

Timed Versus Untimed Testing Conditions and Student Performance

• • • • • • • • • •

Thomas E. Brooks, Ph.D.

Betsy J. Case, Ph.D.

Michael J. Young, Ph.D.

June 2003

(Revision 2, May 2004)



Acknowledgements

This Pearson Report was prepared with the assistance of:

Greg Ayres
Nancy Severance
Nathan Wall
Holly Zhang

Timed Versus Untimed Testing Conditions and Student Performance

Introduction

The adherence to fixed time limits during group administrations of norm-referenced standardized tests has been an accepted practice from the early part of the 20th century (Anastasi, 1976; pp. 32–34).

According to Ebel and Frisbie (1991), standardized tests or measures provide:

The means for making score comparisons among examinees who attempt the same tasks under the same testing conditions and time constraints, and whose responses are scored with the same procedures (p. 286).

As Ebel and Frisbie (1991; pp. 286–288) also point out, standardized testing procedures serve the same function in education as they do when applied to goods and services encountered every day. Without standardized weights and measures, things as simple as purchasing fresh produce or a gallon of gasoline would be quite chaotic.

For example, if Environmental Protection Agency’s published miles-per-gallon (MPG) ratings were not the result of tests conducted under standard conditions, a distorted picture of true fuel efficiencies would emerge. A full-sized sedan tested at a constant, moderate speed under ideal weather conditions could appear (falsely) to have a better MPG rating than a sub-compact car tested at high speed under extremely hot conditions.

Until recently, research examining differences in student performance under fixed time versus untimed testing conditions has been relatively infrequent. An early empirical study (Baxter, 1931) was repeated in 1952 by Myers. Both studies looked at the relationship of speededness and power to the level of difficulty of the test. The conclusion was that the scores on these timed tests were a function of two unrotated (orthogonal) underlying factors: test-takers’ ability and the rate of answering, with ability being more influential. Lord (1956) continued this line of research. Virtually all of the other research conducted from 1931 to 1987 on timed versus untimed testing conditions included only high school or college students.

Timed Versus Untimed Testing Conditions and Student Performance

The need to reconsider of the use of fixed time limits on standardized tests has grown in recent years. Tindal and Fuchs (1999) identified ten empirical studies conducted from 1987 through 1999 that examined the effects of timed versus untimed testing conditions on students in kindergarten through grade 12. All of these studies included students with disabilities and varying ability levels.

As a result of legislation such as

the *Elementary and Secondary Education Act* (ESEA) as reauthorized under the *No Child Left Behind Act* of 2001 (NCLB) (Pub. L. 107–110),

the *Individuals with Disabilities Education Act*, as amended (IDEA) (20 U.S.C. §§ 1400 et seq.), and

Title II of the *Americans with Disabilities Act* of 1990 (ADA) (42 U.S.C. §§ 12101 et seq.),

the impetus for a re-examination of timed versus untimed testing conditions has become stronger and includes new concerns about fairness in testing (i.e., accommodations). Accommodations, as opposed to modifications, are changes in standardized assessment conditions that are incidental to the construct being measured. Traditionally, untimed administrations of norm-referenced tests, when they did occur, meant the sacrifice of the associated norms and other derived score information. With the current focus on accommodations and inclusive testing conditions, the appropriateness of fixed testing times has become an important aspect of research conducted by Pearson Inc. (Pearson).

This Assessment Report presents background information on timed versus untimed conditions as they relate to standardized testing. It also summarizes the results of research studies examining student performance under timed and untimed testing conditions conducted as part of the Stanford 10 standardization program.

What Does the Research Say?

The research conducted to date by others that specifically addresses the consequences of timed versus untimed testing conditions for students in kindergarten through high school is primarily focused on test validity.

Validity is the ability of the test to provide a fair and accurate assessment of what has been learned. At stake are the NCLB-driven efforts of states, districts, and schools to account for students' performance as well as improved knowledge and skills over time.

Each state has developed, or is developing, a set of instructional standards or criteria that may or may not match those developed by other states. Educational assessments are then administered by the states to determine the degree to which (below basic, basic, proficient, and advanced, for example) students have achieved these instructional standards.

Standards-based tests typically have high stakes for students, teachers, and state education agencies. When a standards-based or criterion-referenced score as opposed to a

Timed Versus Untimed Testing Conditions and Student Performance

norm-referenced score is the objective of an assessment, some would advocate the relaxation or abandonment of traditional fixed time administrations. The thinking is that administration procedures such as fixed times have no relevance to the standards-based performance being assessed and may unfairly disadvantage some students. In contrast, however, objections may be raised against any variation in testing conditions thought to provide an advantage to some students over others.

This is where the desire to administer a test under the same conditions under which it was standardized (so that norm-referenced scores can be generated and used) comes into conflict with the desire to offer reasonable accommodations to students with disabilities who are otherwise unable to fully demonstrate their achievement of relevant instructional standards. The ideal middle ground in this conflict would seem to be one where students with disabilities are accommodated—without changing the educational content or process skills being assessed—so that they may show their abilities without the negative influence of factors relevant to their disabilities, (AERA, 1999).

A relatively small number of studies have investigated the impact of timed versus untimed conditions specific to the performance of students with disabilities (Alster, 1997; Halla, 1988; Harker, 1991; Hill, 1984; Munger and Loyd, 1991; Perlman, Borger, Collins, Elenbogen and Wood, 1996; Runyan, 1991a, 1991b).

Huesman and Frisbie (2000) studied the performance of students with and without disabilities on a standardized norm-referenced test and found that students with disabilities perform significantly better under extended time conditions than they do under fixed time conditions.

Pearson's Research

As part of the 2002 standardization of Stanford 10, which was administered for the most part under untimed conditions, Pearson conducted research that compared the performance of students taking the test under timed and untimed testing conditions.

Standardization

The following important note to teachers appeared in the spring and fall standardization editions of the Stanford 10 *Directions for Administering* below a table of proposed testing times:

Times are included for planning purposes only. Stanford 10 is to be administered so that all children have sufficient time to complete it. The time allotment indicated for each subtest should be adequate for many students. However, if necessary, additional time must be provided for a student to complete the test. Allow any student to have as much time as he or she needs to complete the test while working productively.

Timed Versus Untimed Testing Conditions and Student Performance

(NOTE: The above notice now appears in the Directions for Administering for all levels of the currently published Stanford 10 series.)

For the *Stanford Achievement Test Series*, Ninth Edition (Stanford 9) and its predecessors, untimed norms (separate norms for students who took the test using extended time accommodations), though requested, were not available because the tests were not standardized under untimed conditions. The standardization of Stanford 10 under untimed conditions has eliminated any need for separate untimed norms, while allowing for the reporting of both standards-based and norm-referenced scores for students with and without disabilities.

During the Stanford 10 standardization process, teachers were provided with suggested administration times (based on the fixed times used for the administration of Stanford 9) for each subtest. Teachers were also instructed to allow additional time beyond the suggested times as necessary for students to complete each subtest while still working productively. In addition, teachers were asked to record the starting and ending times as each subtest was administered. These data, obtained from about 500 classrooms for each grade tested, were used to determine the extent to which additional time was allowed and to evaluate the extent to which extended time varied from grade to grade.

The Study

The Pearson research was designed to determine if administering Stanford 10 under both timed and untimed conditions would affect test results differentially.

The approximately 360,000 students in the Stanford 10 Standardization Sample were tested under untimed conditions. A separate, additional group of students (about 150 classrooms at each grade) was tested under timed conditions. The “timed” group was selected to represent the same demographic characteristics as the larger untimed group. In addition, students from the larger, untimed standardization group were selected at random to form an “untimed” group similar in size (about 150 classrooms at each grade) and demographic characteristics to the timed group. The grade-appropriate Stanford 10 battery was administered to all students.

The table of proposed testing times contained in the standardization editions of the Stanford 10 *Directions for Administering* were used as fixed times for the group tested under timed conditions.

Results

Data from the study were processed at the Pearson Scoring Center in San Antonio, TX. All data are reported as raw scores and are summarized in Table 1.

Differences in the average raw scores of students tested under timed and untimed conditions were typically very small. In the majority of cases, the differences amounted to less than one raw score point. Although the differences were quite small and non-significant for the most part, all students tested under untimed conditions showed improved performance through grade 6. However, above grade 6, students actually performed slightly better under timed conditions.

Timed Versus Untimed Testing Conditions and Student Performance

Table 1. Difference Between Average Raw Scores for Students Tested under Timed and Untimed Conditions

Grade	Subtest	Timed	Untimed	Difference	Significant?
1	Word Study Skills	–	–	–	–
	Word Reading	24.2	24.8	0.6	No
	Sentence Reading	55.7	57.1	1.4	No
	Mathematics Problem Solving	–	–	–	–
	Mathematics Procedures	–	–	–	–
	Language	–	–	–	–
	Spelling	–	–	–	–
	Environment	–	–	–	–
	Listening	–	–	–	–
	Average Difference			1.0	
2	Word Study Skills	–	–	–	–
	Reading Vocabulary	22.2	22.4	0.2	No
	Reading Comprehension	25.6	26.4	0.7	No
	Mathematics Problem Solving	–	–	–	–
	Mathematics Procedures	–	–	–	–
	Language	–	–	–	–
	Spelling	28.8	28.7	–0.2	No
	Environment	–	–	–	–
	Listening	–	–	–	–
	Average Difference			0.2	
3	Word Study Skills	19.8	20.1	0.2	No
	Reading Vocabulary	20.2	20.7	0.5	No
	Reading Comprehension	34.9	35.2	0.3	No
	Mathematics Problem Solving	29.4	30.3	0.8	No
	Mathematics Procedures	17.7	17.8	0.1	No
	Language	51.4	54.4	3.1	Yes
	Spelling	20.3	22.0	1.7	Yes
	Science	25.0	25.5	0.5	No
	Social Science	24.2	24.6	0.4	No
	Listening	–	–	–	–
	Average Difference			0.8	

(– indicates that data are not included because the test was dictated to students.

Significance level for differences between means is .05.)

Timed Versus Untimed Testing Conditions and Student Performance

Table 1. Difference Between Average Raw Scores for Students Tested under Timed and Untimed Conditions (continued)

Grade	Subtest	Timed	Untimed	Difference	Significant?
4	Word Study Skills	19.4	19.8	0.4	No
	Reading Vocabulary	20.2	21.6	1.4	Yes
	Reading Comprehension	31.3	32.7	1.4	No
	Mathematics Problem Solving	28.0	27.9	-0.1	No
	Mathematics Procedures	19.7	19.9	0.2	No
	Language Mechanics	14.8	15.1	0.3	No
	Language Expression	15.2	15.4	0.2	No
	Spelling	23.2	24.4	1.2	No
	Science	25.5	25.7	0.2	No
	Social Science	23.6	24.3	0.7	No
	Listening	–	–	–	–
Average Difference				0.6	
5	Reading Vocabulary	21.4	21.8	0.4	No
	Reading Comprehension	31.7	32.2	0.5	No
	Mathematics Problem Solving	27.6	28.6	1.1	No
	Mathematics Procedures	18.3	18.7	0.5	No
	Language Mechanics	16.0	15.8	-0.2	No
	Language Expression	15.1	15.2	0.1	No
	Spelling	25.5	25.4	-0.1	No
	Science	22.5	22.8	0.3	No
	Social Science	22.6	24.3	1.8	Yes
	Listening	–	–	–	–
Average Difference				0.5	
6	Reading Vocabulary	21.1	21.7	0.6	No
	Reading Comprehension	32.6	33.2	0.6	No
	Mathematics Problem Solving	25.0	25.5	0.5	No
	Mathematics Procedures	16.8	17.7	0.9	No
	Language Mechanics	16.0	16.3	0.3	No
	Language Expression	15.5	16.0	0.5	No
	Spelling	25.3	25.5	0.2	No
	Science	24.3	24.7	0.4	No
	Social Science	22.1	21.8	-0.3	No
	Listening	–	–	–	–
Average Difference				0.4	

(– indicates that data are not included because the test was dictated to students.
Significance level for differences between means is .05.)

Timed Versus Untimed Testing Conditions and Student Performance

Table 1. Difference Between Average Raw Scores for Students Tested under Timed and Untimed Conditions (continued)

Grade	Subtest	Timed	Untimed	Difference	Significant?
7	Reading Vocabulary	20.1	19.5	-0.6	No
	Reading Comprehension	32.0	31.5	-0.5	No
	Mathematics Problem Solving	22.5	22.0	-0.4	No
	Mathematics Procedures	14.9	16.1	1.2	Yes
	Language Mechanics	16.1	15.7	-0.4	No
	Language Expression	15.6	14.9	-0.8	No
	Spelling	24.8	24.0	-0.8	No
	Science	23.2	22.2	-1.0	No
	Social Science	20.7	19.9	-0.8	No
	Listening	–	–	–	–
	Average Difference			-0.4	
8	Reading Vocabulary	21.3	21.3	0.0	No
	Reading Comprehension	23.4	20.0	-3.4	Yes
	Mathematics Problem Solving	22.1	21.1	-1.0	No
	Mathematics Procedures	17.1	14.7	-2.4	Yes
	Language Mechanics	16.3	13.9	-2.4	No
	Language Expression	16.0	14.7	-1.3	Yes
	Spelling	25.3	24.9	-0.4	No
	Science	19.6	20.2	0.6	No
	Social Science	18.6	17.6	-0.9	No
	Listening	–	–	–	–
	Average Difference			-1.2	
10	Reading Vocabulary	22.2	21.8	-0.4	No
	Reading Comprehension	35.9	34.2	-1.7	Yes
	Mathematics	23.4	21.5	-1.9	Yes
	Language	56.6	54.5	-2.1	No
	Language Mechanics	14.5	13.8	-0.7	No
	Language Expression	16.7	15.4	-1.3	Yes
	Spelling	25.5	25.3	-0.2	No
	Science	21.4	20.3	-1.1	No
	Social Science	16.9	15.9	-1.0	No
		Average Difference			-1.2

(– indicates that data are not included because the test was dictated to students.
Significance level for differences between means is .05.)

Timed Versus Untimed Testing Conditions and Student Performance

The Effect of Untimed Conditions on Non-disabled Students

Another aspect of Pearson's research investigated the effect of extended time limits on the performance of non-disabled students. Research by others examining students in grades K through 12 showed that students with disabilities generally performed better under untimed conditions. Furthermore, research by others (Bridgman, *et al*, 2002; Alster, 1997) showed that the extra time did not enhance the performance of non-disabled students. Pearson's research supported the prior research in that differences in the performance of non-disabled students under timed or untimed conditions during the standardization of Stanford 10 were very small, as shown in Table 1.

Summary

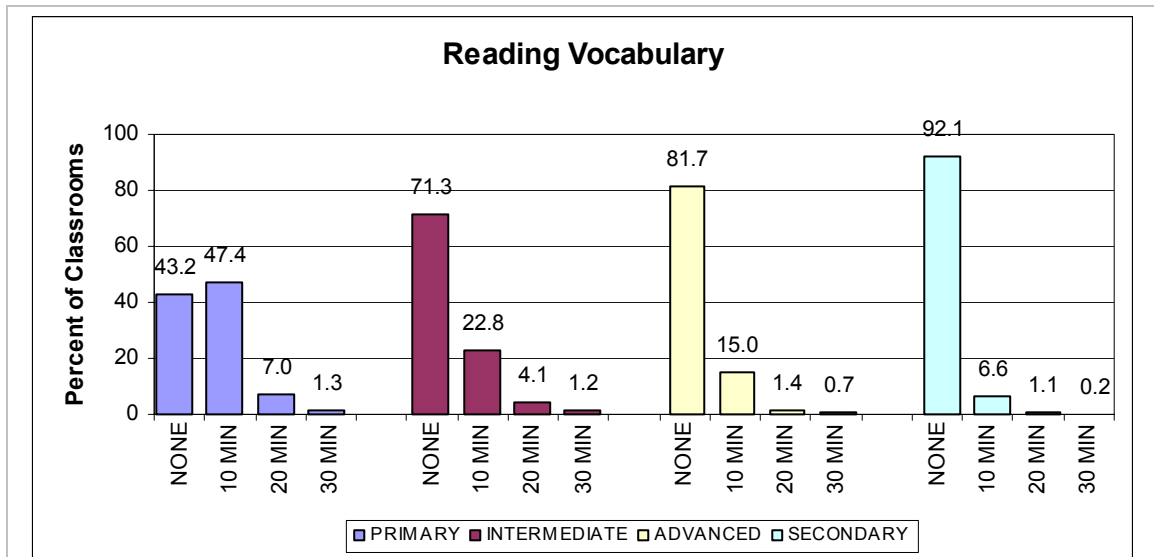
Interestingly, the small differences in performance shown in Table 1 indicate that traditional fixed time limits allow ample time for the great majority of students to complete the test. This demonstrates that procedures established by Pearson over the past 80 years for setting time limits have allowed adequate time for non-disabled students to complete the test without undue strain or errors owing to time pressure. More important, however, is the evidence that the allowance of extended times accommodates disabled students so that they may demonstrate what they have learned while not unfairly inflating the scores of non-disabled students.

How Much Extended Time was Typically Allowed?

Extensions beyond the testing times proposed in the spring and fall standardization editions of the *Directions for Administering* were allowed by teachers in fewer than 50 percent of classrooms participating in the Stanford 10 standardization program. As can be seen in Figures 1a through 1i, the larger amounts of additional time were allowed more frequently in grades 1 through 3 and less frequently in grades 9 through 11. For nearly all classrooms, no more than 20 additional minutes were needed to finish any subtest. Results for nine Stanford 10 subtests are depicted in Figures 1a-1i.

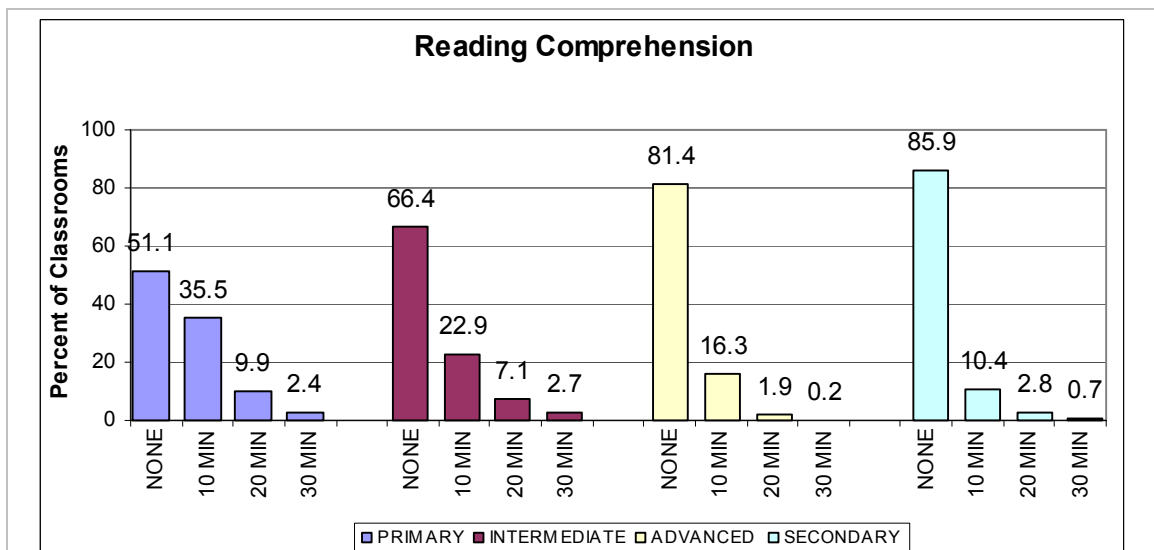
More additional time was allowed on the Reading Vocabulary subtest than on the Reading Comprehension subtest, and more additional time was allowed on the Mathematics Procedures subtest than on the Mathematics Problem Solving subtest. The proposed testing times for the Reading Vocabulary and Mathematics Procedures subtests are considerably less than the proposed testing times for Reading Comprehension and Mathematics Problem Solving; therefore, this greater additional time allowance may reflect a perceived need on the part of the teachers. In any case, because at least 95% of the classrooms finished testing with no more than 20 additional minutes allowed on any subtest, concerns about untimed tests taking significantly more time to administer may be unfounded.

Timed Versus Untimed Testing Conditions and Student Performance



Note: Percentage values shown subject to rounding.

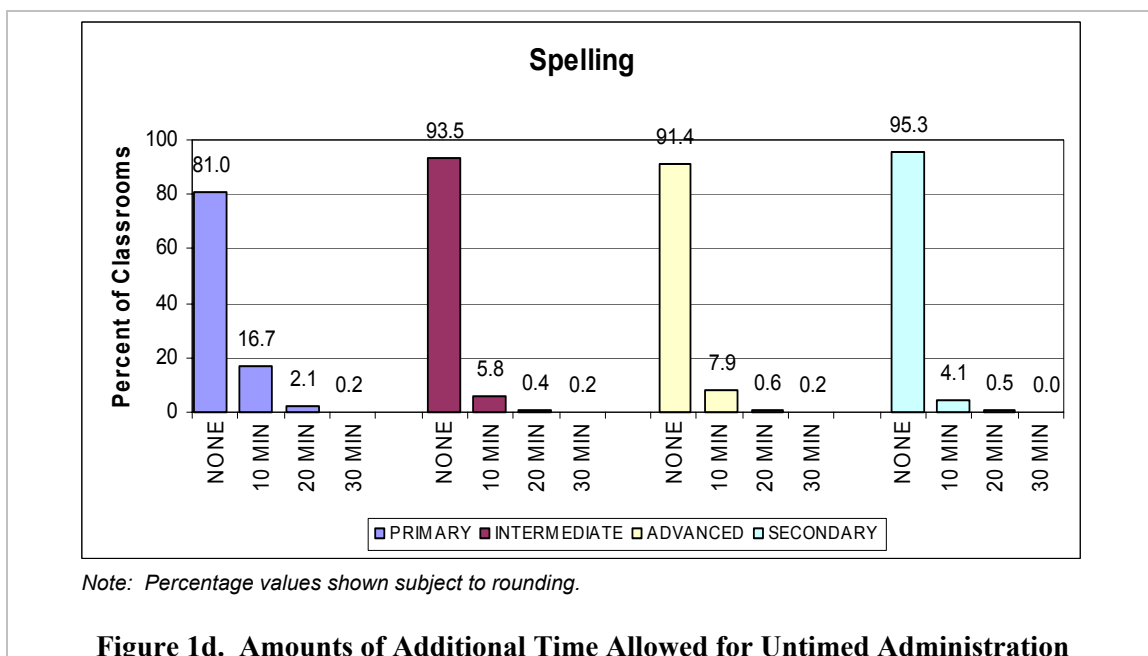
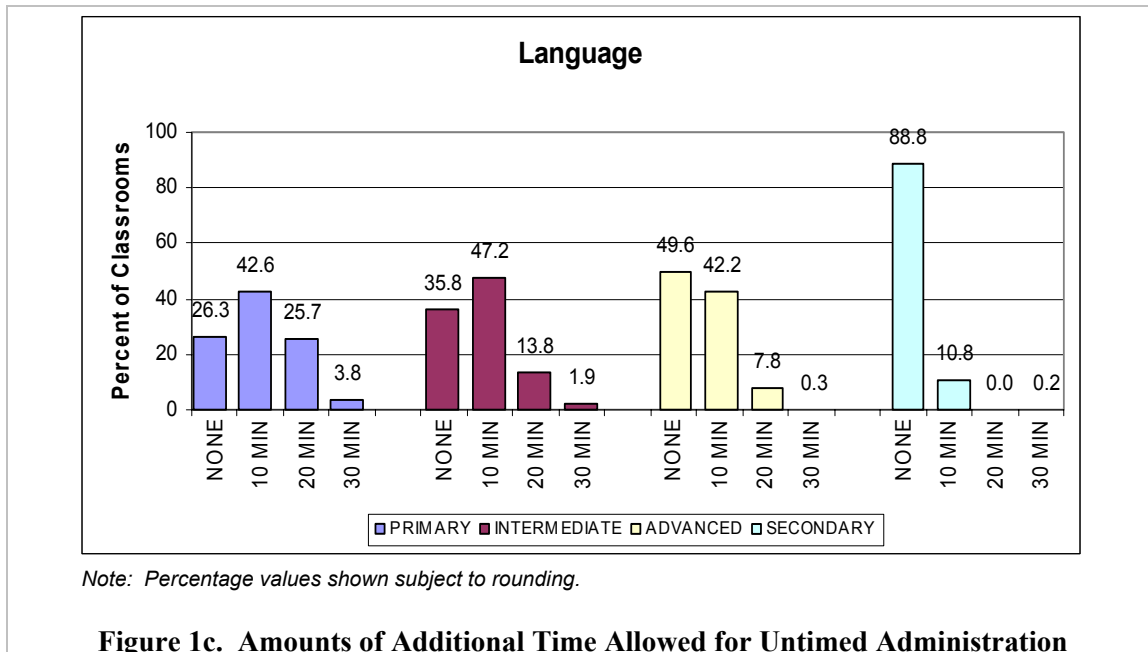
Figure 1a. Amounts of Additional Time Allowed for Untimed Administration



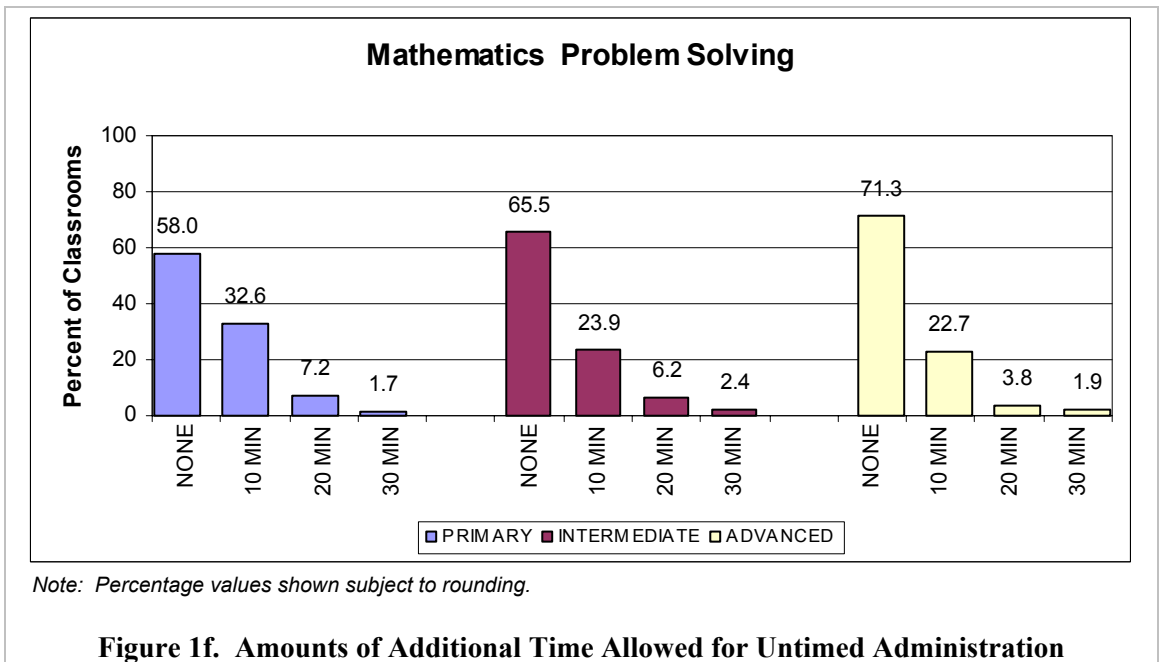
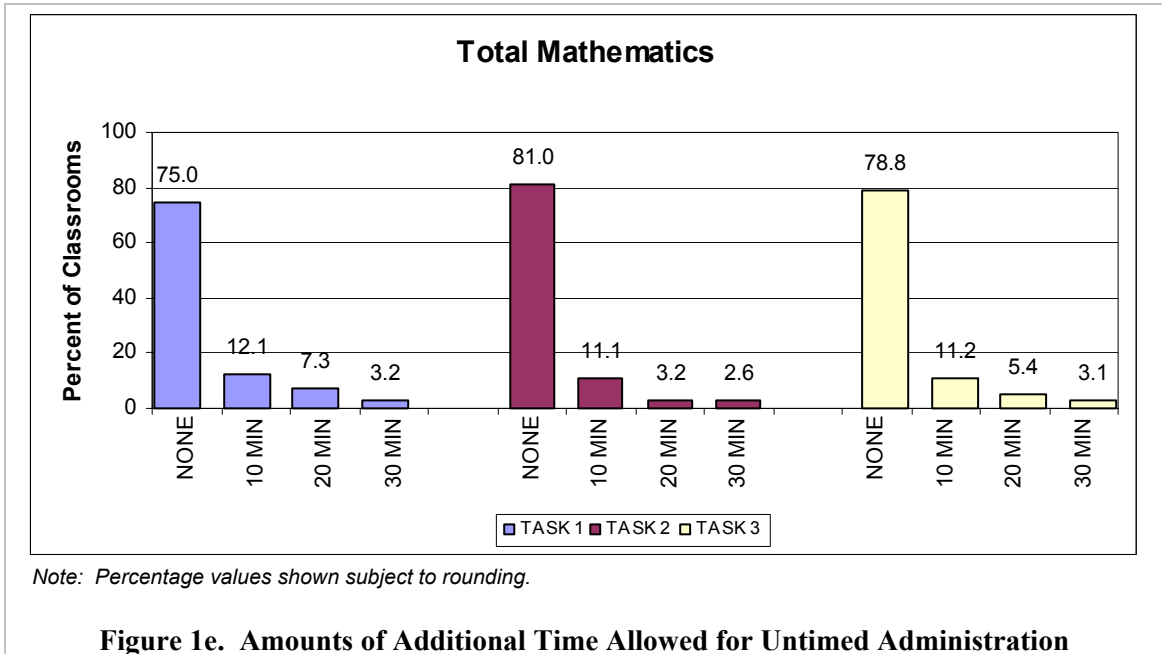
Note: Percentage values shown subject to rounding.

Figure 1b. Amounts of Additional Time Allowed for Untimed Administration

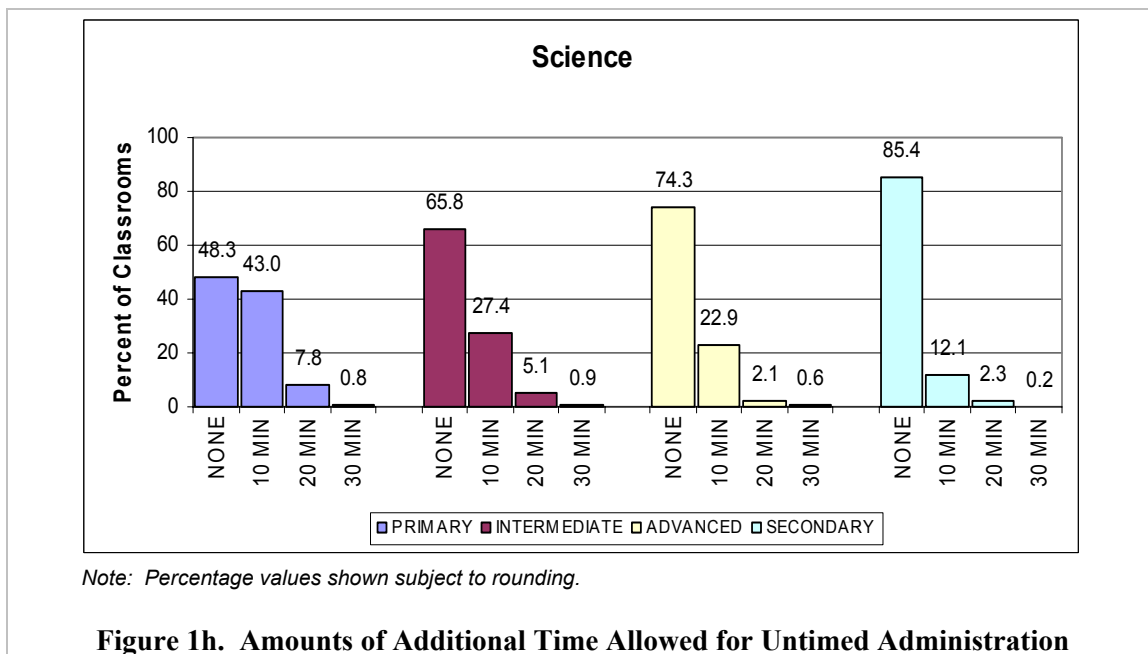
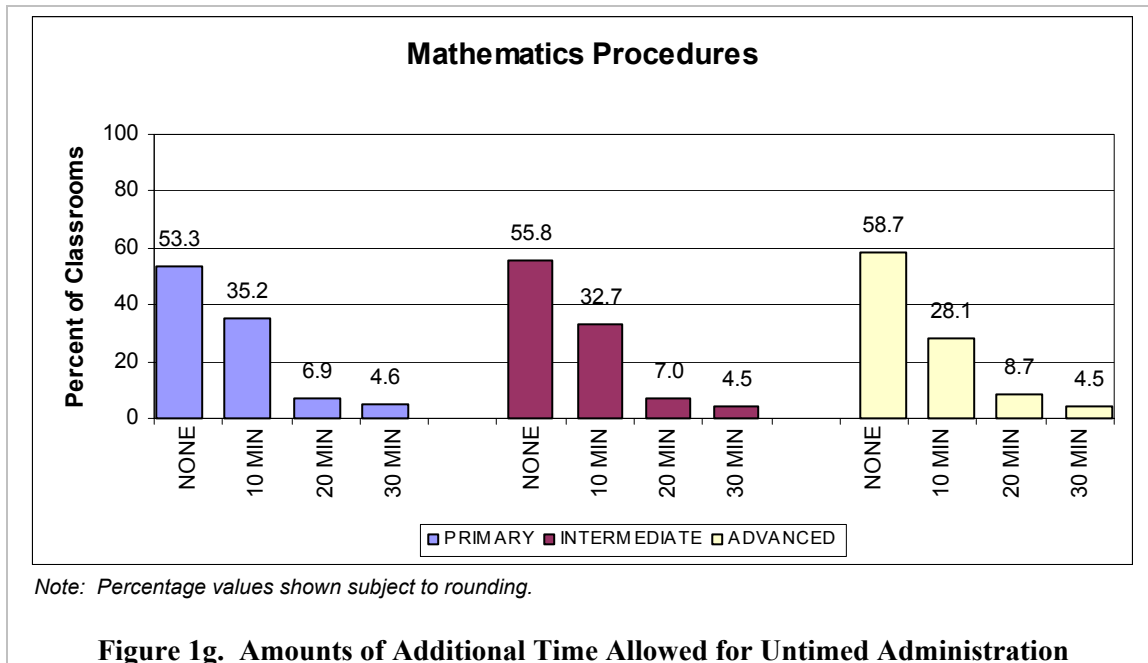
Timed Versus Untimed Testing Conditions and Student Performance



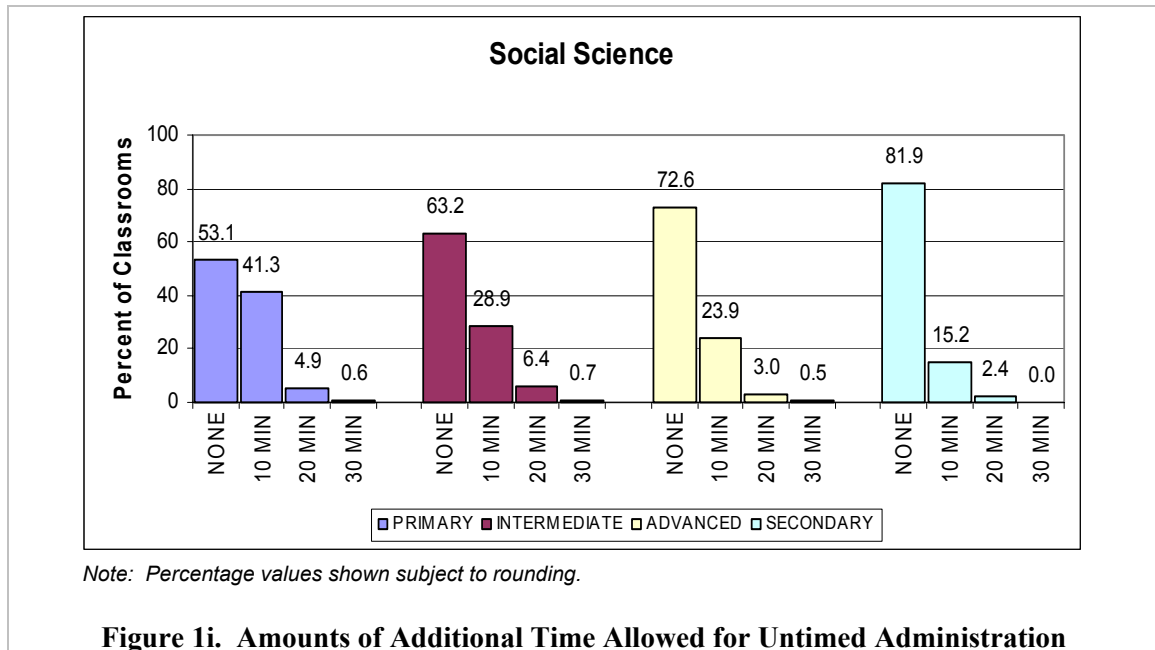
Timed Versus Untimed Testing Conditions and Student Performance



Timed Versus Untimed Testing Conditions and Student Performance



Timed Versus Untimed Testing Conditions and Student Performance



Impact of Additional Testing Time on Non-Disabled Student Performance

The results of Pearson's study showed that allowing additional time on tests gave no demonstrable advantage to non-disabled students. However, there was a concern that some teachers may have allowed excessive amounts of additional time in an effort to maximize their students' scores. To investigate this possibility, the actual testing times recorded by the teachers were merged with actual test scores of the non-disabled students in their classes. Thus, it was possible to calculate the mean raw score for all classrooms at each grade that received a certain amount of additional time.

The results of this analysis for the Reading Comprehension and Mathematics Problem Solving subtests for grades 1 through 8 and grade 10 are presented in Table 2. It is clear that, for a particular grade, there is no consistent increase or decrease in mean raw scores for non-disabled students receiving from none to 25 minutes of additional time. There is also no consistent advantage gained by allowing what might be considered an excessive amount of additional time. Mean raw scores for classes that allowed more than 25 minutes of additional time are not consistently higher or lower than scores for classes that were allowed any other amount of additional time. Therefore, the data show that it is very unlikely that excessive time extensions for non-disabled students would be problematic in practice.

Timed Versus Untimed Testing Conditions and Student Performance

Table 2. Mean Raw Score as a Function of Additional Time Allowed

Reading Comprehension									
Additional Time Allowed	Primary 1	Primary 2	Primary 3	Intermed. 1	Intermed. 2	Intermed. 3	Adv. 1	Adv. 2	TASK
none	30.9	27.3	34.5	30.9	32.0	34.6	33.4	36.0	32.8
5 min.	29.8	25.8	38.5	35.8	32.9	33.6	38.4	33.7	34.9
10 min.	28.5	24.4	35.9	36.1	35.1	36.9	28.2	37.5	31.6
15 min.	28.7	27.9	35.1	28.4	32.1	31.5	34.3	40.4	44.0
20 min.	31.8	25.6	36.3	39.9	30.7	32.4	35.6	NC	32.9
25 min.	28.5	22.3	34.9	33.8	35.4	33.5	36.0	NC	NC
over 25 min.	29.2	29.1	NC	32.8	36.5	36.0	NC	NC	NC
Mathematics Problem Solving									
Additional Time Allowed	Primary 1*	Primary 2*	Primary 3	Intermed. 1	Intermed. 2	Intermed. 3	Adv. 1	Adv. 2	TASK**
none			30.4	28.4	28.1	25.6	24.3	24.1	21.01
5 min.			31.7	30.6	29.2	25.9	30.8	24.3	17.79
10 min.			31.2	25.7	28.6	25.8	22.0	23.9	20.11
15 min.			24.8	34.5	27.6	25.6	19.1	30.8	17.00
20 min.			29.4	28.0	31.4	26.4	25.0	27.1	23.29
25 min.			32.6	28.8	25.3	28.3	32.5	26.6	19.00
over 25 min.			38.0	27.1	19.0	24.6	NC	29.3	NC

NC = No classrooms at this test level allowed additional time.

* = The Mathematics Problem Solving subtest is dictated to students at this test level. Time limits are not applicable.

** = Since there is no separate Mathematics Problem Solving subtest at this level, the Total Mathematics score was used.

Conclusion

Although there was considerable variability in the numbers of classrooms in which additional testing time was allowed, as well as the amount of additional time allowed, teachers did not appear to have allowed excessive amounts of additional testing time during the standardization of Stanford 10. Differences in the performance of non-disabled students under timed and untimed conditions on Stanford 10 are essentially non-existent. Although this may be an unexpected result, it is not inconsistent with the results of studies conducted by other researchers in the field. This result tends to confirm that Pearson's procedures for determining suggested testing times, developed over the past 80 years, have set accurate suggested times for non-disabled students to complete the test without undue strain or errors owing to time pressure. More important is the observation that allowing additional testing time will not boost the scores of non-disabled students over students with disabilities. Therefore, there is little cause for concern that flexible

Timed Versus Untimed Testing Conditions and Student Performance

timing guidelines will be abused by test administrators. The single most crucial result of all this work is the knowledge that Stanford 10 norms are unaffected by the administration of tests under untimed conditions. Thus, students with disabilities who are allowed additional testing time can receive unqualified norm-referenced scores.

References

- Alster, E. H. (1997). The effects of extended time on algebra test scores for college students with and without learning disabilities. *Journal of Learning Disabilities, 30*(2), 222–227.
- American Education Research Association, American Psychological Association, National Council of Measurement in Education. 1999 *Standards for education and psychological testing*. Washington, D.C.: American Education Research Association.
- Americans with Disabilities Act* of 1990 (ADA) (Pub. L. 101–336), as amended.
- Anastasi, A. (1976) *Psychological Testing* (pp. 32–34). New York: Macmillan.
- 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.
- Bridgman, B., Curley, E., & Trapani, C. (2002). *Impact of fewer questions per section on SAT[®] I scores*. Manuscript submitted for publication.
- Cahalan, C., Mandinach, E., & Camara, W. (2002). *Predictive validity of SAT[®] I: Reasoning test for test-takers with learning disabilities and extended time accommodations*. (College Board Research Report No. 2002-5). New York: College Entrance Examination Board.
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement*, pp. 286–288. Englewood Cliffs, NJ: Prentice-Hall.
- Halla, J. E. (1998). A psychological study of psychometric differences in graduate record examinations general test scores between learning disabled and non-learning disabled adults (Doctoral dissertation, Texas Tech University, Lubbock). *Dissertation Abstracts International, 49*, 3341A.
- Harker, J. K. (1991). A comparison of achievement test performance under silent reading and reading plus listening modes of administration (Doctoral dissertation, The University of Iowa, Iowa City). *Dissertation Abstracts International, 52*, 4304A.
- Henderson, C. (2001). *College freshman with disabilities: Statistical year 2000*. Washington DC: American Council on Education.
- Hill, G. A. (1984). Learning disabled college students: The assessment of academic aptitude (Doctoral dissertation, Texas Tech University, Lubbock). *Dissertation Abstracts International, 46*, 642A.
- Huesman, R. L., & Frisbie, D. A. (2000, April). *The validity of ITBS reading comprehension test scores for learning disabled and non-learning disabled students under extended time conditions*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA.
- Kronholz, J. (2002, November 20). Pros debate whether standardized tests should be timed. *The Wall Street Journal*, p. B1.

Timed Versus Untimed Testing Conditions and Student Performance

- Munger, G. G. & Loyd, B. H. (1991). Effect of speededness on test performance of handicapped and nonhandicapped examinees. *Journal of Educational Research*, 85(1), 53–57.
- Perlman, C., Borger, J., Collins, C., Elenbogen, J., & Wood, J. (1996, April). *The effect of extended time limits on learning-disabled students' scores on standardized reading tests*. Paper presented at the annual meeting of the National Council on Measurement in Education, New York, NY.
- 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.
- Runyan, M. K. (1991a). The effect of extra time on reading comprehension scores for university students with and without learning disabilities. *Journal of Learning Disabilities*, 24(2), 104–108.
- Runyan, M. K. (1991b). Reading comprehension performance of learning-disabled and non-learning disabled college and university students under timed and untimed conditions (Doctoral dissertation, University of California, Berkeley). *Dissertation Abstracts International*, 52(2), 2875A.
- Tindal, G., & Fuchs, L. (1999). *A summary of research on test changes: An empirical basis for defining accommodations*. Lexington, KY: University of Kentucky, Mid-South Regional Resource Center.
- 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: College Entrance Examination Board.
- 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.). (1998). *Testing handicapped people*. Boston, MA: Allyn and Bacon.

Additional copies of this and related documents are available from:
Pearson Inc.
19500 Bulverde Rd.
San Antonio, TX 78259
1-800-211-8378
1-877-576-1816 (fax)
<http://www.pearsonassess.com>