

Reading Grade 3

Standard Setting Report

Prepared by Data Recognition Corporation





Introduction

Data Recognition Corporation conducted a standard setting meeting in Columbia, SC, February 15-16, 2017. The purpose of the meeting was to develop a Read to Succeed achievement standard for the reading subscale of the SC READY ELA test for the Act 284 Read to Succeed legislation.

The Read to Succeed Act requires that, beginning with the 2017-2018 school year, a student must be retained in the third grade if the student fails to demonstrate reading proficiency at the end of third grade as indicated by scoring at the lowest achievement level on the state summative reading assessment that equates to Not Met 1 on the South Carolina Palmetto Assessment of State Standards (SC PASS). A student may be exempt for good cause from the mandatory retention but shall continue to receive instructional support and services and reading intervention appropriate for their age and reading level. The Read to Succeed Team is aware of research around third grade retention and is currently researching models and solutions in other states and among stakeholders in South Carolina.

Policy Statement for **Not Met 1**

Student scores in the Not Met 1 performance level “substantially fail to demonstrate third-grade reading proficiency” by the end of the third grade. These scores represent the lowest achievement level on the grade 3 SC READY assessment.

The PLD development for **Not Met 1** was completed on February 7, 2017 in a meeting with South Carolina educators in Columbia.

The achievement level cut score (the transition between *Not Met 1/ Above Not Met 1*) were recommended by the standard setting panelists through three rounds of an Item Mapping method commonly referred to as the Bookmark method (Lewis, Mitzel, Green, and Patz, 2000). Panelists worked through an ordered item booklet (OIB) to make these recommendations. The number of items for the OIB used was 49. The items in the OIB represented all of the operational reading items on the test forms, and past SC PASS reading items that aligned to the new South Carolina College-and Career-Ready standards with the intent to flesh out any gaps in item difficulty. Panelists placed recommendations for the achievement level cut score based on the progression of item difficulty. These cuts (booklet page numbers) were mapped onto item difficulty (Rasch) estimates that were adjusted using the response probability criterion of .50 (RP50).

The SC READY items used for standard setting were operational items from the Spring 2016 test.

Standard Setting Panelists

A total of 25 panelists participated over the two-day meeting to provide recommendations for the achievement level cut score for *Not Met 1/ Above Not Met 1*. These panelists were recruited to represent various stakeholders within the state of South Carolina. A summary of the panel’s demographic information is provided in Table 2. The median number of years of educational experience for the panelists was 15 years, with as few as 2 year and as many as 35 years.

Table 1. Summary of Panel Demographic Information

Gender			Ethnicity				
Male	Female	Missing	White	Black	Asian	Multi-racial/ethnic	Missing
4.0%	96.0%	0.0%	56.0%	40.0%	0.0%	4.0%	0.0%

Current Assignment			Work Setting			Years Teaching		
Classroom Teacher	Educator (non-teacher)	Other	Urban	Suburban	Rural	Minimum	Maximum	Median
68.0 %	24.0%	8.0%	4.0%	36.0%	60.0%	2	35	15

The panelists were seated across five tables. Each of these tables had an appropriate mix of curriculum experience and ethnicity representation. Panelists were recruited by SCDE and the table leaders were also selected by SCDE.

Standard Setting

The agenda for the standard setting is found in Appendix A. Prior to beginning the standard setting process, a practice round was facilitated to ensure that the panelists understood the process of making achievement level cuts, using the definitions of *threshold* students. As with the actual standard setting rounds, the panelists were asked the following question: *What is the last item that 50 out of 100 **Passing into Above Not Met 1** students will answer correctly?* The panelists were instructed to place their bookmark on that item. After the practice exercise the panelists answered questions from a ‘readiness survey’ to indicate they were ready for the standard setting rounds. This was to ensure that all panelists felt comfortable and were ready to recommend achievement level cuts.

Following the practice round, the panelists reviewed and discussed the definitions of the achievement levels (PLDs) used in the standard setting (*Not Met 1/ Above Not Met 1*), indicating the skills that distinguished each achievement level from the others. From this discussion, the

panelists were asked to develop individual profiles for students just passing out of *Not Met 1/ Above Not Met 1*. This task was critical for the standard setting process as these definitions were used by the panelists during the process of indicating achievement level cuts.

Following the PLD review, participants worked through the Spring 2016 SC READY operational reading section of the ELA test as if they were students taking the test. During this task, the panelists could develop an understanding of the nature of the items and the range of item difficulties that will appear in the OIB.

After a break for lunch, the panelists then proceeded to Round 1 of the Bookmark Standard Setting. Panelists reviewed each item in the ordered item booklet, to select the item that marked the transition between the two achievement levels *Not Met 1* and *Above Not Met 1*

After Round 1, each table of panelists was provided the Bookmark placements and summary statistics of those placements determined by each panelist at their table. The panelists were asked to review the information reported in the table summary to be certain that this accurately reflected the OIB page they selected. Following this check, the panelists discussed their selected OIB choices at their table, providing rationales according to the definitions of their *threshold* students.

Once the panelists felt comfortable with their table discussions, they proceeded to Round 2 of standard setting, again making individual achievement level cut score recommendations as they did in Round 1, but with additional knowledge gained from the table discussion.

After Round 2, each table of panelists was provided the Bookmark placements and summary statistics of those placements specified by each panelist of that table. The panelists were asked to review the information in the table summary to be certain that this accurately reflected the OIB page they selected. Following the completion of the table discussion, each table leader presented a summary of their table's recommended cut score and the table's rationale to the entire panel. At the conclusion of this discussion, the entire panel was presented with the recommended cut (median OIB page numbers) and the distribution of students within each achievement level (*Not Met 1/ Above Not Met 1*), using the spring 2016 test population performance on the operational items as a reference. The impact data was shown for the overall test population.

Round 3 occurred in the same manner as Round 2. Panelists provided another set of individually recommended cut scores using the additional knowledge gained from the panel's discussion of the Round 2 results.

The Round 3 cut scores and resulting impact data were shown to the panel as part of the final debriefing of the standard setting meeting. The median of the table medians of the Bookmark placements, specified as cuts by the panelists, were mapped to the corresponding item difficulty estimate from the item difficulties used for the OIB. The distribution of students within each achievement level (*Not Met 1/ Above Not Met 1*) was provided, using the spring 2016 test population performance on the operational items as a reference.

These mapped theta estimates were used as the recommended cut scores provided to the South Carolina Department of Education (SCDE). The results from Round 3 are discussed in the next section.

Table 3 contains the descriptive statistics for the OIB median of the table median page numbers from each of the standard setting rounds.

Table 3. OIB Median, Minimum, and Maximum Page Number by Round

Reading			
Grade	Round	Statistic	Achievement Level Cut
			<i>Not Met 1/ Above Not Met 1</i>
3	Round 1	Median of Table Mdns	12
		Minimum (panelists)	2
		Maximum (panelists)	16
		Median of Room	11
	Round 2	Median of Tables	10
		Minimum (panelists)	5
		Maximum (panelists)	14
		Median of Room	9
	Round 3	Median of Tables	6
		Minimum (panelists)	5
		Maximum (panelists)	11
		Median of Room	6

Following the discussion of the Round 3 results, the panelists revisited the PLDs for *Not Met 1*, and *Above Not Met 1*, making recommendations for revision to the recommended performance level descriptions for approval. The revised PLDs are found in Appendix B.

Analyses

All of the operational items for the reading section of the test form were included in the OIB to allow for a thorough spread of item difficulty estimates for the impact data. The range and distribution of item difficulties was reviewed and approved by SCDE before the OIB was developed. The item parameters are found in Appendix C.

Recommended Cut Scores

The recommended theta cut score for the *Not Met 1/ Above Not Met 1* performance levels are found in Table 4, as is the impact data associated with the recommended cut scores. These panelist recommended theta cut scores were calculated by the median of the table medians at the end of Round 3.

**Table 4a. Recommended theta cut score and associated impact data
(Median of Table Medians)**

Subject / Grade	Cut Score	Associated Impact Data	
	<i>Not Met 1 / Above Not Met 1</i>	<i>Not Met 1</i>	<i>Above Not Met 1</i>
Reading Grade 3	-1.5534	5.3%	94.7%

**Table 4b. Impact Data from recommended cut score, including ELA performance levels
(Median of Table Medians)**

Subject / Grade	Achievement Levels					Total
	NM1	NM2	AE	ME	EE	
Reading Grade 3	5.3%	16.9%	34.1%	29.3%	14.4%	100.00%

Panelist Variability

Estimates of panelist variability can be used to evaluate the stability of the cut score recommendations, considering that the standard setting could be replicated using a different collection of panelists. In order to estimate and describe the variability in panelist's judgments, the method described in Lewis, et al. (2000) was implemented. The calculations were based on the page numbers selected in the second round.

The standard error was computed using the formula below. These standard error estimates were based on the median of the logit values associated with the OIB page:

$$SE_{cut} = \sqrt{\frac{\sigma_{Judges}^2}{N_{Judges}} + [1 + (n - 1)r]}.$$

When the ratings of the panelists are independent of each other ($r = 0$), the value of the SE_{cut} reduces to:

$$SE_{cut} = \sqrt{\frac{\sigma_{Judges}^2}{N_{Judges}}}.$$

It is common for policy-makers to consider the total error associated with cut scores prior to making final decisions, taking into account the uncertainty associated with the recommended cut scores. Total error in this case is conceptualized as the sum of the measurement error associated with the instrument and the error associated with the cut score procedures described above. The standard errors were adjusted by the usual 1.253 to account for the use of the median (SCDE Communication, March 28, 2016). The total error SE_{Total} was calculated as follows:

$$SE_{Total} = 1.253\sqrt{(CSEM_{Cut})^2 / N + (SE_{Cut})^2},$$

where $CSEM_{Cut}$ is the conditional standard error of measurement for the theta cut, and SE_{Cut} is the standard error of the estimated cut score computed according to Lewis et al. (2000). For this analysis, Winsteps (Linacre, 2016) was used to generate a raw-score-to-theta conversion table based on the number of items presented in each of the OIBs. From this, the standard error of measurement values for the recommended theta cut scores (theta estimates closest to, but not higher than the recommended theta cuts) were obtained for computing the total standard error with the above equation.

Table 6 summarizes the panelist recommendations from the standard setting meeting, using the median of the table medians method, including the standard error associated with each cut score.

Table 6. Standard Error Indices with Recommended Theta Cut Scores

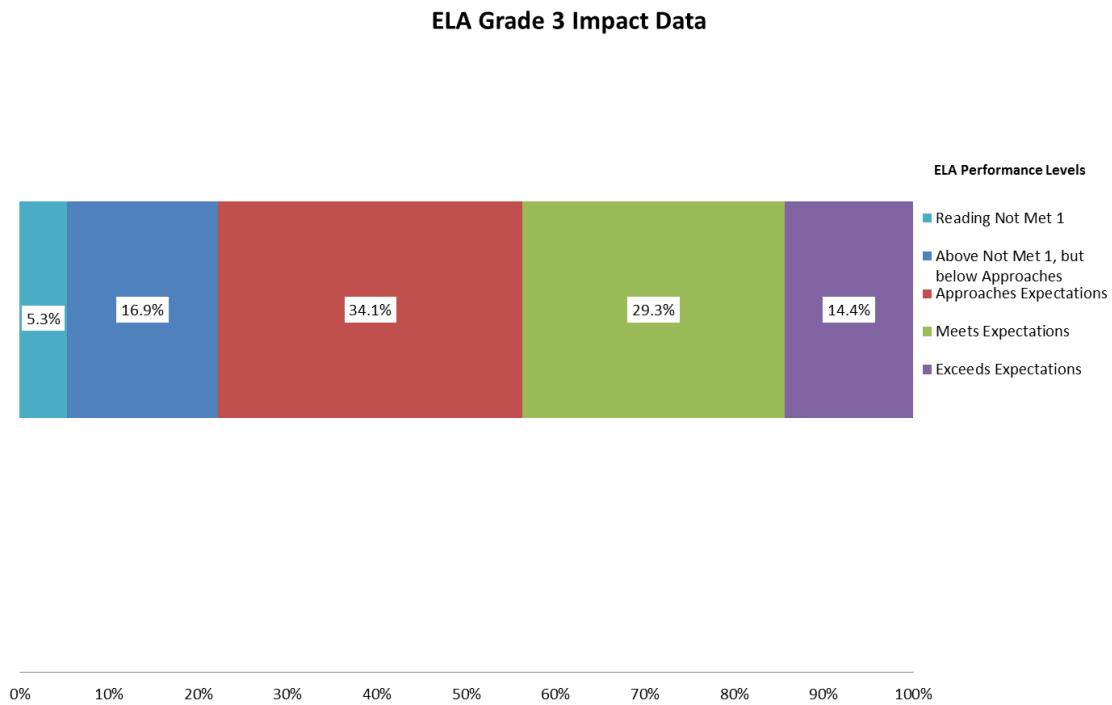
Grade / Subject	Cut	Recommended Cut Score	SE _{cut}	CSEM _{cut}	SE _{total}
Reading Grade 3	Above Not Met 1	-1.5534	0.0298	0.4307	0.1142

The proportion of students, based on the 2016 test, for the standard setting panel recommended performance levels are shown in Figure 1. Figure 1 used the median of the table medians to estimate the recommended cut scores. Figure 2 contains the same information presented using the overall ELA test performance levels.

Figure 1. Graphical presentation of proportion in each Recommended Performance Level – Median of Table Medians



Figure 2. Graphical presentation of proportion in each Recommended Performance Level – Median of the Room



A summarization of an end of standard setting evaluation filled out by the panelists can be found in Appendix D.

References

Lewis, D. M., Mitzel, H. C., Green, D. R., and Patz, R. J. (2000). *The Bookmark standard setting procedure*. Monterey, CA: CTB/ McGraw Hill.

Linacre, J. M. (2016). WINSTEPS: Rasch measurement computer program (Computer software). Beaverton, OR: Winsteps.com.

Last Modified: February 24, 2017

APPENDIX A
SC READY
Reading Grade 3
Standard Setting Agenda

Tuesday, February 14, 2017

Hotel check-in for those traveling long distances

Wednesday, February 15, 2017 – Saluda Shoals Park – River Center

(Times are approximate depending on work completion)

8:00 – 8:30	Check-in – Lobby: Continental breakfast available
8:30 – 8:45	Introduction (SCDE)
8:45 – 10:15	Bookmark training
10:30 – 11:30	PLD and standards review – Grade 3 Reading
11:30 – 12:00	Panelists take the operational test – Grade 3 Reading
12:00 – 1:15	Lunch Table leaders will meet during lunch for a review of roles and responsibilities
1:15 – 3:00	Complete Round 1 of bookmark process – Grade 3 Reading
3:15 – 4:15	Presentation of Round 1 results and table discussion

Thursday, February 16, 2017 – Saluda Shoals Park – River Center
(Times are approximate depending on work completion)

8:00 – 8:30	Check-in – Lobby: Continental breakfast available
8:30 – 9:15	Complete Round 2 of bookmark process – Grade 3 Reading
9:30 – 10:15	Presentation of Round 2 results, impact data, and room discussion
10:15 – 11:00	Complete Round 3 of bookmark process – Grade 3 Reading
11:15 – 12:00	Presentation of Round 3 results, impact data, and room discussion
12:00 – 1:15	Lunch
1:15 – 2:00	Complete Round 4 of bookmark process – Grade 3 Reading (If necessary)
2:15 – 3:00	Presentation of Round 4 results, impact data, and room discussion (If necessary)
3:15 – 3:45	Revision of Grade 3 Reading PLDs

APPENDIX B

Grade 3 Reading of the English Language Arts Assessment

PLD	Standard	Does Not Meet Expectations	
Policy		The student Does Not Meet Expectations as defined by the grade level content standards. The student <i>needs substantial academic support</i> to be prepared for the next grade level and to be on track for college and career readiness.	
Range		A student who performs at the Does Not Meet Expectations level tends to read and comprehend informational texts and literature that do not meet the demands of grade level texts that would signal this student is on track for college and career readiness and requires substantial instructional support to improve reading skills.	
		Not Met 1	Above Not Met 1
	RL.MC.5.1	Answers simple literal questions to demonstrate understanding of texts.	
	RL.MC.6.1		Identifies explicitly stated themes by recalling details.
	RL.MC.7.1	Identifies how simple illustrations generally contribute to emphasize aspects of literature.	
	RL.MC.7.2		Identifies how an author uses characters to develop theme and simple plot in different texts within a series.
	RL.MC.8.1		Uses text evidence to: <ul style="list-style-type: none"> • Identify characters' fundamental traits and feelings, and explain how their actions contribute to plot and • Identify the influence of cultural or historical context on literature.
	RL.LCS.9.1		Identifies how the author uses simple idioms, metaphor, or personification to shape meaning.
	RL.LCS.9.2		Identifies how the author's choice of simple words, illustrations, and conventions combine to contribute meaning and emphasize basic aspects of a character or setting.
	RL.LCS.10.1		Uses sentence-level, explicit context to determine the meaning

			of simple words and phrases.
	RL.LCS.10.2	Determine the meaning of some simple words when common affixes are added to base words.	
	RL.LCS.10.5	Consults print or multimedia resources to find the pronunciation or determine the meaning of common words.	
	RL.LCS.10.6	Acquires and uses some simple general academic words and phrases.	
	RL.LCS.11.1		Identifies clearly stated first or third person points of view.
	RL.LCS.11.2		Identifies the clearly stated point of view of the narrator or a character.
	RL.LCS.12.1		Identifies basic text structure of various genres using the terms paragraph, chapter, scene, and stanza.
	RL.LCS.12.2	Identifies text structures such as a simple collection of photographs or simple poetry texts, texts with a series of short, simple memoirs.	
	RI.MC.5.1	Answers simple literal questions to demonstrate understanding of texts.	
	RI.MC.6.1		Summarizes simple, multi-paragraph texts using some details.
	RI.MC.7.1		Identifies the same topic, idea, or concept in diverse texts.
	RI.8.1		Identifies when the author uses simple words and phrases to inform, explain, or describe.
	RI.LCS.8.2	Uses knowledge of simple appendices, timelines, maps, and charts to locate information;	Identifies how these features (appendices, timelines, maps, and charts) generally contribute to a text.
	RI.LCS.9.1		Uses sentence-level, explicit context to determine the meaning of simple words and phrases.
	RI.LCS.9.2	Determine the meaning of some simple words when common affixes are added to base words.	
	RI.LCS.9.4	Consults print or multimedia resources to find the pronunciation or determine the meaning of simple	

		words or phrases.	
	RI.LCS.9.5	Acquires and uses some simple general academic words and phrases.	
	RI.LCS.10.1		Identifies clearly stated author's purpose and/or perspective.
	RI.LCS.11.1		Identifies simple problem and solution, description, and question and answer structures to locate information.
	RI.LCS.11.2		Identifies simple structures an author uses to support specific points.

APPENDIX C
Ordered Item Book Maps

SC READY Grade 3 Reading Ordered Item Map

Item Map Order	Item Parameter
1	-1.9355
2	-1.7794
3	-1.6904
4	-1.6668
5	-1.5905
6	-1.5534
7	-1.3913
8	-1.3312
9	-1.2662
10	-1.2471
11	-1.2385
12	-1.2365
13	-1.2013
14	-1.1641
15	-1.1417
16	-1.1335
17	-0.8171
18	-0.7426
19	-0.7071
20	-0.6942
21	-0.5547
22	-0.5111
23	-0.4413
24	-0.3773
25	-0.3603
26	-0.238
27	-0.1847
28	-0.1523
29	-0.1244
30	0.0309
31	0.0424
32	0.0448
33	0.2519
34	0.2634
35	0.3863
36	0.4561
37	0.5748
38	0.5979
39	0.6435
40	0.6781
41	0.6924
42	0.7588

NM1/ANM1 Cut

Item Map Order	Item Parameter
43	0.9018
44	0.9004
45	0.9989
46	1.3393
47	1.5831
48	1.8921
49	3.3217

APPENDIX D

Standard Setting Evaluation Results

1. Please rate the extent of your agreement with each statement regarding the opening session:

	Valid N	Percent Selecting Category			
		Strongly Disagree	Disagree	Agree	Strongly Agree
a. The training provided a clear description of the meeting’s goals.	25	4%	0%	4%	92%
b. The training session leader clearly explained the procedures.	25	4%	0%	4%	92%
c. The training session leader clearly explained the materials used in the bookmark process.	25	4%	0%	4%	92%
d. The training helped me understand my tasks.	25	4%	0%	4%	92%
e. The training addressed many of my questions and concerns.	25	4%	0%	12%	84%
f. The training materials were effective in preparing for subsequent tasks.	25	4%	0%	0%	96%
g. The practice exercises were useful.	25	4%	0%	0%	96%

2. Please rate the extent of your agreement with each statement regarding the performance level descriptions (PLDs):

	Valid N	Percent Selecting Category			
		Strongly Disagree	Disagree	Agree	Strongly Agree
a. Adequate information was provided to panelists regarding the PLDs.	25	4%	0%	20%	76%
b. Adequate time was provided for panelists to gain understanding of the PLDs.	25	4%	0%	12%	84%
c. The PLDs communicate a reasonable profile of students’ achievement at each level.	25	4%	0%	4%	92%

3. Please indicate your opinion regarding the usefulness of the following materials used:

	Valid N	Percent Selecting Category			
		Not Useful	Somewhat Useful	Useful	Very Useful
a. Performance level descriptions	25	0%	0%	16%	84%
b. Operational test book	25	0%	0%	4%	96%
c. Ordered item booklet	25	0%	0%	4%	96%
d. Item separation chart	25	0%	0%	0%	100%
e. Item information sheet	25	0%	0%	0%	100%
f. Student score distribution	25	0%	8%	0%	92%
g. Statistical impact data	25	0%	0%	4%	96%

4. Please indicate the extent of your satisfaction with the following roles:

	Valid N	Percent Selecting Category			
		Very Satisfied	Satisfied	Partially Satisfied	Not Satisfied
a. DRC psychometric lead	25	100%	0%	0%	0%
b. DRC content specialist	25	100%	0%	0%	0%
c. Other DRC Staff	25	100%	0%	0%	0%

5. Please indicate your opinion regarding the amount of time allotted for each activity:

	Valid N	Percent Selecting Category		
		Too Little Time	About Right	Too Much Time
a. Training	25	0%	96%	4%
b. PLD discussion	25	4%	96%	0%
c. Round 1 ratings	25	0%	96%	4%
d. Round 1 discussion	25	0%	100%	0%
e. Round 2 ratings	25	0%	100%	0%
f. Round 2 discussion	25	4%	96%	0%
g. Round 3 ratings	24	0%	100%	0%
h. Round 3 discussion	24	0%	100%	0%

6. Please indicate the level of confidence you had in placing the bookmark location for each assessment cut score:

Grade		Valid N	Percent Selecting Category			
			Not Confident	Partially Confident	Confident	Very Confident
3	a. Not Met 1 / Above Not Met 1	25	0%	0%	40%	60%

7. Please rate the extent of your agreement with each statement regarding the processes and results:

	Valid <i>N</i>	Percent Selecting Category			
		Strongly Disagree	Disagree	Agree	Strongly Agree
a. The processes and methods used will produce appropriate results.	25	0%	0%	28%	72%
b. My bookmark placements accurately represent the PLDs.	25	0%	0%	24%	76%