." Others commented: "I've never heard of anyone using the strategy." "I thought it could be done, but it takes time and is painstaking." "I think it's important and interesting to realize that you can make educated guesses without even having to read the passage." Finally, one student advised fellow test takers as follows: "Instead of learning stupid strategies, just learn to read."

Discussion

The results indicated that, as expected, students were able to attain scores that exceeded a chance level. Thus the findings are generally consistent with those obtained for other measures of reading comprehension, including the current verbal portion of the SAT (Katz et al. 1990). The findings also extend the results of earlier studies to the kind of reading questions that will be used in the modification of the SAT.

It is clear from the results that some helpful information can be obtained from reading only the questions themselves. It is equally apparent, however, that resorting to answering questions exclusively on this basis is neither efficient nor effective. Students used nearly as much time per question without reading the passages as is currently allotted for reading SAT passages *and* answering questions. And, although most students did better than chance without the passages, they did not perform substantially better, given what might be expected of the very capable students in the study sample. Nor was the use of a no-passage strategy very dependable. According to indices of internal consistency, the resulting performance varied considerably among items. Thus, any concern about testwise examinees using this strategy to gain an unfair advantage over less savvy test takers seems unfounded. Students' comments reinforced this conclusion.

In addition, better-than-chance performance was not uniform over individual items. A few questions seemed barely more difficult without the passages than with them. However, without the passages, performance on some items was actually lower than would be expected by chance. Items that involved the meaning of a word or phrase tended to be quite easy when the meaning was a relatively common one that might be known without reference to the context in which it occurred. Performance on some other vocabulary questions, however, was worse than would be expected by random guessing, usually because a less common meaning of a word was signaled by the context. When the more usual meaning of the word appeared as one of the alternatives, examinees were likely to choose this typical meaning instead of the correct answer.

Other factors were implicated in difficulty differentials for other items. In one particular set, questions were relatively easy, apparently because examinees were familiar with the author or with the particular passage on which the questions were based. Each of the questions in this set was even easier for examinees who took the test with the passage. Nonetheless, this example illustrates that awareness of an author or of an author's works may make questions easier, both with and without the passages, than when questions are based on less familiar sources. Given current research on reading comprehension, this finding should come as no surprise.

With respect to establishing links between the use of particular strategies and performance on test items, examinees' comments strongly suggested that some strategies were more heavily implicated in success than others. The strategies most frequently mentioned as being helpful involved verbal reasoning rather than skills that rely

on characteristics of questions or answer choices. The two strategies mentioned most often as being helpful involved attending to consistencies among questions and attempting to reconstruct the theme of a missing passage from all the available questions and answer choices. The use of this second strategy was linked to successful performance on particular items more often than any other strategy. These findings seem to run counter to the earlier research by Katz, Blackburn, and Lautenschlager (1991), who effectively eliminated these strategies from examinees' repertoires by scrambling questions from several passages. Even under these conditions, however, examinees were still able to answer questions at a greater-than-chance level without the passages, leading the authors to conclude that information obtained from other questions associated with a passage has relatively little effect on performance without passages. This apparent discrepancy between our results and those of Katz et al. (1991) is consistent with the hypothesis (discussed below) that students invoke different strategies with and without passages.

The use of strategies involving reasoning was not unexpected. That reading and reasoning are intertwined, perhaps inextricably, has been asserted, quite literally, for generations (E. L. Thorndike 1917; R. L. Thorndike 1973–74). More recently, Stanovich and Cunningham (1991) discussed the evolution of reading theory, noting that although reading may indeed be reasoning, it is reasoning of a particular sort. They suggested that a better conception might be reading as "constrained reasoning," since comprehension is at least partially determined by both a reader's expectations and his or her use of world knowledge to supplement text. (They also argued, however, that there are situations in which we would hope that readers do not impose their own meanings on what they read, for example, when a physician consults a medical reference to prescribe an appropriate treatment.)

The study results appear to have definite implications for test development. It is clear that a test development strategy could be implemented that would foil test takers who rely heavily on answering reading comprehension questions without reading the passages. There were several examples of test items for which failure to read the passages resulted in performance that was actually worse than would be attained by random guessing. It appears that items of this sort could be constructed relatively easily. However, resorting to such a strategy seems unnecessary, and the effects on the validity of test scores would need to be established before any such plan was implemented.

With regard to construct validity, Katz et al. (1990, 122) concluded that currently used SAT reading comprehension questions "substantially measure factors unrelated to reading comprehension." This conclusion rests on sound empirical evidence, but it also relies on several major premises and on a particular conception of reading. For instance, the investigators asserted that if SAT reading passages are "essential" to the task, then "no relation" should exist between performance exhibited with and without the passages. Finding correlations between performance with and without the passages, the authors concluded that "the passage is not a necessary component of the task" and that factors other than the ability to understand the passages "contribute strongly to differences among SAT-V scores" (Katz et al. 1990, 126). There is no disputing that, for many questions, examinees can indeed perform at a better-than-chance level without the passages. Thus, in a strictly logical sense, the passages are not essential to attain performance that is better than that expected from random guessing. This is not, however, a level of performance with which most motivated SAT takers would be satisfied. Consulting the information available in the passages would be considered extremely desirable, if not absolutely necessary, by most examinees wishing to achieve test scores that reflect their true verbal reasoning abilities and are competitive in college admission.

A second questionable premise on which Katz et al. (1990) base their conclusions is that a correlation between performance with and without the passages implies that the same strategies are used with and without the passages. The only truly warranted conclusion is that test takers who perform better than their peers when they have the passages also tend to do better without them. When confronted with an unconventional task like the one posed here, verbally able students are apparently able to shift gears, invoking other strategies that seem not only beneficial with regard to test taking but also relevant to verbal reasoning. This flexibility—to adapt strategies to the particular situation—is exactly the kind of trait that characterizes skilled readers (see, e.g., Foertsch 1992). This is probably not surprising, for it is well established that many different kinds of logically distinct abilities are relatively highly correlated. For example, the verbal and mathematical abilities measured by the SAT correlate in the .60s (Donlon 1984). And even such quasi abilities as testwiseness are not completely unrelated to other academic abilities (see, e.g., Diamond and Evans 1972).

Further support for a shifting-gears hypothesis comes from Freedle (1990), who showed that different factors determine the difficulty of reading comprehension questions according to whether or not the passages are available. When passages are accessible, difficulty seems to depend to a much greater extent on the characteristics of

the passages than on the features of the questions. When they are not available, item characteristics exert a greater influence.

Finally, Katz et al. (1990) seem, at least implicitly, to endorse a particular conception of reading that puts a premium on the extraction of literal information from text. Current views of reading, however, stress that readers are "active, constructive, motivated learners" (Garner 1987, 1). Cognitively based conceptions of reading comprehension in particular emphasize the interactive and constructive nature of reading, in which readers' background knowledge is all important. As Dole, Duffy, Roehler, and Pearson (1991, 241) assert: "All readers, both novices and experts, use their existing knowledge and a range of cues from the text and the situational context in which the reading occurs to build, or construct, a model of meaning from the text."

Given the results of the current study and the prevailing conceptions of reading comprehension, what are the implications for advising test takers? The study results do not suggest any particular need to modify the information that is now provided to SAT takers (College Board 1991). Currently, the following information is conveyed to prospective test takers:

> The verbal questions test your verbal reasoning and understanding of what you read. [p. 3]
>
> . . . no specialized knowledge in science, social studies, literature, or other fields is needed. [p. 7]
>
> The reading comprehension questions on the SAT measure your ability to read and understand a passage. Each passage or pair of related passages contains all of the information you'll need to answer the questions that follow. [p. 11]
>
> A passage with a subject that is familiar or interesting to you may be easier for you than a passage that is about an unfamiliar subject. [p. 15]
>
> Answer questions on the basis of what is *stated* or *implied* in the passage. Don't answer questions on the basis of your personal opinion or knowledge. [p. 15]

None of this information seems inconsistent with the results of the study discussed here. First, the new SAT reading passages do indeed appear to depend on one's ability to read and understand passages. This does not imply, however, that other kinds of skills do not also come into play, at least to some degree and under some circumstances. Second, the new SAT reading questions can be answered from the information contained in the passages. Again, this does not mean that personal experiences and knowledge cannot also provide a useful backdrop. As is currently suggested, interest in and familiarity with a topic may facilitate understanding. Perhaps the only necessary fine-tuning of information about the new SAT concerns the suggestion that test takers answer questions solely on the basis of the passages. Certainly this should be students' first resort. But, if this fails, another potentially useful tactic may be to invoke personal opinion or knowledge. Students should be advised, however, that this approach may in some instances be ineffective and even counterproductive.

In conclusion, performance on the kinds of reading comprehension questions that will make up the revised SAT does not appear to depend exclusively on information contained in the reading passages on which the questions are based. However, the importance of nonpassage factors appears to be relatively limited, especially in relation to the influence exerted by the reading passages. Furthermore, the other factors most heavily implicated in test performance without access to the passages are not unrelated to the verbal reasoning skills involved in academic success. The desired interpretation of reading scores based on the new SAT reading comprehension questions does not seem unduly threatened by examinees' ability to benefit from information contained in the test questions themselves.

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. 1985. *Standards for Educational and Psychological Testing*. Washington, D.C.: American Psychological Association.

College Board. 1991. *Taking the SAT 1991–92*. New York: College Entrance Examination Board.

College Board. 1992. *College Bound Seniors: 1992 Profile of SAT and Achievement Test Takers*. Princeton, N.J.: Educational Testing Service.

Conlan, G. 1990. "Text and Context: Reading Comprehension and the Mechanics of Meaning." *College Board Review* 157:19–25.

Diamond, J.J., and W.J. Evans. 1972. "An Investigation of the Cognitive Correlates of Test-Wiseness." *Journal of Educational Measurement* 9:145–150.

Dole, J.A., G.G. Duffy, L.R. Roehler, and P.D. Pearson. 1991. "Moving from the Old to the New: Research in Reading Comprehension Instruction." *Review of Educational Research* 61:239–264.

Donlon, T.F., ed. 1984. *The College Board Technical Handbook for the Scholastic Aptitude Test and Achievement Tests*. New York: College Entrance Examination Board.

Farr, R., R. Pritchard, and B. Smitten. 1990. "A Description of What Happens When an Examinee Takes a Multiple-Choice Reading Comprehension Test." *Journal of Educational Measurement* 27:209–226.

Foertsch, M.A. 1992. *Reading in and out of School.* Washington, D.C.: Office of Educational Research and Improvement.

Freedle, R. 1990. "Evidence for the Construct Validity of the SAT Reading Comprehension Items." Unpublished report. Princeton, N.J.: Educational Testing Service.

Garner, R., ed. 1987. *Metacognition and Reading Comprehension.* Norwood, N.J.: Ablex.

Katz, S., A.B. Blackburn, and G.J. Lautenschlager. 1991. "Answering Reading Comprehension Items without Passages on the SAT When Items Are Quasi-Randomized." *Educational and Psychological Measurement* 51:747–54.

Katz, S., G.J. Lautenschlager, A.B. Blackburn, and F.H. Harris. 1990. "Answering Reading Comprehension Items without Passages on the SAT." *Psychological Science* 1:122–27.

Lawrence, I.M. 1992. "Comparison of Statistical Specifications for Current and New SAT-V." Unpublished report. Princeton, N.J.: Educational Testing Service.

Millman, J., C.H. Bishop, and R. Ebel. 1965. "An Analysis of Testwiseness." *Educational and Psychological Measurement.* 25:707–26.

Preston, R.C. 1962. "A New Approach to Judging the Validity of Reading Comprehension Tests: Summary of an Investigation." In *Challenge and Experiment in Reading*, ed. J.A. Figurel. International Reading Association Conference Proceedings 7:166–67.

Preston, R.C. 1964. "Ability of Students to Identify Correct Responses before Reading." *Journal of Educational Research* 58:181–83.

Pyrczak, F. 1972. "Objective Evaluation of the Quality of Multiple-Choice Test Items Designed to Measure Comprehension of Reading Passages." *Reading Research Quarterly* 8:62–71.

Pyrczak, F. 1974. "Passage-Dependence of Items Designed to Measure the Ability to Identify the Main Ideas of Paragraphs: Implications for Validity." *Educational and Psychological Measurement* 34:343–48.

Pyrczak, F. 1975. "Passage Dependence of Reading Comprehension Questions: Examples." *Journal of Reading* 18:308–11.

Stanovich, K.E., and A.E. Cunningham. 1991. "Reading as Constrained Reasoning." In *Complex Problem Solving: Principles and Mechanisms*, ed. R.J. Sternberg and P.A. Frensch, pp. 3–60. Hillsdale, N.J.: Erlbaum.

Thorndike, E.L. 1917. "Reading as Reasoning: A Study of Mistakes in Paragraph Reading." *Journal of Educational Psychology* 8:323–32.

Thorndike, R.L. 1973–74. "Reading as Reasoning." *Reading Research Quarterly* 9:135–47.

Tuinman, J.J. 1973–74. "Determining the Passage Dependency of Comprehension Questions in 5 [*sic*] Major Tests." *Reading Research Quarterly* 9:206–23.

Weaver, W.W., and A.C. Bickley. 1967. "Sources of Information for Responses to Reading Test Items." In *Proceedings of the 75th Annual Convention of the American Psychological Association*, 293–94.