

Predictive Validity of SAT[®] I: Reasoning Test for Test-Takers with Learning Disabilities and Extended Time Accommodations

Cara Cahalan, Ellen B. Mandinach, and Wayne J. Camara

Predictive Validity of SAT[®] I: Reasoning Test for Test-Takers with Learning Disabilities and Extended Time Accommodations

Cara Cahalan, Ellen B. Mandinach, and Wayne J. Camara

Cara Cahalan is a senior research associate at Educational Testing Service (ETS).

Ellen B. Mandinach is a senior research scientist at Educational Testing Service (ETS).

Wayne J. Camara is vice president of Research and Development at the College Board.

Researchers are encouraged to freely express their professional judgment. Therefore, points of view or opinions stated in College Board Reports do not necessarily represent official College Board position or policy.

The College Board: Expanding College Opportunity

The College Board is a national nonprofit membership association whose mission is to prepare, inspire, and connect students to college and opportunity. Founded in 1900, the association is composed of more than 4,200 schools, colleges, universities, and other educational organizations. Each year, the College Board serves over three million students and their parents, 22,000 high schools, and 3,500 colleges through major programs and services in college admissions, guidance, assessment, financial aid, enrollment, and teaching and learning. Among its best-known programs are the SAT[®], the PSAT/NMSQT[®], and the Advanced Placement Program[®] (AP[®]). The College Board is committed to the principles of excellence and equity, and that commitment is embodied in all of its programs, services, activities, and concerns.

For further information, visit www.collegeboard.com.

Additional copies of this report (item #994216) may be obtained from College Board Publications, Box 886, New York, NY 10101-0886, 800 323-7155. The price is \$15. Please include \$4 for postage and handling.

Copyright © 2002 by College Entrance Examination Board. All rights reserved. College Board, Advanced Placement Program, AP, SAT, and the acorn logo are registered trademarks of the College Entrance Examination Board. Admitted Class Evaluation Service and ACES are trademarks owned by the College Entrance Examination Board. PSAT/NMSQT is a registered trademark jointly owned by both the College Entrance Examination Board and the National Merit Scholarship Corporation. Other products and services may be trademarks of their respective owners. Visit College Board on the Web: www.collegeboard.com.

Printed in the United States of America.

Acknowledgments

We express our gratitude to the individuals at the admissions and disability service offices of all the colleges and universities that participated in this research effort. This study could not have been conducted without their time and effort. In addition, we wish to thank Randy Bennett, Brent Bridgeman, Loring Brinckerhoff, Carol Dwyer, Sheree Johnson-Gregory, Cathy Wendler, and June Zumoff for their input into the study and their insightful reviews. Finally we thank both the College Board and Educational Testing Service for providing funding for this research.

Contents

<i>Abstract</i>	1
I. <i>Introduction</i>	1
<i>Testing Accommodations</i>	1
<i>Validity of SAT for Test-Takers with Learning Disabilities</i>	2
<i>Changes: Population, Item Type, and Time Limits</i>	3
<i>Validity of Revised and Recentered SAT I</i>	4
<i>Summary</i>	4
II. <i>Methods</i>	5
<i>Subjects</i>	5
<i>Variables</i>	6
<i>Procedures</i>	7
III. <i>Results</i>	7
<i>Adjusted Correlations</i>	7
<i>Over- and Underprediction of FGPA</i>	8
IV. <i>Conclusion</i>	9
<i>References</i>	10
<i>Appendix A: Examples of SAT Accommodations</i>	11
<i>Appendix B: Guidelines for Documentation</i>	12

Tables

1. Means and Standard Deviations of Scores and GPAs by Gender and Services Provided for Disability for Sample 1	5
2. Means and Standard Deviations of Scores and GPAs by Gender and Test Accommodation for Sample 2	6
3. Correlations with FGPA, Adjusted for Range Restriction	8
4. Over- (+) and Underprediction (-) of FGPA by Gender and Disability for Sample 2	9

Abstract

The predictive validity of the SAT® I: Reasoning Test was examined for students who took the test with an extended time accommodation for a learning disability. The sample included college students with learning disabilities who took the SAT I between 1995 and 1998 with extended time accommodations. First year grade point average (FGPA) was used as a measure of student performance. Although positive, the adjusted correlation between FGPA and SAT scores was lower for test-takers with a learning disability than has been shown in prior research on test-takers without disabilities. In addition, the SAT scores obtained with an extended time accommodation appear to overpredict FGPA for male test-takers with a learning disability and accurately predict FGPA for female test-takers with a learning disability. When the same students were examined using both SAT I test scores and self-reported high school grade point average (HSGPA) to predict FGPA, the scores and grades of male test-takers did not under- or overpredict while the scores of female test-takers underpredicted FGPA. Due to the relatively small sample size, additional research is required to examine group differences (e.g., type of learning disability, severity of disability) and the impact of differential support received from college disability service offices during the first year of college.

Key Words: SAT, Learning Disability, Extended Time, Validity, Accommodations

I. Introduction

Twenty-seven years after passage of the Education for All Handicapped Children Act of 1975 (Public Law 94-142), students with disabilities are receiving services and applying to college in record numbers. Since students with learning disabilities currently account for the largest percentage of college freshmen with disabilities (approximately 40 percent at four-year colleges and universities) and approximately half of all students with disabilities between the ages of 3 and 21 years, it is not surprising that the most common accommodation on college admissions tests is extended time for a learning disability (Cahalan, 2000; Henderson, 2001; U.S. Department of Education, 2000). Between 1990 and 1995, the percentage of students with a learning disability taking the SAT I: Reasoning Test (with an accommodation) increased an average of 14 percent per year and has since stabilized to account for approximately 1.5 percent of all SAT I test-takers (Cahalan,

2000; Camara, Copeland, and Rothschild, 1998). During a similar time period (1991–1996), the percentage of college freshmen reporting to have a learning disability increased an average of 9 percent per year and then stabilized to account for approximately 2.4 percent of all college freshmen at four-year colleges and universities (Henderson, 2001). Although the percentage of students and test-takers with a learning disability has increased, the percentage of college freshmen reporting a learning disability is substantially larger. This rise of accommodations on standardized tests and college students with reported learning disabilities has led to a greater interest in the predictive validity of test scores that are obtained with an accommodation. Although each accommodation is intended to provide equal access by removing unnecessary challenges (construct irrelevant variance), some types of accommodations appear to change the test's construct and may thus be expected to alter the predictive validity of test scores. Research in this area is necessary but difficult to conduct due to the (a) multiple types of accommodations, (b) variety and severity of disabilities, and (c) controversy regarding how each accommodation changes the tests' constructs.

Testing Accommodations

Testing accommodations are commonly grouped into four categories: presentation, response, timing, and setting. Presentation accommodations provide students with an alternative presentation of testing materials such as Braille, large print, and audiocassettes. Some test-takers with disabilities request response accommodations that include providing a scribe to record answers and the use of a computer for essay tests. Timing accommodations include extended time, the most widely requested accommodation, as well as frequent breaks and multiple testing sessions. On some tests, such as the SAT I, the extended time accommodation also includes the removal of breaks between test sections; thus allowing examinees to complete test sections in any order they wish and return to prior test sections to review and change answers. Setting accommodations consist of a private room, screens to block out distractions, and other changes to the test-taker's surroundings. See Appendix A for a more complete list of testing accommodations.

Just as there are many types of testing accommodations, the variety and severity of disabilities is broad. Since accommodations are reviewed on a case-by-case basis, students with the same overarching disability may require different accommodations, and in some cases no accommodation is necessary. For example, one student with dyslexia (a learning disability) may be granted extended time, while another student with dyslexia may be granted

extended time and an audiocassette. These differences are due to the severity of the disability, the accommodations requested by the student, and prior history of accommodations both in-class and on standardized tests. In most cases a student with a disability is granted all reasonable accommodations requested. Currently students who request accommodations on the SAT I and other College Board tests submit requests through their high school if (a) the student has an academic history of a disability, (b) documentation of the disability meets the College Board's *Guidelines for Documentation* (see Appendix B), and (c) the student has received the requested accommodation (due to the disability) on school-based tests. If one or more of the three criteria is not met, requests for accommodations are directed through an appeals process, which requires the examinee to provide documentation to verify the test-taker's disability and need for a specific testing accommodation. All appeals are reviewed by disability specialists who determine if the student has a disability and if the disability necessitates the specific accommodation requested.

Most accommodations do not appear to change a test's construct, so the interpretation of test scores is not influenced. For example, a reading test written in large print should not change the test's construct because reading is still being assessed. Several accommodations, however, are suspected to change the construct of some tests. Allowing the administrator to read aloud a reading test, for instance, would change the construct of the test from a reading test to a listening comprehension test. Although extra time could not change a reading test to a listening comprehension test, there is some debate regarding the impact of extended time on the predictive validity of tests that are speeded. Extra time allows students to reach questions at the end of a speeded test that they may not have attempted under standard timing conditions, but no general consensus has been reached regarding how extended time changes what the test purports to measure or the predictive validity of test scores.

Phillips (1994) argues that measurement specialists should consider the impact of modifications on changes in the constructs measured and the test's validity. Once modifications have changed test constructs for some individuals, the users of the test can no longer rely on the ability of the test to determine qualifications for admission, employment, certification, or licensure. Assuming that examinees are truly disabled and are incapable of adapting to the standard testing administration, Phillips argues that any changes to testing conditions should be avoided if the change would (a) alter the skill being measured, (b) preclude the comparison of scores between examinees that received accommodations and those that did not, or (c) allow examinees without disabilities to

benefit (if they were granted the same accommodation). This last criterion is debatable and recently several researchers have argued that accommodations should only be provided if they offer a "differential" boost to students with disabilities (Elliott and McKeivitt, 2000; Fuchs and Fuchs, 1999; Pitoniak and Royer, 2001).

For the purpose of this paper we will examine Phillips' second criterion (score comparison) as it pertains to extended time. This study will isolate a single, albeit varied, disability (learning) and test (SAT I: Reasoning Test) in order to examine the predictive validity of scores taken with an extended time accommodation. Although extended time is a commonly used accommodation for most college admissions and state achievement tests (Thurlow, Elliott, and Ysseldyke, 1998), research on extended time is limited. Since extended time does not alter test questions or response format it is unlikely that extended time violates Phillips' first criterion (alters the skill being measured). The third criterion (allows examinees without disabilities to benefit) has recently been investigated by Bridgeman, Curley, and Trapani (2002) and will be discussed later.

Validity of SAT for Test-Takers with Learning Disabilities

The research on extended time for test-takers with learning disabilities is limited, but there are indications that scores on some tests taken with extended time may over-predict performance (Wightman, 1993; Willingham, Ragosta, Bennett, Braun, Rock, and Powers, 1988; Zuriff, 2000). Extended time, however, can differentially impact each test because tests differ in terms of format, content, populations of test-takers, and how the test is used. Two studies have specifically focused on the predictive validity of the SAT for test-takers with learning disabilities (Ragosta, Braun, and Kaplan, 1991; Braun, Ragosta, and Kaplan, 1988). Both of these studies were conducted with a portion of students who took the SAT between 1979 and 1983 and attended a four-year college or university.

The first study by Braun, Ragosta, and Kaplan (1988) investigated the predictive validity of the SAT for students who took the test with and without an accommodation. Students were divided into five groups based upon their disability status: no disability, hearing, learning, physical, and visual. The use of accommodations separated test-takers further; those that received an accommodation and those that took standard test administrations. No distinction was made between students who received different types of accommodations (e.g., extended time or audiocassettes). For the purpose of this study the following three groups will be reviewed: (1) test-takers without a disability who took a standard administration

($n=6,255$), (2) test-takers with a learning disability who took the SAT with an accommodation ($n=437$), and (3) test-takers with a learning disability who took a standard administration ($n=99$). All students had data on high school grade point average (HSGPA), SAT verbal score (SAT-V), SAT math score (SAT-M), and first year college grade point average (FGPA). In order to make FGPA consistent across each institution, standardized values were computed so that the mean FGPA for each institution was zero with a standard deviation of one.

Adjusted correlations were computed between SAT-M, SAT-V, and actual FGPA for the three groups. SAT-M and FGPA correlated the highest for the standard administrations ($r=.25$ for Group 1 and $r=.26$ for Group 3) and lowest for test-takers with learning disabilities who took the test with an accommodation ($r=.10$ for Group 2). SAT-V and FGPA correlated highest for test-takers without disabilities ($r=.26$ for Group 1) and lowest for test-takers with learning disabilities who received an accommodation ($r=.09$ for Group 2). Verbal scores and FGPA correlated at .14 for Group 3. In addition to correlations between actual grades and test scores, correlations between predicted FGPA and actual FGPA were examined. Predicted FGPA was calculated using combined SAT scores with and without HSGPA. Correlations between actual and predicted college performance was lower for test-takers with learning disabilities than test-takers with no disability. When only SAT scores were used to calculate predicted FGPA, actual and predicted FGPA correlated lowest for test-takers with learning disabilities who received an accommodation ($r=.22$ for Group 2) and highest for students without disabilities ($r=.37$ for Group 1). Predicted and actual FGPA correlated at .29 for Group 3 (test-takers with a learning disability and standard administration). When predicted FGPA was calculated using SAT scores and HSGPA, the correlations between actual and predicted FGPA were higher. Although correlations were still highest for students without a disability ($r=.49$), no discrepancy was found between test-takers with a learning disability who received accommodation ($r=.34$) and those that had a standard administration ($r=.33$).

In 1991, Ragosta, Braun, and Kaplan completed a follow-up study with a portion of the students in their 1988 study. The second study included 256 students with learning disabilities (203 with testing accommodations) and 2,472 students without a disability. The purpose was to examine the predictive validity of the SAT on FGPA, overall college grade point average (OGPA), and graduation rate. The adjusted correlation between actual and predicted grade point average (from SAT scores alone) for students with learning disabilities who took the SAT with an accommodation and gradu-

ated was substantially higher for OGPA ($r=.36$) than FGPA ($r=.26$). The relationship between grades and SAT scores was stronger for students without disabilities; but virtually no differences were found between the correlations of SAT with OGPA ($r=.52$) and SAT with FGPA ($r=.51$). A similar pattern was found when predicted GPA was calculated using SAT and HSGPA; OGPA correlated at .46 and .62 and FGPA correlated at .40 and .59, respectively, for students accommodated for a learning disability and students without a disability.

These two studies indicate that the SAT does not predict FGPA as well for students with learning disabilities who take the SAT with an accommodation as it does for students who do not have a disability. Although both studies provided information on the predictive validity of the SAT, several changes that came about during the mid 1990s necessitate that this issue be revisited.

Changes: Population, Item Type, and Time Limits

Since the completion of these studies, two significant changes have occurred that may alter the predictive validity of the SAT. The first change relates to the increase in the percentage of test-takers with learning disabilities that occurred in the early and mid 1990s. As was described earlier, the percent of SAT I test-takers with a learning disability increased by 14 percent per year between 1990 and 1995 and then stabilized to include 1.5 percent of all current SAT I test-takers. This change may be the result of increased awareness and diagnosis of learning disabilities, greater advocacy by the disability community, or a combination of increased awareness, diagnosis, and advocacy. Regardless of the cause, one should not assume that all students who took the SAT with accommodations prior to 1995 perform the same as test-takers who currently receive testing accommodations for a learning disability.

In addition to the larger number of test-takers with learning disabilities, several changes were made to both the math and verbal sections of the SAT in March 1994. On the verbal test, time limits were extended by 15 minutes, an antonyms section was removed, and a set of questions based on a pair of related passages was added. On the math test, time limits were extended by 15 minutes and a new item type was introduced. In addition to the 50 standard multiple-choice items, 10 items require examinees to enter their answers on a grid. Last, the scoring scale was revised so that the average score on both math and verbal scales would be 500. Under current standard time conditions, test-takers have three hours to complete the test; the test is divided into seven sections with separate time limits ranging

from 15 to 30 minutes. Extended time can vary with the specifics of an individual's disability, but in most cases extended time means time and a half (4.5 hours) to complete the test and no section time limits.

Validity of Revised and Recentered SAT I

Some research has examined the predictive validity of the revised and recentered SAT I: Reasoning Test for students without disabilities. Bridgeman, McCamley-Jenkins, and Ervin (2000) examined the predictive validity of the SAT I (for students *without* disabilities) using data from operational test administrations. Subjects included students from the 1994 and 1995 entering classes at 23 colleges and universities. Students entering college in 1994 took the original SAT, while students entering college in 1995 took the new SAT I: Reasoning Test. No significant difference between the predictive validity of the original and new tests was found. Findings indicated the revised SAT is significantly correlated ($r=.35, p < .05$) with first year grade point average (FGPA) for students *without* disabilities who take the test under standard conditions. Predicted FGPA was calculated using the revised SAT scores and HSGPA while controlling for course difficulty. Under- and overpredictions were reported in relation to a "4-point scale" where an F=0.00 and an A=4.00. Results indicated scores from the revised SAT underpredicted FGPA for females (-.10) and overpredicted FGPA for males (.11). When HSGPA *and* the revised SAT were used FGPA was still underpredicted for females (-.07) and overpredicted for males (.08), but the degree of under/overprediction was less.

Other recent research on test-takers *without* disabilities has examined the impact of additional time per test question on test performance by decreasing the number of test items per section (Bridgeman, Curley, and Trapani, 2002). Although this approach is not the same as extending time, it does provide an idea of how students without disabilities would perform with extended time. Researchers created versions of a math and a verbal SAT I subtest that varied in terms of the number of items (e.g., a verbal section that would normally have 35 items was administered with 23, 27, or 35 questions). Items at each difficulty level were removed so that all the shortened versions were of comparable difficulty. The variations in the number of items were roughly equivalent to standard time, time and a quarter, and time and a half per item. Each test version was administered to at least 8,000 test-takers during the experimental section of an operational test administration. The performance of students who took the time-and-a-half equivalent was only slightly better than the performance of students who took the

experimental subtest with standard time. Bridgeman et al. (2002) concluded that if results were extrapolated to a full SAT, test-takers who received the time-and-a-half equivalent test would have a small increase in test scores; 10 points on the verbal scale and 20 points on the math scale. Although further research is needed, this study provides an indication that extended time may not substantially improve performance of test-takers without disabilities. Additional research is currently being conducted that will experimentally manipulate the amount of time students are allocated. This research will also compare test-takers with and without disabilities in order to provide greater insight into how time differentially impacts performance (Mandinach et al., in press).

One study has examined the performance of test-takers with learning disabilities on the revised and recentered SAT I. Research by Camara et al. (1998) examined the difference in performance of test-takers with learning disabilities who took the SAT I: Reasoning Test with and without accommodations. Test scores used for analyses were from the first 13 operational administrations of the revised and recentered SAT I administered during 1994 and 1995. All subjects took the test once during the spring of their junior year and once during the fall of their senior year. Test-takers were separated into four groups based upon their accommodation status during the fall and spring test administrations. The groups included (a) nondisabled test-takers, (b) test-takers with a learning disability who completed two administrations with extended time, (c) test-takers with a learning disability who took a standard administration and retested with extended time, and (d) test-takers with a learning disability who took an extended time administration and retested with standard time. Camara et al. report that the extra time and removal of section timing appears to increase the percentage of questions reached and students' overall score. Score improvement was significantly greater when test-takers with disabilities took a standard administration followed by an extended time condition (standard-extended), than two other comparison groups; students with a learning disability who took an extended time condition followed by a second extended time condition (extended-extended) and test-takers without a learning disability who took two standard time administrations (standard-standard). Scores for test-takers in the fourth group (extended-standard) decreased between their junior and senior year.

Summary

Due to an increase in the diagnosis of learning disabilities and recentering of the SAT, it is now time to reexamine the predictive validity of the SAT I for students with learning disabilities. Earlier research indicated that the original

SAT did not predict FGPA and OGPA as well for students with learning disabilities (who received any accommodation) as it did for test-takers without a disability who took a standard administration of the test. In addition, the revised and recentered SAT I is comparable to the original SAT in terms of the test's ability to predict FGPA for students without disabilities. Although research on the revised SAT indicates that the performance of students with disabilities may increase with extended time, no research has investigated validity. The purpose of this study is to examine if the revised and recentered SAT I is predictive of FGPA for students with learning disabilities who take the test with a specific accommodation: extended time (i.e., time and a half or double time). These findings will help determine the best use of test scores for students with learning disabilities who receive extended time.

II. Methods

Subjects

Subjects in this study were obtained from two sources that differ in terms of data collection methodology and in some of the variables used for analyses. In both samples colleges were selected based on unusually high numbers of students who submitted test scores from nonstandard administrations. Subjects who took the SAT I: Reasoning Test with an accommodation for a learning disability will be referred to as “test-takers with disabilities,” “learning disabled,” or “students with learning disabilities.” It is important to keep in mind that some students with learning disabilities may have been excluded from this sample because they (a) opted to take a standard administration, (b) do not require an accommodation, or (c) for whatever reason did not take

the test with an accommodation (e.g., undiagnosed disability, unaware of testing accommodations). All identifying information was removed from the database so that students' anonymity would be retained.

Sample 1. The first sample consists of test-takers with learning disabilities from the 1997 and 1998 entering classes at 14 institutions in 10 states. Twenty-five colleges were contacted initially, but 11 chose not to participate due to time constraints or lack of personnel. Although an equal number of public and private colleges participated, the public college students accounted for approximately 70 percent of the sample (the private schools in this sample had an average enrollment of 3,037 undergraduates, while the public schools had an average enrollment of 14,570 undergraduates). Each institution that agreed to participate was sent a list of students who had (a) requested their SAT I test scores be sent to that specific college or university, (b) reported that they graduated from high school in 1997 or 1998, and (c) took the SAT I: Reasoning Test with an accommodation for a learning disability and no other disability (i.e., physical, hearing, visual, temporary, or other). Of the original 3,471 test-takers meeting these three selection criteria, 463 had been accepted and attended one of the colleges in the sample for the first year; these comprised Sample 1. In the final sample, 58 percent of the students were male, 88 percent were white, 64 percent were receiving services for their disability from the college disability service office, and none reported speaking another language better than English. Variables in Sample 1 included first year college grade point average (FGPA), SAT scores, disability information, self-reported descriptive data, high school grade point average (HSGPA), and information on if the student received services from the college office for students with disabilities. These variables will be explained in detail in the next section. Table 1 displays descriptive statistics for test scores and grades by gender and services received.

TABLE 1

Means and Standard Deviations of Scores and GPAs by Gender and Services Provided for Disability for Sample 1

	N	SAT I-V		SAT I-M		HSGPA		FGPA	
		M	SD	M	SD	M	SD	M	SD
No Service Provided for Disability									
Male	83	550	88	570	102	2.96	.57	2.38	.81
Female	165	526	89	553	107	2.92	.43	2.33	.80
Total	248	534	89	559	105	2.93	.47	2.35	.80
Services Provided for Disability									
Male	71	533	90	530	91	3.19	.47	2.53	.90
Female	103	524	89	513	101	3.05	.43	2.56	.61
Total	174	528	89	520	97	3.10	.45	2.55	.75
Total									
Male	271	531	91	556	104	2.90	.48	2.32	.81
Female	192	526	90	517	96	3.07	.47	2.56	.74
Total	463	529	90	540	103	2.97	.48	2.42	.79

TABLE 2

Means and Standard Deviations of Scores and GPAs by Gender and Test Accommodation for Sample 2

Group	N	SAT I-V		SAT I-M		HSGPA ¹		FGPA	
		M	SD	M	SD	M	SD	M	SD
Accommodation ² for a Learning Disability									
Female	107	525	95	522	106	3.10	.51	2.62	.63
Male	134	521	101	538	104	3.01	.54	2.48	.78
Total	241	523	99	531	105	3.05	.53	2.54	.72
No Accommodation									
Female	17,594	555	92	551	90	3.56	.51	2.84	.71
Male	16,177	558	94	590	91	3.46	.54	2.69	.77
Total	33,771	556	93	570	93	3.51	.53	2.77	.74

¹For ease of interpretation self-reported HSGPA was recoded from the original 12-point scale (0=F and 12=A+) to a more traditional 4-point scale (0.0=F and 4.3=A+).

²Accommodation for a learning disability refers to any amount of extended time for a learning disability.

Sample 2. The second sample consisted of test-takers with and without learning disabilities from the 1995 or 1997 entering classes at eight institutions in seven states. In order to maintain consistency with Sample 1 an attempt was made to acquire data from the 1996 entering class but FGPA data was not readily available. Sample 2 was compiled by merging data from prior validity studies with a historical database that contained background information, disability status, and test scores. Sample 2 included test-takers with learning disabilities ($n=241$) as well as test-takers who received no testing accommodations ($n=33,771$). Five of the colleges were public, three were private, and all were coeducational. The private and public schools had average enrollments of 11,532 and 24,846, respectively. The five public colleges accounted for 79 percent of the entire student sample, but only 61 percent of the students with a learning disability. Subjects from each institution included a minimum of 10 students with a learning disability. Approximately 48 percent of the students without disabilities were male, 74 percent were white, and 92 percent reported speaking English better than any other language. Demographics on the students with learning disabilities are similar to the first sample; 58 percent were male, 89 percent were white, and 92 percent speak English better than any other language. Information on services provided to students from the college disability service office was not available. Variables in Sample 2 included FGPA, SAT scores, disability information, and self-reported descriptive data. These variables will be described in greater detail in the following section. Table 2 displays descriptive statistics for test scores and grades by gender.

Variables

Institutions in Sample 1 were provided with a list of students who took the SAT I: Reasoning Test with an

accommodation for a learning disability (and no other disability) and who sent test scores to their institution during the 1997 and 1998 admissions cycles. In turn, colleges were asked to provide information regarding high school grade point average (HSGPA), acceptance, matriculation, FGPA, and services provided by the colleges' disability service office for each student. Most colleges were unable to provide students' complete HSGPA, but did provide a slightly adjusted HSGPA. For example, several colleges recalculated HSGPA after removing physical education courses. The recalculated HSGPA was used for the analyses in this study. Two colleges in the study were unable to provide HSGPA. In addition, data on services provided (or not provided) by the disability service office were unavailable for 26 of the subjects at three different schools. To protect students' confidentiality, colleges removed all identifying information before returning data files to the researchers.

In Sample 2, variables from the Admitted Class Evaluation Service™ (ACES™) database and historical data files were merged to produce data comparable to Sample 1. The ACES database, which provided FGPA, contains college transcript data on students in participating institutions. The historical data provided previous SAT scores, information on accommodations for a disability, self-reported HSGPA, and demographic information. HSGPA and demographic information are self-reported information from the Student Descriptive Questionnaire (SDQ), which is completed by 95 percent of students when registering to take the SAT Program tests.

There are four notable differences between variables and students included in the two samples. First, Sample 1 included a variable that indicated if students with disabilities received services from their college's disability service office, and this information was not available for students in Sample 2. Second, Sample 2 may be less

consistent across colleges than Sample 1 because institutions may omit students from the ACES sample for any reason (e.g., FGPA is less than 1.0; student receives services from college disability office). In addition, Sample 2 does not include any students who had a FGPA lower than 1.0 and 6 percent of the students in Sample 1 have FGPA's lower than 1.0. Last, HSGPA was reported by colleges in Sample 1 and by students in Sample 2. College-reported HSGPA is on a conventional continuum from 0.00 (F) to 4.00 (A), while student-reported HSGPA is reported in terms of letter grades from "F" (0) to "A+" (12). In prior research student- and college-reported HSGPA correlate at .79 (Freeberg, 1988). A similar relationship was found between student- and college-reported HSGPA in Sample 1 ($r=.76, p < .001$). Other research examined the relationship between HSGPA and FGPA, and showed that correlations between school-reported HSGPA and FGPA were .04 points larger than the correlation between student-reported HSGPA and FGPA (Donlon, 1984; Freeberg, Rock, and Pollack, 1989). Although these relationships are not identical, they are close and for the purpose of this study self-reported HSGPA will be used as a proxy for actual HSGPA in Sample 2.

Procedures

Adjusted Correlations. Consistent with the approach taken in prior research (Braun et al., 1988; Bridgeman et al., 2000; and Ragosta et al., 1991), all correlations were computed within colleges, adjusted for range restriction, weighted by the number of students at the college, and averaged across colleges. Adjustment for range restriction has been conducted because the range of SAT scores in this sample is narrower than the range of SAT scores in the national sample. Range restriction occurs as students select colleges and colleges select students, thereby decreasing the range of students' SAT scores at an individual college. Because restriction of range decreases the strength of the relationship (expressed as a correlation coefficient) between test scores and FGPA, correlation coefficients are adjusted using the Pearson-Lawley multivariate correction to represent the correlation more accurately (Gulliksen, 1950).

Predicted FGPA. Predicted FGPA was not computed for Sample 1 because data on the performance of test-takers without disabilities had not been collected and the sample of students with disabilities per institution was too small to make accurate predictions of FGPA. For each institution in Sample 2, over/underprediction of FGPA was analyzed by making predictions based on all students. The difference between predicted and

actual FGPA (Predicted FGPA–Actual FGPA) for students with disabilities was computed and is represented in grade-point units on a 4-point scale. Positive values indicate overprediction, while negative values represent underprediction. Prior research (Bridgeman et al., 2000) reported modest differences between the accuracy of predicted FGPA for male and female test-takers.

III. Results

Adjusted Correlations

In the current study correlations for both samples varied across colleges, but overall were generally similar to prior studies of test-takers with any accommodations for learning disabilities. Table 3 displays correlations for both samples.

Sample 1. Initial analysis of this sample revealed that the adjusted correlation between combined SAT scores and FGPA was .12 for the group of all students that matriculated. When adjusted correlations were computed separately for SAT–V and SAT–M, the math test scores showed a weaker relationship with FGPA than verbal test scores did. The SAT–V test correlated with FGPA at .21, while SAT–M and FGPA correlated at .04. Adjusted correlations were also computed for students who had FGPA greater than or equal to 1.0 (D-). As Table 3 indicates, students with FGPA greater than or equal to 1.0 had a higher correlation between combined SAT and FGPA (.26) than did the sample of all matriculated students.

The divergence between the verbal and math subscales remains with the truncated sample (SAT–V=.34 and SAT–M=.12). Although 94 percent of the students were included in all of these adjusted correlations, the relationship between test scores and FGPA is noticeably weaker for the group of all matriculated students than for the truncated sample (i.e., students with FGPA greater than or equal to 1.0). The stronger relationship for the truncated sample is due to the removal of a few extreme cases (e.g., students with high and average test scores and HSGPA, but extremely low FGPA). These extreme cases indicate that nonacademic factors are very likely to contribute to extremely low FGPA.

The adjusted correlations for all SAT scores (SAT–M, SAT–V, and combined SAT) and FGPA were higher for students who did not receive services for their disability than those that did. This pattern was true for all matriculated students as well as the students with FGPA greater than or equal to 1.0. It is worth noting that the difference in adjusted correlations between

TABLE 3

Correlations with FGPA, Adjusted for Range Restriction

Sample		Accommodation for Learning Disability			No Accommodation
		Services	No Services	Total	
SAT I Verbal					
1	FGPA \geq 1.0	.32	.35	.34	NA
1	All matriculated	.22	.34	.21	NA
2	All matriculated	NA	NA	.28	.37
SAT I Math					
1	FGPA \geq 1.0	.10	.20	.12	NA
1	All matriculated	.05	.15	.04	NA
2	All matriculated	NA	NA	.28	.39
SAT I (V+M)					
1	FGPA \geq 1.0	.22	.29	.26	NA
1	All matriculated	.10	.27	.12	NA
2	All matriculated	NA	NA	.35	.48
Number of Students ¹					
1	FGPA \geq 1.0	269	143	438	NA
1	All matriculated	283	154	463	NA
2	All matriculated	NA	NA	310	36,081

¹Data on services provided students were unavailable for 26 of the students in Sample 1. NA indicates that data were not available on test-takers without accommodations from Sample 1 or information regarding any services provided to students in Sample 2.

students who received services and those who did not was greater for the entire sample (.17) than for the truncated sample (.07). Another difference between the two groups was that students who received services tended to have higher high school and college grades ($F [1,346] = 11.45, p \leq .001$ for HSGPA, and $F [1,406] = 6.62, p \leq .01$ for FGPA), but lower SAT I math scores ($F [1,421] = 15.04, p \leq .001$) than students that did not receive services from the college disability service office. No significant difference on SAT I verbal scores was found between students who received services and those that did not. Unfortunately, the small sample size precluded further analyses to examine differences between the students with low FGPA and those with high FGPA.

Sample 2. Because no students in Sample 2 had a FGPA lower than 1.0, this sample is more comparable with Sample 1's truncated sample of students with FGPA greater than or equal to 1.0. As Table 3 indicates, the adjusted correlation between combined SAT and FGPA for Sample 2 was .35; separately, SAT-V and SAT-M were each correlated .28 with FGPA. Although SAT-V and FGPA correlations were fairly consistent across truncated Sample 1, Sample 2, and prior research on test-takers with learning disabilities; inconsistencies exist in the SAT-M and FGPA correlations.

The adjusted correlation for students *without* a disability was similar for the Math and Verbal tests (SAT-V=.37 and SAT-M=.39). Although these correlations are higher than those of test-takers *with* learning disabilities, they are lower than those shown in prior

research on students *without* disabilities (Bridgeman et al., 2000; Pennock-Román, 1994). The adjusted correlation between combined SAT and FGPA ($r = .48$) for students without a disability was similar to values found in prior research on students without disabilities.

Over- and Underprediction of FGPA

Because the number of students (per institution) in Sample 1 was too small to reliably predict FGPA, the following results are only for Sample 2. No adjustments were made for course difficulty because grades per course were not available. It is possible, however, that students with different learning styles sort themselves into courses that vary in terms of difficulty. For this reason all interpretations of predicted FGPA should be made with caution.

Predicted FGPA from SAT scores. Actual FGPA and predicted FGPA for students with a learning disability correlate at .35 ($p < .001$) when predicted FGPA is calculated using SAT scores. For students without disabilities, the correlation between actual and predicted FGPA is .40 ($p < .001$). These correlations are slightly higher than the correlation reported before the SAT was recentered and revised, which indicated correlations of .22 and .37, respectively, for students with learning disabilities who received accommodations and students without a disability (Braun et al., 1988).

Table 4 displays differences between actual and predicted FGPA (on a traditional 4-point scale where

0.00=F and 4.00=A) for students with and without learning disabilities. A difference of +1.00, for example, would indicate that the predicted FGPA is a full grade level higher than the actual FGPA (e.g., predicted FGPA is 3.50 and actual FGPA is 2.50). When SAT scores were used to predict FGPA, students with learning disabilities had a small overprediction of .12. A modest overprediction of .21 was observed for male test-takers with learning disabilities, and female test-takers with learning disabilities had a trivial overprediction of .02.

Predicted FGPA from SAT scores and HSGPA. When FGPA was predicted using HSGPA and SAT scores, the correlation between actual and predicted FGPA for students with a learning disability was .49 ($p < .001$). For students without disabilities the correlation between actual and predicted FGPA was also .49 ($p < .001$). Before the SAT was recentered, predicted FGPA (using SAT and HSGPA) and actual FGPA correlated at .34 and .49, respectively, for students with learning disabilities who received accommodations and students without a disability (Braun et al., 1988). When SAT scores and HSGPA were used to predict FGPA, students with learning disabilities had a trivial underprediction of -.02. Female test-takers with learning disabilities demonstrated a slight underprediction of FGPA by -.08, but male test-takers with learning disabilities showed a trivial overprediction of FGPA (+.03).

Test-takers without disabilities. Students without a disability had under/overpredictions of FGPA similar to those found by Bridgeman et al. (2000). In both studies, predicted FGPA was slightly higher than actual FGPA for female students and slightly lower than actual FGPA for male students. In this study female students' FGPA was underpredicted by -.11 (SAT) and -.07 (SAT

and HSGPA); for male students, FGPA was over-predicted by .12 (SAT) and .08 (SAT and HSGPA). These are nearly identical to those discussed in the earlier review of Bridgeman et al. (2000).

IV. Conclusion

Although the degree of correlation varies, the revised SAT is positively correlated with college achievement (as measured by first year grade point average) for students taking the test with extended time accommodations for a learning disability. SAT scores were fairly accurate predictors of FGPA for students with learning disabilities. In most cases, however, when HSGPA is used along with SAT test scores, the predictive validity of FGPA is increased. Female test-takers with learning disabilities were the only group that had closer predicted FGPA when SAT scores were used alone to predict FGPA, but the differences between actual and predicted FGPA are small regardless of which of the two formulas were used to predict FGPA; SAT alone underpredicted by -.08, and SAT combined with HSGPA overpredicted by +.02. Because these differences are small and prior research has demonstrated the value of HSGPA as an admissions criteria, college admissions officers should not omit or decrease the value of HSGPA as a selection criteria for female applicants with learning disabilities.

Although the correlation between SAT scores and FGPA was positive for both samples, the degree of relationship varied from .10 (all Sample 1 students who matriculated and received services) to .35 (all students in Sample 2). A similar range was found when examining the correlation between SAT-M and FGPA (ranging from .04 to .28), but the correlation between SAT-V and FGPA was fairly consistent (ranging from .21 to .35). In addition, the relationship between the SAT-M and FGPA was not consistent with prior research that indicated a stronger relationship between these two variables. Some differences between the correlations of SAT-M with FGPA may be due to a propensity for students in this sample to take more courses that rely on verbal reasoning skills (e.g., literature) and fewer mathematics courses.

More research is needed to investigate which factors contribute to the varied correlations in this study. Some of these variations may be due to different populations of students used in the second sample. In Sample 2, colleges and universities were permitted to omit students for any reason including receiving services for a learning disability and having a FGPA less than 1.0. Since

TABLE 4

Over- (+) and Underprediction (-) of FGPA by Gender and Disability for Sample 2.

	N	Predicted-Actual FGPA	
		SAT I	SAT I + HSGPA
Accommodation for LD			
Total	241	.12	-.02
Female	107	.02	-.08
Male	134	.21	.03
No Accommodation			
Total	33,771	.00	.00
Female	17,594	-.11	-.07
Male	16,177	.12	.08

Note. Prior research on students without disabilities found overpredictions of .11 (SAT) and .08 (SAT+HSGPA) for male test-takers and underpredictions of -.10 (SAT) and -.07(SAT+HSGPA) for female test-takers (Bridgeman, McCamley-Jenkins, and Ervin, 2000).

students in Sample 1 with FGPA higher than 1.0 and no services provided for their disability had correlations closest to Sample 2, it is worth investigating the relationship between services provided by college disability service offices and FGPA. It is important to remember that the ranges of services vary between schools, while some schools provide minimal support services to students with learning disabilities, and other institutions provide a highly structured fee-based program that includes intensive tutoring, assistive technology, and test accommodations (Harris and Robertson, 2001).

Limitations. There are several limitations to this research primarily due to the relatively small number of students with learning disabilities at each institution. The small sample size precluded the examination of group differences such as ethnicity, language fluency, undergraduate major, level of SAT test score (high, medium, and low), or type and severity of learning disability. Obtaining a large enough sample to examine group differences will always be a challenge due to the nature and diversity of disabilities. In addition, the lack of control over course selection, due to small sample size, may be one cause of the differential correlation between FGPA and the math and verbal subtests in Sample 1. It is possible that students with specific learning disabilities select, avoid, or are exempt from courses that challenge skills and abilities different from the courses that students without disabilities select, avoid, or are exempt from. In other studies of students without disabilities, the increased sample size and data on course grades allows for greater controls to be placed on course difficulty by comparing predicted and actual grade point average within each course or by comparing students within intended undergraduate major (Bridgeman et al., 2000).

Since all of the colleges in the samples were selected because they had a large percentage of applicants with test scores flagged for nonstandard administrations, the results may not be generalizable to all colleges and universities. In addition, the samples used in this study were largely comprised of more selective colleges and universities that agreed to participate in validity research, so results may not be applicable to less selective institutions or those that do not participate in this type of research. The data from the second sample should be treated with particular caution because the colleges may not have provided data on all students. Because less than 1 percent of students in the second sample took the SAT with an accommodation for a learning disability (substantially lower than the national average of 1.5 percent [Cahalan, 2000]), there is reason to hypothesize that some students with disabilities were present on campus but were not included in these analyses.

Future research. Additional research is required to investigate factors that contribute to the correlations between FGPA and the SAT for test-takers with learning disabilities. Although such research will be difficult, due to relatively small numbers of students with disabilities, it is important to examine the validity of standardized admissions and achievement tests for different learning disability subgroups (dyslexia, ADHD, dyscalculia, visual/spatial) as well as for different types of test accommodations (timing, presentation, response, and setting). This paper has focused on the use of self-paced extended time. Other timing accommodations such as extended time with section timing, frequent breaks, multiple testing sessions, and extended time for some (but not all) sections of the test should be examined in future research.

Recent research by Bridgeman et al. (2002) indicates that decreasing the number of test items while retaining the standard test time does not substantially increase SAT test scores for students without disabilities. Mandinach et al. (in press) are examining the impact on test scores of increasing the amount of time allotted to test-takers with and without disabilities. Both of these studies will provide information on whether extended time differentially increases test scores for students with disabilities and the appropriate amount of extended time for students with learning disabilities. Regardless of the impact of these studies on the appropriateness of extended time as a test accommodation, this study provides evidence that the SAT I: Reasoning Test is a valid tool for helping admissions officers select students with learning disabilities (who receive extended time accommodations) for college admission. As with nondisabled students, the SAT is most appropriately used as one of several selection criteria and not as the sole measure of a student's abilities.

References

- Bridgeman, B., Curley, E., & Trapani, C. (2002). *Impact of fewer questions per section on SAT I scores*. Manuscript submitted for publication.
- Bridgeman, B., McCamley-Jenkins, L., & Ervin, N. (2000). *Predictions of freshman grade-point average from the revised and recentered SAT I: Reasoning Test*. (College Board Report No. 2000-1). New York: College Entrance Examination Board.
- Braun, H., Ragosta, M., & Kaplan, B. (1988). Predictive validity. In W. W. Willingham, M. Ragosta, R. E. Bennett, H. Braun, D. A. Rock, & D. E. Powers (Eds.), *Testing handicapped people* (pp. 109-132). Boston, MA: Allyn and Bacon.

- Cahalan, C. (2000, April). *Geographic clusters of learning disabled test takers in the United States*. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED 443 841)
- Camara, W. J., Copeland, T., & Rothschild, B. (1998). *Effects of extended time on the SAT I: Reasoning test score growth for students with learning disabilities*. (College Board Report No. 98-7). New York: College Entrance Examination Board.
- College Board (2001a). Services for students with disabilities: Accommodations. Retrieved December 4, 2001, from <http://www.collegeboard.org/disable/counsel/html/accom.html>
- College Board (2001b). Services for students with disabilities: Guidelines for documentation. Retrieved December 4, 2001, from <http://www.collegeboard.org/disable/counsel/html/document.html>
- Donlon, T. F. (Ed.) (1984). *The College Board technical handbook for the scholastic aptitude and achievement tests*. New York: College Entrance Examination Board.
- Educational Testing Service. (1999). *Information about testing accommodations*. Retrieved November 13, 2001, from <http://www.ets.org/distest/info.html>
- Elliot, S. N. & McKevitt, B. C. (2000, April). *Testing accommodations decisions: Legal and technical issues challenging educators*. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.
- Freeberg, N. E. (1988). *Analysis of the Revised Student Descriptive Questionnaire, Phase I*. (College Board Report No. 88-5). New York: College Entrance Examination Board.
- Freeberg, N. E., Rock, D. A. & Pollack, J. (1989). *Analysis of the revised Student Descriptive Questionnaire: Phase II, predictive validity of academic self-report* (College Board Report No. 89-8). New York: College Entrance Examination Board.
- Fuchs, L.S., & Fuchs, D. (1999). Fair and unfair testing accommodations. *School Administrator*, 56, 24–29.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: John Wiley and Sons.
- Harris, R., & Robertson, J. (2001). Successful strategies for college-bound students with learning disabilities. *Preventing School Failure*, 45, 125–131.
- Henderson, C. (2001). *College freshman with disabilities: Statistical year 2000*. Washington, DC: American Council on Education.
- Mandinach, E. B., Bridgeman, B., Cahalan, C., & Trapani, C. (in press). *The impact of extended time on SAT test performance*. Princeton, NJ: Educational Testing Service.
- Pennock-Román, W. (1994). *College major and gender differences in prediction of college grades: Sex, language, and ethnic group*. (College Board Report No. 94-2). New York: College Entrance Examination Board.
- Phillips, S. E. (1994). High-stakes testing accommodations: Validity versus disabled rights. *Applied Measurement in Education*, 7, 93–120.
- Pitoniak, M.J. & Royer, J.M. (2001). Testing accommodations for examinees with disabilities: A review of psychometric, legal and social policy issues. *Review of Educational Research*, 71, 53–104.
- Ragosta, M., Braun, H., & Kaplan, B. (1991). *Performance and persistence: A validity study of the SAT for students with disabilities*. (College Board Report No. 91-3). New York: College Entrance Examination Board.
- Thurlow, M. L., Elliott, J. L., & Ysseldyke, J. E. (1998). *Testing students with disabilities: Practical strategies for complying with district and state requirements*. Thousand Oaks, CA: Corwin Press.
- U.S. Department of Education. (2000). *Twenty-second annual report to congress on the implementation of the individuals with disabilities education act* (Office of Special Education Programs). Jessup, MD: Author.
- Wightman, L.F. (1993). *Test takers with disabilities: A summary of data from special administrations of the LSAT*. (Research Report No. 93-03). Newton, PA: Law School Admission Council.
- Willingham, W.W., Ragosta, M., Bennett, R.E., Braun, H., Rock, D.A., and Powers, D.E. (Eds.). (1988). *Testing handicapped people*. Boston, MA: Allyn and Bacon.
- Zuriff, G.E. (2000). Extra examination time for students with learning disabilities: An examination of the maximum potential thesis. *Applied Measurement in Education*, 21, 261–267.

Appendix A: Examples of Accommodations

Individuals with disabilities that require testing accommodations are provided the following accommodations when appropriate (College Board, 2001a; Educational Testing Service, 1999).

Computer-based testing (CBT) and paper-and-pencil testing accommodations

- Additional rest breaks
- Alternate test formats (e.g., computer for writing essays, large-print, audiocassettes, or Braille)
- Extended testing time (All tests are timed)
- Permission to bring snacks and medications during testing period
- Reader
- Recorder/writer of answers
- Sign language interpreter (for spoken directions only)
- Large print answer sheets

CBT accommodations

- Intellikeys keyboard
- HeadMaster Plus mouse
- Kensington Trackball mouse
- Selectable background and foreground colors
- ZOOMTEXT

Paper-and-pencil testing accommodations

- Braille
- Enlarged print (14 point)
- Large print (greater than 14 point)
- Large print answer sheet
- Audiocassette with large-print figure supplement
- Audiocassette with Braille figure supplement

Appendix B: Guidelines for Documentation

Students that are eligible for an accommodation must have documentation on file at their school or submit documentation to the College Board that supports the

student's request. College Board Guidelines (College Board, 2001b) requires that documentation must:

- (1) state the specific disability, as diagnosed;
- (2) be current (in most cases, the evaluation should be completed within three years of the request for accommodations);
- (3) provide relevant educational, developmental, and medical history;
- (4) describe the comprehensive testing and techniques used to arrive at the diagnosis (including evaluation date[s] and test results with subtest scores from measures of cognitive ability, current academic achievement, and information processing);
- (5) describe the functional limitations supported by the test results;
- (6) describe the specific accommodations requested, and state why the student's disability qualifies the student for such accommodations on standardized tests; and
- (7) establish the professional credentials of the evaluator, including information about license or certification and area of specialization.



