
ADDENDUM TO THE
STANDARDS-BASED TESTS IN SPANISH
(STS)
TECHNICAL REPORT
SPRING 2011 ADMINISTRATION

The purpose of this addendum to the Standards-based Tests in Spanish Technical Report for the spring 2011 administration is to provide the raw-score-to-scale-score conversions and analyses for estimating the reliability of classification decisions. These analyses were completed following the approval of the STS performance levels for reading/language arts in grades eight through eleven and end-of-course mathematics tests in Algebra I and Geometry by the State Board of Education in July 2012.

Score Conversion Tables

In November 2011, a standard setting for the Standards-based Tests in Spanish (STS) was conducted to establish performance-level cut scores for reading/language arts (RLA) in grades eight through eleven and end-of-course tests in Algebra I and Geometry. These cut scores were approved, were adopted in July 2012 by the State Board of Education, and will be implemented for spring 2013 operational administration. In this addendum, data from the spring 2011 operational administration were used to assess the impact of these cut scores.

In Table 1 through Table 6, the cut scores are identified in the raw-score-to-scale-score conversion tables for these tests. Also shown are the percentages of students in the STS target population in the spring 2011 administration in each performance level defined by the cut scores. The percentages were based on P2 data which contained almost 100 percent of the test results of the entire target population.

Table 1. 2011 Standard Setting Performance Levels and Impact Results: RLA, Grade Eight

Raw Score	Scale Score	CSEM	Performance Level	% Students at Performance Level
75	600	81		
74	589	63		
73	545	45		
72	518	37		
71	499	33		
70	484	30		
69	471	27		
68	460	25	Advanced	10.56%
67	450	24		
66	442	23		
65	434	22		
64	426	21		
63	419	20		
62	413	20		
61	407	19		
60	401	19		
59	396	18		
58	391	18		
57	386	18		
56	381	17		
55	376	17		
54	371	17	Proficient	23.72%
53	367	17		
52	363	16		
51	358	16		
50	354	16		
49	350	16		

Raw Score	Scale Score	CSEM	Performance Level	% Students at Performance Level
48	346	16		
47	342	16		
46	338	16		
45	334	16		
44	330	16		
43	326	15		
42	323	15	Basic	32.41%
41	319	15		
40	315	15		
39	311	15		
38	308	15		
37	304	15		
36	300	15		
35	296	15		
34	292	15		
33	289	16		
32	285	16		
31	281	16		
30	277	16		
29	273	16	Below Basic	26.48%
28	269	16		
27	265	16		
26	261	16		
25	257	16		
24	252	16		
23	248	17		
22	243	17		
21	239	17		
20	234	17		
19	229	18		
18	224	18		
17	219	18		
16	214	19		
15	208	19		
14	202	20		
13	196	20		
12	189	21		
11	182	21	Far Below Basic	6.82%
10	175	22		
9	166	23		
8	158	24		
7	150	26		
6	150	28		
5	150	30		
4	150	33		
3	150	38		
2	150	46		
1	150	64		
0	150	71		

Note: the percentages may not sum to 100% due to rounding.

Table 2. 2011 Standard Setting Performance Levels and Impact Results: RLA, Grade Nine

Raw Score	Scale Score	CSEM	Performance Level	% Students at Performance Level
75	600	80		
74	587	62		
73	543	44		
72	517	37		
71	498	32		
70	483	29		
69	470	27		
68	459	25		
67	450	24	Advanced	9.58%
66	441	22		
65	433	21		
64	426	21		
63	419	20		
62	413	19		
61	407	19		
60	401	18		
59	396	18		
58	391	18		
57	386	17		
56	381	17		
55	376	17		
54	371	17	Proficient	23.73%
53	367	16		
52	363	16		
51	358	16		
50	354	16		
49	350	16		
48	346	16		
47	342	16		
46	338	15		
45	334	15		
44	330	15		
43	326	15		
42	323	15	Basic	33.17%
41	319	15		
40	315	15		
39	311	15		
38	308	15		
37	304	15		
36	300	15		
35	296	15		
34	292	15		
33	289	15		
32	285	15		
31	281	15		
30	277	15		
29	273	16	Below Basic	25.42%
28	269	16		
27	265	16		
26	261	16		
25	257	16		
24	253	16		
23	248	16		

Raw Score	Scale Score	CSEM	Performance Level	% Students at Performance Level
22	244	17		
21	240	17		
20	235	17		
19	230	17		
18	225	17		
17	220	18		
16	215	18		
15	209	19		
14	204	19		
13	198	19		
12	191	20		
11	184	21	Far Below Basic	8.09%
10	177	22		
9	169	22		
8	160	24		
7	151	25		
6	150	27		
5	150	29		
4	150	32		
3	150	37		
2	150	44		
1	150	62		
0	150	74		

Note: the percentages may not sum to 100% due to rounding.

Table 3. 2011 Standard Setting Performance Levels and Impact Results: RLA, Grade Ten

Raw Score	Scale Score	CSEM	Performance Level	% Students at Performance Level
75	600	86		
74	600	67		
73	559	48		
72	531	40		
71	510	35		
70	494	31		
69	480	29		
68	469	27		
67	458	25		
66	449	24	Advanced	10.61%
65	440	23		
64	432	22		
63	425	22		
62	418	21		
61	412	20		
60	406	20		
59	400	20		
58	394	19		

Raw Score	Scale Score	CSEM	Performance Level	% Students at Performance Level
57	389	19		
56	383	19		
55	378	18		
54	373	18		
53	368	18	Proficient	25.00%
52	364	18		
51	359	17		
50	354	17		
49	350	17		
48	346	17		
47	341	17		
46	337	17		
45	333	17		
44	329	17		
43	324	17	Basic	33.03%
42	320	16		
41	316	16		
40	312	16		
39	308	16		
38	304	16		
37	300	16		
36	296	16		
35	292	16		
34	288	16		
33	284	17		
32	280	17		
31	275	17		
30	271	17	Below Basic	25.68%
29	267	17		
28	263	17		
27	258	17		
26	254	17		
25	249	17		
24	245	17		
23	240	18		

Raw Score	Scale Score	CSEM	Performance Level	% Students at Performance Level
22	235	18		
21	230	18		
20	225	18		
19	220	19		
18	215	19		
17	209	19		
16	204	20		
15	198	20		
14	191	21		
13	185	21		
12	178	22		
11	170	23	Far Below Basic	5.68%
10	162	23		
9	154	24		
8	150	26		
7	150	27		
6	150	29		
5	150	31		
4	150	35		
3	150	40		
2	150	48		
1	150	64		
0	150	64		

Table 4. 2011 Standard Setting Performance Levels and Impact Results: RLA, Grade Eleven

Raw Score	Scale Score	CSEM	Performance Level	% Students at Performance Level
75	600	97		
74	600	76		
73	590	54		
72	558	45		
71	534	39		
70	516	35		
69	500	32		
68	487	30		
67	476	29		
66	465	27	Advanced	8.71%
65	455	26		
64	447	25		
63	438	24		
62	431	24		
61	423	23		
60	416	22		
59	410	22		
58	403	21		
57	397	21		

Raw Score	Scale Score	CSEM	Performance Level	% Students at Performance Level
56	392	21		
55	386	20		
54	380	20		
53	375	20		
52	370	20	Proficient	24.44%
51	365	19		
50	360	19		
49	355	19		
48	350	19		
47	345	19		
46	341	19		
45	336	19		
44	331	18		
43	327	18		
42	322	18	Basic	33.85%
41	318	18		
40	313	18		
39	309	18		
38	304	18		
37	300	18		
36	296	18		
35	291	18		
34	287	18		
33	282	18		
32	278	18		
31	273	18		
30	269	19	Below Basic	26.26%
29	264	19		
28	259	19		
27	254	19		
26	250	19		
25	245	19		
24	240	19		
23	235	20		

Raw Score	Scale Score	CSEM	Performance Level	% Students at Performance Level
22	229	20		
21	224	20		
20	218	20		
19	213	21		
18	207	21		
17	201	22		
16	195	22		
15	188	22		
14	181	23		
13	174	24		
12	166	24		
11	158	25	Far Below Basic	6.74%
10	150	26		
9	150	27		
8	150	29		
7	150	30		
6	150	33		
5	150	35		
4	150	39		
3	150	45		
2	150	54		
1	150	59		
0	150	59		

Table 5. 2011 Standard Setting Performance Levels and Impact Results: Algebra I

Raw Score	Scale Score	CSEM	Performance Level	% Students at Performance Level
65	600	86		
64	600	67		
63	576	48		
62	547	40		
61	527	35		
60	510	31		
59	497	29		
58	485	27		
57	474	26		
56	465	25	Advanced	2.40%
55	456	23		
54	448	23		
53	441	22		
52	434	21		
51	427	21		
50	421	20		
49	415	20		
48	409	19		
47	404	19		
46	398	19		

Raw Score	Scale Score	CSEM	Performance Level	% Students at Performance Level
45	393	19		
44	388	18		
43	383	18		
42	378	18		
41	373	18	Proficient	9.71%
40	368	18		
39	364	18		
38	359	18		
37	355	17		
36	350	17		
35	345	17		
34	341	17		
33	336	17		
32	332	17		
31	328	17		
30	323	17	Basic	28.92%
29	318	17		
28	314	17		
27	309	18		
26	305	18		
25	300	18		
24	295	18		
23	290	18		
22	285	18	Below Basic	30.34%
21	280	18		
20	275	19		
19	270	19		
18	265	19		
17	259	20		
16	253	20		
15	247	20		
14	241	21		
13	234	21		
12	227	22		
11	219	23		
10	211	24		
9	203	25	Far Below Basic	28.63%
8	193	26		
7	183	27		
6	171	29		
5	157	32		
4	150	35		
3	150	40		
2	150	48		
1	150	68		
0	150	92		

Table 6. 2011 Standard Setting Performance Levels and Impact Results: Geometry

Raw Score	Scale Score	CSEM	Performance Level	% Students at Performance Level
65	600	116		
64	600	90		
63	600	64		
62	600	53		
61	600	47		
60	581	42		
59	563	39		
58	547	36		
57	533	35		
56	520	33	Advanced	6.63%
55	508	32		
54	497	30		
53	487	29		
52	478	29		
51	469	28		
50	460	27		
49	452	27		
48	444	26		
47	436	26		
46	429	26		
45	422	25		
44	415	25		
43	408	25		
42	401	24		
41	394	24		
40	388	24	Proficient	19.16%
39	381	24		
38	375	24		
37	369	24		
36	362	24		
35	356	24		
34	350	23		
33	344	23		
32	338	23		
31	331	23		
30	325	24	Basic	28.26%
29	319	24		
28	313	24		
27	306	24		
26	300	24		
25	294	24		
24	287	24		
23	280	24		
22	274	25	Below Basic	30.47%
21	267	25		
20	260	25		
19	252	26		

Raw Score	Scale Score	CSEM	Performance Level	% Students at Performance Level
18	245	26		
17	237	26		
16	229	27		
15	221	27		
14	212	28		
13	203	29		
12	193	30		
11	183	31		
10	172	32		
9	160	33	Far Below Basic	15.48%
8	150	35		
7	150	37		
6	150	39		
5	150	42		
4	150	47		
3	150	53		
2	150	65		
1	150	68		
0	150	68		

Decision Classification Analyses

The methodology used for estimating the reliability of classification decisions is described in Livingston and Lewis (1995) and is implemented using the ETS-proprietary computer program RELCLASS-COMP (Version 4.14).

Decision accuracy describes the extent to which examinees are classified in the same way as they would be on the basis of the average of all possible forms of a test. Decision accuracy answers the following question: How does the actual classification of test takers, based on their single-form scores, agree with the classification that would be made on the basis of their true scores, if their true scores were somehow known? RELCLASS-COMP also estimates decision accuracy using an estimated multivariate distribution of reported classifications on the current form of the exam and the classifications based on an all-forms average (true score).

Decision consistency describes the extent to which examinees are classified in the same way as they would be on the basis of a single form of a test other than the one for which data are available. Decision consistency answers the following question: What is the agreement between the classifications based on two non-overlapping, equally difficult, forms of the test? RELCLASS-COMP also estimates decision consistency using an estimated multivariate distribution of reported classifications on the current form of the exam and classifications on a hypothetical alternate form using the reliability of the test and strong true-score theory.

In each case, the proportion of classifications with exact agreement is the sum of the entries in the diagonal of the contingency table representing the multivariate distribution. Reliability of classification at a cut score is estimated by collapsing the multivariate distribution at the passing score boundary into an n by n table (where n is the number of performance levels) and summing the entries in the diagonal. Figure 1 and Figure 2 present the two scenarios graphically.

Figure 1 Decision Accuracy for Achieving a Performance Level

		Decision made on a form actually taken	
		Does not achieve a performance level	Achieves a performance level
True status on all-forms average	Does not achieve a performance level	Correct classification	Misclassification
	Achieves a performance level	Misclassification	Correct classification

Figure 2 Decision Consistency for Achieving a Performance Level

		Decision made on the alternate form taken	
		Does not achieve a performance level	Achieves a performance level
Decision made on the form taken	Does not achieve a performance level	Correct classification	Misclassification
	Achieves a performance level	Misclassification	Correct classification

The results of spring 2011 STS analysis are presented in Table 7 through Table 12.

Each table includes the contingency tables for both accuracy and consistency of the various performance level classifications. The proportion of students being accurately classified is determined by summing across the diagonals of the upper tables; these proportions ranged from 0.68 to 0.74 across all of the STS. The proportion of students that were classified consistently (diagonals of the lower tables) was from 0.58 to 0.64 across all proficiency levels for these STS.

When the decisions are collapsed to below proficient versus proficient and above, the proportion of students that were classified accurately ranged from 0.90 to 0.94 across all STS. Similarly, the proportion of students that are classified consistently ranged from 0.86 to 0.91 for students classified into below proficient versus proficient and advanced.

Please note that there might be inconsistencies in data that appear in the “Total” due to rounding.

Table 7. Reliability of Classification for RLA, Grade Eight

	Placement Score	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Category Total
Decision Accuracy	0–22	0.04	0.03	0.00	0.00	0.00	0.07
	23–35	0.02	0.20	0.04	0.00	0.00	0.26
	36–48	0.00	0.04	0.24	0.04	0.00	0.32
	49–59	0.00	0.00	0.04	0.18	0.02	0.24
	60–75	0.00	0.00	0.00	0.03	0.08	0.11
Estimated Proportion Correctly Classified: Total = 0.74, Proficient & Above = 0.92							
Decision Consistency	0–22	0.04	0.03	0.00	0.00	0.00	0.07
	23–35	0.03	0.18	0.06	0.00	0.00	0.26
	36–48	0.00	0.06	0.21	0.06	0.00	0.32
	49–59	0.00	0.00	0.05	0.15	0.04	0.24
	60–75	0.00	0.00	0.00	0.03	0.07	0.11
Estimated Proportion Consistently Classified: Total = 0.64, Proficient & Above = 0.89							
Alternate Form							

Table 8. Reliability of Classification for RLA, Grade Nine

	Placement Score	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Category Total
Decision Accuracy	0–22	0.05	0.03	0.00	0.00	0.00	0.08
	23–35	0.02	0.20	0.04	0.00	0.00	0.25
	36–48	0.00	0.04	0.25	0.04	0.00	0.33
	49–58	0.00	0.00	0.04	0.17	0.02	0.24
	59–75	0.00	0.00	0.00	0.03	0.07	0.10
Estimated Proportion Correctly Classified: Total = 0.74, Proficient & Above = 0.92							
Decision Consistency	0–22	0.05	0.03	0.00	0.00	0.00	0.08
	23–35	0.03	0.17	0.05	0.00	0.00	0.25
	36–48	0.00	0.06	0.22	0.06	0.00	0.33
	49–58	0.00	0.00	0.06	0.14	0.04	0.24
	59–75	0.00	0.00	0.00	0.03	0.06	0.10
Estimated Proportion Consistently Classified: Total = 0.64, Proficient & Above = 0.89							
Alternate Form							

Table 9. Reliability of Classification for RLA, Grade Ten

	Placement Score	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Category Total
Decision Accuracy	0–22	0.04	0.02	0.00	0.00	0.00	0.06
	23–36	0.01	0.20	0.04	0.00	0.00	0.26
	37–48	0.00	0.04	0.25	0.04	0.00	0.33
	49–57	0.00	0.00	0.05	0.17	0.03	0.25
All-forms Average	58–75	0.00	0.00	0.00	0.03	0.07	0.11
	Estimated Proportion Correctly Classified: Total = 0.73, Proficient & Above = 0.91						
Decision Consistency	0–22	0.03	0.02	0.00	0.00	0.00	0.06
	23–36	0.02	0.18	0.06	0.00	0.00	0.26
	37–48	0.00	0.06	0.21	0.06	0.00	0.33
	49–57	0.00	0.00	0.06	0.14	0.05	0.25
Alternate Form	58–75	0.00	0.00	0.00	0.04	0.07	0.11
	Estimated Proportion Consistently Classified: Total = 0.63, Proficient & Above = 0.87						

Table 10. Reliability of Classification for RLA, Grade Eleven

	Placement Score	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Category Total
Decision Accuracy	0–22	0.04	0.03	0.00	0.00	0.00	0.07
	23–36	0.01	0.21	0.04	0.00	0.00	0.26
	37–47	0.00	0.04	0.24	0.05	0.00	0.34
	48–56	0.00	0.00	0.05	0.18	0.02	0.24
All-forms Average	57–75	0.00	0.00	0.00	0.04	0.04	0.09
	Estimated Proportion Correctly Classified: Total = 0.71, Proficient & Above = 0.90						
Decision Consistency	0–22	0.04	0.03	0.00	0.00	0.00	0.07
	23–36	0.02	0.19	0.05	0.00	0.00	0.26
	37–47	0.00	0.06	0.20	0.07	0.00	0.34
	48–56	0.00	0.00	0.06	0.14	0.04	0.24
Alternate Form	57–75	0.00	0.00	0.00	0.04	0.04	0.09
	Estimated Proportion Consistently Classified: Total = 0.61, Proficient & Above = 0.86						

Table 11. Reliability of Classification for Algebra I

	Placement Score	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Category Total
Decision Accuracy	0–18	0.20	0.07	0.01	0.00	0.00	0.29
	19–24	0.05	0.17	0.08	0.01	0.00	0.30
	25–35	0.00	0.03	0.25	0.00	0.00	0.29
	36–45	0.00	0.00	0.04	0.06	0.00	0.10
All-forms Average	46–65	0.00	0.00	0.01	0.01	0.00	0.02
	Estimated Proportion Correctly Classified: Total = 0.68, Proficient & Above = 0.94						
Decision Consistency	0–18	0.19	0.07	0.02	0.00	0.00	0.29
	19–24	0.07	0.13	0.09	0.01	0.00	0.30
	25–35	0.00	0.05	0.21	0.02	0.00	0.29
	36–45	0.00	0.00	0.04	0.05	0.00	0.10
Alternate Form	46–65	0.00	0.00	0.00	0.01	0.00	0.02
	Estimated Proportion Consistently Classified: Total = 0.59, Proficient & Above = 0.91						

Table 12. Reliability of Classification for Geometry

	Placement Score	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Category Total
Decision Accuracy	0–18	0.12	0.03	0.00	0.00	0.00	0.15
	19–25	0.04	0.17	0.08	0.01	0.00	0.30
	26–33	0.00	0.04	0.20	0.04	0.00	0.28
	34–45	0.00	0.00	0.03	0.17	0.00	0.19
All-forms Average	46–65	0.00	0.00	0.01	0.03	0.03	0.07
	Estimated Proportion Correctly Classified: Total = 0.68, Proficient & Above = 0.90						
Decision Consistency	0–18	0.12	0.04	0.00	0.00	0.00	0.15
	19–25	0.06	0.13	0.08	0.02	0.01	0.30
	26–33	0.00	0.06	0.16	0.06	0.00	0.28
	34–45	0.00	0.00	0.04	0.15	0.01	0.19
Alternate Form	46–65	0.00	0.00	0.01	0.03	0.03	0.07
	Estimated Proportion Consistently Classified: Total = 0.58, Proficient & Above = 0.86						

Reference

Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classification based on test scores. *Journal of Educational Measurement*, 32, 179–97.