



SLEP[®]

Technical Supplement

**For use with Forms
4, 5, and 6 only**



*Listening.
Learning.
Leading.*

Secondary Level English Proficiency Test

TABLE OF CONTENTS

1.	Introduction.....	1
2.	Pilot Study Design	1
3.	Equating and Scaling	1
4.	Statistical Characteristics of SLEP Forms	2
5.	Table 1: Summary Statistics for the New SLEP Forms.....	2
6.	Table 2: Correlations Among Scale Scores	2
7.	Relating Student Characteristics to Performance on SLEP	3
8.	Table 3: SLEP Performance by Type of Program	3
9.	Table 4: SLEP Performance by Students' U.S. Citizenship Status.....	4
10.	Table 5: SLEP Performance by Length of Time in United States	4
11.	Table 6: SLEP Performance by Length of Time Enrolled in U.S. School.....	5
12.	Table 7: SLEP Performance by Length of English Study in U.S.	5
13.	Table 8: SLEP Performance by Length of English Study Outside U.S.	6
14.	Table 9: SLEP Performance by Native Language	6
15.	Table 10: SLEP Performance by Grade.....	7
16.	Table 11: SLEP Performance by Gender	7
17.	Table 12: Scaled Scores of Native English-Speaking Students by Grade	8
18.	Relationship of SLEP Scores to Teachers' Ratings of Students' English Language Proficiency.....	9
19.	Table 13: SLEP Performance by Teachers' Evaluation of Students.....	9
20.	Figure 1: Listening Scale	10
21.	Figure 2: Reading Scale	10
22.	SLEP Descriptors for Listening and Reading.....	11

SLEP Technical Supplement

This supplement summarizes the statistical characteristics of newly developed SLEP Forms 4, 5, and 6 and provides additional comparative data for score interpretation. The document is organized into five sections:

- Pilot Study Design
- Equating and Scaling
- Statistical Characteristics of SLEP Forms
- Relating Student Characteristics to Performance on SLEP
- Relationship of SLEP Scores to Teachers' Ratings of Students' English Language Proficiency

Pilot Study Design

The purpose of the pilot study was to collect data to determine the statistical characteristics of the test forms, equate the forms to the current SLEP scale, and develop norms for interpretative and comparative purposes.

The pilot study was conducted during February and March of 2003. Forms 4, 5, and 6 were administered to 1,650 nonnative English-speaking students in grades 7 - 12 in several public school districts in Florida, New Jersey, Louisiana, Maryland, Ohio, Texas, Utah, and Virginia. Each student was administered one of the new forms together with an old form. The test pairs were administered in counterbalanced order, such that approximately half the students took the new form followed by the old and half took the old form followed by the new. Test administrations took place over two to four days, depending on the number of tests each teacher administered on any one day.

To gather validity evidence, one of the new forms, Form 4, was also administered to a small sample of 175 native English speakers in grades 9 - 12. This group of examinees was not included in the equating, but the data were used for comparative purposes.

In addition, teachers were asked to rate all participating students on their listening and reading skills, independent of their performance on SLEP. The ratings were then compared with the students' actual SLEP listening and reading scores to provide additional evidence that SLEP is measuring important listening and reading skills.

Equating and Scaling

The SLEP listening scale ranges from 10 to 32; the reading scale ranges from 10 to 35; and the total scale score, which is composed of the sum of the scale scores for Listening and Reading, ranges from 20 to 67. To maintain the current scale for the newly developed forms, they were equated to scale through a previously developed form.

The equating used a single-group design in which each student was administered both tests, a new and an old, in counterbalanced order. The frequency distributions for the listening and reading raw scores were then smoothed and the equipercentile equating was performed. The results of the equatings, that is the score conversion tables, are provided in the *SLEP Test Manual*.

Statistical Characteristics of SLEP Forms

Table 1 presents summary statistics for the three new SLEP forms 4, 5, and 6 including the means, standard deviations, reliabilities and standard errors of measurement for the listening, reading and total scale scores. These data are based on the 1,650 students who participated in the SLEP pilot study.

The slight differences in means and standard deviations reflect the differences in language ability among pilot samples taking each form. Because of test administration logistics, the forms could not be randomly assigned across the entire group of participating students. Forms were randomly assigned to intact classrooms, resulting in slight differences within the sample taking each form.

The reliabilities were estimated using Cronbach's coefficient alpha, which is an internal consistency measure of reliability. This type of reliability takes into account variation due to content sampling.

Table 1			
Summary Statistics for the New SLEP Forms			
	Form 4 (N=559)	Form 5 (N=590)	Form 6 (N=501)
Listening			
Mean	21.18	20.45	20.70
Standard Deviation	5.88	5.80	5.96
Reliability	.94	.93	.94
Standard Error of Measurement	1.5	1.6	1.5
Reading			
Mean	20.51	20.06	20.31
Standard Deviation	5.11	4.84	5.17
Reliability	.88	.91	.90
Standard Error of Measurement	1.8	1.4	1.6
Total			
Mean	41.69	40.51	41.01
Standard Deviation	10.25	9.99	10.47
Reliability	.95	.95	.96
Standard Error of Measurement	2.3	2.1	2.2

Table 2 presents the correlations among the scale scores. The observed correlations are presented above the diagonal; and the corrected below. The corrected correlations provide estimates of the correlations between the scores if the tests were perfectly reliable.

Table 2			
Correlations Among Scale Scores			
	Listening	Reading	Total
Listening		0.76	0.95
Reading	0.82		0.93
Total	1.00	1.00	

Relating Student Characteristics to Performance on SLEP

Students participating in the pilot study were also asked to complete a background information questionnaire. Tables 3 - 13 present summary statistics of student performance on SLEP broken down by the various demographic variables. Each table provides the means and standard deviations for the listening, reading, and total scale scores. The number of students in each response category was also reported. The results in these tables provide information useful for comparing the performance or progress of students with similar characteristics.

Table 3 provides a summary of SLEP performance organized by the type of remedial English instruction programs in which the students were enrolled. The results indicate that for students enrolled in either an ESL or a bilingual education program, those receiving part-time instruction had higher mean scores than did their counterparts receiving full-time instruction. While the results suggest mainstreamed students performed better than students receiving remedial instruction, the differences in scores between mainstreamed students and students in part-time ESL programs were very small.

It should also be noted that the data presented here reflect many different ESL and bilingual programs. No attempt was made to ensure the comparability of program types across participating schools. It is highly probable that program types varied considerably from school to school.

		Listening		Reading		Total	
Program Type*	N	Mean	SD	Mean	SD	Mean	SD
ESL, full time	501	18.4	5.6	18.6	4.7	37.0	9.6
ESL, part time	906	22.5	5.3	21.6	4.7	44.1	9.3
Bilingual, full time	146	14.5	4.8	16.9	4.5	31.3	8.8
Bilingual, part time	56	16.2	4.9	18.2	3.2	34.3	7.3
Mainstream class	37	23.2	7.4	22.4	7.1	45.6	14.0

*Four examinees were excluded because of non response.

Table 4 provides a comparison of SLEP performance in terms of U.S. citizenship status. The results show that students with U.S. citizenship had higher mean scores than those of non-U.S. citizens. This difference is probably a result of the fact that, on average, the students with U.S. citizenship were in the country for a longer period of time and had more opportunities to improve their English language proficiency.

U.S. Citizenship Status*	N	Listening		Reading		Total	
		Mean	SD	Mean	SD	Mean	SD
U.S. Citizen	205	21.9	5.8	21.3	5.0	43.2	10.3
Non U.S. Citizen	1,192	19.8	6.0	19.8	5.0	39.6	10.3
Other	220	21.2	5.7	20.5	4.9	41.7	9.9

*Thirty three examinees were eliminated because of non response.

For nonnative English-speaking students, the length of time in the United States plays a significant role in the improvement of their English language proficiency. The results in Table 5 indicate that the longer students were in this country, the higher their average SLEP scores. Listening proficiency, in particular, benefits from a longer stay. For example, students who were in the United States from two to five years outperformed those who were in the country less than six months by 5.5 scale points in listening; this compares with a 3.3 scale score difference in their average reading scores.

Length of Time in U.S.	N	Listening		Reading		Total	
		Mean	SD	Mean	SD	Mean	SD
Less than 6 months	131	16.5	5.8	17.7	5.0	34.2	10.4
6 months to 1 year	229	17.5	5.9	18.7	5.0	36.2	10.4
1 year to 2 years	471	18.8	5.8	19.5	4.8	38.3	10.0
2 years to 5 years	636	22.0	5.3	21.0	4.9	43.0	9.4
Over 5 years - less than lifetime	148	24.9	4.5	22.4	4.7	47.3	8.5
Lifetime	35	24.7	4.4	23.4	4.3	48.1	8.0

English language proficiency is also related to the length of time students are enrolled in U.S. schools. Table 6 summarizes performance on SLEP by the length of time students were enrolled in U.S. schools. The results indicate that students enrolled in U.S. schools for more than one year had higher mean scores than those of students enrolled in U.S. schools for less than six months.

Length of Time in US School*	N	Listening		Reading		Total	
		Mean	SD	Mean	SD	Mean	SD
Less than 6 months	204	16.9	5.7	17.9	4.8	34.7	10.1
6 months to 1 year	260	17.8	6.1	18.8	5.1	36.6	10.6
More than 1 year	1185	21.5	5.6	20.9	4.9	42.4	9.8

*One student was eliminated because of non response.

Table 7 presents a summary of SLEP performance by the length of time students studied English within the United States. As one might expect, both listening and reading SLEP mean scores increase as the length of study in the United States increases.

Length of English Study in U.S. *	N	Listening		Reading		Total	
		Mean	SD	Mean	SD	Mean	SD
None	50	16.2	5.8	17.0	4.1	33.2	9.4
Less than 1 year	408	17.0	5.8	18.0	5.0	35.0	10.2
Over 1 year, but less than 2 years	437	19.8	5.7	20.4	4.8	40.2	9.9
2 years or more	753	22.7	5.2	21.4	4.8	44.2	9.2

*Two students were eliminated because of non response.

Table 8 presents a summary of performance on SLEP by the length of time students studied English outside the United States. Based on these data, there is no clear relationship between SLEP scores and length of study outside the United States. However, these results are likely to be affected by the small sample reporting that they had not studied English in the country. It is therefore difficult to tease out the effects of English study outside the United States on SLEP scores.

Length of English Study outside U.S. *	N	Listening		Reading		Total	
		Mean	SD	Mean	SD	Mean	SD
None	776	20.2	6.3	19.9	5.0	40.1	10.7
Less than 1 year	314	19.0	6.0	19.1	4.8	38.1	10.2
Over 1 year, but less than 2 years	159	21.1	5.7	20.7	5.2	41.8	10.3
2 years or more	395	21.4	5.5	21.3	5.0	42.7	9.8

*Six students were eliminated because of non response.

Table 9 presents a summary of performance on SLEP based on the students' native languages; and Table 10 provides additional comparative data on SLEP performance by grade level.

Native Language*	N	Listening		Reading		Total	
		Mean	SD	Mean	SD	Mean	SD
Arabic	21	21.9	6.1	20.9	5.3	42.8	10.5
Chinese	35	21.2	6.0	20.5	4.6	41.7	10.0
Creole (Haitian)	341	15.8	5.2	17.5	4.3	33.3	8.9
French	18	20.8	6.7	21.3	5.2	42.1	11.5
Hindi **	8	--	--	--	--	--	--
Korean	28	24.9	3.5	23.4	3.4	48.3	6.3
Punjabi	2	--	--	--	--	--	--
Russian	13	25.9	4.4	25.1	3.3	51.0	6.6
Spanish	919	21.1	5.8	20.7	5.1	41.8	10.2
Serbo-Croatian **	4	--	--	--	--	--	--
Tagalog	11	22.9	4.5	22.9	4.3	45.8	8.4
Urdu	18	19.8	5.4	18.7	4.2	38.5	8.7
Vietnamese	48	22.5	4.5	21.8	4.1	44.3	8.2
Other	173	22.6	5.2	20.7	5.0	43.3	9.5

* Eleven students were eliminated because of non response.

** No summary statistics are reported if the sub-group N is less than 10.

The results in Table 10 indicate that students' scores do not seem to be related to grade level, i.e., students at higher grade levels did not achieve higher mean scores.

Table 10 SLEP Performance by Grade							
Grade*	N	Listening		Reading		Total	
		Mean	SD	Mean	SD	Mean	SD
7 th	112	23.4	5.3	22.0	4.7	45.4	9.4
8 th	122	23.4	5.5	21.8	5.2	45.2	10.1
9 th	399	18.2	6.0	18.1	4.5	36.3	9.9
10 th	413	19.8	6.0	20.1	5.0	39.9	10.4
11 th	368	20.6	5.9	20.5	5.0	41.2	10.3
12 th	233	21.5	5.3	21.5	4.8	43.0	9.2
Other**	1	--	--	--	--	--	--

*Two students had missing values on this variable.

** No summary statistics were reported if the subgroup N is less than 10.

SLEP performance by gender is presented in Table 11. The results show that male and female nonnative students performed equally well on the new SLEP forms; the mean scores and standard deviations were almost identical for each section as well as for the for the total score.

Table 11 SLEP Performance by Gender							
Gender*	N	Listening		Reading		Total	
		Mean	SD	Mean	SD	Mean	SD
Male	858	20.4	6.1	20.2	5.1	40.6	10.4
Female	787	20.3	6.0	20.1	5.0	40.5	10.4

*Five students were eliminated due to non-response.

As part of the pilot study, Form 4 of SLEP was administered to a small group of 175 native English-speaking students in grades 9 - 12. This was done to illustrate that SLEP does not discriminate language ability among native speakers and is therefore useful only for students for whom English is a second language. Table 12 presents these data. Means and standard deviations of SLEP scores of native English-speaking students are presented by grade. Although the number of students at each grade level is small, the results do suggest that, although native English-speaking students on average outperform the nonnative English speakers group, their scores on SLEP are not related to grade level.

Table 12
Scaled Scores of Native English-Speaking Students by Grade

Grade*	Listening		Reading		Total		
	N	Mean	SD	Mean	SD	Mean	SD
9 th	47	30.8	2.0	31.8	3.5	62.6	5.0
10 th	48	29.7	4.1	29.9	6.5	59.6	9.5
11 th	49	30.1	4.4	32.2	5.5	62.3	8.3
12 th	30	30.5	2.4	33.3	3.7	63.9	5.3

*One student was eliminated due to a missing grade indicator.

Relationship of SLEP Scores to Teachers' Ratings of Students' English Language Proficiency

As part of the pilot study teachers of participating schools were asked to independently rate the listening and reading skills of their students on a scale that ranged from 1 to 4, with level 1 the lowest and level 4 the highest. A copy of the rating scale used by the teachers, including the definition of each level, is provided at the end of this document.

Table 13 presents a summary of the SLEP means and standard deviations for students by level, based on the teacher evaluation. As indicated in the table, the mean scores for each level increase steadily for both listening and reading.

Ability Level	Listening			Reading*		
	N	Mean	SD	N	Mean	SD
1	338	14.5	4.3	384	15.9	3.6
2	390	19.1	5.0	527	19.3	4.0
3	581	22.4	4.6	504	22.4	4.3
4	341	26.2	3.0	232	25.2	4.0

*Three students were eliminated due to missing teacher evaluations for reading skills.

Figures 1 and 2 provide a visual representation of these data. They are box-and-whisker plots that graph several percentiles of the distribution of scores. The rectangular box extends from the 25th to the 75th percentile and the whiskers extend to the 10th and 90th percentile of the distribution of test scores. The horizontal line within the box represents the 50th percentile.

In Figure 1, the box-and-whisker plots for listening clearly indicate that SLEP can discriminate the four different levels of ability as defined in the rating scale. Although there is certainly overlap in the scores in adjacent levels, both the 50th and 75th percentiles of each higher level is higher than the level below. Figure 2 presents the plot for reading. In this plot, while there is a steady progression in the percentiles of the adjacent levels, there is more overlap in scores among the levels and the differences are not as clear as for listening. This may also be due to the more difficult task of rating students' reading ability which is not as directly observable as listening skills are.

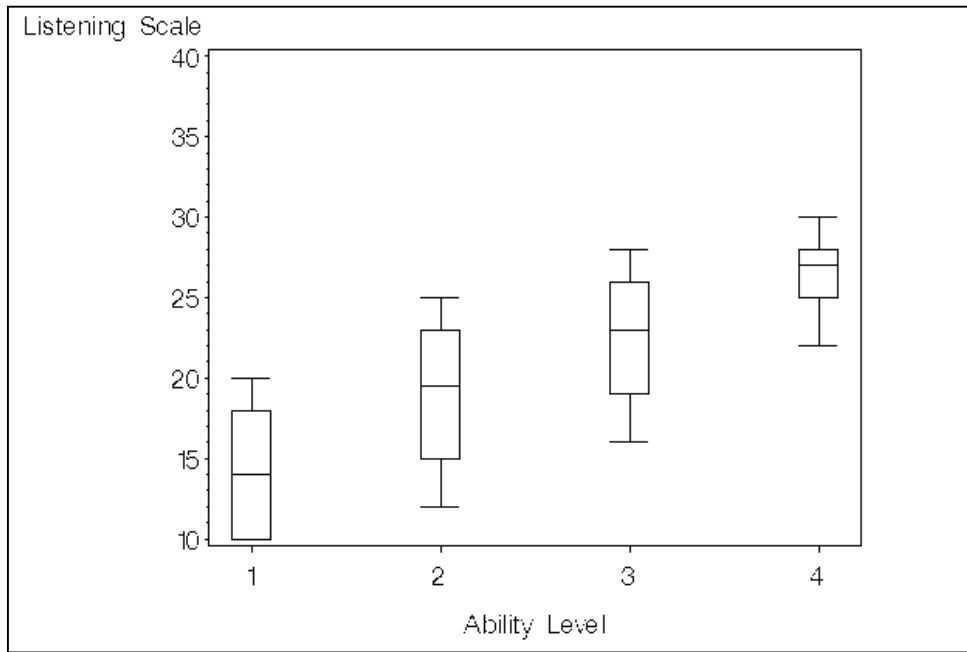


Figure 1: Listening Scale Scores by Teachers' Listening Evaluations

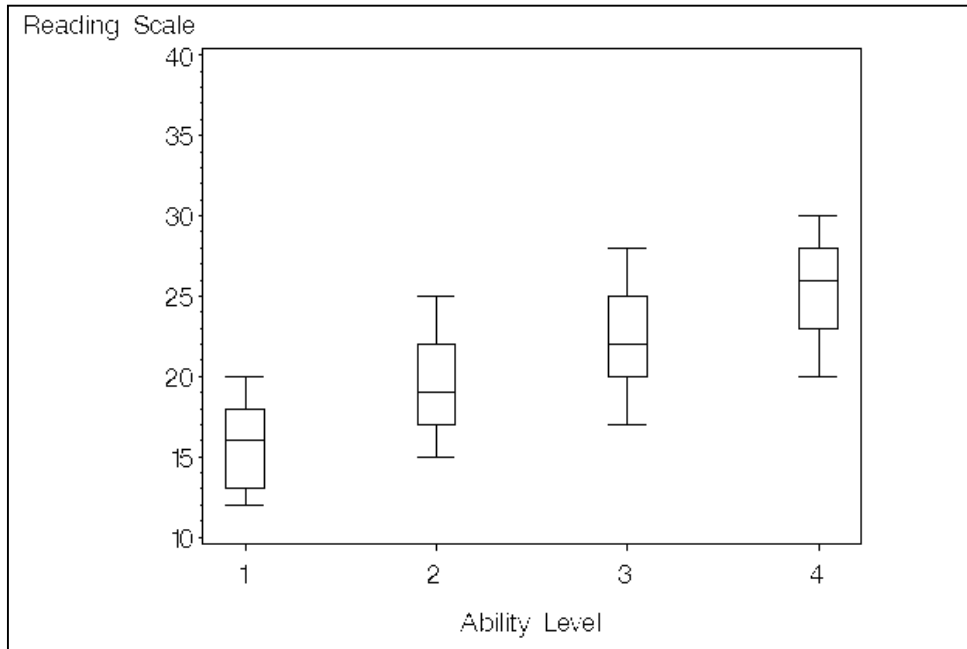


Figure 2: Reading Scale Scores by Teachers' Reading Evaluations

SLEP Descriptors for Listening and Reading

Listening

Level One

Comprehends very little or no spoken English. Can understand single words or phrases, but has difficulty following complete sentences. Struggles to understand simple grammatical constructions even in interpersonal conversational situations.

Level Two

Comprehends little spoken English. Is relatively successful in following and understanding personal conversations, but is unable to follow along in a noninteractive listening situation (such as listening to an announcement or listening to the radio).

Level Three

Comprehends most spoken English. Is very competent in interpersonal situations, but experiences difficulty in fully comprehending in noninteractive listening situations.

Level Four

Comprehends virtually all spoken English. Is not deterred by unfamiliar vocabulary, using the same contextual strategies a native speaker would use in processing unfamiliar words.

Reading

Level One

Possesses a very limited reading vocabulary. Can recognize only words, short phrases, and some simple sentences but cannot read at the paragraph level.

Level Two

Can read and comprehend paragraphs and very short stories that contain only relatively simple sentences and relatively high frequency vocabulary.

Level Three

Can read and comprehend, with difficulty, some grade-level English text that contains high-level vocabulary and complex sentence structures.

Level Four

Can easily read and comprehend all forms of written English, including literature filled with complex sentence structures and high-level, grade-appropriate vocabulary.

