



Smarter Balanced Assessment Consortium: Achievement Level Setting Final Report

January 7, 2015



Contents

Executive Summary	5
Overview of Tasks and Processes.....	5
Online Panel Results.....	5
In-Person Workshop Results	7
Results of Cross-Grade Review	10
Impact by Subgroup.....	12
Action by Chiefs.....	16
Section 1. Introduction	18
Smarter Balanced Background	18
Purpose and Organization of this Report	19
Section 2. Communications Plan	20
Purpose and goals	20
Targets.....	20
Research and Planning.....	21
Environmental scan.....	22
Focus group study.....	23
Recruitment and Outreach	23
Social media campaign/paid promotion.....	24
Media Strategy/Communications Planning	24
Communications planning for key activities and state approval.....	24
Contingency planning	25
Intake interviews: Development of contingency plan, messaging and protocols.....	25
Ebola contingency response.....	26
Digital and Creative Support	26
Expansion of video production scope.....	27
Achievement Level Naming Research	27
Interviews with assessment experts.....	28
Focus groups among teachers.....	29
Online surveys among parents and students.....	29
Section 3. Logistics	31
Schedule.....	31
Facilities.....	32
Hardware and Software Support.....	33
Meetings.....	33

In-person workshop.....	33
Cross-grade review.....	36
Technical Advisory Committee Meetings.....	36
Collaboration Conferences.....	36
Advisory Panel Meetings.....	37
National Conference on Student Assessment.....	37
Other Meetings.....	37
Security.....	38
Section 4. Audit.....	40
Purpose of the Audits.....	40
Selection of the Auditors.....	40
Major Audit Functions.....	41
Review of tools and materials.....	41
On-site review of procedures.....	41
Review of documentation.....	42
Legal review.....	42
Section 5. Panelist Recruitment.....	43
Online Panel for Achievement Level Setting.....	43
Implementation.....	43
Findings and Refinements.....	46
In-Person Panel for Achievement Level Setting.....	46
Implementation.....	47
Selection Process.....	48
Findings and Refinements.....	50
Cross-Grade Review Panel.....	54
Implementation.....	54
Findings and Refinements.....	55
Section 6. Standard Setting Design and Implementation.....	57
The Bookmark Procedure.....	57
Creating ordered item booklets.....	57
Item mapping.....	59
Typical application of the bookmark procedure.....	59
Design and Implementation of the Online Panel.....	60
Software development.....	61
The home page.....	61
Online panel activities.....	70

Conduct and results of the online panel.....	70
Design and Implementation of the In-Person Workshop.....	71
Recruitment and selection of panelists.....	72
Preparation of materials.....	72
Training of facilitators and table leaders.....	73
Orientation and training.....	73
Round-by-round item review and discussion	74
Data analysis and reporting	78
Design and Implementation of the Cross-Grade Review Panel	78
Follow-Up Activities	80
Actions by Chiefs.....	81
Review of ALDs.....	82
Section 7. Field Test.....	84
Field Testing the Online Tool	84
Preparations for field testing.....	84
Implementation.....	85
Feedback.....	86
In-Person Workshop.....	89
Software support.....	90
Vertical Articulation.....	91
Debriefing Session with Panelists.....	91
Software issues.....	92
Presentation/preparatory session usefulness.....	92
Global observations	92
Debriefing Session with Smarter Balanced Staff.....	93
Lessons Learned – Preparations for October	94
Section 8. Management Activities.....	95
Meetings.....	96
Weekly calls.....	96
Twice-monthly calls.....	96
Collaborative Calls and Meetings	96
Technical Advisory Committee	96
Advisory Panel	96
Other Meetings.....	97
Reports	97
References	98

Appendix A – Communication Plan and Related Documents	99
Appendix B – Logistics Documentation.....	312
Appendix C – Audit Report and Related Documentation	488
Appendix D – Recruitment Documentation.....	542
Appendix E – Online Panel Documentation	610
Appendix F – In-Person Workshop Documentation.....	657
Appendix G – Cross-Grade Review Panel Documentation	925
Appendix H – Follow-Up Activities	954
Appendix I – Field Test Documentation	1010

Executive Summary

Achievement level setting for Smarter Balanced assessments in Mathematics and English/Language Arts and Literacy occurred in three phases: online panel, in-person workshop, and cross-grade review. This report documents each of the three phases and provides results and recommendations for cut scores for Smarter Balanced Assessments to be given in the spring of 2015.

Overview of Tasks and Processes

The purpose of the three activities was to derive cut scores using a subset of 2014 field test assessment data and based on a transparent, scientifically rigorous, well documented procedure that would stand up to external scrutiny. For this purpose, Smarter Balanced contracted with Measurement Incorporated (MI) to devise and carry out such a plan; MI worked with McGraw-Hill Education CTB to develop ALS design details. MI presented that plan at the Collaborative Meeting on April 29, 2014, and it was unanimously approved. Essentials of the plan included the following:

- Recruitment of thousands of individuals in Smarter Balanced states to participate online
- Recruitment of 504 individuals in Smarter Balanced states to participate in an in-person workshop
- Recruitment of 64 individuals from among the 504 to participate in cross-grade review
- Use of the Bookmark Procedure to guide participants through the achievement level setting process
- Development of software to deliver assessments and achievement level setting tools to online and in-person participants

Measurement Incorporated and McGraw-Hill Education CTB staff carried out all tasks under the supervision of Smarter Balanced, conducting a pilot test of all activities August 14–20, the online panel October 6–17, the in-person workshop October 13–19, and the cross-grade review on October 20, 2014. Results of the three October activities are provided below.

Online Panel Results

Over 10,000 individuals registered for the online panel and 2,660 participated by entering one cut score—for Level 3—for one test. Participation by state is summarized in Table ES1, while participation by role is summarized in Table ES2. Recommended cut scores are summarized in Table ES3. Additional detail for participation as well as cut scores and impact is provided in the body of the report.

Table ES1. Online Panel Participation by State.

State	Online Participants	State	Online Participants	State	Online Participants
California	459	Missouri	393	South Dakota	68
Connecticut	121	Montana	43	Virgin Islands	1
Delaware	48	North Carolina	10	Vermont	23
Hawaii	57	North Dakota	57	Washington	402
Iowa	15	New Hampshire	48	Wisconsin	132
Idaho	89	Nevada	40	West Virginia	116
Maine	75	Oregon	314	Wyoming	5
Michigan	69	Pennsylvania	4	Other	71
				Total	2,660

Table ES2. Online Panel Participation by Role.

Role	Online Participants	Role	Online Participants
Teachers	1,730	Higher Ed. Administrator	18
Non-Teacher Educators	209	Parents	65
Administrators	272	Business	14
Higher Ed. Faculty	187	Other	165
		Total	2,660

Table ES3. Online Panel Cut Score Recommendations for Level 3

Test	Cut Score	% At or Above
Grade 3 English Language Arts/Literacy	411	46.8
Grade 3 Mathematics	409	53.7
Grade 4 English Language Arts/Literacy	468	42.9
Grade 4 Mathematics	474	42.7
Grade 5 English Language Arts/Literacy	453	62.9
Grade 5 Mathematics	520	36.5
Grade 6 English Language Arts/Literacy	516	46.9
Grade 6 Mathematics	540	37.8
Grade 7 English Language Arts/Literacy	510	54.9
Grade 7 Mathematics	584	27.0
Grade 8 English Language Arts/Literacy	551	47.8
Grade 8 Mathematics	632	17.7
Grade 11 English Language Arts/Literacy	544	54.9
Grade 11 Mathematics	650	22.0

In-Person Workshop Results

The in-person workshop took place in Dallas October 13–17, 2014, in three waves: Grade 11 on October 13–15, grades 6–8 on October 15–17, and grades 3–5 on October 17–19. There were 16 panels in all – one for each subject/grade combination in grades 3–8 and two each for grade 11 English language arts/literacy and mathematics. Dr. Michael Bunch (MI) oriented participants to the

Common Core State Standards and Smarter Balanced Achievement Level Descriptors. Participants then took the Smarter Balanced online Practice Tests before receiving instruction from Dr. Daniel Lewis (CTB) in the Bookmark Procedure. Contractor staff led Bookmark orientation exercises, and then led participants through three rounds of review of ordered item booklets and discussion, during which participants also reviewed results of the online panel. Table ES4 summarizes participation by state, while Table ES5 summarizes participation by role. Results of the final round of cut score recommendations (in scale score values) are summarized in Table ES6.

Table ES4. In-Person Workshop Participation by State.

State	In-Person Participants	State	In-Person Participants	State	In-Person Participants
California	66	Missouri	14	South Dakota	10
Connecticut	26	Montana	23	Virgin Islands	18
Delaware	24	North Carolina	8	Vermont	10
Hawaii	22	North Dakota	20	Washington	32
Iowa	17	New Hampshire	20	Wisconsin	28
Idaho	20	Nevada	15	West Virginia	31
Maine	29	Oregon	21	Wyoming	14
Michigan	14	Pennsylvania	0		
				Total	482

Table ES5. In-Person Workshop Participation by Role.

Role	In-Person Participants	Role	In-Person Participants
Educator: ELL Experience	36	Educator: SWD Experience	31
Educator: General	242	General Public	28
Educator: Non-Teaching	89	Higher Education	56
		Total	482

Table ES6. In-Person Workshop Cut Score Recommendations: Final Round

Test	Level 2 Cut	% At or Above	Level 3 Cut	% At or Above	Level 4 Cut	% At or Above
Grade 3 English Language Arts/Literacy	362	66.5	427	40.1	485	19.1
Grade 3 Mathematics	383	67.3	436	38.9	506	10.8
Grade 4 English Language Arts/Literacy	413	64.4	470	42.0	530	18.9
Grade 4 Mathematics	400	77.6	470	44.7	541	15.6
Grade 5 English Language Arts/Literacy	406	78.6	450	63.9	574	16.9
Grade 5 Mathematics	459	63.5	532	31.4	583	13.8
Grade 6 English Language Arts/Literacy	467	66.2	527	42.2	614	12.2
Grade 6 Mathematics	491	58.3	563	28.7	604	15.3
Grade 7 English Language Arts/Literacy	474	68.3	547	40.1	660	6.6
Grade 7 Mathematics	513	53.2	609	19.4	674	5.8
Grade 8 English Language Arts/Literacy	471	76.4	547	49.4	663	10.2
Grade 8 Mathematics	534	51.3	605	25.6	683	7.4
Grade 11 English Language Arts/Literacy	490	72.7	565	47.4	677	11.6
Grade 11 Mathematics	533	59.0	644	23.6	740	5.8

Results of Cross-Grade Review

On October 20, 2014, 64 pre-selected members of the in-person workshop panels met to review cut scores and impact across grades. Dr. Bunch (MI) led a joint introduction to the review procedure for E/la and mathematics cross-grade review panelists. Following the review, the two content area panels separated to conduct independent review sessions; Dr. Bunch (MI) facilitated the ELA/literacy cross-grade review and Dr. Lewis (CTB) facilitated the mathematics cross-grade review. Both content areas worked with the same materials and followed the same general procedures. Panelists in each content area considered all cut scores and related impact for grades 8 and 11. Afterwards, they followed the same procedure to consider cut scores across grades 3–7. Panelists made a total of 20 changes, 8 in ELA and 12 in Mathematics. Results of the cross-grade review are summarized in Table ES7, which is a reproduction of Table ES6 with changes highlighted. Impact (percent of students at or above each cut score) is summarized in Figures ES1 and ES2.

Table ES7. Cross-Grade Review Results

Test	Level 2 Cut	% At or Above	Level 3 Cut	% At or Above	Level 4 Cut	% At or Above
Grade 3 English Language Arts/Literacy	362	66.5	427	40.1	485	19.1
Grade 3 Mathematics	381	68.3	436	38.9	501	12.1
Grade 4 English Language Arts/Literacy	413	64.4	470	42.0	530	18.9
Grade 4 Mathematics	413	72.3	487	36.5	551	12.6
Grade 5 English Language Arts/Literacy	434	69.7	494	47.1	574	16.9
Grade 5 Mathematics	459	63.5	532	31.4	583	26.1
Grade 6 English Language Arts/Literacy	453	71.3	527	42.2	610	13.1
Grade 6 Mathematics	491	58.3	570	26.1	609	14.0
Grade 7 English Language Arts/Literacy	474	68.2	547	40.1	644	9.5
Grade 7 Mathematics	513	53.1	596	23.2	674	5.8

Test	Level 2 Cut	% At or Above	Level 3 Cut	% At or Above	Level 4 Cut	% At or Above
Grade 8 English Language Arts/Literacy	482	73.1	562	43.3	663	10.2
Grade 8 Mathematics	534	51.3	616	22.1	683	7.4
Grade 11 English Language Arts/Literacy	488	73.3	578	42.6	677	11.6
Grade 11 Mathematics	565	48.3	650	22.0	740	5.8

Figure ES1. Percent of Students At or Above Level 3: English Language Arts/Literacy

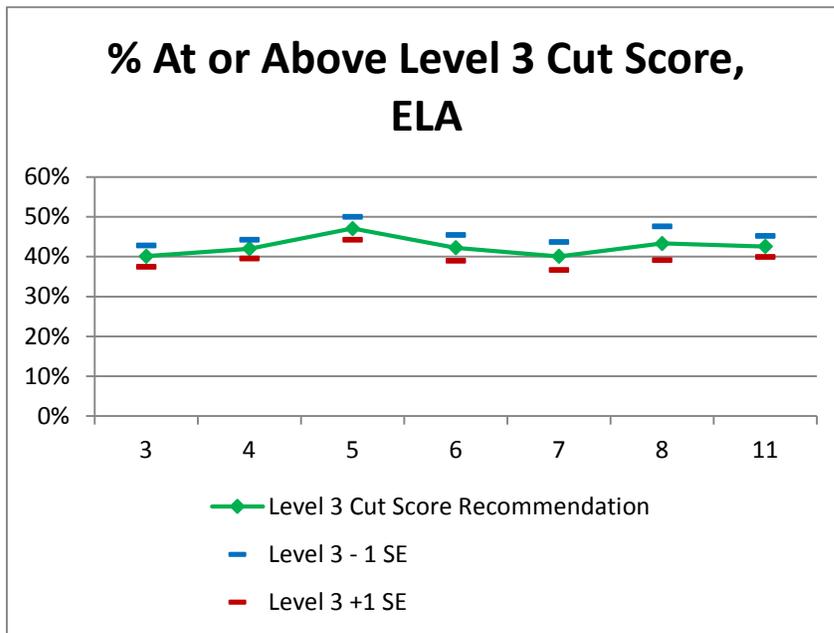
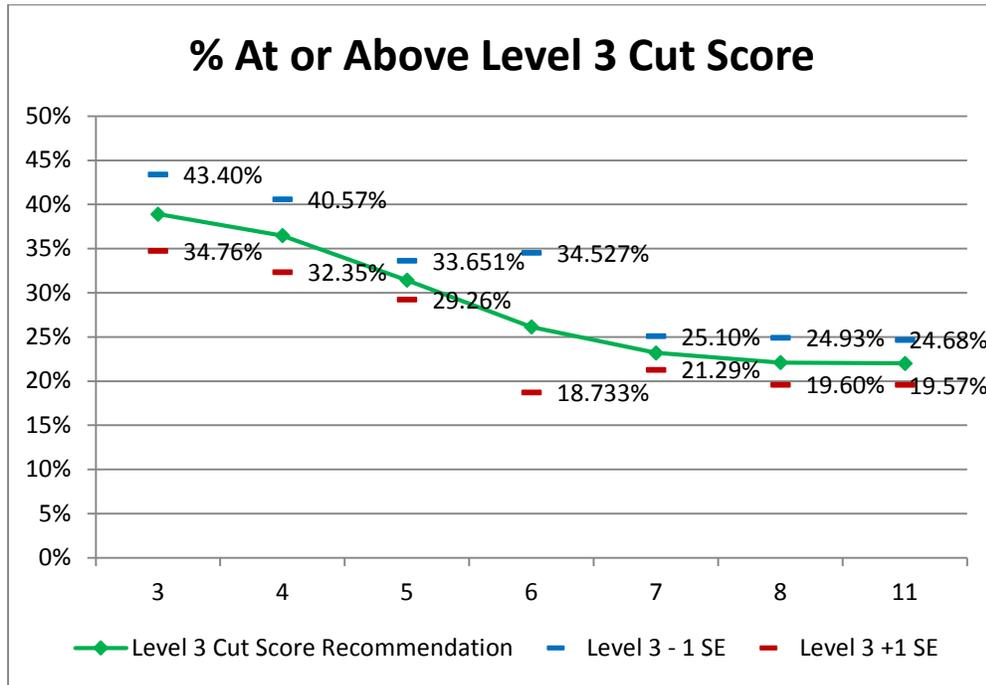


Figure ES2. Percent of Students At or Above Level 3: Mathematics



Figures ES1 and ES2 show that panelists' final recommendations had some degree of flexibility, based on the sampling error associated with panelists' bookmark placements. That is, rather than represent each cut score (and percent of students at or above that cut score) as a point, these figures present cuts and resulting impacts as ranges. Thus, for example, for third grade mathematics, the final cut score would lead to approximately 63 percent of students reaching Level 3 or above, give or take about three percent.

We note that the sampling error presented is an estimate of the variability of cut score recommendations had different samples of panelists been selected to participate in the ALS. The SE is an estimate, based on Round 2 results, which are utilized because independence of judgments across small groups (or tables) within each grade and content area are fostered through Round 2. Round 3 judgments are not independent due to room-level discussion. Thus, the Round 2 SE estimates acknowledge, and provide an estimate of the variability that is expected due to the sampling of panelists.

Impact by Subgroup

Results in Table ES7 are for the total population of students who took Smarter Balanced field tests in the spring of 2014. Data are also available showing the impact of these cut scores for specific categories of students. Tables ES8 through ES13 show the impact of these cut scores by demographic group and program designation.

Table ES8. ELA: Impact of Level 3 Recommendations, by Gender.

Grade	All	Female	Male
3	40.1%	43.5%	36.9%
4	47.1%	52.6%	41.8%
5	40.1%	47.1%	33.3%
6	42.6%	49.2%	36.0%
7	42.0%	47.2%	37.1%
8	42.2%	47.6%	37.1%
11	43.3%	50.8%	36.2%

Table ES9. Math: Impact of Level 3 Recommendations, by Gender

Grade	All	Female	Male
3	38.9%	38.5%	39.3%
4	44.7%	43.3%	46.0%
5	31.4%	30.0%	32.8%
6	29.4%	28.4%	30.3%
7	19.3%	18.6%	20.0%
8	25.6%	25.1%	26.0%
11	23.6%	22.7%	24.5%

Table ES10. ELA: Impact of Level 3 Recommendations, by Race/Ethnicity

Grade	All	Asian	Black	Hispanic	Native American	Pacific Islander/ Alaska Native	White	Multi-Ethnic
3	40.1%	57.7%	23.1%	25.7%	24.3%	28.9%	49.8%	42.3%
4	42.0%	58.5%	25.2%	26.3%	19.4%	27.5%	49.6%	43.8%
5	47.1%	63.0%	30.2%	31.0%	24.7%	34.0%	52.7%	47.8%
6	42.2%	61.4%	26.0%	27.2%	24.4%	27.8%	49.7%	42.9%
7	40.1%	62.9%	24.6%	27.1%	24.7%	27.6%	49.8%	50.2%
8	43.3%	59.1%	24.7%	30.6%	25.3%	28.5%	50.5%	49.0%
11	42.6%	58.2%	23.9%	33.8%	29.1%	28.8%	47.0%	41.8%

Table ES11. Math: Impact of Level 3 Recommendations, by Race/Ethnicity

Grade	All	Asian	Black	Hispanic	Native American	Pacific Islander/ Alaska Native	White	Multi-Ethnic
3	40.1%	60.7%	19.8%	24.1%	21.2%	24.0%	44.2%	41.6%
4	42.0%	66.0%	19.7%	26.8%	25.1%	33.5%	50.4%	46.2%
5	47.1%	54.7%	11.9%	15.7%	14.5%	20.3%	34.6%	32.8%
6	42.2%	52.7%	12.9%	16.3%	15.4%	22.3%	32.8%	30.9%
7	40.1%	41.5%	6.9%	8.2%	8.3%	11.8%	25.1%	22.5%
8	43.3%	52.1%	11.8%	13.4%	12.9%	15.7%	32.0%	29.5%
11	42.6%	49.6%	10.2%	12.5%	13.0%	10.5%	27.3%	23.1%

Table ES12. ELA: Impact of Level 3 Recommendations, by Special Needs

Grade	All	SWD	ELL	Section 504	Low SES
3	40.1%	16.7%	14.9%	29.0%	26.4%
4	42.0%	17.1%	10.5%	36.2%	27.7%
5	47.1%	14.8%	10.2%	45.8%	31.9%
6	42.2%	10.8%	5.3%	33.4%	28.5%
7	40.1%	8.0%	4.3%	40.5%	27.9%
8	43.3%	9.3%	5.6%	44.8%	30.6%
11	42.6%	9.5%	5.8%	37.4%	33.8%

Table ES13. Math: Impact of Level 3 Recommendations, by Special Needs

Grade	All	SWD	ELL	Section 504	Low SES
3	40.1%	18.4%	16.5%	38.3%	26.0%
4	42.0%	16.4%	15.3%	48.5%	28.5%
5	47.1%	9.1%	5.6%	28.5%	17.4%
6	42.2%	8.4%	4.2%	26.8%	16.7%
7	40.1%	4.0%	2.7%	22.3%	9.6%
8	43.3%	5.5%	3.9%	24.9%	14.6%
11	42.6%	4.0%	3.3%	21.3%	14.3%

Action by Chiefs

Smarter Balanced staff prepared options to present to the Chiefs on November 14, 2014, incorporating evidence from recent studies conducted by the National Assessment Governing Board (NAGB). In addition, Smarter Balanced staff created a new reporting scale, replacing the temporary scale used throughout achievement level setting and cross-grade review. They also recommended ELA cut scores that were somewhat higher and mathematics cut scores that were somewhat lower than those recommended by the cross-grade review committee. These modifications kept recommended cut scores within or very close to the one SEM range, approximated NAGB results, and brought ELA and mathematics impacts into closer alignment with each other. The Chiefs voted unanimously (with two abstentions) on November 14 to approve the modified cut scores, presented in Table ES14.

Table ES14. Final Cut Scores Approved by Chiefs, With Impact Data.

Test	Level 2 Cut	% At or Above	Level 3 Cut	% At or Above	Level 4 Cut	% At or Above
Grade 3 English Language Arts/Literacy	2367	65%	2432	38%	2490	18%
Grade 3 Mathematics	2381	68%	2436	39%	2501	12%
Grade 4 English Language Arts/Literacy	2416	63%	2473	41%	2533	18%
Grade 4 Mathematics	2411	73%	2485	37%	2549	13%
Grade 5 English Language Arts/Literacy	2442	67%	2502	44%	2582	15%
Grade 5 Mathematics	2455	65%	2528	33%	2579	15%
Grade 6 English Language Arts/Literacy	2457	70%	2531	41%	2618	11%
Grade 6 Mathematics	2473	65%	2552	33%	2610	14%
Grade 7 English Language Arts/Literacy	2479	66%	2552	38%	2649	8%
Grade 7 Mathematics	2484	64%	2567	33%	2635	13%

Test	Level 2 Cut	% At or Above	Level 3 Cut	% At or Above	Level 4 Cut	% At or Above
Grade 8 English Language Arts/Literacy	2487	72%	2567	41%	2668	9%
Grade 8 Mathematics	2504	62%	2586	32%	2653	13%
Grade 11 English Language Arts/Literacy	2493	72%	2583	41%	2682	11%
Grade 11 Mathematics	2543	60%	2628	33%	2718	11%

Section 1. Introduction

In a request for proposals (RFP) issued in October 2013, Smarter Balanced called for a contractor to provide services for a multi-phase standard setting process and to plan and execute a comprehensive Communication Plan. The standard setting plan was to be executed in five phases from conducting distributed standard setting to finalizing achievement level descriptors. The communication plan was to “proactively explain the rationale for setting common achievement standards tied to the Common Core State Standards, describe the standard-setting process in layman’s terms, and make the case for approval of the performance standards derived from the standard setting process.” (RFP, p. 10).

Smarter Balanced Background

The Smarter Balanced Assessment Consortium (Smarter Balanced) was established in 2010 under a grant from the U.S. Department of Education, authorized by the American Recovery and Reinvestment Act of 2009 (ARRA). The purpose of Smarter Balanced has been to create rigorous assessments in English language arts/literacy and mathematics for grades 3-8 and high school, aligned with the Common Core State Standards.

Prior to the awarding of a contract for standard setting (henceforth referred to as achievement level setting), Smarter Balanced had awarded several other contracts for program management, test development, and development of achievement level descriptors (ALDs). There are, or will be, four sets of ALDs (from Egan, Schneider, & Ferrara, 2012):

- Policy – brief statements that articulate policy makers’ vision of goals and rigor for the final performance standards;
- Range – guidelines created by test developers to identify which aspects of items align to a particular performance level in regard to the cognitive and content rigor that has been defined;
- Threshold (Target) – detailed statements created in conjunction with the Range ALDs and are used by achievement level setting panelists to represent the knowledge and skills of a student just at the threshold of a given level;
- Reporting – relative brief statements developed by a sponsoring agency once cut scores are finalized, to define the appropriate and intended interpretations of test scores.

The first three sets have been developed; the Reporting ALDs are under review as of this writing.

Achievement level setting is the culminating set of activities in a four-year enterprise to create, field test, and implement a set of rigorous computer-adaptive assessments closely aligned to the Common Core and to provide guidance to educators regarding the achievements of their students, with particular reference to college and career readiness.

As Smarter Balanced shifts from a developmental to an operational mode, additional research on cut scores is planned. The final task under this contract is to prepare a long-range research agenda that will allow Smarter Balanced to test the validity of the cut scores against various external criteria.

Purpose and Organization of this Report

This report documents work performed between February and December 2014 in fulfillment of Smarter Balanced Contract 21: Standard Setting and Communication Services. In its proposal, dated December 18, 2013, Measurement Incorporated spelled out seven task clusters, listed below:

1. Communication Plan - create a communication strategy that will garner the approval of political and education leaders of the new achievement levels of the Smarter Balanced assessments and widespread public acceptance of the Common Core State Standards (CCSS). These tasks are addressed in Chapter 2.
2. Standard Setting Logistics (with standard setting henceforth referred to as achievement level setting) - implement the logistics for the kick-off meeting, in-person standard setting for the operational event and the field test, and vertical articulation committee (VAC) as well as other meetings, both in-person and online. These tasks are addressed in Chapter 3.
3. Audit of Standard Setting Processes - provide Smarter Balanced with a selection of candidates for Auditor, and propose a method for maintaining a firewall between the auditor and the contractor's standard setting team. These tasks are addressed in Chapter 4.
4. Panelist Recruitment - conduct panelist recruitment for the DSS and the in-person standard setting (field test and actual), and the Cross-Grade Review Panel. These tasks are addressed in Chapter 5.
5. Standard Setting Design and Implementation - design and carry out a distributed achievement level setting activity (online) as well as an in-person activity and a vertical articulation activity. These activities are addressed in Chapter 6.
6. Standard Setting Field Test Methods - field test key components of the distributed, in-person, and vertical articulation activities, report findings, and modify procedures and materials in accordance with findings. These tasks are addressed in Chapter 7.
7. Project Management and Records of Decision Making - provide program plan documentation, human resource documentation, communication plan documentation, distributed standard setting methodology process and documentation, in-person standard setting methodology process and documentation, field test methodology process and documentation, program outcomes, and technical reports; the work plan also incorporates standard program reporting including regular program evaluation reports, status reports, schedule and progress reports, dispensation reports, and invoicing. These tasks are addressed in Chapter 8.

Work to be performed under each task cluster was further broken down into discrete tasks and described in some detail in the proposal. The purpose of this report is to describe how each of those tasks was actually completed, account for any deviations from the proposal, and explain the results. The sections of this report correspond to the seven task clusters enumerated above. In addition, there are four appendices:

- A. Participants
- B. Software
- C. Forms and Training Materials
- D. Analyses and results

Section 2. Communications Plan

Measurement Incorporated partnered with Hager Sharp, Inc., a full-service communications firm with experience serving education clients and working on issues in assessment. Working with members of the Contract 21 team (Measurement Incorporated, CTB, WestEd and Smarter Balanced Staff, Hager Sharp developed a communications plan, which proposed an overarching communications strategy for achievement levels and specific tactics for two key phases.

Phase 1 focused on building awareness about the achievement level setting process with each of our target audiences. Following the launch of recruitment for the Online Panel, we began a series of focused outreach activities. The goal of this effort was to create targeted awareness through a proactive media strategy delivering to major news outlets and beginning the process of educating members of the media, key stakeholders, and allies about the achievement level setting process.

Phase 2 focused on synchronizing messaging across the Consortium and explaining the process to key audiences. Beginning with recruitment and through the final vote on achievement level score recommendations, messaging was disseminated to help prepare states and communities for anticipated drops in the share of students performing at grade-level or above. The plan provided recommendations for identifying the needs of a range of audiences (i.e. educators, higher education faculty and leaders, business and community leaders, policymakers, general public) and researching key messages for primary stakeholders groups.

Purpose and goals

This plan identified communications challenges confronting Smarter Balanced, specifically as it related to developing achievement level scores in the fall, and proposed an overarching strategy for addressing those challenges. Through a mix of communications tactics, we helped the Consortium navigate confusion and potential sensitivity about the achievement level setting process and the rollout of the new assessments. We proposed a communications plan comprised of two phases:

Targets

Phase 1:

- Gain buy-in and support for the achievement levels in advance of launch through the unique, broadly inclusive and scientific process.
- Generate interest in, and support for, the assessment beyond the K-12 education community by engaging parents, the business community, the higher education community, and others interested in achievement levels.
- Assist Smarter Balanced states in driving recruitment for the Online Panel for Achievement Level Setting taking place October 2014.
- Ensure diversity and inclusiveness of participants; achieve balance of participants per state and per category (e.g., teachers, parents, school administrators, business community, higher education community, education reporters).
- Mitigate risk of “hostile takeover” or biasing of the Online Panel by opponents.
- Educate reporters and constituencies on the Online Panel, ensuring transparency and openness.

Phase 2:

- Develop compelling messages on the performance standards to present to K-12 and higher education decision makers (e.g., Chief State School Officers and State Higher Education Executive Officers) in advance of states voting to approve the performance standards.
- Identify and deliver compelling messages on the Common Core State Standards, the Smarter Balanced Assessment System, the standard setting design, and the performance standards to key Smarter Balanced constituents.
- Respond to contingency situations arising at the Online Panel and In-Person Panel.
- Respond to potential crises as a result of new achievement level recommendations and state consideration of those recommendations.
- Promulgate info on the new performance standards to state audiences.

Summary of Major Activities

Research and Planning

We spent the first period of our engagement with Smarter Balanced and Measurement Incorporated learning about the technical process for setting achievement levels. For the benefit of reviewers, we recapped discussion points from our SWOT (strengths, weaknesses, opportunities, and threats) analysis and included findings from an informal parent focus group.

In collaboration with the Consortium staff, Hager Sharp analyzed the relevant internal strengths and weaknesses as well as external opportunities and threats facing the Consortium. In close coordination with Measurement Incorporated, the Smarter Balanced team, and other partners, a series of startup meetings was held with the team, allowing extensive inquiry into communications needs for the standard-setting work. Hager Sharp worked with Smarter Balanced to evaluate current communications materials, look to the challenges ahead, and determine specific action steps and tasks that would contribute to the success of the project. Considerable work had already gone into developing messages and materials for Smarter Balanced; to leverage that existing work, Hager Sharp suggested revisions to existing documents and developed new materials to address needs and assist with recruitment efforts.

This information helped pinpoint how Smarter Balanced can achieve its operational goals in developing cut scores for each assessment, while also communicating with clarity about the new assessment and achievement levels. The table below illustrates some of the considerations identified in March 2014 as contributing factors for the communications plan.

Strengths	Weaknesses
<ul style="list-style-type: none"> • Inclusive process • Tried and true • Research-based • Transparent • Built on what we've learned over years • K-12 and higher education leaders join forces • Accessible by more students than ever before • 508-compliant, language aides • Built by states, for states 	<ul style="list-style-type: none"> • State affiliates vary in communications ability and willingness • Political savviness varies across states • Testing time requires significant commitment • Six hours for elementary; eight hours for high school • Perceived loss of local control • Federal funding = perception of federal involvement • Reliance on local technology, which varies by location • Broadband issues • Technical capabilities of younger students—can they

<ul style="list-style-type: none"> • Public sector halo—not commercial • Broad coalition • Test is comparable in time length to some of the longer tests currently in states • Not just a test; it’s an assessment system • Cost savings in the aggregate • Consistency across states; bigger data set • “Adaptive” test—challenging for most 	<p>reliably/accurately complete online assessment?</p>
<ul style="list-style-type: none"> • Leverage higher education experts as thought leaders • Develop playbook for states, plus collateral materials to strengthen communications process and make it more consistent • Educate staff on education committees in state legislatures • Give states good arguments against commercial rivals • Research comparing to ACT may demonstrate Smarter Balanced assessment is superior • Conduct “gut check” commitments with states • “What if” scenarios: need to make sure audiences understand what could happen with “score cliff,” etc. • Use Kentucky model for seeding environment • Showcase testimonials from participants • Connect with trades community, trade unions, etc. to generate support from non-college community • Connect with minority organizations to generate support (e.g., La Raza, etc.) • Online discussion board • Champion the people who participate in the standards setting, and leverage their testimonials 	<ul style="list-style-type: none"> • Commercial testing products looking for weaknesses will try to discredit Smarter Balanced assessment • South Carolina, Alabama moving to ACT—will this cause chain reaction? • Concern around rigor of test • Defending subjectivity in the process • Managing the Online Panel process: Rogue panelists/gate crashers/managing outliers • States decide not to use Smarter Balanced cut scores • How do scores compare to SAT, others? • “Losers” in “score cliff” → will we lose students through lack of confidence? • Will the assessment exacerbate inequities? Middle-upper class parents can afford tutoring, disadvantaged parents cannot • Local education leaders need support, cover—they may be blamed for “score cliff” • We need to be prepared for all communications to go public—opponents taking things out of context

Environmental scan.

Hager Sharp developed audience specific messaging geared toward building understanding of the achievement level setting process and preparing communities for upcoming shifts in student performance levels. We can classify and prioritize key audiences for Smarter Balanced and stakeholders in K-12 standards and assessments.

Initial messaging was created building on a literature review of existing research and our understanding of best practices of communicating the complexity of assessments and academic standards. The primary source materials for these message came from the “Common Core State Standards, Public Opinion Research” primer from the Bill and Melinda Gates Foundation, the “The Language of Deeper Learning in America” focus group research conducted by Luntz Global on behalf of the William and Flora Hewlett Foundation, and the polling conducted by the Tarrance Group and David Binder Research on behalf of the Collaborative for Student Success (“Findings From National

Survey On Common Core Standards”). Tested messages were created and based on our review of research from the Gates Foundations, Hewlett Foundation, and other sources (See Communications Plan, section on Key Messages).

Focus group study.

Building on our conversations in the kickoff meeting and follow-up meeting with CTB/McGraw Hill, we began the process of conducting thorough research on public attitudes toward the achievement level setting process and possible outcomes from that process.

To clearly communicate with audiences concerned with achievement level setting and the Smarter Balanced assessments in general, those audiences were put at the center of all activities and outreach. Because parents have such a large stake in testing and score results, it was critical that communications effort begin with a formal message testing and audit of the audience to understand their attitudes and preferences. To gather feedback from key audience segments on messages and materials for the Smarter Balanced standard setting activities, we proposed cost-effective online focus groups among parents, teachers and school administrators from two Smarter Balanced governing states, California and Michigan. These states were selected to represent a balance of political, geographic and cultural diversity, and they represent significant numbers of our target audiences. The research focused on parents, teachers and school administrators.

Feedback from the focus groups provided valuable insights into the messages and information audiences find most appealing and relevant. These insights informed the team’s outreach efforts for the standard-setting activities May through October 2014. Findings from the research study were used to better understand and prepare for various contingencies leading up to and following the standard setting workshops, including changes in student performance standards on test scores.

Hager Sharp also gathered detailed feedback on messages and materials for the Smarter Balanced achievement level setting activities using cost-effective online focus groups among parents, teachers, and school administrators from two Smarter Balanced Governing States: California and Michigan. Feedback from the focus groups allowed us to integrate relevant insights from parents into materials. Staff, state leads and public information officers were later briefed on findings from the focus group study and its implications.

Recruitment and Outreach

Similar to the resources provided for the field test, a recruitment communications toolkit was developed for states to use to raise awareness about the Online Panel opportunity and conduct outreach to stakeholder groups on the topics of achievement level setting. These customizable templates were developed to help states with outreach and communication with statewide organizations, districts, colleges and universities and stakeholder groups such as education associations and PTA affiliates to inform them about the opportunity and encourage them to participate. Throughout the duration of standard-setting, editorial support was provided to Smarter Balanced on content related to achievement level setting, to ensure that key messages about the process were integrated into planned outreach and communications. Examples include stories in *SmarterNews*, presentations to AFT/NEA teacher ambassadors and several presentations to chiefs.

Outreach to partner organizations was conducted at the state and national level, using materials created by Hager Sharp and approved by Smarter Balanced. The primary motivation for this outreach was recruitment for the Online Panel, but these efforts were also intended to help prepare stakeholder groups for the assessment rollout with more challenging expectations for students. A list of partner organizations was created and segmented by target audience. Hager Sharp drafted a

variety of outreach communications, including recruitment letters and Dear Colleagues to help create awareness among national K-12 and higher education groups, advocates, foundations and others.

Social media campaign/paid promotion.

As part of our work to assist Smarter Balanced efforts to recruit parents, teachers and interested community members for the Online Panel, we provided social media support for both Smarter Balanced Facebook and Twitter channels. A variety of social media posts were created highlighting the Online Panel opportunity. These posts used messages intended to resonate with different target audiences, including teachers, higher education faculty, parents and content experts.

As an extension of these recruitment efforts, contractors worked closely with Education Week, an online magazine focusing on education, in order to create a banner ad publicizing registration for the Online Panel. We worked internally to design and layout a banner ad that could run on the site and would bring interested parties to the Smarter Balanced Online Panel registration page. The banner ad ran for 4 days and resulted in increased awareness about the Online Panel. A combination of social media and paid promotion helped to nearly double the number of participants registered for Online Panel windows.

Media Strategy/Communications Planning

Throughout the project, Hager Sharp worked with earned media coverage and social media activity related to the Common Core State Standards, Smarter Balanced and testing. As part of our work to build awareness about the process behind achievement level setting, Hager Sharp proposed holding a “background briefing” for trade media and other national partners who would be interested in the achievement level setting process. The goal of this briefing was to educate trade media on the process for developing new performance standards in advance of the Online and In-Person Panels in October and the vote in November.

While many education beat writers possessed detailed knowledge of assessments—perhaps even of the Common Core-aligned assessments—Hager Sharp believed that even experts may have little pre-existing knowledge of the highly technical psychometric process for developing cut scores. The publication of blog posts by anti-Common Core scholars like Sandra Stotsky and Diane Ravitch and anti-testing principal Carol Burris showed the importance of providing additional context and detail to education media. These writers set the agenda for coverage of Common Core and assessment issues nationwide and are hungry for details about the Consortium’s planned next steps.

For this briefing, Hager Sharp facilitated a webinar briefing on September 10, with Joe Willhoft and Mike Bunch as featured speakers. Smarter Balanced leadership provided an overview of Smarter Balanced plans for assessment launch in 2014-2015 with a focus on the need to establish a new baseline for student achievement. Measurement Incorporated provided a walk-through of achievement level setting, helping reporters and opinion leaders understand the contours of the process and why it is occurring. To support the briefing, Hager Sharp produced materials, including an advisory for targeted reporters, talking points, and PowerPoint deck, and created a targeted list of trade journalists covering the Consortium. Several key reporters from trade publications attended the briefing and a recording of the presentation was later distributed to reporters.

Communications planning for key activities and state approval.

Extensive communications planning also went into the vote of approval and chiefs’ meeting originally planned for Nov. 6. Draft materials for alternate scenarios were developed, along with a presentation

for chiefs to provide an ore view of messaging and communications materials. A series of external briefings were planned and a run-of-show for a press conference and staffing plans for the day's events. Due to the delayed vote, the media briefing planned for trade media and other external presentations was subsequently cancelled.

Contingency planning

Intake interviews: Development of contingency plan, messaging and protocols.

The contingency communications plan was intended to enhance understanding of operational vulnerabilities; maintain credibility and reputation; involve appropriate cross-divisional staff through effective preparation and tasks; and protect participating states, employees, and a range of Smarter Balanced stakeholders. The plan was developed to help implement best practices during a contingency situation as well as incidents that may occur before, during, and after the achievement level setting process. The document proposed protocols and messages for responding to key audiences through a variety of channels.

In order to assess the different types of external threats and internal weaknesses, Hager Sharp conducted research on Smarter Balanced operations and planned achievement level setting activities using several primary methods, including:

- A kickoff meeting and subsequent planning meetings with Smarter Balanced and representatives from WestEd (PMP), Measurement Incorporated, CTB/McGraw-Hill, and GMMB.
- In-depth phone interviews with key employees and contractors including leadership staff, the lead facilitators for achievement level setting and project leads from PMP.

The plan was designed to help Smarter Balanced respond rapidly to opposition and potential threats, especially those related to the achievement level setting process and acceptance of the new cut scores. In preparation for what will likely be a challenging rollout of the new achievement levels and cut scores, we worked concurrently with our general communications planning to develop response protocols and topline messages for the most likely contingency scenarios. The contingency communications plan included protocols to ensure a timely media response in fast-moving scenarios, such as:

- Technical or operational problems at the In-Person Panel or involving the administration of the Online Panel;
- A disagreement or negative exposure following the In-Person Panel;
- A disagreement or divided vote at the fall Collaboration Conference when states vote on a motion to endorse the new achievement levels and cut scores;
- Concerns about the achievement level setting process from Smarter Balanced internal groups or stakeholders; and
- The release of a third-party review questioning the validity of the achievement level setting process (e.g., the National Academy of Sciences report criticizing National Assessment of Education Progress [NAEP] achievement levels) or validity of cut scores in relation to external measures such as NAEP, PISA, SAT, or ACT.

On September 23, a contingency training session was held for key personnel and contractors, senior leadership and communications staff. The session focused on team-building, rehearsing messages and team roles, and discussion of best practices during a contingency situation.

Ebola contingency response.

Prior to the in-person panel for achievement level setting, the first case of Ebola was confirmed in Dallas, Texas. Due to potential health concerns regarding the disease, Hager Sharp monitored media outlets throughout the crisis including daily review of stories in the media, social media sites and continued contact with the CDC and local Dallas health officials. While in Dallas, second and third cases of Ebola were confirmed. It was also noted that the third person to contract the virus had recently flown from Texas to Ohio, and back to Texas. In order to protect the safety of all In-Person Panel participants, Hager Sharp conducted an in-depth search to identify the airline name, flight number and exact flight pattern the plane took prior to being decontaminated. It was confirmed that no participants of the In-Person panel had flown on the airline.

Hager Sharp's monitoring of the Ebola crisis also included crafting messaging to send out to panelists both prior to the In-Person Panel and during the event. In order to ease any potential fears participants might have had, Hager Sharp crafted a message alerting participants that Smarter Balanced was aware of the health situation in Dallas and would keep them apprised of any urgent updates as they arose. Hager Sharp also drafted language during the In-Person Panel when it seemed as though the city of Dallas would declare a State of Emergency in order to keep anyone who had come in contact with the virus from travelling. Hager Sharp drafted this language for Joe Willhoft to announce during the opening sessions for the grade 3-8 panels.

Fortunately, despite the public health crisis in Dallas, only a handful of participants opted out of attending the panel and all others continued with the Achievement Level Setting Process. Upon the conclusion of the In-Person panel, Hager Sharp continued to monitor the Ebola crisis. It was discovered that one teacher from Maine, although she had not left the hotel where the Achievement Level Setting process was taking place, had been put on administrative leave out of concern that she may have been in contact with the disease while in Dallas. Hager Sharp assisted Smarter Balanced in drafting a message to her thanking her personally for her participation in the In-Person panel.

Digital and Creative Support

Hager Sharp provided a variety of design and web consulting services to assist with recruitment for the online panel. The registration page was redesigned to ensure a smoother public-facing presence, using plain language that would be understood by educators. Creative products such as web buttons and flyers were also created to build awareness.

Substantial resources were invested in the planning of a video package for use by Smarter Balanced and member states. The video of the achievement level setting process was developed to explain the purpose and structure of the achievement level setting process through interviews with multiple panelists. The video also served to preview the cut scores and the impact with states and communicate broader messages about the value of the assessment.

The final product was a video package that states and Chiefs may use in order to reinforce claims for the validity of the process, the seriousness of the discussion, and the inclusive approach. In addition to the master edit (or national version), Smarter Balanced requested video presentations to support individual state chiefs after the Nov. 6 vote as they sought endorsement, ratification and approval of the achievement level scores within their own states.

Once completed, we created a master video explaining the Achievement Level Setting process that can be used both on the Smarter Balanced website and as a piece for states to bring to their local legislature. In addition, we also created more than 18 state-specific videos for participating Smarter Balanced states represented at the In-Person Panel.

Expansion of video production scope.

In addition to the video package document the standard-setting process for purposes of peer review, Hager Sharp was contracted to develop 18 additional versions of the video customized for use by Smarter Balanced member states. The video production crew spent three days on-site at the In-Person Panel, collecting some 60 interviews from panelists from all of the Smarter Balanced members and a range of categories, including teachers, higher education faculty, parents, representatives from business and the community, staff and advisors.

Customization of the videos was accomplished by inserting interview clips from state panelists. For example, Missouri's video included an interview with a Missouri panelist, while California's video featured a California panelist. All videos will be finalized before the end of 2014.

Prior to the trip to Dallas, Hager Sharp compiled a list of states whom we were looking to interview participants from, created questions to ask participants and designed a sign-up sheet for participants to use to register for interviews. Hager Sharp also ensured that a media release form was in place so that Smarter Balanced would have the rights to all footage and that all interviews could potentially be used in the final products.

Achievement Level Naming Research

Additional research was commissioned to test language options for the names of the achievement levels, particularly the label for level 3. Our objective was to gather feedback from assessment experts, educators, parents and students to develop descriptors that will effectively communicate a student's degree of readiness for college and/or career beyond high school as a function of performance on the assessment. Hager Sharp proposed and executed a three-part approach to the work, which included in-depth interviews with educational researchers/experts, focus groups with teachers, and a survey of parent and student attitudes with a sample size of n=500. These three approaches, in combination, will provide Smarter Balanced with the needed primary research from their key audiences for the Achievement Level names. On behalf of Smarter Balanced, Hager Sharp conducted qualitative and quantitative research to develop descriptors for achievement levels that may appear on the score reports of the new assessment. The research consisted of the following components:

- Key informant interviews with five assessment experts
 - The 20-minute interviews took place September 29 through October 3, 2014
- Two 60-minute online focus group discussions among teachers in Smarter Balanced states
 - Teachers in Eastern and Midwestern states, October 9, 2014, 6:00-7:00 p.m. EST
 - Teachers in Western states, October 9, 2014, 7:30-8:30 p.m., EST
- An online survey among parents of children 18 years old or younger
 - Fielded by Ipsos, a leading market research firm, on October 17-24, 2014
 - National sample of 615 parents, with 271 men and 344 women
- An online survey among students ages 12-17
 - Fielded by Ipsos, a leading market research firm, on October 17-24, 2014
 - National sample of 505 children ages 12 to 17, with 258 boys and 247 girls

Through this iterative research process, we explored initial concepts for descriptors in discussions with assessment experts, and then tested and refined them with teachers. We then conducted a final test of selected sets of performance level descriptors with larger samples of parents and students using online surveys. Detailed summaries of all research components are included in this report.

Interviews with assessment experts.

Assessment experts determined the descriptors should meet three criteria:

- They should communicate measurement against the Common Core State Standards (CCSS).
- They should communicate the degree of college and/or career readiness appropriate to the achievement level.
- They should convey positive and negative tone as appropriate to the achievement level.

Specifically, Achievement Level 3 designates a student performance that is ‘proficient’ or on track for college- and career-readiness.

After discussing these criteria, the experts explored specific terms to be used as potential descriptors. Detailed comments about the terms are included in this report.

After discussing the strength and weaknesses of various terms, the experts recommended the following descriptor options to be explored in the second phase of the research, the focus groups among teachers.

<p><u>Option 1</u></p> <ul style="list-style-type: none"> • Superior command of knowledge and skills • Solid command of knowledge and skills • Sufficient command of knowledge and skills • Partial command of knowledge and skills • Limited command of knowledge and skills 	<p><u>Option 2</u></p> <ul style="list-style-type: none"> • Superior mastery of knowledge and skills • Sufficient mastery of knowledge and skills • Partial mastery of knowledge and skills • Limited mastery of knowledge and skills
<p><u>Option 3</u></p> <ul style="list-style-type: none"> • Advanced • College ready (for grade 11 only; lower grades would be ‘on track for college readiness’) • Progressing to college readiness or partially college ready • In danger of not meeting goals 	<p><u>Option 4</u></p> <ul style="list-style-type: none"> • Exceeding expectations (for college and career readiness) • Meeting expectations (for college and career readiness) • Needs improvement (for college and career readiness) • Not meeting expectations (for college and career readiness)
<p><u>Option 5</u></p> <ul style="list-style-type: none"> • Succeeding • Progressing appropriately • Needing improvement • Intervention needed (or strong support needed) 	<p><u>Option 6</u></p> <ul style="list-style-type: none"> • Progressing above expectation • At expectation • Needing improvement (or progressing below expectation) • Needing intervention
<p><u>Option 7</u></p> <ul style="list-style-type: none"> • Advanced • Proficient • Partially proficient • Unsatisfactory 	

The interview guide for the expert interviews (Appendix A) and a scan of assessment descriptors used in U.S. states (Appendix A) are included in this report.

Focus groups among teachers.

We refined the suggested options from the experts, and explored the following options in the two online focus groups with teachers in Smarter Balanced states:

- Option 1: Thorough, Adequate, Partial, Basic
- Option 2: Superior, Sufficient, Partial, Limited
- Option 3: Advanced, On Track for College and Career Readiness, Progressing to College and Career Readiness, In Danger of Not Meeting Goals
- Option 4: Exceeding Expectations, Meeting Expectations, Partially Meeting Expectations, Not Meeting Expectations
- Option 5: Advanced, Proficient, Partially Proficient, Unsatisfactory
- Option 6: Distinguished Command, Strong Command, Moderate Command, Partial Command
- Option 7: Achievement Level 4, Achievement Level 3, Achievement Level 2, Achievement Level 1

In both groups, teachers discussed these options at length, and details of the discussion are provided in this report. When asked to pick their favorite from among the seven options, half of the teachers in Eastern and Midwestern states preferred Option 5 above. Other favorites included Options 4 and 6. Teachers in Western states preferred Options 3 and 4. One teacher in the Western state group suggested refining Option 4 to explicitly reflect measurement of the Common Core standards:

- Option 4: Exceeding Grade Level Standards, Meeting Grade Level Standards, Partially Meeting Grade Level Standards, Not Meeting Grade Level Standards

The moderator's guide for the teacher focus groups (Appendix A) is included in this report.

Online surveys among parents and students.

Based on the feedback from teachers and following informal discussions with parents, we refined the descriptors to determine final options for quantitative testing among parents and students. We tested the following four options through online surveys among parents of children 18 years old or younger as well as students ages 12-17.

<p><u>Option 1</u></p> <ul style="list-style-type: none"> • Level 4: Advanced • Level 3: Proficient • Level 2: Basic • Level 1: Minimal 	<p><u>Option 2</u></p> <ul style="list-style-type: none"> • Level 4: Advanced • Level 3: On track • Level 2: Needs improvement • Level 1: At risk
<p><u>Option 3</u></p> <ul style="list-style-type: none"> • Level 4: Exceeding expectations • Level 3: Meeting expectations • Level 2: Partially meeting expectations • Level 1: Not meeting expectations 	<p><u>Option 4</u></p> <ul style="list-style-type: none"> • Level 4: Exceeding grade level standards • Level 3: Meeting grade level standards • Level 2: Partially meeting grade level standards • Level 1: Below grade level standards

When asked to select the single most effective set of descriptors from among the four options, parents selected the following by a substantial margin over the other three:

- Level 4: Exceeding grade level standards
- Level 3: Meeting grade level standards
- Level 2: Partially meeting grade level standards
- Level 1: Below grade level standards

The weakest descriptor in this set is the Level 2 descriptor. An alternate descriptor to consider for this level may be ‘Needs improvement in grade level standards,’ as ‘needs improvement’ tested well in Option 2 and the addition of ‘grade level standards’ makes it consistent with the other descriptors in this set.

Students also preferred the same set of descriptors. As with parents, the weakest descriptor in this set is the Level 2 descriptor. An alternate descriptor to consider for this level may be ‘Needs improvement in grade level standards,’ as ‘needs improvement’ tested well in Option 2 among students and the addition of ‘grade level standards’ makes it consistent with the other descriptors in this set.

Section 3. Logistics

Schedule

Planning for the achievement level setting activities began with the kickoff meeting in February, 2014. The achievement level setting culminated in the Chiefs approving the recommended cut scores on November 14, 2014. Key meetings attended or hosted by the contractor, or for which the contractor prepared materials, are listed in Table 3.1.

Table 3.1. Key Meetings Related to Achievement Level Setting

Meeting	Date(s)
Kickoff Meeting	February 24–25, 2014
Technical Advisory Committee	March 6, 2014
Hager Sharp Planning Meeting	March 19, 2014
AERA/NCME Conference	April 3–7, 2014
Collaboration Conference	April 29–May 2, 2014
Technical Advisory Committee	April 30, 2014
Technical Advisory Committee (virtual)	May 30, 2014
National Conference on Student Assessment	June 25–27
Technical Advisory Committee	July 16–17, 2014
Achievement Level Setting Advisory Panel	July 17–18, 2014
Technical Advisory Committee (virtual)	August 27, 2014
Background Briefing for Trade Media	September 10, 2014
Achievement Level Setting Advisory Panel	September 22, 2014
Contingency Communications Training	September 23, 2014
Online Panel (virtual)	October 6–18, 2014

Meeting	Date(s)
In-Person Workshop	October 13–19, 2014
Vertical Articulation (Cross-grade Review)	October 20, 2014
External Review of Achievement Level Setting	October 13–20, 2014
Achievement Level Setting Advisory Panel (virtual)	October 28, 2014
Technical Advisory Committee	October 30, 2014
United States Education Department Briefing	October 31, 2014
Chiefs’ Meeting	November 6, 2014
Chiefs’ Dinner Meeting	November 14, 2014

Facilities

The Smarter Balanced achievement level setting was held at the Hilton Anatole hotel in Dallas, Texas, from October 13 to 20, 2014. The contractor staff performed an extensive review of more than one hundred potential venues located in Smarter Balanced and non-consortium states prior to the selection of the Hilton Anatole. The contractor staff understood that the venue would be paramount to the success of the achievement level setting. The Hilton Anatole was thoroughly vetted, including an onsite visit by contractor staff prior to signing a contract in April 2014. Many important factors were taken into consideration prior to selecting the Hilton Anatole as the site for the achievement level setting.

Location: An ideal venue would be both geographically central and located in close proximity to a large airport hub. Dallas, Texas, is a central city in the United States to which panelists could travel on a flight of four hours or less. There are two major airports located in Dallas: Dallas/Fort Worth International Airport, an American Airlines hub; and Dallas Love Field, a Southwest Airlines hub.

Meeting space: One of the determining factors in selecting the Hilton Anatole was the size, flow, and functionality of its 600,000 square feet of event space. It was crucial to secure space that was functional foremost and that allowed for ease of movement throughout the day to accommodate such an extensive agenda. The achievement level setting workshop called for 12 breakout rooms, a Smarter Balanced staff workroom, a room for contractor staff to process data, a large ballroom for general sessions, and a space for more than 400 panelists and staff to dine for breakfast and lunch daily. The Hilton Anatole was able to secure the majority of the Atrium section of the hotel, keeping all the achievement level setting activities in close proximity to one another and allowing both panelists and staff to move quickly and efficiently from one space to the next.

Audio/visual and Internet capabilities: Imperative to the success of the achievement level setting activities was having cutting edge technology and Internet connectivity available throughout all meeting spaces. The contractor's network operations team worked closely with the venue staff to ensure the necessary infrastructure was in place to accommodate 400 computers being online simultaneously.

Overnight accommodations: The Hilton Anatole has more than one thousand guestrooms, which ensured panelists would be comfortable after a rigorous day of achievement level setting activities. The Hilton Anatole is accustomed to hosting large events and conferences and ensured a seamless check-in/check-out process for all panelists.

Hardware and Software Support

The contractor staff arrived onsite several days in advance of the start of the achievement level setting in order to properly set up and test all hardware and software. Nelson Androes, Measurement Incorporated's Vice President of Information Technology, and Eric Lewis, Measurement Incorporated's Director of Network Operations, were both onsite for the duration of the achievement level setting workshop to ensure the process ran smoothly and efficiently, and to identify and resolve any hardware or software issues that arose.

In addition, Mr. Androes and Mr. Lewis had a staff of five other network operations professionals who continuously monitored breakout rooms should the hardware or software fail at any time. In order to guarantee the achievement level setting activities would never be slowed down due to a failed computer or dropped Internet connection, all facilitators were given a comprehensive phone list in order to reach a member of the network operations staff immediately should they need assistance.

In one instance a visually impaired panelist was having difficulty viewing the content on a standard-sized monitor. In response, the contractor's network operations staff purchased and installed a large television monitor to assist the visually impaired panelist move through the content more quickly.

The contractor's network operations staff continually monitored Internet usage and increased bandwidth when necessary. A network server, along with a backup, was stored and monitored in a secure meeting space.

Meetings

In-person workshop.

There were two distinct parts to the logistics of the in-person achievement level setting workshop. The first part will focus on panelist logistics, and the second part will concentrate on the logistics of the facilities and materials.

After panelists were nominated and selected and had confirmed their intent to participate, the contractor's meetings manager began communicating via e-mail with each panelist regarding travel and meeting logistics.

The contractor identified Cardinal Travel Service, with whom the contractor has a long-standing professional relationship, as the travel agency to book airline tickets for all panelists and staff. The contractor instructed the agency to arrange for all panelists to arrive in Dallas one day prior to their specific panel's start date, and to have them depart Dallas late in the afternoon of the day their panel was scheduled to conclude. Panelists were given contact information for Cardinal Travel Service in the first e-mail communication so they could begin making flight arrangements soon after

being chosen as panelists. The contractor's meeting manager received copies of all final itineraries in order to coordinate ground transportation for panelists in Dallas in an efficient manner.

Panelists were sent, via e-mail, copies of all forms they would eventually receive onsite in Dallas. These forms included: non-disclosure agreement, panelist information form, travel reimbursement form, substitute reimbursement form, and information on honoraria/stipends. Panelists had one week to return the completed non-disclosure agreement and panelist information form to the contractor's meetings manager. The panelist information form was essential in gathering details about panelists to ensure appropriate accommodations were made. Information such as food allergies and ADA accommodations were all noted.

All forms were explained to panelists in the initial communication. In addition, the contractor outlined the policy for travel reimbursement along with eligibility for substitute reimbursement and honoraria or stipends. Panelists who were teachers and needed a substitute for their classes in their absence were eligible to have their school or district directly reimbursed for the actual cost of the substitute teacher. An honorarium or stipend of \$150 per day for a panel session (\$450 maximum for the 3-day panel) was available for those panelists who were not employed, were self-employed, or who took an unpaid leave from work to attend. Educators did not receive a stipend for any days for which their district was reimbursed for substitute costs.

Panelists who failed to make travel arrangements by the deadline were sent multiple reminders via e-mail until all airline travel had been booked. Once panelists' flight arrangements were finalized, the contractor's meeting manager reserved hotel accommodations at the Hilton Anatole on behalf of all panelists.

Two weeks prior to travel, a reminder containing flight, ground travel, and hotel information was sent to all panelists. This communication also included web links for panelists to access Smarter Balanced practice tests, Achievement Level Descriptors, and other reference material on the Smarter Balanced website so panelists could appropriately prepare for the achievement level setting workshop.

Two days prior to travel, a final message was sent to panelists, providing them with the contractor's meeting manager's cell phone number should any issues arise en route to or upon arriving in Dallas.

Panelists convened for the first time onsite at breakfast and registration on the first morning of their panel. The first session took place in the Carpenter Ballroom adjacent to the breakfast area. At that time, introductions and opening remarks were made and the workshop commenced.

The onsite logistics of the achievement level setting workshop were coordinated by the contractor's meeting manager. The facilities for the in-person workshop included 12 breakout rooms, two offices, a large ballroom for general sessions, and a dining area. Panels were assigned to breakout rooms based on size and proximity to other panels taking place simultaneously. Specific assignments are displayed in Table 3.2. A map of the entire meeting space is included in Appendix B, along with documentation of logistical support for lodging, air travel, ground transportation, meals, and other requirements.

Four of the 12 breakout rooms contained six tables with six computer setups at each table to accommodate the high school panels, comprising 36 panelists. The remaining eight breakout rooms contained five tables with six computer setups at each table to accommodate the grade 3–8 panels. The grade 3–8 panels each had 30 panelists. Pods were created by pushing together two six-foot-by-eighteen-inch tables. The pod configuration encouraged collaboration and conversation among table members and created space for facilitators to easily move about the room. Breakout rooms were all set with extra chairs in the rear to accommodate observers and Smarter Balanced staff.

All computers were set up and tested by the contractor’s network operations staff prior to each panel going live in any breakout room.

In addition to the computer pods, each breakout room was equipped with a projector and projection screen. Facilitators were able to control the projector from the head table with a wireless keyboard and mouse.

Table 3.2. Breakout Room Assignments

Grade/Content	Breakout Room
Grade 3 English Language Arts/Literacy	Batik A
Grade 3 Mathematics	Batik B
Grade 4 English Language Arts/Literacy	Cardinal A
Grade 4 Mathematics	Cardinal B
Grade 5 English Language Arts/Literacy	Steuben
Grade 5 Mathematics	Wyeth
Grade 6 English Language Arts/Literacy	Lalique
Grade 6 Mathematics	Rosetta
Grade 7 English Language Arts/Literacy	Fleur-de-Lis A
Grade 7 Mathematics	Fleur-de-Lis B
Grade 8 English Language Arts/Literacy	Edelweiss
Grade 8 Mathematics	Dardanelles
Grade 11–A English Language Arts/Literacy	Batik A
Grade 11–A Mathematics	Cardinal A
Grade 11–B English Language Arts/Literacy	Batik B
Grade 11–B Mathematics	Cardinal B

Panelists were served breakfast, lunch, and two coffee breaks daily. All meals were served in the Atrium directly below the mezzanine level where the breakout rooms were located.

Cross-grade review.

The Cross-Grade Review Panel was selected from among the in-person panelists based on nominations from Smarter Balanced member states in August. Notifications were sent to 64 nominated panelists via e-mail. Panelists were instructed on steps to take should they need to revise previously arranged travel.

The cross-grade review took place at the Hilton Anatole in Dallas, Texas, on Monday, October 20, 2014. The 64 panelists were split by content area into two groups of 32. Two breakout rooms were utilized. Specific assignments are displayed in Table 3.3.

Table 3.3. Cross-Grade Review Panel Breakout Room Assignments

Content	Breakout Room
English Language Arts/Literacy	Batik A/B
Mathematics	Cardinal A/B

The cross-grade breakout rooms utilized the same computer pod setup that was used earlier in the workshop. Airwalls in both the English Language Arts/Literacy and Mathematics rooms were opened to allow more room for observation and collaboration.

Technical Advisory Committee Meetings

The contractor attended and presented achievement level setting information at five TAC meetings, and prepared presentation materials for a sixth TAC meeting. The purpose of these presentations was to inform the TAC of achievement level setting plans, procedures, and results, and to seek the committee’s advice and approval.

The TAC unanimously endorsed the achievement level setting at the October 30 meeting:

“The Technical Advisory Committee concludes that the Smarter Balanced achievement level setting design and implementation reflect contemporary professional practice and represent a valid process that supports the defensibility of the content-based performance standards.”

Collaboration Conferences

The contractor attended the spring Collaboration Conference in Minneapolis, Minnesota, and participated in sessions relevant to achievement level setting, including data exchanges, test map discussion, and recruitment. The contractor presented the Achievement Level Setting Plan to the TAC and subsequently to the Chiefs. The Chiefs accepted the Plan with certain revisions, which became the Achievement Level Setting Plan Supplement.

The Consortium did not hold a fall Collaboration Conference, but the contractor attended the November 6 Chiefs’ Meeting and presented the achievement level setting results.

Advisory Panel Meetings

The contractor hosted three Achievement Level Setting Advisory Panel meetings. Two of the meetings were held in person for one day each. The third meeting was a half-day webinar. The Advisory Panel advised Smarter Balanced and Measurement Incorporated on a variety of psychometric and policy issues related to the design and conduct of the achievement level setting and related tasks.

The Advisory Panel consisted of a dozen experts in education, assessment, and psychometrics. (One member dropped off the panel after the first meeting due to a job change, leaving the total number of panelists at eleven.)

The Advisory Panel unanimously endorsed the achievement level setting at the October 28 meeting:

“After reviewing the information about the Achievement Level Setting activities and the auditors’ report, the Advisory Panel confirms the design and procedures for the Achievement Level Setting and the Vertical Articulation were implemented as planned, represent a valid process that is consistent with best practices in standard setting, and support the defensibility of the content-based performance standards.”

National Conference on Student Assessment

In June, MI underwrote travel and related support for Smarter Balanced staff to present the achievement level setting plans at the annual meeting of the National Conference on Student Assessment. Dr. Willhoft presented Smarter Balanced: Innovative Approaches to Achievement Level Setting with support from MI staff. The presentation is included in Appendix B.

Other Meetings

- Hager Sharp Planning Meeting: Measurement Incorporated, CTB/McGraw-Hill Education, and Smarter Balanced staff met at the Washington, D.C., offices of Hager Sharp to plan achievement level setting communications. Another communications firm that works with Smarter Balanced, GMMB, joined via telecom to coordinate messages and to avoid duplication of materials.
- AERA/NCME Conference: Dr. Bunch attended the 2014 AERA/NCME conference in Chicago.
- Background Briefing for Trade Media: Hager Sharp hosted a webinar briefing for trade media in order to provide background information reporters would need to understand and write about the achievement level setting.
- Contingency Communications Training: Measurement Incorporated, Hager Sharp, and Smarter Balanced staff attended a half-day training on the Contingency Communications Plan. The group discussed a variety of scenarios and the procedures that would be followed should those or other situations arise.
- Online Panel: 10,096 people from across the nation registered to participate in the virtual achievement level setting. Of that number, more than 5,000 logged into the site during the participation windows they had selected during registration. Panelists were provided with

numerous training materials and extensive instructions to help them understand the purpose of the activity and the procedure for setting a Level 3 bookmark. Bookmarks were set and the activity completed by 2660 panelists.

- External Review of Achievement Level Setting: States were invited to send an observer to the in-person workshop and Cross-Grade Review Panel meeting in Dallas. These observers were given access to all panels and provided with seating in the breakout rooms.
- United States Education Department Briefing: Drs. Willhoft, Alpert, and Bunch participated in a face-to-face meeting with representatives of the U.S. Department of Education on October 31 to provide an update on achievement level setting.
- Chiefs' meeting on November 6: Measurement Incorporated prepared materials for and provided on-site support for this meeting, at which the chiefs passed two resolutions: one to suspend rules for allowing affiliate member states to vote and another endorsing the process used by the contractor to conduct achievement level setting.
- Chiefs' Dinner Meeting: Measurement Incorporated hosted the Chiefs' Dinner Meeting at the CCSSO Conference on November 14. Smarter Balanced and Measurement Incorporated staff presented adjusted cut scores recommended by Smarter Balanced, which the Chiefs approved.

Security

The security of the achievement level setting was of vital importance to the success of the achievement level setting. Throughout the planning process the importance of security was communicated to all staff and panelists. The location of the achievement level setting was not revealed to the panelists until it was necessary to begin booking travel arrangements. At that time the contractor staff stressed the significance of the achievement level setting and that it was in the best interest of all parties not to divulge the location. The location was never published or announced publically prior to the beginning of the in-person workshop.

Once panelists committed to participating in the achievement level setting, the next step was to sign non-disclosure agreements, which stated that panelists were personally responsible for maintaining strict confidentiality of all information related to the achievement level setting and that violation of the agreement could lead to disciplinary or legal action. Contractor staff distributed and collected non-disclosure agreements electronically at the beginning of the process and then required a hand-signed hard copy of the agreement onsite in Dallas as a back-up measure.

Security was a top priority from the very beginning of the setup process upon arrival at the achievement level setting workshop in Dallas. The contractor staff had the Hilton Anatole staff "re-key" all meeting space to be used for the achievement level setting workshop. This precaution allowed for only contractor staff to have access to all meeting space where secure achievement level setting materials were being stored. No hotel staff was able to enter any of the achievement level setting meeting space. Room keys were assigned to the contractor's project manager, meetings manager, information technology staff, and network operations staff only. Facilitators were instructed during training to contact one of these key holders when the session in their room concluded to allow for the meeting room to be secured. No meeting room was ever left unlocked or unattended. Once panelists left the room at the conclusion of the session they were not able to gain access inside their meeting room until their next session was set to begin and their facilitator was present. In the instance of a panelist leaving behind a personal article at the end of a session or at

the end of the day, the panelist was escorted back to the meeting room with a key holder from the contractor's staff to collect his or her belongings.

In communications with panelists leading up to the in-person achievement level setting, workshop panelists were told that security would be a top priority and that they should bring only essential personal belongings inside the room with them. Panelists were instructed to leave all other belongings in their sleeping rooms. There was a strict no cell phone policy; cell phones were not to be displayed or used at any time inside a meeting room. If it was necessary for a panelist to make or receive a telephone call, they were required to step out into the hallway.

Registration took place on the first morning of each panel. Panelists received a packet with forms, general information, reference materials, and a secure notepad for taking notes inside meeting rooms. During his opening remarks, Dr. Bunch once again stressed the importance of security and outlined what panelists could and could not do once inside their meeting rooms. At the conclusion of each session, facilitators were responsible for collecting all paper materials from all panelists before the first panelist could exit the room. At the conclusion of the workshop, all paper materials were collected for a final time and were securely shredded onsite at the Hilton Anatole. Panelists were not allowed to take any notes or reference materials with them upon the completion of the workshop.

Section 4. Audit

Two external, independent audits of the achievement level setting process were conducted to verify that the process was executed as designed and approved by the Smarter Balanced governing states and fulfilled legal requirements. This chapter describes the purpose of the two audits, the selection of the auditors, and the major audit functions performed by the auditors.

Purpose of the Audits

Smarter Balanced requested that an independent, external auditor verify that the achievement level setting process was executed as designed and presented to the Smarter Balanced Technical Advisory Committee and according to standard setting best practices. The technical audit process included the review the materials and tools used to support the achievement level setting process, observation of the In-Person Panel sessions from October 13-19, 2014, and the Cross-Grade Review Panel on October 20, 2014, and the review of the documentation of the achievement level setting.

In addition, Smarter Balanced requested a legal audit to verify that the process complied with existing federal and state laws and to make sure that the results of achievement level setting could withstand legal challenges. This audit was to be performed independently of the technical audit.

Selection of the Auditors

Since the auditor was required to be an independent, third party, employees of Measurement Incorporated (MI) and CTB/McGraw-Hill were not considered for the role of conducting the audit.

Industry leaders with extensive experience in the standard setting process were considered for this position. The contractor contacted Dr. Gregory J. Cizek, who was interested in the role but, due to a prior commitment, would not be available on October 17, 2014. Dr. Cizek executed to an NDA with CTB/McGraw-Hill Education on May 24, 2014, to discuss the opportunity further.

On June 13, 2014, the contractor proposed to Dr. Gregory Cizek that, due to his schedule conflict on October 17, 2014, Dr. Heather H. Koons support the audit in coordination with Dr. Cizek, who agreed that she would be an excellent candidate for this role.

The contractor contacted Dr. Heather Koons about the possibility on June 13, 2014. Dr. Koons executed an NDA with CTB/McGraw-Hill Education on June 16, 2014.

Dr. Gregory J. Cizek and Dr. Heather H. Koons, have expertise and extensive experience in the area of setting performance standards (see, e.g., Cizek, 2001, 2012; Cizek & Bunch, 2007; Cizek, Bunch, & Koons, 2004). Dr. Cizek is also a member of the Smarter Balanced Technical Advisory Committee.

It was agreed that Dr. Cizek would be lead auditor and would observe the In-Person Panel sessions on October 13-16, 2014 and 18-19, 2014 as well as the vertical articulation session on October 20th. Dr. Koons would observe the In-Person Panel sessions on October 17, 2014, and coordinate the reporting of her observations with Dr. Cizek.

Dr. Cizek executed a contract to audit the Smarter Balanced Achievement Level Setting with CTB/McGraw-Hill Education on September 21, 2014. Dr. Koons executed a contract to audit the Smarter Balanced Achievement Level Setting with CTB/McGraw-Hill Education on September 22, 2014.

On July 23, Dr. Susan Phillips executed a contract to perform the legal audit. Dr. Phillips is a psychometrician who also earned a law degree and has been practicing educational assessment law since 2000. She is also a member of the Smarter Balanced Achievement Level Setting Advisory Panel.

Major Audit Functions

To verify that the Smarter Balanced Achievement Level Setting was executed as designed and presented in the Smarter Balanced Level Setting Plan and Supplement, the Auditors reviewed the materials and tools used to support the achievement level setting process, observed the In-Person Panel sessions from October 13-19, 2014, and the Cross-Grade Review Panel on October 20, 2014, and reviewed the documentation of the Smarter Balanced Achievement Level Setting.

Review of tools and materials.

The auditors reviewed the tools and materials used to support the Smarter Balanced Achievement Level Setting, including the scripts and presentations used to support the In-Person Panel, the online software used to facilitate the Online Panel and the In-Person Panel, the software used to support the In-Person Panel. The contractor ensured that all tools and materials were available to the Auditors to review.

On October 6, 2014, the contractor emailed the following documents to Dr. Gregory Cizek:

- Smarter Balanced Achievement Level Setting Plan
- Smarter Balanced Achievement Level Setting Plan Supplement
- OIB Selection Process
- In-person Standard Setting Facilitator Scripts
- In-person Workshop PowerPoint Slides

On October 7, 2014, Measurement Incorporated (MI) demonstrated the Online Panel software and the software developed to assemble Order Item Booklets for Dr. Gregory Cizek at the MI headquarters in Durham, North Carolina. The materials distributed to the Online Panel participants were provided to the Auditor for review.

On October 13, 2014, CTB/McGraw-Hill Education demonstrated the use of Bookmark Pro software used to manage the analysis of achievement level setting data during the In-Person Panel. The demonstration took place in Dallas, Texas.

The Auditor review of the tools and materials used to support the Smarter Balanced Achievement Level Setting Audit Report distributed by Smarter Balanced on October 24, 2014.

On-site review of procedures.

Dr. Gregory Cizek observed the In-Person Panel sessions on October 13-16, 2014 and 18-19, 2014, in Dallas, Texas. Dr. Heather Koons observed the In-Person Panel sessions October 17, 2014.

Drs. Cizek and Koons coordinated their efforts on October 16, 2014, and again on October 18, 2014, in order to encompass Dr. Koons' observations of the In-Person Panel sessions on October 17, 2014, in the Audit Report.

Dr. Gregory Cizek observed the Cross-Grade Review Panel in Dallas, Texas, on October 20, 2014.

The conclusions expressed in the Summary and Recommendations section of the Audit Report are those of Dr. Cizek, in consultation with Dr. Koons.

Review of documentation.

Dr. Gregory Cizek reviewed the documentation of the Smarter Balanced Achievement Level Setting presented to the Technical Advisory Committee on October 30, 2014.

Dr. Cizek notified the contractor on October 23, 2014, that he had submitted the Audit Report to Dr. Willhoft of Smarter Balanced. Dr. Willhoft emailed a copy of the accepted Audit Report to the contractor on October 24, 2014.

The Smarter Balanced Achievement Level Setting Audit Report was posted on the Smarter Balanced website on November 21, 2014 and is included in Appendix C of this report.

Legal review

Dr. Phillips reviewed project documents and conducted interviews with Smarter Balanced and contractor staff. On December 17-18, she presented her findings in a webinar. In general, she found that actions taken by Smarter Balanced and its contractor in the conduct of achievement level setting were sound and legally defensible. Specific recommendations are included in the PowerPoint presentation contained in Appendix C of this report.

Section 5. Panelist Recruitment

The contractor leveraged educator recruitment processes and communication plans standardized under Smarter Balanced contract 16/17 to recruit panelist in all phases of the achievement level setting events. This chapter provides the implementation and findings of each of the recruitment activities for the Smarter Balanced Achievement Level Setting.

Critical elements of the plan included a complete recruitment program packet that provided state leads and TICs with all required information to assist them in implementing the recruitment activities for each of the events (see Appendix D). A single communication to states leads, TICs, and higher education leads for each of the events was provided that included recruitment, application, and nomination procedures. Each recruitment package included:

- Timeline for activities
- Communication plan and sample communications
- Recruitment targets and applicant qualification criteria
- Recruitment materials, including information for educators on time commitments and compensation and sample application
- Panelist/participant nomination and selection process
- Panelist/participant notification communications
- Frequently Asked Questions document
- Contract for approved nominees for each the events
- Method for informing nominees/applicants of status as participant/panelist for each of the events

These recruitment efforts began in mid-April and concluded in early October.

Online Panel for Achievement Level Setting

The goal of the online panel was to allow for participation in the achievement level setting process by as many members of the Smarter Balanced community as can be recruited. With that goal in mind, the contractor developed an application process working through the state leads, TIC, and higher education networks to provide initial communication about the opportunity. This allowed educators the first opportunity to register for the opportunity. Hagar sharp recruitment efforts focused on public communications (see Chapter 2).

Implementation.

The contractor distributed recruitment information packets to TIC's and state and higher education leads on April 16, 2014 (see Appendix D). The information packets included general information about the opportunity, sample recruitment materials and FAQ's. Using the information provided in the recruitment package, each state distributed the request through their networking paths. States were provided recruitment targets and weekly reports on their state's registration status.

Beginning on May 16, 2014 weekly reminders were distributed to state leads, TIC's and higher education leads to encourage continued recruitment of online panelists. Included with each reminder was a report outlining the current registration numbers by state and role as well as the minimum educators needed. Figure 4.1 provides an example of the weekly reports.

Figure 4.1. Sample Weekly Online Registration Report

Summary by State									
State	Teachers	Non-Teacher Educators	Admin	Higher Ed. Faculty	Higher Ed. Admin	Parents	Business	Other	Total
CA	71	14	26	6	1	4	3	7	132
CT	41	11	10	2	0	4	1	5	74
DE	8	1	1	2	2	0	0	3	17
HI	33	6	4	0	0	1	0	1	45
IA	0	0	0	0	0	0	0	0	0
ID	78	5	16	0	0	2	0	0	101
ME	0	0	0	0	0	0	0	0	0
MI	2	0	1	0	0	0	1	0	4
MO	0	1	0	0	0	0	0	0	1
MT	0	1	2	0	0	0	0	0	3
NC	2	1	3	0	0	0	0	2	8
ND	0	0	0	0	0	0	0	0	0
NH	3	0	0	0	0	0	0	0	3
NV	0	0	0	0	0	0	0	0	0
OR	6	0	2	3	0	0	1	0	12
PA	1	0	0	0	0	0	0	0	1
SD	19	3	1	3	0	0	0	3	29
VI	0	0	0	0	0	0	0	0	0
VT	0	0	0	0	0	0	0	0	0
WA	69	3	3	0	0	3	0	3	81
WI	37	5	6	1	0	0	1	4	54
WV	0	0	0	4	0	0	0	0	4
WY	0	0	0	0	0	0	0	0	0
Other	5	2	0	0	0	1	1	1	10
Total	370	51	75	21	3	14	7	28	579

Registration.

Interested participants were asked to visit a link which directed them to an application page. Once on the application page, participants chose the grade and content area (English language arts/literacy or mathematics) in which they wanted to participate. The Application process collected participants email address, role, and demographic information. To confirm registration, applicants were asked to verify their email address and select a 2-day window in which to participate in the online panel for achievement level setting, as shown in Figure 4.2.

Figure 4.2. Online panel application page.

Online Panel for Achievement Level Setting


Enter your email address and contact information

Step 2 of 4

* Primary Email Address

Confirm Email Address

* Primary email address will only be used to transmit registration/account information. This email must be active to confirm registration and receive periodic email reminders before the event.

State

ZIP Code

Occupation or Role

Race

Ethnicity

Gender





[Privacy & Terms](#)

The panels occurred online from October 6 to 17, 2014. Panelists selected a 2-day window for participation during the registration process. It was estimated that panelists would need up to 3 hours to complete the orientation process, review test questions, and recommend a score for Achievement Level 3.

During the registration process, panelists provided their email address, role, and demographic information. The timeline for the Online Panel is summarized in Table 4.1.

Table 4.1: Online Panel Recruitment Timeline

Activity	Responsible Party	Date(s)
Contact Participants	TICs, K-12 Leads, Higher Ed Leads	April 22–September 19, 2014
Register for Online Panel	Participants	April 22–September 26*, 2014
Complete Online Achievement Level Setting	Participants	October 6–17, 2014 (two-day window selected during registration)

*Note: Registration originally was planned to end on September 19, 2014. An extension was allowed to obtain a full and representative online panel.

Findings and Refinements

Over 10,000 educators, higher education faculty and administration, parents, and community leaders registered for the online panel for achievement level setting. Of those registered, 2,660 completed the online achievement level setting and submitted results. Table 4.2 provides the number of online panelist that registered for the opportunity and submitted results by state.

Table 4.2: Online Panel Registration and Submission Counts.

Online Panel Registration		
State	Registered	Submitted
CA	2271	459
CT	471	121
DE	142	48
HI	247	57
IA	33	15
ID	325	89
ME	220	75
MI	299	69
MO	1442	393
MT	120	43
NC	66	10
ND	185	57
NH	143	48
NV	156	40
OR	1132	314
PA	17	4
SD	226	68
VI	29	1
VT	77	23
WA	1310	402
WI	424	132
WV	305	116
WY	29	5
Other	430	71
Total	10099	2660

In-Person Panel for Achievement Level Setting

The successful recruitment and participation of the in-person panelists is of central importance to the appropriateness and defensibility of performance standards. Therefore, the design of the panels

included participants with a diverse set of experiences, expertise, and points of view. The goal for the In-Person Panel was to select 504 educators with the following representation:

- Representation from every Governing State
- Minimum of one on-grade general education teacher per grade level from every Governing State
- Diversity of teaching experiences among panelists, including panelists who have worked with special populations
- Diversity of demographic characteristics among panelists
- Representation from two- and four-year colleges and universities for Grade 11 panels

To meet these goals and select participants, the consortium implemented a nomination process in which Governing States nominated potential panelists based on grade level and subject panel targets specific for their state (see Appendix D). Tables 4.3 and 4.4 outline the overall panel targets by role used to recruit and select panelists.

Table 4.3. Target Counts by Role for ELA In-Person Panels

English Language Arts (ELA)								
Role	3	4	5	6	7	8	11	Total
Educator: ELL experience	2	2	2	2	2	2	4	16
Educator: General	17	17	17	16	17	17	18	119
Educator: Non-Teaching	6	6	6	7	6	6	10	47
Educator: SWD experience	2	2	2	2	2	2	4	16
General Public	3	3	3	3	3	3	6	24
Higher Education							30	30
Grand Total	30	30	30	30	30	30	72	252

Table 4.4. Target Counts by Role for Mathematics In-Person Panels

Mathematics								
Role	3	4	5	6	7	8	11	Total
Educator: ELL experience	2	2	2	2	2	2	4	16
Educator: General	17	17	17	16	17	17	18	119
Educator: Non-Teaching	6	6	6	7	6	6	10	47
Educator: SWD experience	2	2	2	2	2	2	4	16
General Public	3	3	3	3	3	3	6	24
Higher Education							30	30
Grand Total	30	30	30	30	30	30	72	252

Implementation.

On April 16, 2014 State Leads, Teacher Involvement Coordinators, and Higher Education Leads of member states were provided informational packages that described the achievement level setting process and designated the minimum number of potential nominees requested for each of the

segments. The package included resources to support each state's recruitment efforts and defined the following steps for nominating panelist:

3. The TIC and/or State Lead and Higher Education Lead identified potential nominees based on a minimum number of requested panelists from their state who met approved qualifications.
4. The state contacted potential nominees (sample recruitment emails for educators and parents/community members and FAQs were provided in the information packet).
5. Once the member state determined their final list of nominees, those individuals were referred to the online Nominee Registration Form at <https://www.surveymonkey.com/s/IPSSRE>.
6. The TIC then sent a list of nominees to CTB_SmarterBalanced21@ctb.com using a provided template.

Selection Process

To select participants for the In-Person Panel, the contractor used the information from the online Nominee Registration Form. Only nominees with complete registrations (those who were nominated by the state and completed online registration form) were considered for selection. Each panelist was selected to achieve coverage of the necessary content expertise and state representation. Pennsylvania and South Carolina did not provide nominees for selection.

The selection process began on May 27, 2014 after states submitted their nominees. The contractor incorporated a multi-stage selection process to ensure first and foremost state representation. For example, stage 1 criteria for panelist selection included matching content expertise, grade level, role, and state representation to the targets included in the in-person panel plan. After all qualified nominations were filled using stage 1 criteria, the contractor selected based on stage two criteria which may not have had an exact match by role; however, targeted state representation by subject and grade. In Stage 3 selection was based on role, expertise, and grade level regardless of which state the panelist represented. All 504 panelist and 10 alternates were selected and confirmed by August 5th; however panelist participation continued to decline. Throughout the selection process states were notified of their nominee's selection status and when applicable were requested to provide additional nominees or support to obtain a representative panel. Table 4.5 provides the criteria used and notifications provided for each stage of the selection process.

Table 4.5 Selection Criteria

Selection Stage	Matching Criteria	Notifications
Stage 1: Selection of panel targets within state and role	Matched state representation, content expertise, role, and grade level	June 9, states were notified of their selected nominees and requested to provide additional nominees for any identified gaps.
Stage 2: Selection of panel targets within state and across roles	Matched sample targets for state representation, content expertise, and grade level across roles	June 30 states were notified of their nominee’s selection status and asked to help obtain confirmation for any non-confirmed selections.
Stage 3: Selection of panel targets across states.	Matched sample targets across states for content expertise, role, and grade level	September 5 states were informed that all selected and non-selected panelists had been notified and Contractor had a complete panel confirmed to participate.
Stage 4: Replacement	Matched sample targets first within state and then across states for any vacated panel positions	October 30, states were provided a final In-Person Panel Report listing all panelists who attended from their state.

Once participants were selected, they received an invitation to confirm their intent to participate. On multiple occasions the contractor had secured a full panel of 504 confirmed panelists; however, due to a variety of circumstances some of these panelists later declined participation. Therefore the contractor continued to select newly vacant panel positions until September 29, 2014. Table 4.6 summarizes key dates during the recruitment process.

Table 4.6. Recruitment Timeline

Process	Responsible Party	Date(s)
Recruitment	State Teacher Involvement Coordinator	4/16/2014 – 5/23/2014
Deadline to Submit Recommended Participants https://www.surveymonkey.com/s/IPSSRE	State Teacher Involvement Coordinator	5/23/2014
Participant Selection and TIC Notification	Contractor	5/27/2014 – 9/5/2014
Participant Notification	Contractor	6/10/2014 – 8/29/2014

Findings and Refinements

The nomination process for the in-person panel generated 1098 complete nominations, from 22 of the member states, eligible for selection. However, throughout the selection process panelist declined their opportunity to attend the event. A total of 119 nominees declined for a variety of reasons including health concerns due to the presence of Ebola in Dallas, Texas at the time of the achievement level setting panels. Excluding those who declined, the contractor had a pool of 979 nominees eligible for selection.

A total of 601 nominees were invited to attend, and of those 482 participated in the in-person panel for achievement level setting. Table 4.7 presents the number of panelists selected by race and ethnicity from complete nominations. Tables 4.8-4.10 show the actual numbers of panelists participating, by sex and race, along with a comparison to national and Consortium norms. Table 4.11 shows the actual number of panelists participating, by role.

Table 4.7. Eligible and Selected Panelists by Race and Ethnicity from All Complete Nominations

State	Hispanic		American Indian or Alaska Native		Asian		Black or African American		Multiple Race		Native Hawaiian or Other Pacific Islander		White		Missing Race	Grand Total	
	*E	**S	E	S	E	S	E	S	E	S	E	S	E	S		E	S
CA	46	13			21	7	16	6	17	5	3	1	254	34		357	66
CT	1	1			1	1	3	3	1				35	21		41	26
DE									1	1			26	23		27	24
HI					15	12			6	4			13	6		34	22
IA							1	1					20	16		21	17
ID									1	1			22	19	9	32	20
ME									2	1			79	28		81	29
MI													16	14		16	14
MO							1	1					22	13		23	14
MT					1	1							28	22		29	23
NC	1				1				1	1			10	7		13	8
ND							1	1	2	2			21	17	1	25	20
NH													49	20	2	51	20
NV													19	15	7	26	15
OR			1	1									22	20	1	24	21
SD	1	1							1	1			9	8	5	16	10
VI	2	2			1	1	18	13					2	2		23	18

VT	1	1											17	9		18	10
WA	2	1					2	2					37	29		41	32
WI							1	1	1	1			30	26		32	28
WV							3	3					30	28		33	31
WY													16	14		16	14
Grand Total			1	1	40	22	46	31	32	17	3	1	777	391	26	979	482

*E - Represents eligible nominees
 **S - Represents selected nominees

Table 4.8. Demographic Representation for ELA In-Person Panels

State	Total	Female	Males	Hispanic	Multiple Race, not Hispanic	American Indian, not Hispanic	Asian, not Hispanic	Black, not Hispanic	Native Hawaiian, not Hispanic	White, not Hispanic
In- Person Panel		80.9%	19.1%	3.8%	4.3%	0.0%	3.0%	8.1%	0.4%	80.4%
Nation*		75.9%	24.1%	7.8%	1.0%	0.5%	1.8%	6.8%	0.1%	81.8%
Consortium**				6.7%	1.2%	0.0%	2.1%	3.3%	0.0%	85.0%

Table 4.9. Demographic Representation for Mathematics In-Person Panels

State	Total	Female	Males	Hispanic	Multiple Race, not Hispanic	American Indian, not Hispanic	Asian, not Hispanic	Black, not Hispanic	Native Hawaiian, not Hispanic	White, not Hispanic
In- Person Panel		72.5%	27.5%	4.0%	2.8%	0.4%	6.1%	4.9%	0.0%	81.8%
Nation*		75.9%	24.1%	7.8%	1.0%	0.5%	1.8%	6.8%	0.1%	81.8%
Consortium**				6.7%	1.2%	0.0%	2.1%	3.3%	0.0%	85.0%

Table 4.10. Demographic Representation for All In-Person Panels

State	Total	Female	Males	Hispanic	Multiple Race, not Hispanic	American Indian, not Hispanic	Asian, not Hispanic	Black, not Hispanic	Native Hawaiian, not Hispanic	White, not Hispanic
In- Person Panel		76.6%	23.4%	3.9%	3.5%	0.2%	4.6%	6.4%	0.2%	81.1%
Nation*		75.9%	24.1%	7.8%	1.0%	0.5%	1.8%	6.8%	0.1%	81.8%
Consortium**				6.7%	1.2%	0.0%	2.1%	3.3%	0.0%	85.0%

Table 4.11. Panelist Representation by Role

Panelist Representation by Role								
Role	3	4	5	6	7	8	11	Grand Total
Educator: ELL experience	3	5	7	5	4	5	7	36
Educator: General	33	34	32	31	31	33	48	242
Educator: Non-Teaching	12	8	12	16	11	14	17	90
Educator: SWD experience	3	5	3	4	5	4	7	31
General Public	5	4	2	4	6	3	4	28
Higher Education							55	55
Grand Total	56	56	56	60	57	59	138	482

Cross-Grade Review Panel

The cross-grade review panel consisted of exemplary panelists who attended the in-person achievement level setting. Panel members were nominated by state and higher education leads after panelist had confirmed their attendance for the in-person panel for achievement level setting.

Implementation.

During the selection process nominees' indicated their interest and availability in participating in the panel. On August 15, 2014 the contractor provided each state with their confirmed nominees and requested them to consider the following attributes in recommending cross-grade review panel members/Table Leaders:

- Ability to facilitate discussions among small groups
- Ability to find diplomatic middle-ground solutions between panelists
- Distinguished service to education
- Superior knowledge of Smarter Balanced Achievement Levels and/or Common Core State Standards

The contractor used the recommendations received from State Leads to create a cross-grade review panel as depicted in Table 4.12.

Table 4.12 Desired Composition of Cross-Grade Review Panel by Content Area.

	ELA & Math Grades 3–8	ELA & Math Grade 11
On-Grade Special Education	1–2	1–2
On-Grade ELL	1–2	1–2
On-Grade General Education	6–8	6–8
Above-Grade General Education	2–4	0
Non-Teacher Educators	1–2	1–2
Higher Education Faculty	0	2–4
General Public	1–2	1–2
Total per Content Area	16	16

The goal was to create the panel, as designed, with each Smarter Balanced Member State having at least one representative on either the ELA or Mathematics panel.

The majority of committee members selected to participate in the panel also served as table leaders at the In-Person Panel. The exception was at the high school level where there were more cross-grade review panel members than table leader positions available. Cross-grade review panel members had the option to stay in Dallas to attend the meeting on October 20 or to return to Dallas, if preferred.

Findings and Refinements

The cross-grade review panel included a total of 64 members from across the member states. All panel members also attended the In-Person Panels for Achievement Level Setting. Each state had at least one panelist also serving as a cross-grade review panel member. Table 4.13 presents the number of panel members from each state by Grade level, while and Table 4.14 presents the number of panelists by role.

Table 4.13. Numbers of Cross-Grade Review Panel Members by State and Grade.

Cross-Grade Review Panelists: Mathematics and ELA Combined								
State	3	4	5	6	7	8	11	Total
CA			2				3	5
CT			1	1			2	4
DE						1		1
HI	1						1	2
IA				1	1		1	3
ID	1					1	1	3
ME			1		1			2
MI	2						1	3
MO		1				1	1	3
MT		1	1				2	4
NC							2	2
ND				2				2
NH					1		1	2
NV				1			1	2
OR		1			1		2	4
SD			1				3	4
VI					1			1
VT		1						1
WA					1		4	5
WI							2	2
WV	1	2					5	8
WY						1		1
Grand Total	5	6	6	5	6	4	32	64

Table 4.14. Numbers of Cross-Grade Review Panelists by Role and Grade.

Cross-Grade Review Panelists by Role								
Role	3	4	5	6	7	8	11	Total
Educator: ELL experience		1		1			2	4
Educator: General	4	3	3	4	3	2	13	32
Educator: Non-Teaching	1	1	1		2	1	6	12
Educator: SWD experience		1	2				3	6
General Public					1	1	1	3
Higher Education							7	7
Grand Total	5	6	6	5	6	4	32	64

Section 6. Standard Setting Design and Implementation

This chapter documents a wide range of activities, from design and implementation of an online panel to the conduct of the Cross-Grade Review Panel. All activities were carried out in accordance with a detailed plan (see Appendix E) prepared by Measurement Incorporated (MI), and submitted to Smarter Balanced for review and approval. That plan was subsequently approved at the April 30 Collaboration meeting.

The online panel and in-person workshop used a Bookmark procedure (Lewis, Mitzel, Mercado, & Schultz, 2012), while the vertical articulation (cross-grade review) employed a procedure described by Cizek & Bunch (2007, Chapter 14). Details of both procedures are described in the sections below.

The Bookmark Procedure

The Bookmark standard setting procedure (Lewis et al., 2012) is an item response theory-based item mapping procedure developed in 1996 in response to the need for a robust standard setting procedure for high-stakes assessments of mixed format. Since 1996, it has become the most widely used procedure for setting cut scores on statewide assessments and other high stakes educational assessments. Its psychometric foundation is well documented (e.g., Cizek & Bunch, 2007), and its usefulness has been well established through adoption of cut scores produced by Bookmark-based standard-setting activities.

Creating ordered item booklets.

The most important consideration when creating an ordered item booklet (OIB) is to ensure appropriate content coverage. Psychometricians and content specialists from MI worked together closely to construct content specifications that matched Smarter Balanced's guidelines with respect to targets and claims used by item writers to inform item and test development. The OIBs contained a required 70+ pages and their content was weighted according to the specifications. In addition, each OIB contained an entire performance task, that is, a set of 5-6 items/tasks all related to a set of stimuli. In order to minimize the reading load of the panelists, the ELA booklets included items related to reading passages only when there was a minimum of three items from the given passage.

Statistical considerations must also be considered when building the OIBs. Thus, the booklets contained items that had a wide range of difficulty across the score scale, exhibited acceptable classical statistics, and showed no differential functioning. Combining the content and statistical constraints decreased the number of items for selection, but the final OIBs were very representative of the specified test content.

All OIBs were reviewed by MI, CTB, and Smarter Balanced's content and measurement experts. The reviews resulted in the removal and insertion of several items within each grade-content area until Smarter Balanced staff gave their final approval.

In a typical Bookmark procedure, each item in an OIB is mapped to an underlying construct in terms of the amount of that construct the examinee must possess in order to have a reasonable chance of answering the item correctly (in the case of a selected-response item) or obtaining a given score point or higher (in the case of an open-ended item or performance task).

In the three-parameter logistic (3PL) model, the Bookmark procedure relies on the basic relationship between person ability (θ) and item difficulty (b), discrimination (a), and pseudo-guessing (c), where

the probability of answering a dichotomously scored item correctly (P) can be expressed as shown in equation (6.1).

$$P_j(X=1|\theta) = c_j + (1 - c_j)/\{1+\exp[-1.7a_j(\theta - b_j)]\} \quad (6.1)$$

where P_j is the probability of answering correctly, θ is the ability required, a_j is the item discrimination index, \exp is the exponential function, and b_j is the item difficulty index. The way that guessing is accounted for is critical to the mapping. For most bookmark procedures, the c (pseudo-guessing) parameter is set to zero, so that the response probability specified is associated with the likelihood of a student knowing the correct response without guessing, as shown in equation (6.2). For this project, this two-PL model was used.

$$P_j(X=1|\theta) = 1/\{1 + \exp[-1.7a_j(\theta - b_j)]\} \quad (6.2)$$

For items with two or more score points, the probability of achieving any score k point or better given student ability $P_{jk}(\theta)$ in a 2-parameter logistic model can be expressed as shown in equation 6.3 from Mitzel, Lewis, Patz & Green (2001).

$$P_{jk}(\theta) = \frac{\exp(z_{jk})}{\sum_{i=1}^{m_j} \exp(z_{ji})}, \quad (6.3)$$

where m_j is the number of score points or steps for item j , and $z_{jk} = (k - 1)\alpha_j - \sum_{i=0}^{k-1} \gamma_{ji}$; α_j is the discrimination index of item j , k is the number of this score point or step, and γ_{ji} is the step value for item j at step i . Thus, the probability of scoring at step k is a joint function of examinee ability, item discrimination, and the likelihood of obtaining any of the $k - 1$ other scores. In this formulation, the value for a score of 0 (step 0) is set equal to zero; i.e., $\gamma_{j0} = 0$ for all items.

In practice, item maps show each item ordered in terms of the underlying value of θ required to answer dichotomously scored items correctly and the value of θ required to obtain at least each score point for multi-point items. Such maps may also contain other data, such as content domain, or other item metadata. It is also possible to show validation data.

In the Bookmark procedure, panelists are typically asked to find an item that a certain percentage of examinees at a critical threshold will be able to answer correctly. The cut score is identified at the point in an ordered item booklet beyond which panelists can no longer say that the target group would have the specified likelihood of answering correctly. The choice of that percentage is critical not only to defining the group of examinees but to defining the threshold between adjacent ability groups. This percentage is commonly called the RP value. In practice, users of the Bookmark procedure have employed 50 percent, 60 percent, 67 percent, and other values. For this project, upon the advice of the Technical Advisory Committee (TAC), RP50 was used.

Solving equation (6.2) for θ produces equation (6.4):

$$\theta = b_j + \ln(1/P_j - 1)/(-1.7a_j) \quad (6.4)$$

where \ln is the natural logarithm and other values are as defined above. For any value other than 50%, the value for $\ln(1/P_j - 1)$ is nonzero. However, when $P_j = .50$, the value of $\ln(1/P_j - 1)$ reduces to $\ln(1)$, which is 0, and the value of θ reduces to the item difficulty b_j , and item discrimination plays no part in the determination of the threshold ability level. Solving equation 6.3 for θ a bit more complex and is explained in Cizek & Bunch (2007, Chapter 10).

Item mapping.

Item mapping allows individual items to be located along the scale score continuum so that interpretations about what students know and can do at individual scale score points may be facilitated. Item mapping is a component in the process of setting performance standards in the Bookmark Standard Setting Procedure (BSSP; Lewis, Mitzel, & Green, 1996). Though item mapping is based in item response theory, it requires the human judgmental process because the response probability (RP; the likelihood that a person answers the item correctly) must be determined in order to align an item with a specific score point.

In addition to purely psychometric information, item maps may also contain item metadata (content standard, depth of knowledge, etc.) and other information. For this project, the contractor developed item maps that contained the content standard to which each item was aligned, the depth of knowledge associated with that item, ability level (expressed in scale score units), and for the grade 11 tests a region corresponding to the college and career scale score levels of the ACT Assessment.

External data.

Some of the items in the OIBs for grades 4, 8, and 11 are not Smarter Balanced items but actually come from other tests such as the National Assessment of Educational Progress (NAEP) and the Programme for International Student Assessment (PISA). These items were embedded in the spring 2014 field test because stakeholders and the public may wish to draw comparisons between the performances of students on the Smarter Balanced tests with the performance of students on other tests.

In addition, for both Math and ELA in grade 11, panelists could see an area of the item map where ACT benchmark scores were projected. These benchmarks are estimates of scores students need to attain on the ACT in order to be considered ready to enter credit-bearing coursework at the postsecondary level.

Facilitators presented and discussed the external data rather briefly. Because many factors differentiate the Smarter Balanced tests from these other assessments, the facilitators maintained the focus of the panelists on the Smarter Balanced ALDs, relevant claims and targets, and the items in the OIBs.

Typical application of the bookmark procedure.

In a typical application of the bookmark procedure, panelists receive extensive training in the content standards, the achievement level descriptors, the test to be reviewed, and the bookmark procedure itself. This training typically takes a day or more. Panelists are then organized into small groups of 5-6 and instructed to review the OIB and place one or more bookmarks in accordance with the training procedures. Each such small group is led by a panelist serving as a table leader. Several such small groups make up a panel of 15 or more panelists, led by a facilitator in addition to the several table leaders. The facilitator provides ongoing instruction and leads discussions between rounds of item review. There are typically two or three rounds of item review.

After training in the bookmark procedure, panelists typically complete a practice round, setting a single bookmark in a very short OIB (usually 6-9 pages) and discuss the experience among themselves with leadership by the facilitator. Once all panelists confirm that they understand the process and the task, they begin Round 1.

In Round 1, panelists review the items in the OIB with a series of questions in mind:

1. *What do you know about a student who responds successfully to this item; that is, what skills must a student have in order to know the correct answer?*
2. *What makes this item more difficult than preceding items?*
3. *Would a student at the threshold have at least a 50% chance of earning this point?*
 - *Yes: Move on to the next item.*
 - *No: Place your bookmark here.*

Panelists then place a bookmark on the first page in the OIB where they believe the student at the threshold for that level would NOT have at least a 50% chance of answering correctly. They complete this task once for each cut score.

After Round 1, bookmarks are tallied and shared among panelists for a given table. Those five or six panelists compare their Round 1 bookmark placements, discuss their rationales and understandings of the threshold student at each level, and review the procedures for placing bookmarks. After this discussion, they answer a brief questionnaire indicating readiness to begin Round 2.

In Round 2, panelists once again review the OIB, this time bypassing pages that clearly did not contribute to bookmark placement, continuing to discuss the contents of the items but placing their own bookmarks, using the same set of guiding questions they used in Round 1, placing a single bookmark for each cut score.

After Round 2, bookmarks are tallied, and a median bookmark for each cut score is calculated. These results are shared with the entire panel, along with impact data – percentages of students who would be classified at each level as well as percentages classified at or above all but the lowest level. Panelists, led by their facilitator, discuss the bookmark distributions as well as the impact data. After the discussion, panelists complete a brief questionnaire indicating their readiness to begin Round 3.

In Round 3, panelists once again review the OIB as in Round 2, but with the knowledge of the impact of their bookmark placements. Each panelist enters a bookmark for each cut score and submits his or her final bookmarks. After receiving the final median bookmark placements and associated impact data, panelists complete a final questionnaire and evaluation form.

Design and Implementation of the Online Panel

The purpose of the online panel was to broaden the input into the process of making decisions about cut scores. In addition, the online panel allowed thousands of people to examine the tests and to express their opinions. The original proposal called for an online panel of 840 individuals; subsequent negotiations increased that number significantly. The final plan called for the contractor to support up to 250,000 online panelists. The intent was to have these individuals review a single ordered item booklet (OIB) and place a bookmark to indicate the location of the Level 3 cut score. The OIBs and support materials were the same as those used in the in-person workshop, but without the extensive training, interaction, and support provided to in-person workshop panelists.

Software development.

MI staff consulted with Smarter Balanced staff create a detailed development schedule defining essential tasks and timelines for the online standard-setting web site. Using the approved requirements documentation, MI developers designed the application, continuing to work closely with Smarter Balanced staff in accordance with the timeline shown in Table 6.1.

Table 6.1 Software Development Timeline.

Software Development Task/Deliverable	Begin	End
Gather requirements/modify application design	2/3/14	3/7/14
Develop online tool	3/10/14	4/25/14
QA application	4/28/14	5/16/14
Receive additional SBAC feedback	5/19/14	5/30/14
Implement changes/make updates	6/2/14	8/1/14
Deploy and field test application	8/4/14	8/15/14
Address issues	8/18/14	9/19/14
Demonstrate for Smarter Balanced	9/22/14	10/3/14
Go live	10/6/14	10/20/14

The basic elements of the system were the home page, item map, and ordered item booklet. The home page contained all instructions, links to external resources (e.g., the Smarter Balanced website to allow panelists to take practice tests), and links to internal resources (instructions on applying the bookmark procedure, Common Core State Standards, and Achievement Level Descriptors). The item map had many features that could be turned on or off, depending on the round and nature of the task to be performed. The OIB contained the items as well as metadata, sample responses, and links to the ALDs.

The home page.

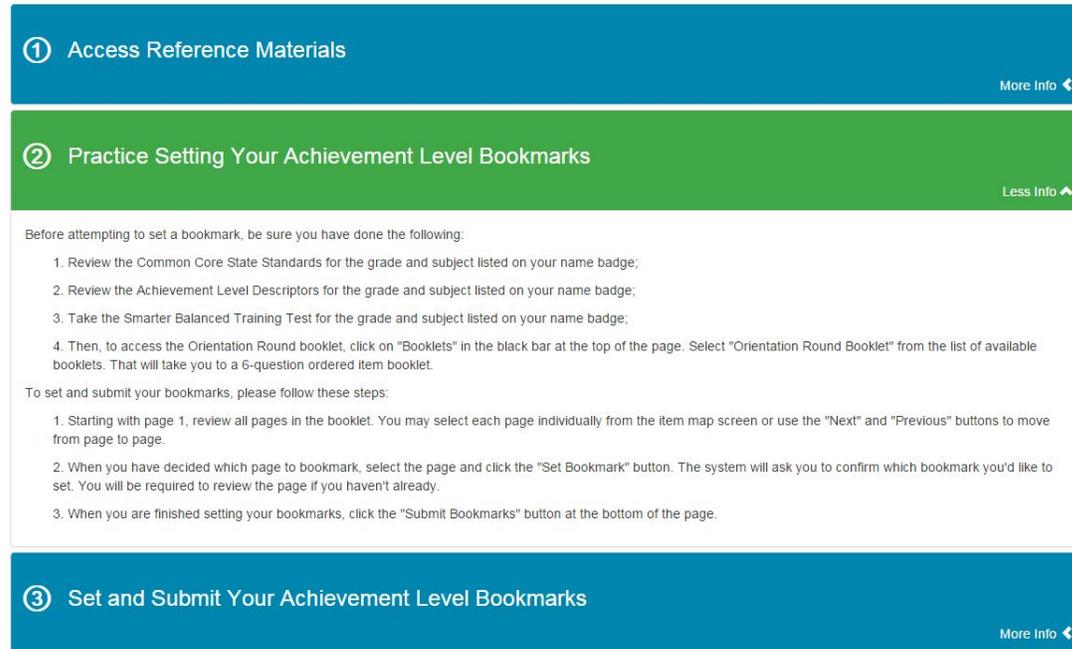
The home page contained all instructions plus links to additional resources. It consisted of four numbered, horizontal bars that could be expanded to reveal detailed information about each step of the process, as shown in Figure 6.1

Figure 6.1. Home Page With One Instruction Bar Expanded.

How to Use the Tool

During the In-Person Panel for Achievement Level Setting, you will need to complete the three steps listed below. To learn more about a step, click on the "More Info" button next to each one.

At certain points in this process, the system will present you with a short questionnaire. You must enter a response to all questions, and submit the questionnaire, before moving on to the next task. You may save your questionnaire at any time by clicking the "Save" button.



① Access Reference Materials More Info ◀

② Practice Setting Your Achievement Level Bookmarks Less Info ▲

Before attempting to set a bookmark, be sure you have done the following:

1. Review the Common Core State Standards for the grade and subject listed on your name badge;
2. Review the Achievement Level Descriptors for the grade and subject listed on your name badge;
3. Take the Smarter Balanced Training Test for the grade and subject listed on your name badge;
4. Then, to access the Orientation Round booklet, click on "Booklets" in the black bar at the top of the page. Select "Orientation Round Booklet" from the list of available booklets. That will take you to a 6-question ordered item booklet.

To set and submit your bookmarks, please follow these steps:

1. Starting with page 1, review all pages in the booklet. You may select each page individually from the item map screen or use the "Next" and "Previous" buttons to move from page to page.
2. When you have decided which page to bookmark, select the page and click the "Set Bookmark" button. The system will ask you to confirm which bookmark you'd like to set. You will be required to review the page if you haven't already.
3. When you are finished setting your bookmarks, click the "Submit Bookmarks" button at the bottom of the page.

③ Set and Submit Your Achievement Level Bookmarks More Info ◀

The home page contained a list of all resource materials, accessible through hyperlinks, as shown in Figure 6.2.

Figure 6.2. List of Resources Accessible From Home Page.

In-Person Panel for Achievement Level Setting 

Instructions Reference Materials **17** Booklets **2** Hello MATH11_1! Log off

Achievement Level Setting References

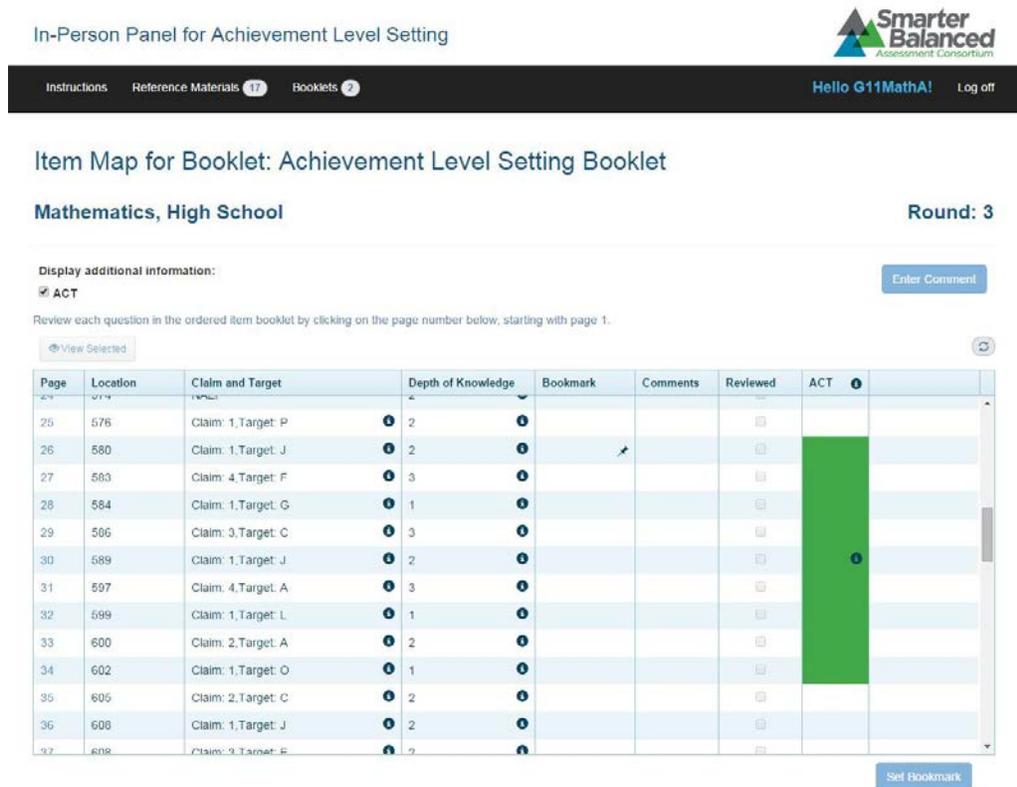
Here is a list of important reference links you may use during the achievement level setting process.

Smarter Balanced Practice and Training Tests Smarter Balanced Practice and Training Tests
Grade 3 English Language Arts/Literacy Achievement Level Descriptors ELA Grade 3 ALDs
Grade 4 English Language Arts/Literacy Achievement Level Descriptors ELA Grade 4 ALDs
Grade 5 English Language Arts/Literacy Achievement Level Descriptors ELA Grade 5 ALDs
Grade 6 English Language Arts/Literacy Achievement Level Descriptors ELA Grade 6 ALDs
Grade 7 English Language Arts/Literacy Achievement Level Descriptors ELA Grade 7 ALDs
Grade 8 English Language Arts/Literacy Achievement Level Descriptors ELA Grade 8 ALDs

The online item map page (see Figure 6.3) allowed panelists to review their progress, navigate through the ordered item booklet pages, access relevant item data, and submit their bookmarks. The Item Map drop down menu allowed panelist to select and view their current results as well as the results from their previous round. Hovering over a comment indicator displayed the comments the submitted for a specific item during a round.

Figure 6.3. Sample Item Map.

In-Person Panel for Achievement Level Setting



Page	Location	Claim and Target	Depth of Knowledge	Bookmark	Comments	Reviewed	ACT
25	576	Claim: 1, Target: P	2				
26	580	Claim: 1, Target: J	2				
27	583	Claim: 4, Target: F	3				
28	584	Claim: 1, Target: G	1				
29	586	Claim: 3, Target: C	3				
30	589	Claim: 1, Target: J	2				
31	597	Claim: 4, Target: A	3				
32	599	Claim: 1, Target: L	1				
33	600	Claim: 2, Target: A	2				
34	602	Claim: 1, Target: O	1				
35	605	Claim: 2, Target: C	2				
36	608	Claim: 1, Target: J	2				
37	608	Claim: 3, Target: F	2				

Each OIB page displayed item specific information including a preview of the item, item statistics, answer key(s), and associated passages and scoring rubrics. Additionally, the OIB page was designed to allow the panelist to make a comment about an item and store that comment for later review. The OIB page included a link to the Achievement Level Descriptor (ALD) for each test. Figure 6.4 shows a sample selected-response item, while Figure 6.5 shows the associated item information page, and Figure 6.6 shows a page for a constructed-response item (in this case, a performance task).

The item map and OIB pages were designed to allow panelists to toggle back and forth. Panelists could gain access to any page in the OIB by clicking that page number in the item map and return to the item map by clicking “Back to Item Map” at the top or bottom of the page. Each OIB page displayed the item, item statistics, rubrics, passages, and sample responses. Additionally, the OIB page was designed to allow the panelist to specify a cut score or navigate to the next or previous OIB page.

All items presented in the OIB were in static, portable data file (pdf) format rather than in interactive format as they had been in the practice tests on the Smarter Balanced website or as administered in the spring 2014 field test. The decision to render items in a static format was based on concerns about the rendering of the interactive versions of items on an uncontrollable array of online panelist devices and browsers. By displaying a static image or PDF of the item, it was possible to ensure that every panelist saw exactly the same rendering of the item for review independent of the platform they were using.

Figure 6.4. Sample OIB Page With Selected-Response Item.

Ordered Item Booklet: Achievement Level Setting Booklet

Mathematics, High School

Page: 06

Round: 2

← Back to Item Map

Set Bookmark Enter Comment

← Previous Next →

Item Question Information Passages and Other Materials 0 Achievement Level Descriptors 1

This item permits calculator use.

12155

The formula for the rate at which water is flowing is $R = \frac{V}{t}$, where

- R is the rate,
- V is the volume of water measured in gallons (g), and
- t is the amount of time, in seconds (s), for which the water was measured.

Select an appropriate measurement unit for the rate.

(A) gs

(B) $\frac{g}{s}$

(C) $\frac{s}{g}$

(D) $\frac{1}{sg}$

← Back to Item Map

Figure 6.5. Item Information Page.

In-Person Panel for Achievement Level Setting

Instructions

Reference Materials 17

Booklets 2

Hello G11MathA!

Log off

Ordered Item Booklet: Achievement Level Setting Booklet

Mathematics, High School

Page: 31

Round: 2

← Back to Item Map

Set Bookmark Enter Comment

← Previous Next →

Item Question Information Passages and Other Materials 3 Achievement Level Descriptors 1

Page	31
Location	597
Claim and Target	Claim 4, Target A
Depth of Knowledge	3
Answer Key	See Passages and Other Material Tab

← Back to Item Map

Figure 6.6. OIB Page For Constructed-Response Item.

In-Person Panel for Achievement Level Setting

Instructions Reference Materials 17 Booklets 2 Hello G11ELAA! Log off

Ordered Item Booklet: Achievement Level Setting Booklet

English Language Arts/Literacy, High School Page: 03 Round: 1

Back to Item Map

Set Bookmark Enter Comment Previous Next

Item Question Information Passages and Other Materials 3 Achievement Level Descriptors 1

62019

Student Directions for Part 2

You will now review your sources, take notes, and plan, draft, revise, and edit your article. You may use your notes and refer to the sources. Now read your assignment and the information about how your article will be scored; then begin your work.

Your assignment:
After completing your research, you share your findings with your teacher. She is impressed with your work. As a final project for your psychology class, everyone must write an article for the Psychology Club's website. Your teacher suggests writing about malleable intelligence, and you decide this is a good idea. The audience for your article will be other students, teachers, and parents.

Using more than one source, craft a thesis to explain the concept of malleable intelligence. Once you have a thesis, select the most relevant information to support your thesis. Then, write a multi-paragraph explanatory article explaining your thesis. Clearly organize your article and elaborate on your ideas. Develop your ideas clearly and use your own words, except when quoting directly from the sources. Be sure to reference the source title or number when quoting or paraphrasing details or facts from the sources.

Explanatory Scoring
Your explanatory article will be scored using the following:

- 1. Organization/purpose:** How well did you state your thesis, and maintain your thesis with a logical progression of ideas from beginning to end? How well did you narrow your thesis so you can develop and elaborate the conclusion? How well did you consistently use a variety of transitions? How effective was your introduction and your conclusion?
- 2. Elaboration/evidence:** How well did you integrate relevant and specific information from the sources? How effective were your elaborative techniques? How well did you clearly state ideas using precise language that is appropriate for your audience and purpose?
- 3. Conventions:** How well did you follow the rules of grammar usage, punctuation, capitalization and spelling?

Now begin work on your article. Manage your time carefully so that you can:

- plan your multi-paragraph article
- write your multi-paragraph article
- revise and edit the final draft of your multi-paragraph article

Word-processing tools and spell check are available to you.

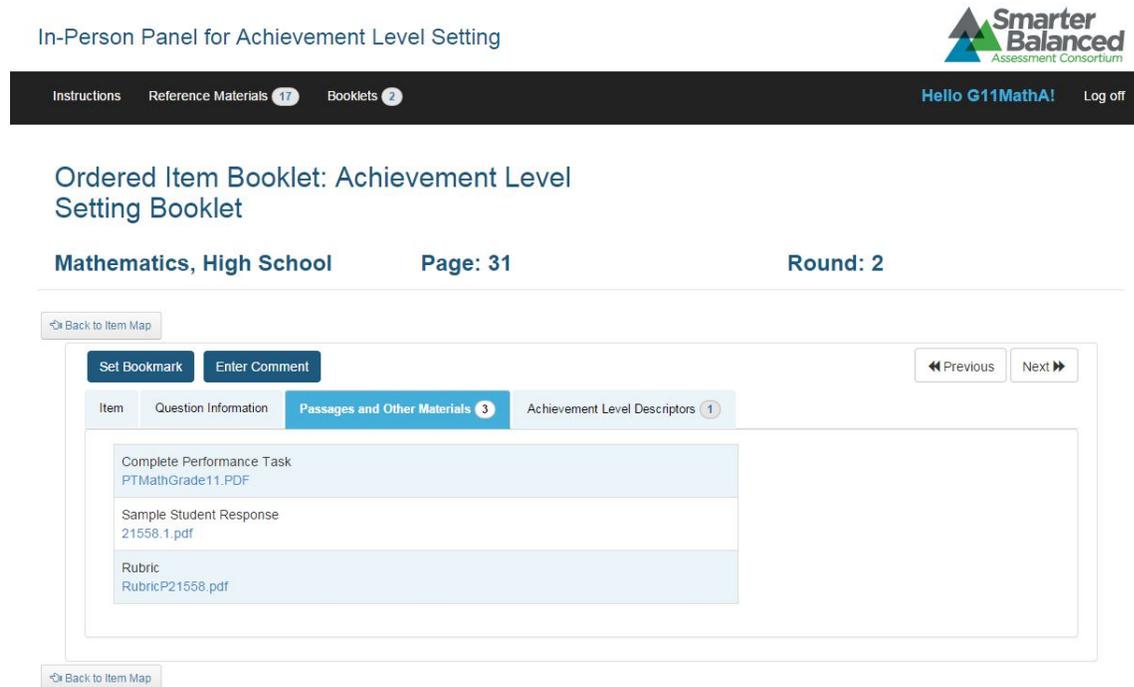
For Part 2, you are being asked to write a multi-paragraph article, so please be as thorough as possible. Type your response in the space provided. The box will expand as you type.

Remember to check your notes and your prewriting/planning as you write and then revise and edit your article.

B I U T_x [bulleted list] [numbered list] [link] [undo] [redo] [bold] [italic] [underline] [text color] [background color] [font size] [font face] [font color] [font style]

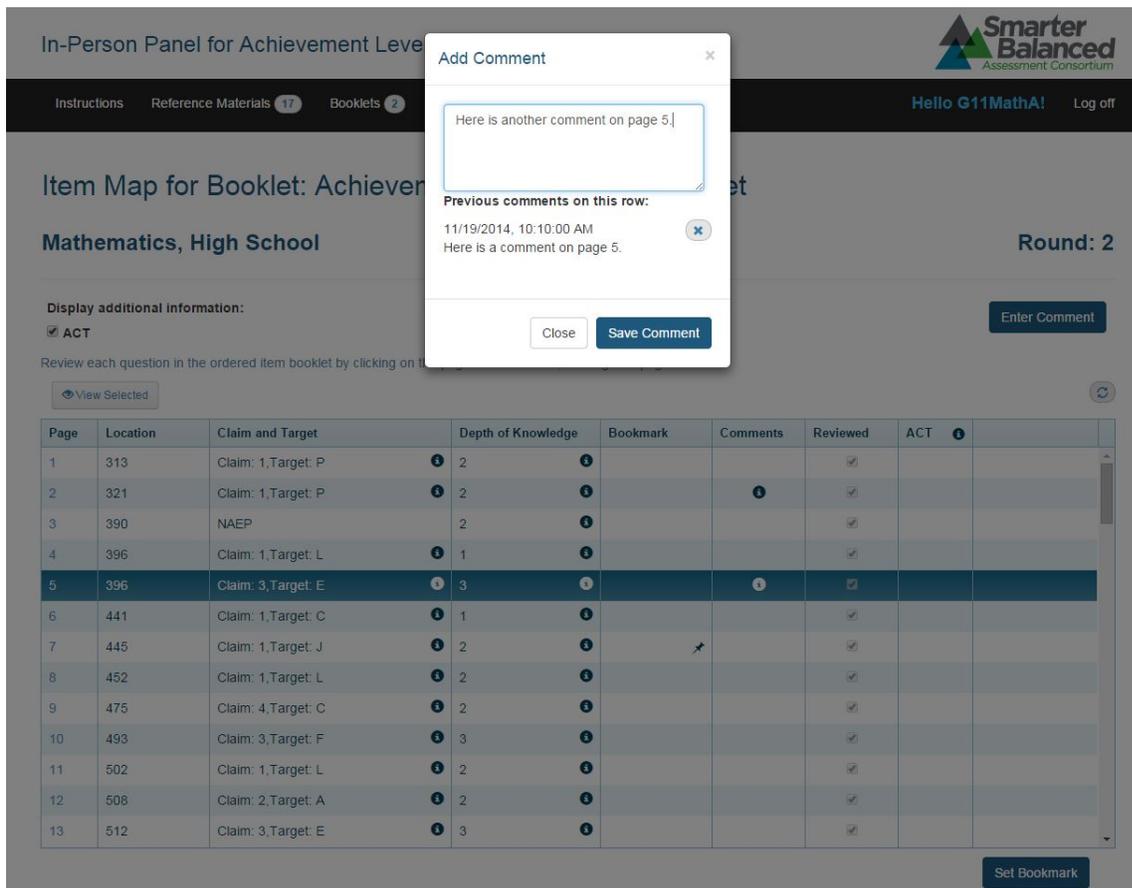
By clicking “Passages and Other Materials,” panelists could see resource materials such as reading or listening passages, sample student responses, and scoring rubrics, as shown in Figure 6.7.

Figure 6.7. OIB Page Showing Links to Performance Task, Sample Student Response, and Rubric.



The system was designed to allow panelists to leave comments on any test item by clicking on “Comments” in the OIB or in the appropriate row of the item map. These comments were intended to be used during inter-round discussions of the items by the in-person panelists or for the online panelists if they needed to leave the task and resume it later. Figure 6.8 illustrates the “Comment” function.

Figure 6.8. Comment.

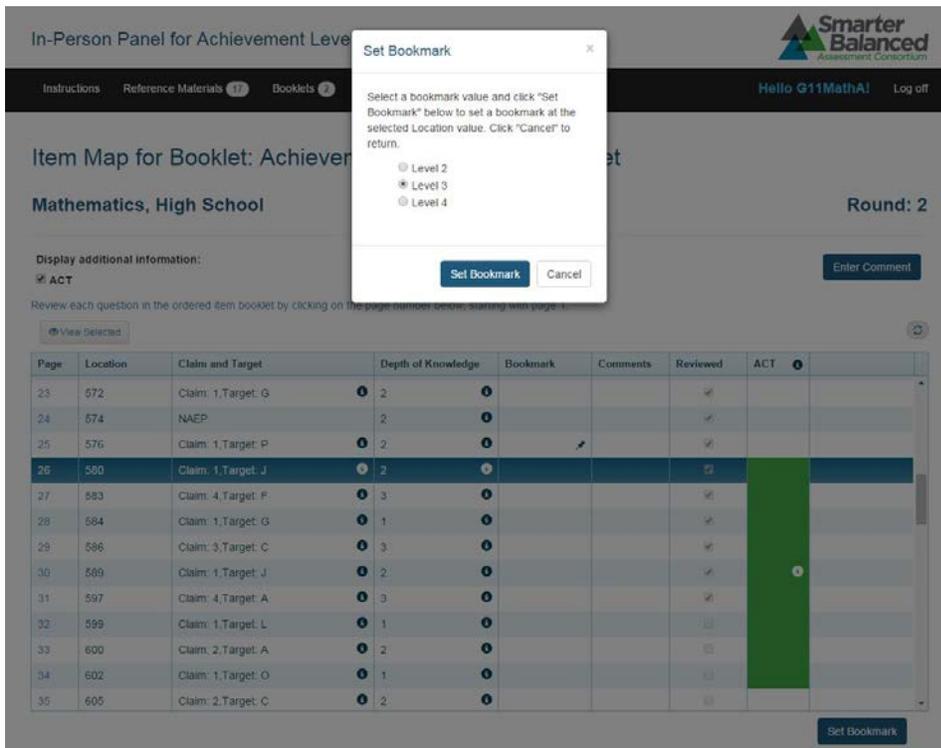


The screenshot shows the 'In-Person Panel for Achievement Level 3' interface. A modal window titled 'Add Comment' is open, displaying a text input field with the text 'Here is another comment on page 5.' Below the input field, it shows 'Previous comments on this row:' with a timestamp '11/19/2014, 10:10:00 AM' and the text 'Here is a comment on page 5.' The modal has 'Close' and 'Save Comment' buttons. In the background, the 'Item Map for Booklet: Achievement Level 3 Mathematics, High School' is visible, showing a table with columns for Page, Location, Claim and Target, Depth of Knowledge, Bookmark, Comments, Reviewed, and ACT. The table has 13 rows, with row 5 highlighted. A 'Set Bookmark' button is visible at the bottom right of the table area.

Page	Location	Claim and Target	Depth of Knowledge	Bookmark	Comments	Reviewed	ACT
1	313	Claim: 1, Target: P	2			<input checked="" type="checkbox"/>	
2	321	Claim: 1, Target: P	2			<input checked="" type="checkbox"/>	
3	390	NAEP	2			<input checked="" type="checkbox"/>	
4	396	Claim: 1, Target: L	1			<input checked="" type="checkbox"/>	
5	396	Claim: 3, Target: E	3			<input checked="" type="checkbox"/>	
6	441	Claim: 1, Target: C	1			<input checked="" type="checkbox"/>	
7	445	Claim: 1, Target: J	2			<input checked="" type="checkbox"/>	
8	452	Claim: 1, Target: L	2			<input checked="" type="checkbox"/>	
9	475	Claim: 4, Target: C	2			<input checked="" type="checkbox"/>	
10	493	Claim: 3, Target: F	3			<input checked="" type="checkbox"/>	
11	502	Claim: 1, Target: L	2			<input checked="" type="checkbox"/>	
12	508	Claim: 2, Target: A	2			<input checked="" type="checkbox"/>	
13	512	Claim: 3, Target: E	3			<input checked="" type="checkbox"/>	

After reviewing items, panelists could enter a bookmark by clicking either on the page in the OIB or in the appropriate row of the item map. Figure 6.9 illustrates the “Enter Bookmark” function. After entering all bookmarks (a single bookmark for Level 3 for the online panel activity or bookmarks for Levels 2, 3, and 4 for the in-person workshop), panelists were prompted to review their work and make sure they were ready to submit their bookmark(s), as shown in Figure 6.10.

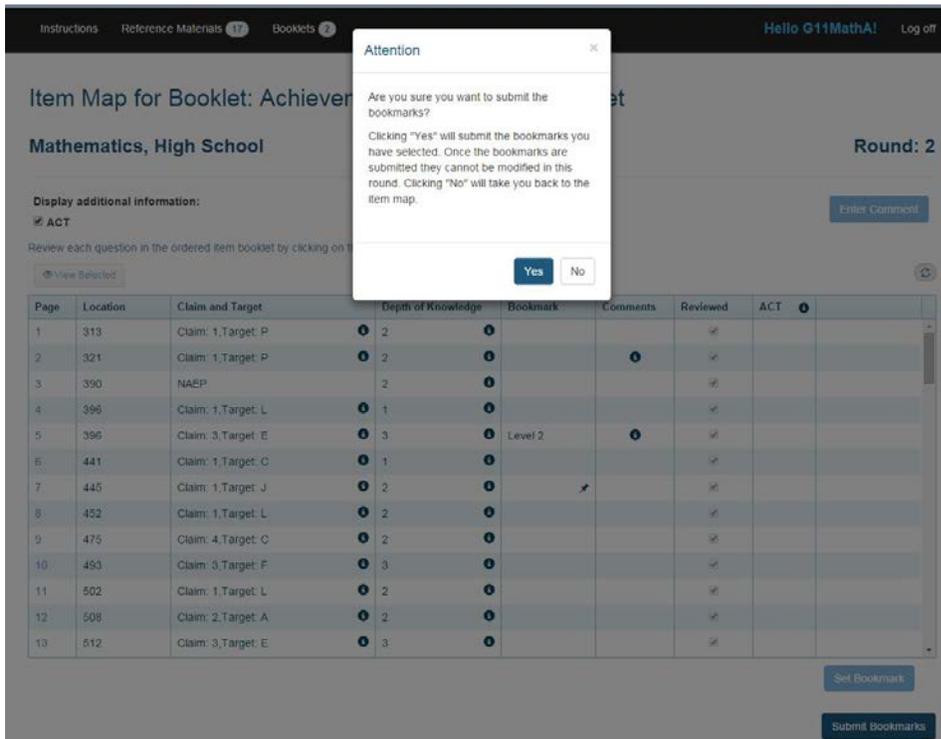
Figure 6.9. Set Bookmark Dropdown Box in the Item Map.



The screenshot shows the 'Set Bookmark' dialog box with the following text: "Select a bookmark value and click 'Set Bookmark' below to set a bookmark at the selected Location value. Click 'Cancel' to return." The options are: Level 2, Level 3, and Level 4. Buttons for 'Set Bookmark' and 'Cancel' are at the bottom.

Page	Location	Claim and Target	Depth of Knowledge	Bookmark	Comments	Reviewed	ACT
23	572	Claim: 1,Target: G	2			✓	
24	574	NAEP	2			✓	
25	576	Claim: 1,Target: P	2			✓	
26	580	Claim: 1,Target: J	2			✓	
27	583	Claim: 4,Target: F	3			✓	
28	584	Claim: 1,Target: G	1			✓	
29	586	Claim: 3,Target: C	3			✓	
30	589	Claim: 1,Target: J	2			✓	
31	597	Claim: 4,Target: A	3			✓	
32	599	Claim: 1,Target: L	1			✓	
33	600	Claim: 2,Target: A	2			✓	
34	602	Claim: 1,Target: O	1			✓	
35	605	Claim: 2,Target: C	2			✓	

Figure 6.10. Submitting Bookmarks.



The screenshot shows the 'Attention' dialog box with the following text: "Are you sure you want to submit the bookmarks? Clicking 'Yes' will submit the bookmarks you have selected. Once the bookmarks are submitted they cannot be modified in this round. Clicking 'No' will take you back to the item map." Buttons for 'Yes' and 'No' are at the bottom.

Page	Location	Claim and Target	Depth of Knowledge	Bookmark	Comments	Reviewed	ACT
1	313	Claim: 1,Target: P	2			✓	
2	321	Claim: 1,Target: P	2			✓	
3	390	NAEP	2			✓	
4	396	Claim: 1,Target: L	1			✓	
5	396	Claim: 3,Target: E	3	Level 2		✓	
6	441	Claim: 1,Target: C	1			✓	
7	445	Claim: 1,Target: J	2			✓	
8	452	Claim: 1,Target: L	2			✓	
9	475	Claim: 4,Target: C	2			✓	
10	493	Claim: 3,Target: F	3			✓	
11	502	Claim: 1,Target: L	2			✓	
12	508	Claim: 2,Target: A	2			✓	
13	512	Claim: 3,Target: E	3			✓	

Online panel activities.

Online panel activities commenced with recruitment (see Chapter 5), which began in April, 2014. Staff of McGraw-Hill Education (CTB), working in concert with Smarter Balanced staff and staff of Hager Sharp (H-S) crafted messages, first for educators and later for the general public, to alert them to the opportunity and explain the logistics.

Staff of Measurement Incorporated (MI) developed the software to support the online experience. That software included a home page, directions, links to reference materials, and digital OIBs, described below. Details of the software development and implementation are included in Appendix E.

Prior to the launch of the online panel on October 6, 2014, MI staff conducted a field test on August 14-15. That activity is described in Chapter 7 and summarized briefly here. Panelists for the online field test were 40 MI readers who logged in to a 30-minute webinar explaining the purpose of the activity and providing a brief introduction to the bookmark procedure. Panelists then had 48 hours to review an OIB for one of four ELA tests (grades 4, 6, 8, or 11) and enter a single bookmark, an activity that was estimated to take about three hours. Most who completed the activity took longer than three hours. Feedback from the panelists was collected via Survey Monkey, analyzed, and used to modify the process for October. Results are presented in Appendix J.

Conduct and results of the online panel.

Online panelists signed up for one of six 48-hour windows, the first of which started on October 6. Ultimately, all windows were extended, and the final date was moved to October 18. By October 6, 10,099 individuals had registered to participate. Of that number, 5,840 logged in, and 2,660 placed a bookmark. Results for online panelists entering a bookmark are presented in Table 6.2. Impact (percent of students who would score at or above the Level 3 cut score) is presented in Table 6.3. Impact is not reported for groups smaller than 25 online panelists. These results were also shared with the in-person workshop panelists and with and cross-grade review panelists.

Table 6.2. Numbers of Online Panelists, by Role, Grade, and Subject

Grade	Teachers		Administrators		Higher Education		Other	
	ELA	Math	ELA	Math	ELA	Math	ELA	Math
3	151	167	67	37	9	5	31	30
4	89	124	31	28	2	4	16	22
5	96	114	31	35	5	5	12	21
6	66	91	11	22	4	8	9	17
7	70	100	12	22	4	5	6	8

	Teachers		Administrators		Higher Education		Other	
8	87	115	27	39	4	7	11	22
11	193	267	55	64	60	83	13	26

Table 6.3. Impact of Online Panel Bookmark Placements: Percent of Students At or Above Level 3

	Teachers		Administrators		Higher Education		Other	
Grade	ELA	Math	ELA	Math	ELA	Math	ELA	Math
3	51%	54%	39%	50%			47%	45%
4	44%	43%	31%	52%				
5	61%	46%	65%	37%				
6	48%	38%						
7	57%	27%						
8	48%	18%	43%	18%				
11	55%	26%	48%	28%	56%	26%	58%	27%

Design and Implementation of the In-Person Workshop

As noted above, the bookmark procedure was used in the in-person workshop. The workshop took place at the Hilton Anatole in Dallas Texas on October 13-19, 2014. There were three waves of panels: the first wave, grade 11, began on Monday morning, October 13 and went through noon October 15; the second wave, grades 6–8, began on Wednesday morning, October 15 and went through noon October 17; the final wave, grades 3–5, began on Friday morning, October 17 and went through noon October 19. Table 6.4 summarizes the numbers of panelists by subject and grade. Table 6.5 summarizes the agenda for each 2.5-day session. Appendix F contains a detailed agenda for each day of the workshop.

Table 6.4. In-Person Workshop Panelists by Subject and Grade

Grade	English Language Arts/Literacy		Mathematics	
	Planned	Obtained	Planned	Obtained
3	1 panel of 30	1 panel of 26	1 panel of 30	1 panel of 30
4	1 panel of 30	1 panel of 27	1 panel of 30	1 panel of 29

5	1 panel of 30	1 panel of 27	1 panel of 30	1 panel of 29
6	1 panel of 30			
7	1 panel of 30	1 panel of 27	1 panel of 30	1 panel of 30
8	1 panel of 30	1 panel of 30	1 panel of 30	1 panel of 29
11	2 panels of 36	2 panels of 34	2 panels of 36	2 panels of 35
Total	252	235	252	247
Grand Total	504	482 (95.6%)		

Table 6.5. High-Level Agenda for Each In-Person Workshop.

Day - Time	Event(s)
Day 1 A.M.	Welcome; overview, training on CCSS, ALDs, tests
Day 1 P.M.	Review of Ordered Item Booklet
Day 2 A.M.	Orientation to the Bookmark Procedure; complete Round 1
Day 2 P.M.	Review Round 1; complete Round 2
Day 3 A.M.	Review Round 2; complete Round 3; evaluate process

Recruitment and selection of panelists.

Recruitment of panelists for the In-Person Workshop began April 15. State K-12 Leads, Higher Education Leads, and Teacher Involvement Coordinators received communication tools developed by the contractor and approved by Smarter Balanced to enable them to recruit teachers (general as well as teachers of English language learners and students with disabilities), school administrators, higher education faculty, business and community leaders, and the general public. Each Smarter Balanced state had 20–25 positions to fill, giving each state an opportunity to have at least one representative for each of the 14 tests.

Preparation of materials.

Staff of MI and CTB prepared the following training materials, which, except for secure documents (so noted) can be found in Appendix F or via the links provided below:

- Introductory PowerPoint presentation to orient panelists to the goals and tasks of the workshop
- Common Core State Standards – up-to-date versions of the subject/grade-specific content standards as well as guidelines to their use in the achievement level setting activity (ELA: <http://www.smarterbalanced.org/wordpress/wp-content/uploads/2011/12/ELA-Literacy-Content-Specifications.pdf> ; Math: <http://www.smarterbalanced.org/wordpress/wp-content/uploads/2011/12/Math-Content-Specifications.pdf>)
- Achievement Level Descriptors – up-to-date versions of the ALDs for the specific subject and grade for each panel (ELA: <http://www.smarterbalanced.org/wordpress/wp->

[content/uploads/2012/11/Smarter-Balanced-ELA-Literacy-ALDs.pdf](http://www.smarterbalanced.org/wp-content/uploads/2012/11/Smarter-Balanced-ELA-Literacy-ALDs.pdf); Math:
<http://www.smarterbalanced.org/wp-content/uploads/2012/11/Smarter-Balanced-Math-ALDs.pdf>

- Practice Test – using the online version of the Smarter Balanced practice tests for each grade and subject (<http://sbac.portal.airast.org/practice-test/>)
- Orientation to the ordered item booklet – PowerPoint presentation designed to show panelists what to look for and questions to ask as they review items in the OIB
- Orientation to the Bookmark procedure – PowerPoint presentation designed to show panelists how Bookmark works and specifically how panelists are to implement the procedure in a computer-based environment
- Bookmark Orientation Round – an exercise involving a 6-page OIB that panelists reviewed prior to entering a single bookmark and discussing their placements in a large-group setting (secure document).
- Readiness Form – a multipart form that asked panelists at several key points during the process how well they understood the process they were implementing and how ready they were to proceed to the next step
- Evaluation Form – a series of statements about the training, environment, and conduct of the workshop that the panelists responded to on a graded scale (such as Strongly Agree to Strongly Disagree)

MI and CTB staff drafted all training materials and submitted them to Smarter Balanced staff and the external auditor for review in advance of the workshop. Final versions of all training materials reflect the comments and recommendations of these reviews and were approved by Smarter Balanced leadership prior to use. All training materials are included in Appendix F.

Training of facilitators and table leaders.

In advance of the in-person workshop, staff of MI and CTB prepared a detailed facilitator script which was reviewed and approved by Smarter Balanced. Staff identified as facilitators studied the scripts and participated in in-house training sessions the week prior to the in-person workshop. In addition, Mr. Ricardo Mercado of CTB conducted a two-hour facilitator training session on Sunday night, October 12, and again on Tuesday night, October 14, and Thursday night, October 16, as facilitators for each wave arrived in Dallas. At the same time, Dr. Jennifer Lord-Bessen of CTB provided a two-hour orientation for table leaders who had been identified in advance by their State Leads. Training materials for those sessions are included in Appendix F

Orientation and training.

Using the training materials approved by Smarter Balanced, MI and CTB staff provided large-group and small-group training. For the opening session, Dr. Joe Willhoft gave the welcome and charge. Dr. Michael Bunch of MI provided specific training on the content standards, ALDs, and practice tests. Dr. Daniel Lewis of CTB provided the orientation to the Bookmark procedure. At the end of each training session, panelists completed a portion of the Readiness Form (see Appendix F).

In-Person Workshop panelists were encouraged to review the appropriate ALDs and CCSS standards prior to coming to the workshop. However, it was not assumed that all had done so, and panelists were given an opportunity not only to review the materials on site but to discuss them in a large-group setting. They had an opportunity to indicate on the Readiness Form just how familiar they were

with those materials. No panelist was permitted to advance to item review without indicating familiarity with the ALDs and content standards and indicating readiness to proceed.

The afternoon of Day 1 was devoted entirely to review of the OIB. In addition to being oriented to the software, panelists were introduced to the test items themselves. They spent the entire afternoon annotating items, using the Comments function of the software, and discussing items with others at their tables in terms of the first two guiding questions. While this activity had been scheduled to end at 5 p.m. on Day 1, all panels required additional time and received from 30 to 60 minutes to complete the task at the beginning of Day 2, following orientation to the bookmark procedure.

At the beginning of Day 2, all panelists assembled in the ballroom for orientation to the bookmark procedure. Dr. Daniel Lewis, Chief Research Advisor at CTB and co-creator of the bookmark procedure, provided the orientation and answered questions. Following the orientation to the Bookmark procedure, panelists adjourned to their small groups to gain first-hand experience in setting a bookmark through a practice exercise. This exercise consisted of a 6-page OIB with items of varying difficulty. Each panel had access to two facilitators who oriented panelists to the computers and software and showed them how to navigate the OIB. Panelists then had several minutes to review the six items and enter a bookmark. The facilitator then led a discussion focusing on how many panelists chose each page to place their bookmarks. Following this discussion, panelists completed a section of their Readiness Forms, indicating their readiness to begin Round 1.

Round-by-round item review and discussion.

Panelists were invited to work through their on-screen OIBs and discuss the items with others at their table. They were able to discuss their opinions with one another at their table as much as they wished, but when they entered a bookmark, it was to be their bookmark, not that of the table. They started by placing a bookmark for Level 3, then Level 4, and finally, Level 2. After placing three bookmarks, panelists were dismissed for lunch, during which time CTB staff tallied bookmarks but did not provide reports to the panelists. Results are shown in Table 6.6 in terms of median bookmark placement for each subject, grade, and level. Complete results, including distributions of bookmark placements, are included in Appendix F

Table 6.6. Results of Round 1 of Bookmark Placement (Entries are Median Page Numbers).

Subject/Grade	ELA			Math		
	Level 2	Level 3	Level 4	Level 2	Level 3	Level 4
ELA 3	16.0	38.0	58.5	22.0	47.0	69.5
ELA 4	20.0	42.0	60.0	12.0	33.0	69.0
ELA 5	13.0	27.0	63.0	21.5	50.0	65.5
ELA 6	15.0	35.0	63.0	18.0	37.5	61.0
ELA 7	16.0	41.0	69.0	21.5	42.5	63.0
ELA 8	19.0	39.5	68.0	18.0	39.0	58.0
EALA 11	21.5	45.0	66.0	19.0	48.5	69.0

Panelists, upon returning from lunch, were directed to share their Round 1 bookmark placements with others at their table, discuss their rationales for placing those bookmarks, and compare approaches as well as comments they had left on the item map. The facilitator then introduced and led a discussion on the bookmark placements of the online panel. Once they completed their

discussions, panelists completed the portion of the Readiness Form that indicated they were ready to begin Round 2.

In Round 2, panelists proceeded as in Round 1, conferring with others at their table but entering their own bookmarks. When they entered three bookmarks and submitted them, they were free to log out for the day. Results of Round 2 are shown in Table 6.7. Complete results, including bookmark distributions and interquartile ranges, are shown in Appendix F.

Table 6.7. Results of Round 2 of Bookmark Placement (Entries are Median Page Numbers).

Grade	ELA			Math		
	Level 2	Level 3	Level 4	Level 2	Level 3	Level 4
3	19.0	38.0	57.5	28.0	49.0	70.0
4	20.0	44.0	63.0	9.0	32.0	71.0
5	14.0	27.0	61.0	20.0	50.5	64.0
6	15.0	36.5	63.0	16.5	37.0	60.0
7	16.0	37.5	69.5	18.0	46.0	61.0
8	17.0	40.5	66.5	17.0	40.0	60.0
11	22.0	42.0	65.0	20.0	50.0	69.0

Panelists returned the morning of the third day to see the results of Round 2. The facilitator led a discussion of the range of bookmark placements, corresponding cut scores, and percentages of students classified at each level, based on the Round 2 cut scores. Once again, the facilitator showed the online panel results, this time in terms of percentages of students at or above Level 3, based on online panelists bookmark placements. A roomwide discussion ensued. Finally, facilitators revealed the impact for the next grade up; i.e., panelists in grade 8 were able to see the final impact of the cut scores set by grade 11 panels, panelists in grade 7 were able to see the Round 2 results for grade 8, and so on down to grade 3. By virtue of being first, grade 11 panelists did not get to see results of any other in-person workshop panels.

After review and discussion of all results, panelists completed the final section of their Readiness Forms and began Round 3. They completed Round 3 as they had Round 2, bypassing many pages which no one had recommended in previous rounds and keeping or changing their bookmark placements depending on their response to the discussion. Each panelist entered three bookmarks and then submitted those bookmarks for analysis. Results of Round 3 are shown in Tables 6.8 (bookmark placement) and 6.9 (scale score cuts and percentages of students at or above each level).

Table 6.8. Results of Round 3 of Bookmark Placement (Entries are Median Page Numbers).

Subject/Grade	ELA			Math		
	Level 2	Level 3	Level 4	Level 2	Level 3	Level 4
3	13.0	33.0	54.0	27.0	47.0	70.0
4	19.0	43.0	62.0	15.0	39.0	71.0

5	11.0	27.0	63.0	19.0	50.0	64.0
6	14.5	34.5	60.5	18.0	45.5	61.5
7	16.0	38.0	66.0	17.0	45.0	61.0
8	18.0	39.5	68.0	16.0	40.0	60.0
11	19.0	42.0	65.0	19.5	48.0	68.0

Table 6.9. Round 3 Cut Score Recommendations: Scale Score Cuts and % At or Above

Test	Level 2 Cut	% At or Above	Level 3 Cut	% At or Above	Level 4 Cut	% At or Above
Grade 3 English Language Arts/Literacy	362	66.5%	427	40.1%	485	19.1%
Grade 3 Mathematics	383	67.3%	436	38.9%	506	10.8%
Grade 4 English Language Arts/Literacy	413	64.4%	470	42.0%	530	18.9%
Grade 4 Mathematics	400	77.6%	470	44.7%	541	15.6%
Grade 5 English Language Arts/Literacy	406	78.7%	450	64.0%	574	16.9%
Grade 5 Mathematics	459	63.5%	532	31.4%	583	13.8%
Grade 6 English Language Arts/Literacy	466	66.6%	527	42.2%	614	12.2%
Grade 6 Mathematics	491	58.3%	561	29.4%	603	15.6%
Grade 7 English Language Arts/Literacy	474	68.2%	547	40.1%	660	6.6%
Grade 7 Mathematics	513	53.1%	609	19.3%	674	5.8%
Grade 8 English Language Arts/Literacy	471	76.4%	543	50.9%	663	10.2%
Grade 8 Mathematics	534	51.3%	605	25.6%	683	7.4%
Grade 11 English Language Arts/Literacy	490	72.9%	565	47.6%	677	12.1%
Grade 11 Mathematics	533	62.6%	644	28.0%	740	8.0%

After entering their Round 3 bookmarks, panelists took a short break and returned to review the final cut scores, complete a final questionnaire, and then evaluate the process, using online evaluation forms. Results of those questionnaires and evaluation Forms are summarized in Tables 6.10 and 6.11. Questionnaire and evaluation results for individual panels are included in Appendix F.

Table 6.10. Round 3 Questionnaire Results: Confidence in Cut Scores Recommended (Discounting Blanks)

How confident are you about the three bookmarks you just entered?

Bookmark	Very Confident	Confident	Uncertain	Very Uncertain	Total
Level 2	222 (47%)	237 (51%)	10 (2%)	0 (0%)	469
Level 3	234 (50%)	220 (47%)	15 (3%)	0 (0%)	469
Level 4	245 (52%)	217 (46%)	7 (1%)	0 (0%)	469

Table 6.11. Summary of Round 3 Evaluation Responses (Discounting Blanks)

Evaluation Statement	Strongly Agree	Agree	Disagree	Strongly Disagree	Total
The orientation provided me with a clear understanding of the purpose of the meeting.	253 (58%)	170 (39%)	13 (3%)	2 (0%)	438
The workshop leaders clearly explained the task.	245 (56%)	161 (37%)	25 (6%)	7 (2%)	438
The training and practice exercises helped me understand how to perform the task.	247 (56%)	174 (40%)	16 (4%)	1 (0%)	438
Taking the practice test helped me to understand the assessment.	231 (53%)	192 (44%)	14 (3%)	1 (0%)	438
The Achievement Level Descriptions were clear and useful.	199 (45%)	216 (49%)	21 (5%)	2 (0%)	438
The large and small group discussions aided my understanding of the process.	300 (68%)	132 (30%)	4 (1%)	2 (0%)	438
The time provided for discussions was appropriate.	230 (53%)	184 (42%)	23 (5%)	1 (0%)	438
There was an equal opportunity for everyone in my group to contribute his/her ideas and opinions.	292 (67%)	135 (31%)	8 (2%)	3 (1%)	438
I was able to follow the instructions and complete the rating tasks accurately.	284 (65%)	151 (34%)	1 (0%)	2 (0%)	438
The discussions after the first round of ratings were helpful to me.	273 (62%)	151 (34%)	12 (3%)	2 (0%)	438
The discussions after the second round of ratings were helpful to me	270 (62%)	156 (36%)	11 (3%)	1 (0%)	438

Evaluation Statement	Strongly Agree	Agree	Disagree	Strongly Disagree	Total
The information showing the distribution of student scores was helpful to me.	220 (50%)	200 (46%)	13 (3%)	4 (1%)	437
I am confident about the defensibility and appropriateness of the final recommended cut scores.	203 (46%)	202 (46%)	27 (6%)	6 (1%)	438
The facilities and food service helped create a productive and efficient working environment.	324 (74%)	104 (24%)	10 (2%)	0 (0%)	438

Data analysis and reporting.

As panelists entered and submitted bookmarks, the data flowed directly from their computers to servers MI had set up prior to the start of the workshop. Staff from CTB, using BookmarkPro software, received the data, analyzed them, and produced reports that facilitators shared at the beginning of the next round. A full set of reports is included in Appendix F.

Design and Implementation of the Cross-Grade Review Panel

The vertical articulation committee was renamed the cross-grade review committee to reflect more clearly the nature of their task, which was to review all cut scores and impact across all grades within a given subject and make adjustments where necessary to prevent or minimize large discontinuities in impact across grades. For example, if 50 percent of students in grades 5, 6 and 8 were at or above Level 3, but only 40 percent of grade 7 students were above Level 3, one might wonder why such a discrepancy existed.

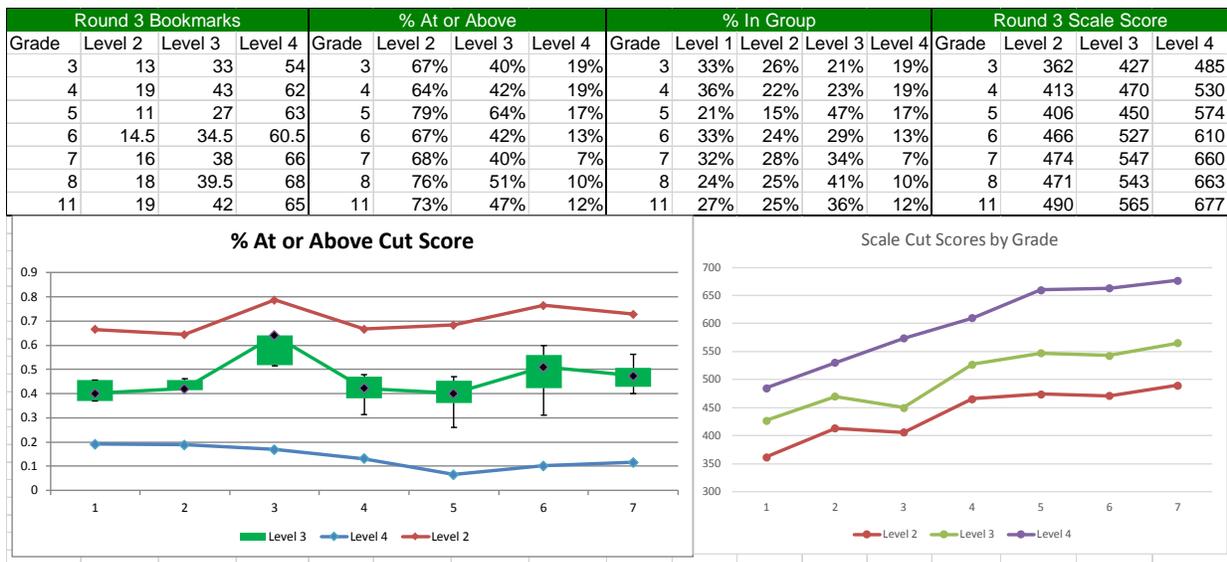
The committees (32 members each for ELA and mathematics) met on October 20, 2014. Dr. Bunch provided an introduction to the tasks and groundrules. The complete presentation is included in Appendix G and is summarized here.

Trends in percentages of students at or above a given level tend to be of one of three types: either more and more students reach a given level over time or across grades (i.e., generally increasing), the same percentages of students reach a given level year after year or from one grade to the next (level), or fewer and fewer students reach a given level over time or across years (generally decreasing). Trends that go up one year and down the next or up for one grade and down the next are much more difficult to explain (though there may be legitimate reasons for such trends). The task of the cross-grade review committee was to investigate any discontinuities and determine whether they were accurate reflections of reality or indications that one or more panels had been overly stringent or overly lenient.

Dr. Bunch explained that the process would include review of actual OIBs and impact data, starting with grades 8 and 11. Any panelist would be welcome to recommend changing any cut score, although a panelist from the grade directly involved or from an adjacent grade would be the preferred initiator of any recommended change. He explained the process for introducing and seconding a motion to change a cut score, to be followed by discussion and a vote. Given that any change would alter the work of a panel of 30 to 36 people, a 2/3 super-majority was required to pass any recommended change.

After the orientation, 32 mathematics panelists reconvened in an adjacent room, while the 32 ELA panelists remained in the room in which the orientation had taken place. In both rooms, computers from the previous week’s in-person workshop were still in place with all software still loaded. For each subject, all seven OIBs and all support materials used by in-person workshop panelists were available. Whenever anyone suggested a change, the facilitator (Dr. Bunch for ELA and Dr. Lewis for mathematics) was able to show on a large screen in the front of the room a projected image of how that change would affect impact. An example of the on-screen graphic is shown in Figure 6.11.

Figure 6.11. Cross-Grade Review Graphic.



In Figure 6.11, the four tables at the top represent the Round 3 median bookmark placements, the percentages at or above Levels 2-4 based on those bookmark placements, the resulting percentages of students classified into each level, and the Round 3 bookmark placements translated into temporary scale scores. The graph on the bottom left reflects the impact in the second table, with the black dots representing the medians, the green boxes representing the interquartile ranges and the black vertical lines representing the 10th and 90th percentiles. Panelists could recommend changing any bookmark placement in the first table, and all other tables, as well as the two graphs at the bottom, would immediately change accordingly.

Panelists began by reviewing cut scores for grades 8 and 11 and then worked their way down through the middle and elementary grades. By the end of the day, the ELA committee had made 8 changes, and the mathematics committee had made 11. Final results for the two committees are shown in Table 6.12, with changes from Round 3 highlighted in yellow.

Table 6.12. Cross-Grade Review Results

Test	Level 2 Cut	% At or Above	Level 3 Cut	% At or Above	Level 4 Cut	% At or Above
Grade 3 English Language Arts/Literacy	362	66.5%	427	40.1%	485	19.1%

Test	Level 2 Cut	% At or Above	Level 3 Cut	% At or Above	Level 4 Cut	% At or Above
Grade 3 Mathematics	381	68.3%	436	38.9%	501	12.1%
Grade 4 English Language Arts/Literacy	413	64.4%	470	42.0%	530	18.9%
Grade 4 Mathematics	413	72.3%	487	36.5%	551	12.6%
Grade 5 English Language Arts/Literacy	434	69.7%	494	47.1%	574	16.9%
Grade 5 Mathematics	459	63.5%	532	31.4%	583	13.8%
Grade 6 English Language Arts/Literacy	453	71.3%	527	42.2%	614	12.2%
Grade 6 Mathematics	491	58.3%	570	26.1%	609	14.0%
Grade 7 English Language Arts/Literacy	474	68.2%	547	40.1%	644	9.5%
Grade 7 Mathematics	513	53.1%	596	23.2%	674	5.8%
Grade 8 English Language Arts/Literacy	482	73.1%	562	43.3%	663	10.2%
Grade 8 Mathematics	534	51.3%	616	22.1%	683	7.4%
Grade 11 English Language Arts/Literacy	488	73.3%	578	42.8%	677	11.6%
Grade 11 Mathematics	565	48.3%	650	26.4%	740	5.8%

Follow-Up Activities

Completion of the cross-grade review was the final step of the in-person workshop, but it was not the final step of achievement level setting. Still remaining were submission of cut scores to Chiefs for review and approval and a final review of achievement level descriptors in the event of changes to the cut scores by Chiefs.

Actions by Chiefs.

Subsequent to the completion of the cross-grade review, Smarter Balanced and contractor staff prepared to present results to the Chiefs for review and approval. On November 6, Chiefs met in Chicago to review the results (see Appendix H). While endorsing the work of the panels, the Chiefs did not vote on the cut scores. A second meeting was scheduled for November 14, in conjunction with the meeting of the Council of Chief State School Officers (CCSSO) in San Diego. Meanwhile, Smarter Balanced staff prepared options to present to the Chiefs at the November 14 meeting (see Appendix H), incorporating evidence from recent studies conducted by the National Assessment Governing Board (NAGB). In addition, Smarter Balanced staff created a new reporting scale, replacing the temporary scale used throughout achievement level setting and cross-grade review. While the temporary scale had a range of 100 to 900, the final scale had a range of 2000 to 3000 and can be easily derived from the temporary scale by adding 2000 to the original scale. Thus, for example, the grade 11 mathematics Level 2 cut score of 565 would translate to a final score of 2565.

The two options presented to the Chiefs at the November 14 meeting consisted of the results shown in Table 6.12 and those same results moderated in the direction of the NAGB results. Specifically, while working within a range of plus-or-minus one standard error of measurement of the cut scores recommended by the cross-grade review committee, Smarter Balanced staff recommended ELA cut scores that were higher and mathematics cut scores that were lower than those recommended by the cross-grade review committee. These modifications kept recommended cut scores within or very close to the one SEM range, approximated NAGB results, and brought ELA and mathematics impacts into closer alignment with each other. The Chiefs voted unanimously (with two abstentions) on November 14 to approve the modified cut scores, presented in Table 6.13.

Table 6.13. Final Cut Scores Approved By Chiefs, With Impact Data.

Test	Level 2 Cut	% At or Above	Level 3 Cut	% At or Above	Level 4 Cut	% At or Above
Grade 3 English Language Arts/Literacy	2367	65%	2432	38%	2490	18%
Grade 3 Mathematics	2381	68%	2436	39%	2501	12%
Grade 4 English Language Arts/Literacy	2416	63%	2473	41%	2533	18%
Grade 4 Mathematics	2411	73%	2485	37%	2549	13%
Grade 5 English Language Arts/Literacy	2442	67%	2502	44%	2582	15%

Test	Level 2 Cut	% At or Above	Level 3 Cut	% At or Above	Level 4 Cut	% At or Above
Grade 5 Mathematics	2455	65%	2528	33%	2579	15%
Grade 6 English Language Arts/Literacy	2457	70%	2531	41%	2618	11%
Grade 6 Mathematics	2473	65%	2552	33%	2610	14%
Grade 7 English Language Arts/Literacy	2479	66%	2552	38%	2649	8%
Grade 7 Mathematics	2484	64%	2567	33%	2635	13%
Grade 8 English Language Arts/Literacy	2487	72%	2567	41%	2668	9%
Grade 8 Mathematics	2504	62%	2586	32%	2653	13%
Grade 11 English Language Arts/Literacy	2493	72%	2583	41%	2682	11%
Grade 11 Mathematics	2543	60%	2628	33%	2718	11%

Review of ALDs.

Given the differences between the cut scores approved by the Chiefs on November 14 and those emerging from the cross-grade review on October 20, a review of the alignment of the final cut scores to the threshold ALDs was deemed in order. Contractor staff translated final cut scores into OIB page numbers and prepared a table showing which final cut scores were within the ranges established by the cross-grade review committee and which were not. Table 6.14 shows those cut scores considered beyond the plausible ranges established by the cross-grade review committee.

Table 6.14. Comparison of Final Cuts to Those Recommended by the Cross-Grade Review Committees.

		OIB Page #			Plausible Range						In Range?			Out of Range by ___ Pages		
		L2	L3	L4	L2		L3		L4		L2	L3	L4	L2	L3	L4
Subject	Grade	L2	L3	L4	From	To	From	To	From	To						
E	3	14	36	57	11.75	15.50	28.00	38.00	53.00	61.75	Yes	Yes	Yes			
E	4	21	44	63	15.00	20.00	37.00	44.00	60.00	63.00	No	Yes	Yes	1		
E	5	21	38	65	10.00	18.00	27.00	37.00	61.00	65.00	No	No	Yes	3	1	
E	6	11	35	61	7.75	19.00	29.00	40.00	52.00	63.25	Yes	Yes	Yes			
E	7	16	39	65	8.00	16.00	34.50	43.50	64.00	74.00	Yes	Yes	Yes			
E	8	22	46	68	14.00	21.50	34.00	46.50	60.00	70.00	No	Yes	Yes	0.5		
E	11	19	46	65	15.25	23.00	40.00	45.00	63.00	66.00	Yes	No	Yes		1	
M	3	26	46	69	26.00	28.00	44.50	53.00	66.00	72.25	Yes	Yes	Yes			
M	4	17	46	72	12.00	18.00	38.00	49.00	71.00	73.00	Yes	Yes	Yes			
M	5	17	49	61	18.25	21.00	50.00	51.00	62.00	64.00	No	No	No	-1.25	-1	-1
M	6	15	40	63	13.00	20.00	32.50	53.50	59.00	63.00	Yes	Yes	Yes			
M	7	9	30	53	13.25	21.00	40.00	51.00	58.75	64.00	No	No	No	-4.25	-10	-5.75
M	8	8	33	50	15.00	18.00	36.50	48.00	57.50	63.00	No	No	No	-7	-3.5	-7.5
M	11	20	44	63	17.50	27.00	44.00	55.25	63.75	69.00	Yes	Yes	No			-0.75

As can be seen, most final cuts are either in range or very close to the plausible range. For those not within the Plausible Range, the final three columns indicate the distance from the edge of the range. Those out-of-range cuts were the primary focus of the ALD review. Cut scores outside these ranges were investigated in the following manner:

1. Final cut scores were translated into OIB page numbers.
2. Two content specialists examined the test items on those pages (items of interest), as well as all items between the item of interest and the edge of the October 20 range.
3. Content specialists noted any discrepancies between the content and skill requirements of the item of interest and the language of the threshold ALD.
4. Content specialists then recommended language changes to the threshold ALD to align with the requirements of the item of interest.

The content specialists reviewed 15 cut scores. In each instance, the item corresponding to the final cut score aligned with the language of the threshold ALD for that level. Therefore, no ALD language changes were recommended. These recommendations were subsequently ratified by Smarter Balanced staff.

Section 7. Field Test

The contractor field tested all critical components of the online, in-person, and cross-grade review process in August 2014, approximately six weeks before the operational events took place. This chapter describes the field test activities, lessons learned, and follow-up actions taken in preparation for the October implementation.

Field Testing the Online Tool

Measurement Incorporated (MI) had proposed a software solution that would permit online presentation of ordered item booklets and support materials and entry of bookmarks to indicate cut scores on Smarter Balanced assessments. Development of the tool is described in some detail in Chapter 6, while field testing and refinement of the tool are described here.

Preparations for field testing.

MI recruited 40 of its readers, most of whom had recently completed scoring of Smarter Balanced field tests and were therefore quite familiar with the tests. The week prior to the online panel field test, panelists, all of whom would be participating in the field test remotely, received an e-mail with hardware and browser specifications, as well as links to the Smarter Balanced website. The 40 panelists were also assigned a grade level and subject: 10 each to English language arts grade 4, 6, 8, and high school.

At the time of field testing, software for both the online and in-person versions of the software were functional and loaded with ordered item booklets consisting only of selected-response items. At that time (August 2014), scoring of open-ended items was still under way, so data from those items were not available. All selected-response items had been scored, however, and their data were available for construction of OIBs. Similarly, since MI had not yet received ACT or SAT comparison data from Smarter Balanced states, no comparison data were available for inclusion in the OIBs.

For the field test, panelists viewed instructions, item maps, and ordered item booklets, as depicted in Figures 7.1 – 7.3.

Figure 7.1. Online Panel Instruction Screen

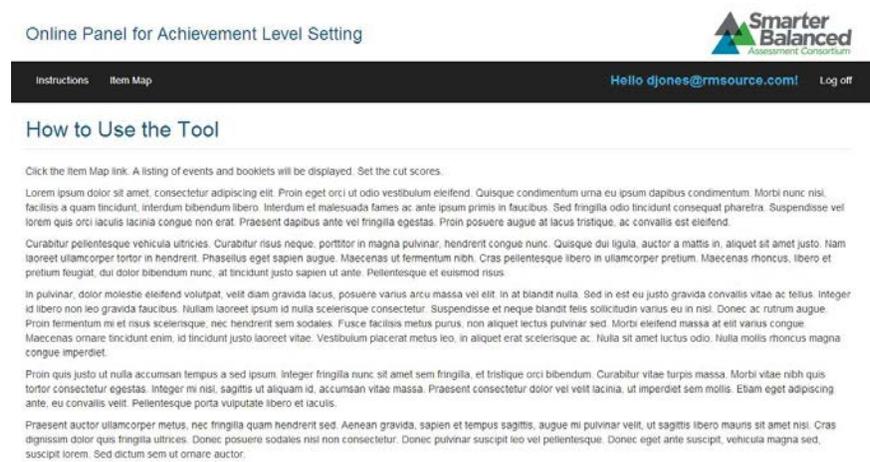


Figure 7.2. Online panel item map

Online Panel for Achievement Level Setting

Instructions Item Map Hello djones@rmsource.com! Log off

Item Map for Booklet: booklet 1

Mathematics, High School Round: 3

display standard comparisons: NAEP PISA ACT SAT Other [Enter Comment](#)

Page	Ability	Standard Code	DOK	Bookmark	Reviewed	Comments	NAEP	PISA	ACT	Other	SAT	At or Above
01	4.14	S.B.123	1		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	98%
02	4.15	S.B.123	1		<input checked="" type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	90%
03	4.16	S.B.123	1	Level 2	<input checked="" type="checkbox"/>				<input type="checkbox"/>		<input type="checkbox"/>	84%
04	4.17	S.B.123	1		<input checked="" type="checkbox"/>		<input type="checkbox"/>				<input type="checkbox"/>	82%
05	4.18	S.B.123	1		<input checked="" type="checkbox"/>			<input type="checkbox"/>			<input type="checkbox"/>	74%
06	4.19	S.B.123	2	Level 3	<input checked="" type="checkbox"/>			<input type="checkbox"/>			<input type="checkbox"/>	70%
07	4.20	S.B.123	2		<input checked="" type="checkbox"/>						<input type="checkbox"/>	65%
08	4.21	S.B.123	2		<input checked="" type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	60%
09	4.22	S.B.123	2		<input checked="" type="checkbox"/>			<input type="checkbox"/>			<input type="checkbox"/>	57%
10	4.23	S.B.123	1	Level 4	<input checked="" type="checkbox"/>			<input type="checkbox"/>			<input type="checkbox"/>	50%
11	4.24	S.B.123	2		<input checked="" type="checkbox"/>						<input type="checkbox"/>	45%
12	4.25	S.B.123	1		<input checked="" type="checkbox"/>						<input type="checkbox"/>	15%

[Set Bookmark](#) [Submit Bookmarks](#)

Figure 7.3. Online panel sample item page

Online Panel for Achievement Level Setting

Instructions Item Map Hello djones@rmsource.com! Log off

Item Map for Booklet: booklet 1

Mathematics, High School Round: 3

Aenon eget dui moutis. Maecenas a felis quam. Maecena ac vehicula justo. Donec dapibus arcu quis lorem ultrices convallis. Pellentesque habitant morbi tristique senectus et netus?

A. 1/8
B. 2/9
C. 2/5
D. 7/12

Key: C

Item info Resource(s) ALD

Ability	4.19
DOK	2
CCSS	sdf

[Enter Comment](#) [Set Bookmark](#)

[Previous](#) [Next](#)

[Back to List](#)

Implementation.

Dr. Bunch prepared a brief webinar which he presented at 1:00 p.m. (EDT) on August 14 (see Appendix I). Dr. Arnold, the contract manager, participated in the webinar as well as the follow-up

activities. After the webinar, participants were directed to the website that contained the materials they would review, listed below.

- How to Use the Tool
- Common Core State Standards (CCSS)
- Threshold Achievement Level Descriptors (ALDs)
- Link to Smarter Balanced Practice Tests
- Setting Achievement Levels Using the Bookmark Procedure
- Bookmark Practice Round Booklet (6 items)
- Ordered Item Booklet
- Item Map

Panelists were instructed to review the materials in the order listed above and complete the final exercise with the ordered item booklet and item map, placing a single bookmark to indicate the Level 3 cut score. Although the presentation ended at 1:30 p.m. (EDT), the chat feature of the webinar remained open until 5:00 p.m. so that MI staff could monitor the questions panelists had about navigation of the system.

MI deployed the Online Achievement Level Setting website software on two Microsoft IIS web servers in a single web farm. An F5 load balancer was used to distribute the users evenly across both web servers. The web servers were connected to a Microsoft SQL Server database cluster for data access. The system was hosted on the internal Measurement Incorporated network. Forty solid state thin client workstations were placed on the same network and configured with a shortcut to the web address.

The field test provided MI's software development staff the opportunity to monitor panelist and facilitator activities as they participated in the bookmark process and identify areas for improvement as well as areas of the software that worked well and required no modification.

Since the Online Panelists were required to self-train on the software, a majority of the issues the Online Achievement Level Setting panelists identified were associated with training material. Based on their input we reorganized the instructional information creating a 4 step process. Each step was color coded and labeled with a step number. Each step was an expandable section detailing the tasks they needed to complete. For example Step 1 described the preparation work they needed to complete before they began the bookmarking process. Step 2 described how to review the orientation booklet and set their level 3 bookmark within the orientation round and so on.

Additionally, the "How to Use the Online Tool" instructional document was updated with additional screenshots and information. Additional hyperlinks were added to the instructional steps providing increased accessibility to the document throughout the instructional process.

Feedback.

At the conclusion of the online panel activity, the 40 panelists received a questionnaire via Survey Monkey (see Appendix I). Responses are summarized in Tables 7.1.

Table 7.1. Summary of Online Panel Field Test Survey Responses

Question	Responses	Count	Percent
Did you have any difficulty logging on to the Measurement Incorporated website?	Yes	2	5%
	No	38	95%
Did you have any difficulty logging on the Smarter Balanced portal to take the practice test?	Yes	10	25%
	No	30	75%
How helpful was the document “Setting Achievement Levels Using the Bookmark Procedure”?	Very Helpful	7	18%
	Somewhat Helpful	17	45%
	Not Very Helpful	10	26%
	Not At All Helpful	4	11%
Did you have any difficulty navigating through the item map?	No	10	26%
	Yes, some	19	49%
	Yes, a great deal	10	26%
Did you have any difficulty navigating through the ordered item booklet?	No	14	36%
	Yes, some	18	46%
	Yes, a great deal	7	18%
How confident are you that the bookmark you placed represents the ability of a student just entering Level 3?	Confident	8	21%
	Unsure	24	62%
	Not At All Confident	7	18%

In addition, online panelists indicated how long it took them to complete preparation (review Common Core State Standards and Achievement Level Descriptors) and complete the OIB exercise (complete the Smarter Balanced practice test, complete the 6-item OIB orientation exercise and the main OIB activity). On average, online panelists spent about an hour and 15 minutes on the

preparation activities and about 3 hours and 15 minutes on the bookmarking activities. Total time for the entire operation ranged from about 20 minutes to over 7 hours, with a median engagement time of about 4 hours and 15 minutes, which may be an underestimate of time required, as only 21 of the 40 panelists actually submitted a bookmark. The comments of the remaining 19 panelists in the open-ended portion of the survey were quite instructive. Some comments and questions indicated that panelists were not using appropriate browsers. MI staff were able to resolve those issues fairly quickly. Other comments and questions revealed shortcomings in the system that needed to be addressed before October. These are summarized in Table 7.2

Table 7.2. Summary of Responses to Online Panel Field Test Survey Open-Ended Questions

Question	Summary of Comments
Did you have any difficulty logging on to the Measurement Incorporated website?	Only two panelists had difficulty, and both reported that the problems were quickly resolved.
Did you have any difficulty logging on the Smarter Balanced portal to take the practice test?	Of the 10 panelists who had difficulty, two received error messages that indicated they were using unsupported browsers. When they installed new browsers, the problems cleared up. Others were confused by the array of practice tests, training tests, and performance tasks. Still others noted that they did not have an event code, which was not required, indicating that they had clicked on the wrong icon at the bottom of the page.
How helpful was the document “Setting Achievement Levels Using the Bookmark Procedure”?	Many panelists found the 6-page document too much information in too little time. Others did not read the document until they had completed two or three steps and found themselves in difficulty. One panelist admitted that the entire process was too complicated and did not complete it.
Did you have any difficulty navigating through the item map?	Most panelists had some or a great deal of difficulty with the item map. The concept itself seemed confusing, and some problems were mechanical (timing out, clicking the wrong icon, etc.).
Did you have any difficulty navigating through the ordered item booklet?	Many were confused by the term “ordered item booklet.” Others never got to that far and so had no comments. Those who got to the OIB seemed to find the directions helpful.

Question	Summary of Comments
<p>How confident are you that the bookmark you placed represents the ability of a student just entering Level 3?</p>	<p>Only about one-fifth of the panelists were confident in their bookmark placements. Since only half actually placed a bookmark, that number represents about two-fifths of those who completed the task, still not a very encouraging figure. There was considerable confusion about the ordering of the items; most felt that the items were not in true difficulty order and were therefore not certain they had placed a bookmark in the right place. Others were confused about the 50% probability, interpreting it as 50% of all students being able to answer correctly, rather than 50% of students at the threshold for Level 3 being able to answer correctly.</p>

In-Person Workshop

On August 18, 40 MI readers gathered at the MI scoring facility in Durham for orientation to the in-person workshop. These readers had been recruited from among MI's Durham scoring staff, with special consideration given to those who had recently scored the Smarter Balanced field test and/or who had some teaching experience. Prior to the workshop, MI and CTB staff had practiced using the software and had met on August 17 for a final planning session in Durham.

Smarter Balanced Staff (Drs. Willhoft, Alpert, Mitchell, and Arnold) were in attendance all three days. Dr. Heather Koons, a member of the audit team, observed the afternoon session of the first day and the morning of the second day.

The field test of the in-person workshop followed the schedule that had been approved for October. Dr. Bunch provided the initial orientation using an adaptation of the PowerPoint that had been approved for October. The adaptations were to acknowledge that the activity was a field test and included only four grades (4, 6, 8, and high school) and one subject (English language arts). All materials used in the in-person field test are included in Appendix I.

Following the orientation, panelists separated into their grade-level breakout rooms, 10 readers per panel. Each panel was led by a CTB or MI staff member who will lead panels in October. In addition, each facilitator had a second CTB or MI staff person as backup. These facilitators led panelists through all preliminary activities including review of the CCSS and ALDs and Smarter Balanced Practice Test.

In the afternoon, Dr. Lewis, co-creator of the Bookmark procedure, provided an overview of the Bookmark procedure. Panelists then returned to their individual breakout rooms for the Bookmark practice round and beginning of Round 1. The first day's session ended at 5:00 p.m.

At 5:30 Dr. Bunch conducted a debriefing with MI and CTB staff. Smarter Balanced staff were also in attendance. The purpose of the debriefing was essentially to acknowledge and discuss any and all procedural issues that did not go smoothly. There had been a few software problems that were quickly resolved. There were also some procedural issues that precipitated changes in scripts and responses. For example, many panelists asked for written copies of the ALDs so they would not have

to toggle back and forth between the OIB and the ALDs. MI staff made copies of the ALDs for the second day's activity.

On August 19, panelists and facilitators returned to their breakout rooms to complete Round 1. This activity concluded at 11:30 a.m., and panelists broke for lunch. During lunch, CTB staff analyzed Round 1 data with BookmarkPro and prepared reports for the Round 2 discussion. Dr. Mercado briefed all facilitators on the results of Round 1 and modeled a brief presentation. Panelists and facilitators returned to their breakout rooms at 1:00 p.m. to discuss Round 1 results and complete Round 2. This session concluded at 5:00 p.m.

At 5:15 p.m., MI, CTB, and Smarter Balanced staff met to discuss the second day's activities. This session was considerably shorter than the previous day's session, as there were fewer issues to discuss. Dr. Bunch solicited from the facilitators and other staff general impressions and specific issues and noted additional changes or updates needed prior to Day 3. The session ended at approximately 6:00 p.m.

On August 20, panelists and facilitators returned to their breakout rooms to discuss the results of Round 2 and complete Round 3. Most panelists completed the activity by 11:00 a.m. CTB staff were able to analyze Round 3 data and produce all reports by noon. Panelists who had been identified as vertical articulation committee (VAC) members were advised to return at noon to begin that activity. All other panelists were advised to return to the main meeting room for a debriefing of the entire process.

Software support.

Before the panelists set a bookmark they were required to complete the Smarter Balanced Practice Test to gain a familiarity with item types and interaction variations. The task of completing the practice test occurred while completing the first panelist questionnaire. During the field test, Measurement Incorporated discovered the questionnaire would timeout while the panelist was completing the practice test. A save feature was added to the software which would allow a panelist to save their questionnaire at any point during the process. If a timeout occurred all of their responses would be retained and displayed when they reopened the survey.

As facilitators followed their scripts and provided instruction to the panelists, it became evident that the placement of some of the panelist questionnaires within the process was hindering their ability to revisit the ordered item booklet and respond to the questionnaire. The software was modified to allow MI to configure a questionnaire at multiple points throughout the process. The modification provided the capability to configure a panelist questionnaire to be presented when they logged in; before they started a round; or when they submitted their bookmarks. This flexibility allowed panelists to review the items; interact with fellow panelists as well as the facilitator; and respond to questionnaires at the appropriate stage in the process.

During the bookmarking process, panelists were asked to add comments to each page of the booklet. The software allowed panelists to make multiple comments and displayed each comment in order by a date time stamp. Throughout the field test multiple panelists requested the ability to delete a comment from a page. A modification was made to the software that allowed them to delete any or all comments for a particular page.

After the field test, both facilitators and panelists were asked what needed improvement as well as what went well. MI addressed the items that needed improvement (as described above) and took

note of what went well to ensure those areas were not impacted by any modifications. Areas identified as effective or easy to use included the process of setting a bookmark, adding comments, and navigating the booklet. Overall, there was positive feedback from panelists on the simplicity of setting a bookmark. The software provided the opportunity to set a bookmark from the item map as well as the ordered item booklet page.

Additionally, panelists were pleased with the ease of accessing item information like DOK, and standard descriptions through the hover over functionality in both the item map and the ordered item booklet pages. Additional information like rubrics and sample responses were accessible through descriptive tabs in the ordered item booklet.

Facilitators responded well to the simplicity of starting and ending rounds as well as viewing the status of each panelist's progress within a round.

Finally, the data team was very pleased with the straightforward method used to export the results between rounds. Their user accounts allowed them to export the round results in a format that could be directly loaded into their system for impact analysis.

Vertical Articulation

Vertical articulation took place the afternoon of August 20. Dr. Bunch provided an orientation to the process at noon (see Appendix I). The focus of the orientation was on the nature and purpose of vertical articulation. In addition, Dr. Bunch explained the groundrules for changing a cut score. Specifically, he explained that all changes would require a motion, a second, and a 2/3 majority vote. He showed examples, projecting an Excel spreadsheet similar to the one shown in Figure 6.5 onto a screen. After a brief question-and-answer period, the vertical articulation activity began at 12:30 p.m., starting with Level 3. Panelists then reviewed Round 3 results for all three levels as well as the Level 3 results of the online panel.

The panel started with Level 3 and focused specifically on the cut score for grade 8, as that cut score resulted in a dip in the trend line from grade 4 to high school. After considerable deliberation regarding the items in the general vicinity of the cut, the panel elected not to change the grade 8 Level 3 cut score. Two other Level 3 cut score changes were considered and rejected.

The panel then turned its attention to Level 4 and quickly decided that no changes were necessary. Finally, the panel turned its attention to Level 2 and modified the cut scores for grades 4, 6, and 8. In all, the vertical articulation committee made three changes, all at Level 2.

During the activity, panelists had access to the four OIBs they had used during the three rounds of achievement level setting. They also had access to all the item maps and other supports that had been available during the in-person workshop.

In general, the vertical articulation went quite well. After three rounds of achievement level setting, panelists were quite familiar with the OIBs, the item maps, and the procedures for identifying cut scores. The instantaneous updating of cut scores and impact each time a new bookmark was entered was well received and kept panelists engaged throughout the process.

Debriefing Session with Panelists

At the same time that the vertical articulation activity was being conducted, all remaining in-person workshop panelists met in an adjacent room for a debriefing. Results of that session are summarized below.

Software issues.

- couldn't see all of the question on screen at once
- don't know which sentences are "clickable" choices on items with sentence selection
- with sentence selection and check box items, you have to unclick if you change your mind, otherwise both (or all) choices stand
- can't erase comments from past rounds, so if your thinking has changed, you have to just add another comment that says to ignore the old one
- too easy to return to list of available item maps (from item map link on menu bar) instead of just returning to the appropriate item map
- placement of page number is too high on each item page, so you have to scroll up to see it
- can we put the standard on the item itself?
- uncomfortable with email address as identification; can there be panelist numbers?
- can they use the "find" function within passages? They're jpgs, not PDFs
- can there be line numbers in the passages?
- for future item development... can we delineate text within an item from the item stem and instructions themselves?
- can answer key be on each item
- can the passage and item screens be presented side to side so that they can look at them at the same time?
- can the "previous" and "next" buttons be in a consistent place on each screen so they can scroll through items?
- Is there a way to get rid of the pages that don't make sense for bookmark placement (i.e., those pages that appear to be anomalies in order of difficulty)?

Presentation/preparatory session usefulness.

- should spend less time on CCSS and more on ALDs so they have opportunity to internalize latter
- it was unclear at the beginning why, specifically, they were there – give more context and more specifics about panelists' roles at the front end
- there needs to be a LOT of clarification (more frequent and more consistent) about where the bookmark goes
- confusion about the ALDs; as threshold ALDs; these describe the abilities of the just barely Level X student
- very difficult for people to understand the 50% concept; goes hand in hand with confusion about setting bookmark

Global observations.

- more time with ALDs needed
- confusion about participant roles
- remind panelists that they will have 3 rounds so not to get too anxious about round 1
- be clear about where/how to place bookmark
- arrange the ALDs so that all Level 2 bullet points are together on one page, all Level 3 bullet points are together on one page, and all Level 4 bullet points are together on one page, since those are the ways the bookmarks are set/reviewed
- More explanation at front end of DOKs and other information – how can/should they be used in the bookmark task?
- put item information on actual item pages

- when walking group through the first few items in the OIB, talk through thought process one would go through about where/how to set bookmark

Debriefing Session with Smarter Balanced Staff

On August 21, staff of Smarter Balanced and MI met for a debriefing on the field test of the online panel, in-person workshop, and vertical articulation. Specifically, Drs. Willhoft, Alpert, and Mitchell of Smarter Balanced and Dr. Bunch of MI met to discuss what had happened, what went well, and what could be improved.

Smarter Balanced staff were highly complementary of the field test process, plan and execution, particularly with MI staff abilities to keep the conversations within panels moving forward. They were also impressed with MI staff skill in debriefing at the end of each day.

It was noted that for operational in-person achievement level setting, a command center would be needed. It was also noted that MI would need to identify and make contingency plans for challenges that might arise, and to include in those plans an escalation process to follow when an emergency arises. As a result, daily debriefings among staff were built into the plan for in-person achievement level setting in October. In addition, the following suggestions were made:

- Lunch hour debrief-opportunity for panelists to “park” questions that can be answered in the last 20 minutes of the lunch hour.
- Lead facilitators to handle any issues from facilitators and escalate policy issue to Dr. Willhoft or Dr. Alpert as necessary and take response back to facilitators to disseminate to panelists
- A list of cell phone numbers for all facilitators, lead facilitators, and Smarter Balanced staff will be available on an as-needed basis to staff in Dallas.
- Concerns about security led to establishment of rules about cell phones, use of social media, and use of paper in breakout rooms in Dallas. These rules were to be further refined after the Advisory Panel meeting in September.
- The use of two facilitators per room (as opposed to the one facilitator per room in the proposal) was hailed as a boon to productivity and continuity, and MI was encouraged to continue the practice in Dallas.
- To avoid confusion in Round 1, there was a suggestion to change the analogy from person vs. person sports (e.g., tennis) to person vs. object (e.g., skiing).
- There was a suggestion to prepare a 2 sentence “script” posted in each room the reiterate the message from opening meeting. This suggestion led to the posting of the two achievement level setting questions (“What skill is required...?” “What makes this item harder...?”) in each breakout room in Dallas.
- There was a suggestion to move the modeling of setting the bookmark to Day 2 morning session for all in ballroom so panelists could be more familiar with the process.
- Dr. Willhoft suggested an emotional drain-off-discussion to allow panelists to raise issues unrelated to achievement level setting but nonetheless important to them and potentially harmful to the overall process if not addressed. Details of the process were to be worked out prior to the meeting in Dallas.

Lessons Learned – Preparations for October

The purpose of the field test was to determine what worked, what did not work, and how to fix those things that did not work and institutionalize or improve those things that worked well. There were four major findings:

1. The online panel will be a challenge. The task is complex, the opportunity to provide instruction and feedback is minimal, and online panelists are likely to become frustrated and fail to finish the task. Therefore, MI prepared a condensed version of the bookmark training materials, expanded some of the instructions on the home page and made them much more explicit, prepared a series of frequently asked questions (FAQ), and made arrangements with intermediaries (e.g., higher education leads) to funnel questions from online panelists to MI staff. As a final measure, the windows for the online panel were extended from two days to three, to allow everyone time to complete the exercise.
2. The software for both the online panel and the in-person workshop works quite well. There were a few bugs, but they were quickly eliminated due to the constant presence of information technology staff. Their contribution to the field test clearly indicated the need for their presence on site in Dallas, and it was quickly determined that IT staff would remain on site throughout the in-person workshop in October.
3. There was considerable variation among facilitators. Although there had been a formal script, different facilitators used it differently. For October's in-person workshop, it was determined that the script should be condensed to a few pages of bullet points with back-up information in an appendix. In addition, all facilitators should undergo common training. This recommendation translated into formal in-house training at both CTB and MI the week prior to the in-person workshop in Dallas and additional two-hour briefing sessions the night before each wave of the event (i.e., October 12, 14, and 16).
4. The vertical articulation worked quite well and required no modification. This was no surprise inasmuch as the activity replicated in terms of both software and procedure similar activities MI had carried out many times before.

Section 8. Management Activities

Management activities included meeting facilitation/attendance and reporting. The achievement level setting effort and related communications required many meetings with Smarter Balanced leadership, advisory committees, and other contractors. Internal meetings within the Measurement Incorporated-CTB/McGraw Hill Education-Hager Sharp Collaborative were also an important part of coordinating the achievement level setting. Management meetings are summarized in Table 8.1.

Table 8.1. Management Meetings

Meeting	Frequency	Role and Number of Meetings
Leadership	Weekly	Facilitated 30 meetings
Cross-Contract	Twice weekly until October, then once per week	Attended 41 meetings
Functional Work Group	Once or twice per month	Facilitated 11 meetings
Communications Committee	Twice per month, then weekly	Facilitated 15 meetings
TAC	On request	Attended 6 meetings
Advisory Panel	July, September, and October	Facilitated 3 meetings
K-12 and/or Higher Ed Leads	On request	Attended 7 meetings
Online Panel Scrum	As needed	Facilitated 4 meetings
Contingency Planning (Ebola)	As needed	Facilitated 2 meetings
OIB Review	As needed	Facilitated 20 meetings
Collaborative Meetings (Measurement Incorporated, CTB/McGraw Hill, Hager Sharp)		
Recruiting (MI, CTB, HS)	Weekly	Facilitated 25 meetings
Item data/Preparation (MI, CTB)	Weekly	Facilitated 27 meetings, including two in-person

Meetings

Weekly calls.

- Leadership: Smarter Balanced, Measurement Incorporated, CTB/McGraw Hill Education, and Hager Sharp leaders, along with the contract manager and project management partner, met to discuss issues and make decisions to keep the project on track and moving forward. Additional participants were invited as needed.
- Cross-contract: Measurement Incorporated and CTB/McGraw Hill Education staff attended the Consortium's cross-contract calls from April through November. These meetings enabled staff to discuss such topics as data handoffs, standard setting sample size, and the ordered item booklet approval schedule with other contractors and Smarter Balanced staff.

Twice-monthly calls.

- Functional Work Group: Members of the TD&V and the contract manager met with Measurement Incorporated twice per month in the early months of the project to offer advice on plans and to review materials and documents such as the Contingency Communications Plan. Later in the project the work group convened once per month.
- Communications Committee: Smarter Balanced's Director of Communications met with Measurement Incorporated, Hager Sharp, and the contract manager twice monthly to discuss and plan communications related to the achievement level setting. Immediately after the in-person workshop, the group began meeting weekly to address communications related to the cut scores and Chiefs' meetings.

Collaborative Calls and Meetings

- Recruiting: Measurement Incorporated, CTB/McGraw-Hill Education, and Hager Sharp met weekly via webinar to plan and develop recruiting materials for the online panel, in-person panel, and vertical articulation committee, and to coordinate distribution of the messages.
- Item data/Preparation: Measurement Incorporated and CTB/McGraw-Hill Education met weekly via phone and twice in person to coordinate data handoffs and analysis related to constructing the ordered item booklets and to capturing the achievement level setting in-person panel results. This time was also used to plan and produce the facilitator scripts.

Technical Advisory Committee

Measurement Incorporated was asked to present information related to achievement level setting at five TAC meetings, and to prepare materials for a sixth meeting to which MI was not invited.

Advisory Panel

The Achievement Level Setting Advisory Panel, which was initially called the Task Force, met twice in person and once via webinar to offer a broader perspective on standard setting than that provided by the established TAC. The panel advised Smarter Balanced and Measurement Incorporated on a variety of psychometric and policy issues related to the design and conduct of the standard setting and related tasks.

The first Advisory Panel meeting was held on the afternoon of July 17 and the morning of July 18, 2014, in Portland, Oregon. The group discussed the Achievement Level Setting Plan and

Supplement; plans for the August field test of the online panel, in-person workshop, and VAC; online panel recruitment; and plans for work beyond the contract's end, among other topics.

The second Advisory Panel meeting was held on September 22, 2014, in Chicago, Illinois. Topics of discussion included a review of the August achievement level setting field test and subsequent changes; the final report outline; a review of the Follow-Up Plans Proposal; and plans for the October achievement level setting, among others.

The final Advisory Panel meeting was a webinar held on October 28, 2014. The panel received updates about the achievement level setting. The results of the online panel, in-person workshop, and VAC were discussed, along with the auditors' report. The panel offered guidance about how to most effectively present the results to the Chiefs.

Other Meetings

The Collaborative prepared materials for and attended meetings as requested by Smarter Balanced. Explaining recruiting materials to the K-12 and Higher Ed Leads is an example.

Circumstances sometimes required impromptu meetings between the Collaborative and Smarter Balanced. Examples include the Online Panel Scrum calls to discuss strategies for increasing participant registration, and the Contingency Planning calls to consider potential effects of the Dallas Ebola case on the in-person workshop and VAC.

The ordered item booklet review process required numerous meetings to facilitate approval on a tight schedule. Measurement Incorporated hosted planning calls with Smarter Balanced staff, office hour webinars to enable Smarter Balanced staff to discuss OIB issues with the Executive Director, and webinars to present the final OIBs to the Executive Director for approval.

Reports

Weekly reports included the Weekly Status Report and the Weekly Issues and Action Items Log. Since Leadership calls occurred weekly, recaps of those meetings were also submitted each week.

Interim reports included the Project Schedule, the Deliverables Matrix, the Milestone Tracker, and the Project Communication Plan for Two Months Out. These reports were updated monthly.

Additional reports included recaps of the Functional Work Group and Communications Committee calls, as well as recaps of the Advisory Panel meetings and webinar.

References

- Cizek, G. J. & Bunch, M. B. (2007). *Standard Setting: A Guide to Establishing and Evaluating Performance Standards on Tests*. Thousand Oaks, CA: Sage.
- Egan, K. A., Schneider, C., & Ferrara, S. (2012). Performance level descriptors. In G. J. Cizek (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations* (2nd Ed.), New York: Routledge.
- Lewis, D. M., Mitzel, H. C., Mercado, R. L., & Schultz, E. M. (2012). The bookmark standard setting procedure. In G. J. Cizek (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*. New York: Routledge.
- Lewis, D. M., & Mitzel, H. C. (1995). An item response theory based standard setting procedure. In D. C. Green (Chair), *Some uses of item response theory in standard setting*. Symposium conducted at the annual meeting of the California Educational Research Association, Lake Tahoe, NV.
- Lewis D.M., Mitzel, H. C., Green, D. R. (1996). *Standard Setting: A Bookmark Approach*. In D. R. Green (Chair), *IRT-Based Standard-Setting Procedures Utilizing Behavioral Anchoring*. Symposium presented at the 1996 Council of Chief State School Officers 1996 National Conference on Large Scale Assessment, Phoenix, AZ.
- Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The Bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting Performance Standards: Concepts, Methods, and Perspectives*. Mahwah, NJ: Lawrence Erlbaum.

Appendix A – Communication Plan and Related Documents

• A01_Communications Plan.....	100
• A02_Focus Group Report.....	129
• A03_Achievement Level Setting Presentation.....	170
• A04_Communications Toolkit.....	185
• A05_Contingency Plan.....	189
• A06_On-Site Interview Intake Form.....	222
• A07_Panelist Signup Form.....	223
• A08_Multi-Media and Interview Release Form.....	224
• A09_Level Naming Survey Results.....	227
• A10_Scale Scores.....	244
• A11_Nov 6 Media Briefing.....	245
• A12_November Media Briefing.....	265
• A13_Achievement Levels Presentation.....	285



Smarter Balanced Assessment Consortium: Communications Plan for Achievement Level Setting

Prepared by Hager Sharp

UPDATED May 27, 2014



Table of Contents

Preface: Engagement Calendar of Key Activities	102
Situation Analysis	104
Audience Analysis	106
Insights from Informal Parent Focus Group	106
Asset Analysis.....	109
Key Messages	110
Audience Specific Messages.....	112
Messaging for Pre-K-12 Educators and Administrators.....	112
Messaging for Parents	112
Messaging for Higher Education.....	113
Messaging for Business.....	113
Phase 1 (April–September).....	113
Strategic Goals	113
Recruitment Goals	114
Schedule	114
<i>Communications Packet: Turnkey Materials for States Pt. 1</i>	114
Strategies and Tactics for Recruitment and Outreach.....	116
Owned Media	116
Earned Media.....	116
Social Media	118
Partner Outreach.....	119
Phase 2 (September–November)	121
Strategic Goals.....	121
Communications Packet: Turnkey Materials for States Pt. 2.	121
Strategies and Tactics for Promoting the Achievement Levels.....	122
Owned Media	122
Earned Media.....	122
Social Media	123
Grassroots Engagement	124
Outreach for Spanish Speaking Audiences.....	124
Spanish-Language Messaging.....	124
Project Conclusion	125
Appendix I: Partnership Targets	126

Preface: Engagement Calendar of Key Activities

April

- Develop and test initial materials and messages
- Conduct initial outreach to teachers, higher education faculty and business community through Governing States, partnerships, and channel marketing
- Begin soliciting registration for Online Panel
- Attend Collaboration Conference in Minneapolis
- Send *SmarterNews* blast with news item to support recruitment
- Draft communication plan

May

- Finalize communication plan
- Develop recruitment toolkit for Online Panel
- Develop infographic to explain Achievement Level Setting process
- Conduct message testing/validation with online focus group of parents, teachers and administrators

June

- Kick off parent outreach for Online Panel through webinar/outreach to National PTA and other parent groups
- Deliver recruitment toolkit to states
- Deliver targeted versions of recruitment toolkit to K-12, parent, higher education and business community groups
- Submit crisis communication plan
- Spanish language translation/adaptation plan submitted
- Message training webinar for public information officers/state leads on focus group findings
- Send *SmarterNews* blast with news item to support recruitment

July

- Send *SmarterNews* blast with news item to support recruitment
- Promoted posts through Facebook and Twitter
- Webinar/webcast briefing with trade media, partners and allies on achievement level setting
- Ongoing outreach to K-12, parent, higher education, business and advocacy groups
- Placement of blog posts and newsletter items in support of recruitment

August

- Send *SmarterNews* blast with news item to support recruitment
- Crisis communications drill/media training
- Paid media/promoted posts to support Online Panel (contingent on recruitment)
- Presentations to the Technical Advisory Committee and Governing States

September

- Back-to-school: public recruitment blitz for Online Panel
- Message training webinar for public information officers/state leads on media coordination during In-Person, VAC and Online Panels
- Send *SmarterNews* blast with news item related to achievement levels

- Turnkey materials on achievement levels (cut score focus) delivered to states

October

- Media coordination during month of achievement level setting panels
- Send *SmarterNews* blast with news item related to achievement levels
- Recommendations for speaking tour

November

- News conference to announce achievement level scores
- Op-ed in major news outlet related to new achievement levels
- Weekly news blast announcing results of achievement level results and cut scores
- Video package released on developing achievement levels

The plan proposes an overarching communications strategy that includes a reliable means of identifying the needs of a range of audiences (i.e. educators, higher education faculty and leaders, business and community leaders, policymakers, general public) and researching key messages for primary stakeholders groups. Considerable work has already gone into developing messages and materials for Smarter Balanced; to leverage that existing work, we have begun suggesting revisions to existing documents and developing new materials to address current needs and assist with recruitment efforts.

This plan identifies communications challenges confronting Smarter Balanced, specifically as it relates to developing achievement level scores this fall, and proposes an overarching strategy for addressing those challenges. Through a mix of communications tactics, we will help the Consortium navigate confusion and potential sensitivity about the achievement level setting process and the rollout of the new assessments. As we currently envision it, the communications plan will:

We propose a communications plan comprised of two phases:

Phase 1 will focus on us building awareness about the achievement level setting process with each of our target audiences. Following the launch of recruitment for the Online Panel, we will begin a series of focused outreach activities. The goal of this effort will be to create targeted awareness through a proactive media strategy delivering to major news outlets and begin the process of educating members of the media, key stakeholders, and allies about the achievement level setting process.

Phase 2 will focus on synchronizing messaging across the Consortium and explaining the process to key audiences. Beginning with recruitment and through the final vote on achievement level score recommendations, we will manage expectations regarding the anticipated drops in score points as a result of the new levels.

Additionally, through the development and implementation of a separate crisis communications plan, we will—and already have begun to—identify challenges specifically related to the achievement level setting process. The contingency plan for crisis communications will establish protocols and messages for responding to key audiences and through a variety of channels, and it will outline plans for providing training and preparation to key personnel at Smarter Balanced.

We spent the first period of our engagement with Smarter Balanced and Measurement Incorporated learning about the technical process for setting achievement levels. For the benefit of reviewers, we recapped discussion points from our SWOT (strengths, weaknesses, opportunities, and threats) analysis and included findings from an informal parent focus group.

Situation Analysis

In collaboration with the Consortium staff, Hager Sharp analyzed the relevant internal strengths and weaknesses as well as external opportunities and threats facing the Consortium. In close coordination with Measurement Incorporated, the Smarter Balanced team, and other partners, we held a series of startup meetings that allowed us to inquire about the current state of your work, evaluate your current communications materials, look to the challenges ahead, and determine specific action steps and projects that will help you succeed in communicating with greater clarity about the new assessment and achievement levels.

This information will help us identify how Smarter Balanced can achieve its operational goals in developing cut scores for each assessment.

The table below illustrates some of the factors we considered when developing the ideas in this plan.

SWOT Analysis

Strengths	Weaknesses
<ul style="list-style-type: none"> • Inclusive process • Tried and true • Research-based • Transparent • Built on what we've learned over years • K-12 and higher education leaders join forces • Accessible by more students than ever before • 508-compliant, language aides • Built by states, for states • Public sector halo—not commercial • Broad coalition • Test is comparable in time length to some of the longer tests currently in states • Not just a test; it's an assessment system • Cost savings in the aggregate • Consistency across states; bigger data set • “Adaptive” test—challenging for most 	<ul style="list-style-type: none"> • State affiliates vary in communications ability and willingness • Political savviness varies across states • Testing time requires significant commitment • Six hours for elementary; eight hours for high school • Perceived loss of local control • Federal funding = perception of federal involvement • Reliance on local technology, which varies by location • Broadband issues • Technical capabilities of younger students—can they reliably/accurately complete online assessment?
<ul style="list-style-type: none"> • Leverage higher education experts as thought leaders • Develop playbook for states, plus collateral materials to strengthen communications process and make it more consistent • Educate staff on education committees in state legislatures • Give states good arguments against commercial rivals • Research comparing to ACT may demonstrate Smarter Balanced assessment is superior • Conduct “gut check” commitments with states • “What if” scenarios: need to make sure audiences understand what could happen with “score cliff,” etc. • Use Kentucky model for seeding environment • Showcase testimonials from participants • Connect with trades community, trade unions, etc. to generate support from non-college community • Connect with minority organizations to generate support (e.g., La Raza, etc.) • Online discussion board • Champion the people who participate in the standards setting, and leverage their testimonials 	<ul style="list-style-type: none"> • Commercial testing products looking for weaknesses will try to discredit Smarter Balanced assessment • South Carolina, Alabama moving to ACT—will this cause chain reaction? • Concern around rigor of test • Defending subjectivity in the process • Managing the Online Panel process: Rogue panelists/gate crashers/managing outliers • States decide not to use Smarter Balanced cut scores • How do scores compare to SAT, others? • “Losers” in “score cliff” → will we lose students through lack of confidence? • Will the assessment exacerbate inequities? Middle-upper class parents can afford tutoring, disadvantaged parents cannot • Local education leaders need support, cover—they may be blamed for “score cliff” • We need to be prepared for all communications to go public—opponents taking things out of context

Opportunities

Threats

Audience Analysis

Building on our conversations in the kickoff meeting and follow-up meeting with CTB/McGraw Hill, we have begun the process of conducting thorough research on public attitudes toward the achievement level setting process and possible outcomes from that process.

To clearly communicate with audiences concerned with achievement level setting and the Smarter Balanced assessments in general, we need to put those audiences at the center of all activities and outreach. Because parents have such a large stake in testing and score results, it is critical that we begin with a formal message testing and audit of the audience to understand their attitudes and preferences. We have submitted a formal focus group proposal to key personnel on Contract 21. We plan to gather detailed feedback on messages and materials for the upcoming Smarter Balanced achievement level setting activities using cost-effective online focus groups among parents, teachers, and school administrators from two Smarter Balanced Governing States: California and Michigan. Feedback from the focus groups will allow us to integrate relevant insights from parents into materials.

Insights from Informal Parent Focus Group

To inform the planning process and development of materials, the Hager Sharp team gathered very preliminary insights and reactions from parents regarding the opportunity to participate in the Online Panel. We conducted an informal 60-minute in-person focus group session among eight parents of public school children in Maryland; Virginia; and Washington, DC. While we understand they do not represent Smarter Balanced states, we believe their reactions to messaging and their general thoughts reflect many of the questions and comments we can expect from a broader group of parents.

Questions and Concerns about Common Core and Assessments

All participants said they were familiar with the Common Core State Standards (CCSS), and some understood it to be “the replacement for *No Child Left Behind*.” Several parents commented that they were aware of debates, skepticism, and negative comments about CCSS.

Participants were not aware of the process for developing an assessment aligned with CCSS, and they were not familiar with Smarter Balanced efforts specifically. They had heard a lot in general about testing, but they complained about “mixed messages and ambiguous messages from schools.” They had no clear idea of timing for the rollout of an assessment to measure CCSS. They also expressed confusion about state and county assessments in relation to other assessments: “As a parent, it is unclear how all these assessments are related.”

In discussing testing in schools, parents also complained about technological challenges. Several parents said assessments had been delayed or had to be conducted a second time because of technological difficulties. They were concerned about the “waste of time” caused by technical problems or a lack of policy planning, especially as they see the school year tightly packed with little margin for error.

One parent expressed concern regarding the experience in New York, “where all the students had to take the CCSS assessment and they all failed.” Parents were concerned about perceptions of school quality related to declining test scores (i.e., the “score cliff”).

Reactions to Opportunity to Participate in Online Panel

After describing the online standard setting opportunity with an invitation for them to participate, parents expressed several questions and concerns about the opportunity. First and foremost, parents were concerned that the standard setting process would not be credible if it is based on the opinions and contributions of thousands of parents, as they believe they are not qualified to determine what their children should know at each grade level. They wanted assurance that the standard setting is based on the authority of experts and a scientific process. They had a great deal of skepticism around taking a crowdsourcing approach, as it sounded “too random.” They also did not understand the role of college faculty in the standard setting process.

In addition, parents were concerned about the length of time required to participate and having to reserve that time so far in advance. Most said they would not be able to pick a time slot in May for an event in October, and most also expressed concern about being assigned a more specific time window closer to the event. They need more flexibility in their schedules.

We later explained that the invitation represented an opportunity for parents to “weigh in” to the assessment development process. They preferred this message frame better, but they had questions about how their opinions/contributions would be used.

Parents were confused about why Smarter Balanced would be looking for input from parents when the “standards had already been set through the CCSS.”

After explaining the full process for the standard setting (i.e., in person, online, and committee as final decision maker), parents had mixed thoughts—on one hand, they want the standard to be based on expert authority and a scientific process; on the other hand, they appreciate the opportunity to weigh in and experience what their children will experience in taking an assessment. However, when they learned that their opinion/contributions to the process may or may not shape the final decision, they expressed skepticism about the Online Panel. Some thought the crowd-sourced event was purely a political exercise. Most thought it would be a waste of their time.

As they discussed the standard setting process, parents expressed a desire for credible, easy-to-understand examples of the process information goes through from inclusion in the standard to how it is taught in the curriculum to how it is measured on an assessment: “I just want you to show me an example—here’s the standard, here’s how it is incorporated into the curriculum, and here is how it is tested.” Some parents expressed skepticism about the quality of assessments in general: “Are these assessment questions really measuring what they are supposed to be measuring?” Others asked how the information from assessments would be used: “Is this to measure the fail/pass criteria for grades or something more general? How does this fit into the much-cluttered landscape of PSAT, SAT, ACT, etc.?” They expressed a desire for an easy-to-understand primer on the educational assessment landscape.

Parents also discussed the importance of the source of information—they trust their local school systems most for information. They said they would not necessarily pay attention to information from Smarter Balanced, as they were unfamiliar with the Consortium and its degree of authority.

Parents were confused about the details of what would be expected of them in the standard setting process: “Are we going to be filling in bubbles for three hours?” After we described the bookmarking process, they reiterated their concern that they were not qualified to determine what their children should know at each grade level. They concluded that other parents would feel similarly, which would leave the process vulnerable to the participation of only parents with extreme views who would attempt to undermine the process.

By the end of this part of the discussion, it appears that a better motivation for parents to participate would be more along the lines of “experience what your children will experience” rather than “help us set the standards your children will have to meet.” They were curious about the format of questions—e.g., selected-response item, constructed response, etc. Some parents suggested it would be better if parents have the opportunity to take the test to experience it and then provide feedback, rather than framing it as an opportunity to participate in the bookmarking process and recommend an achievement level score.

Takeaways from Parent Focus Group

- Parents are confused about the plethora of assessments; they need easy-to-understand information about the objectives of the various assessments.
- They are interested in specific examples of how the content in the standards is incorporated into curriculum, and then how it is tested.
- Parents trust their local schools—they rely on teachers and school administrators for information. We need to make sure these intermediaries are fully prepared for the questions they will get from parents.
- Parents want assurance that achievement level scores (and the assessment in general) is based on the authority of experts and a scientific process.
- Parents are concerned about the length of time required to participate and having to reserve that time so far in advance. They need more flexibility in their schedules. Moreover, they are concerned about what they would contribute to the process and how their contribution would be used.
- In terms of messaging for the online opportunity, the best motivation for parents to participate would be more along the lines of “experience what your children will experience” rather than “help us set the standards your children will have to meet.”
- Parents have questions and skepticism about technological challenges; these need to be addressed in the messaging.
- Parents are concerned about perceptions related to declining test scores for their schools; they want informative and encouraging messaging about this.
- As part of our recruitment and outreach strategy, we recommend focusing on the National PTA instead of “rank and file” parents, who may lack the time and sufficient context to participate in the Online Panel.

Parent Word Preferences

After discussing communication objectives and plans for outreach to parents, we asked for opinions about specific words, summarized as follows:

- **Assessment vs. test:** “Test” is an ordinary classroom test that is part of the student’s grade. “Assessment” works better to describe Smarter Balanced, as it is different from a classroom test.
- *Preference:* Assessment

- **Smarter Balanced Assessment Consortium vs. Smarter Balanced:** “Smarter Balanced Assessment Consortium” is more credible and delivers more context. First reference should always be the full name.
 - *Preference: Smarter Balanced Assessment Consortium*
- **English and math vs. mathematics and English Language Arts:** “Mathematics and English Language Arts” is too stiff and formal.
 - *Preference: English and math*
- **Computer-adaptive vs. computer-based:** “Computer-adaptive” is an unfamiliar term.
 - *Preference: Computer-based*
- **State-of-the-art vs. next generation:** Neither term conveys meaning or appeals to audiences. “State of the art” is overused and “next generation” is too vague.
 - *Preference: Neither*
- **Core knowledge and skills vs. real-world skills:** These are very different concepts, and they cannot be used interchangeably. Parents find “real world skills” desirable. “Core knowledge and skill” requires further explanation—e.g., what is considered “core” and how was this determined?
 - *Preference: Real-world skills*
- **College and career vs. success after graduation:** Neither term was entirely satisfactory. “Success after graduation” is too vague. “College and career” implies only college-bound students.
 - *Suggested alternatives: Ready for the next step, job ready, ready for the job market, college and career readiness*
- **Scientific process vs. inclusive process:** These are very different concepts, and they cannot be used interchangeably. “Inclusive process” is a better motivator for parents—they appreciate the opportunity to experience what their children will experience in the assessment and the opportunity to weigh in. However, they will want to know that their contribution/opinion will be used in the process and how it will be used.
 - *Preference: Inclusive process*
- **Fair and rigorous vs. fair and accessible to all abilities:** “Fair and accessible to all abilities” is too vague and suggests lowest common denominator. “Rigorous and fair” or “rigorous but fair” is preferred.
 - *Preference: Rigorous and fair*

Asset Analysis

As a critical part of launching this campaign, we have requested more information about and access to specific assets already available to the Consortium and your states. As we build out the tactical elements of our plan and move into implementation, we will ensure that all of our suggestions take advantage of existing assets. Primary examples of assets include:

- Owned media (channels you control)
- Partners
- Spokespeople
- Conference schedule
- Public facing materials

To date, we have requested a list of partners maintained by GMMB and the Consortium and will look to leverage existing relationships and opportunities and to add new ones.

To prepare for the challenging task of explaining the achievement level setting process and defending test results from the new cut scores, Smarter Balanced will require supporters and spokespeople from member states who can convey key messages on behalf of the Consortium and will also inspire trust and confidence among our target audiences. We suggest working with key members of the Consortium staff and your states to identify spokespeople from stakeholder groups impacted by the assessments and the achievement level setting process. This should include representatives who can speak to achievement levels from a national perspective as part of a “speaker’s bureau.”

- Executive staff
- Work group members
- State chiefs
- Higher education experts and faculty
- Educators and administrators

Key Messages

Message development is essential to effectively positioning and promoting the value of the Common Core State Standards and the Smarter Balanced assessment. It is also critical to clarifying the roles they play in improving the college- and career-readiness of American students. Once we return with the results of our focus groups, we will develop a detailed messaging matrix that reflects overarching and audience-specific messages for national and state/local use, geared toward the problem of the “score cliff” and explaining achievement level setting to the general public.

We can classify and prioritize key audiences for Smarter Balanced and stakeholders in K-12 standards and assessments: What should our member states say about the standards and assessment, in which ways, and through which media? Who can help us get the word out? How will we define success? To answer these questions, we have developed initial messaging, building on a literature review of existing research and our understanding of best practices of communicating the complexity of assessments and academic standards. The primary source materials for these message came from the “Common Core State Standards, Public Opinion Research” primer from the Bill and Melinda Gates Foundation, the “The Language of Deeper Learning in America” focus group research conducted by Luntz Global on behalf of the William and Flora Hewlett Foundation, and the polling conducted by the Tarrance Group and David Binder Research on behalf of the Collaborative for Student Success (“Findings From National Survey On Common Core Standards”).

We have created test messages using three levels to frame the debate beginning with an appeal to universal values, followed by a clear definition of the issue, and concluding with a specific policy or program to help remedy the problem (Lakoff’s Level of Analysis). Based on our review of research from the Gates Foundations, Hewlett Foundation, and other sources, an outline of initial messaging is as follows:

Level 1: Nothing is more important than giving all students an education that prepares them for success.

Nothing is more important than giving all students an education that prepares them for success. Yet independent research shows America's schools are failing our children. Though many assume that poverty in America is pulling down education scores, research shows that middle-class U.S. students are far behind their peers in other countries (Hewlett: 55 percent of Americans agree).

Appealing phrases (Hewlett):

- ALL children deserve an exceptional education.
- By raising standards for ALL, we can challenge them to do better.
- America should have the best results in the world.
- Preparing students for success in college, careers, and life.
- Teaching students to THINK and COMMUNICATE.
- Giving families access to effective education options.
- Equipping students for the REAL WORLD.
- An equal opportunity to succeed.

Level 2: Consistent, high expectations are the means for reaching higher standards.

Common Core is “an effort to ensure that every student in every state has the opportunity to learn the same topics and master the same skills through a common national standard, so that no student is denied basic learning” (Hewlett: 65 percent of Americans agree).

Consistent, high expectations (Gates):

- Are supported by teachers;
- Are already delivering results;
- Ensure students are prepared for college and career;
- Ensure students can compete for jobs in today's economy;
- Ensure ALL students have access to a high-quality education; and
- States should work together to set minimum standards for learning and curriculum. We must hold all states accountable to ensure standards do not drop too low (Hewlett: 56 percent of Americans agree).

It is better for all states to have the SAME STANDARDS at each grade level in math and English so students have to meet the same expectations (Gates: 67 percent of Americans agree).

Appealing phrases (Hewlett):

- Proven, effective education strategies.
- Delivering real, measurable results.
- States should WORK TOGETHER to set standards.

Level 3: Accurate assessments are essential for achieving high expectations.

The Smarter Balanced assessments are a key part of implementing CCSS and preparing all students for success in college and careers. The assessments will replace existing tests and provide an “academic checkup” for students, parents, and teachers by measuring real-world skills such as critical thinking and problem solving. In addition, they will provide information during the year to give

teachers and parents a better picture of where students are succeeding and where they need help (Smarter Balanced website).

It is better for all states to have the SAME TESTS at each grade level in math and English so test scores can be compared across states (Gates: 57 percent of Americans agree).

Appealing phrases (Hewlett):

- We need to make the tests we HAVE more effective.
- Reward effective schools and hold failing ones accountable.

Audience Specific Messages

Messaging for Pre-K-12 Educators and Administrators

- The Common Core State Standards will have a positive effect on students' ability to think critically and use reasoning skills (Gates: 77 percent of teachers).
- Consistent standards will provide clear guidelines and improve efficiency, especially when integrating students from other states.
- This is a test by states and for states, developed with ongoing input from teachers.
- Participating in the Consortium will help educators share and learn from best practices through collaboration with other participating states.
- Educators have the opportunity to see actual test items and experience the test as their students will experience it.
- Educators will have widespread opportunity for input via the Online Panel, in addition to the in-depth opportunity through the In-Person Panel.
- Consistent standards may ease their burden and improve efficiency, especially when integrating students from other states.

Messaging for Parents

- High standards will help more children succeed. Assessments are a tool for ensuring we reach those standards.
- By raising standards for all schools and students, we can challenge them to do better (Hewlett: 58 percent of Americans agree).
- America should have the highest standards and best results in the world.
- Assessments need to be more effective, accurately measuring true knowledge and skills—not just memorized information (Hewlett: 47 percent of Americans agree that we do not have too many tests).
- We need to ensure that our public education systems prepare students for success after high school (Gates: 46 percent of parents think public schools are not preparing students to be successful, but 42 percent think they are preparing students),
- The Common Core State Standards and assessments will help reduce inequality and ensure we hold ourselves accountable for improving education (Gates research on communities of color and less educated audiences).
- (Hewlett: Americans rank math, tech, and science education as most important for the country's future,)

Messaging for Higher Education

- High academic standards will ensure that students who enter college are ready for credit-bearing coursework. Higher K-12 standards and the Common Core-aligned assessments are part of the college completion agenda.
- Higher academic standards and assessments ensure that high school graduates experience rigor in K-12 and require less remedial coursework in college.
- Less remediation will free up resources and enable to focus on innovation and instruction of higher level courses.
- This multi-state effort contributes to readiness and competitiveness beyond the priorities and resources of a single state.
- Alignment with the Common Core will promote college access and success and make the admissions process more efficient.
- The next generation of assessments and a larger data pool will enable an evidence-based research agenda for education schools in colleges and universities.
- Higher education faculty and leadership will have input to the process and can contribute expertise through the Online and In-Person Panels.

Messaging for Business

- Too many students today show up at college or the workplace lacking foundational skills.
- Business leaders have knowledge of the skills it takes to succeed in the workforce. Your input is critical for ensuring that the test measures skills needed for today's jobs.
- Help ensure these tests are rigorous and fair and give us a clear picture of whether students are ready for success after high school.
- Higher academic standards and alignment with the Common Core will make workforce recruitment process steadier and more reliable.
- Improving student achievement will reduce the need for remedial education in company training programs.
- Business and community leaders will have input through the Online Panel, ensuring that they can get the quality standards they want.

Phase 1 (April–September)

Strategic Goals

- Gain buy-in and support for the achievement levels in advance of launch through the unique, broadly inclusive and scientific process
- Generate interest in and support for the assessment beyond the K-12 education community by engaging parents, the business community, the higher education community, and others interested in achievement levels.

Recruitment Goals

- Assist Smarter Balanced states in driving recruitment for the Online Panel for Achievement Level Setting taking place October 2014.
- Ensure diversity and inclusiveness of participants; achieve balance of participants per state and per category (e.g., teachers, parents, school administrators, business community, higher education community, education reporters).
- Mitigate risk of “hostile takeover” or biasing of the Online Panel by opponents.
- Educate reporters and constituencies on the Online Panel, ensuring transparency and openness.

Schedule

April–May

- Materials development and testing.
- Initial outreach to teachers, higher education faculty, and business community through Governing States, partnerships, and channel marketing.
- Solicit registration for Online Panel; intensive outreach before end of school year.
- Op-ed in early May to drive teachers/higher education community to registration pages.
- Direct outreach to higher education community through colleges/universities to raise awareness of Online Panel before end of school year.
- Outreach to teachers and higher education community through conferences and selected media channels to raise awareness and encourage registration for Online Panel.
- Outreach to parents through webinar with National PTA, targeted media and social media channels to raise awareness and encourage registration for Online Panel.

June–August

- Ongoing outreach to business community, through professional organizations and selected media channels to raise awareness and encourage enrollment for Online Panel
- Ongoing outreach to all stakeholders through earned media and digital media
- Paid advertising in select educator/industry publications

September–October

- Media coordination leading up to and during month of Online and In-Person Panels
- Webinar for briefing friends and partners immediately before Panels

Communications Packet: Turnkey Materials for States Pt. 1

- Recruitment email templates for target audience (parents, higher education, teachers, business)
- Stock newsletter content for target audiences (parents, higher education, teachers, business) for use in-house or in a publication
- Media release for top-tier media

- FAQ website (beta language to be revised following focus groups)
- Content for social media channels (Smarter Balanced specific and generic)
- Content for individual newsletters
- A “Facts and FAQ” document explaining achievement level setting
- Customizable talking points for PIO and chiefs
- Op-ed/letter to the editor (LTE) for submission to local newspapers
- Talking points for use as a packet stuffer
- PowerPoint presentation

Hager Sharp is proposing an *overlapping outreach* approach to registration and recruitment for the Online Panel, beginning with teacher, higher education, and business audiences and moving into parent audiences more slowly. It is critical that we engage teachers and higher education faculty before the end of the academic year. This will allow us to focus on the stakeholder groups with professional incentives to participate and with detailed understanding of academic standards and assessments in general. This will give us additional time to develop a message that differentiates between the gatekeeper groups (teachers, business/community leaders, PTA) and a more general parent audience.

We will conduct outreach for the Online Panel in two stages, the first focusing on contacting stakeholders through a “soft touch” using direct targeting (Teacher Involvement Coordinators, K-12 and higher education leads, the Smarter Balanced newsletter) and the second focusing on targeted public outreach (partnerships with associations and influencers, social media, organizational newsletters). Hager Sharp proposes moving forward immediately with outreach to teachers, school administrators and higher education faculty, since they are insiders to the process and nomenclature.

Based on our preliminary research and the need for additional research on parent attitudes, we recommend that Smarter Balanced begin with outreach to educators and higher education faculty, followed by parents and the general public. An overlapping approach will also allow us to reach “gatekeepers” (teachers, business/community leaders) before pushing a message to parents. This strategy has been used successfully with many public education and public health campaigns, by reaching gatekeepers first and ensuring their trust is granted before engaging with parents and families.

The strategies and tactics below outline timing, tactics, and messengers/partnerships to consider as part of communications activities to support recruitment and stakeholder outreach. Outreach should begin in late April and continue through June. K-12 leads will receive an email to send to their Teacher Involvement Coordinators (TICs) with a request to invite other teachers to register for the Online Panel. We will send follow-up emails to TICs on a weekly or biweekly basis offering additional tools. Higher education leads will receive an email with a similar request to invite their networks to register for the Online Panel.

The next stage of the recruitment process will build on initial awareness in the states, expanding into a more public outreach phase focused on partnerships, common messaging, and involving stakeholder groups. This stage will launch in late May, following the results of our focus groups, and continue over the course of the summer through earned media and targeted outreach to relevant associations, organizations, state-based advocacy groups, and others.

Strategies and Tactics for Recruitment and Outreach

Owned Media

Beginning in April and May, the outreach strategy for the Online Panel will begin with outreach to the Consortium's owned media, i.e., channels that you control. We will launch the formal recruitment push by providing information and registration links to target audiences through channels internal to Smarter Balanced, including your K-12 and higher education leads, work group members, contacts at school districts, etc.

Primary examples of owned media include:

- Smarter Balanced social media channels
- Weekly News email blast
- *Smarter News*
- K-12 and Higher Ed Leads update calls

Registration for the Online Panel formally begin on April 22. Once K-12 and Higher Ed Leads have had sufficient time to focus on recruitment for the In-Person Panel, we will supply them with additional outreach resources through the recruitment communications toolkit to help raise awareness about the Online Panel opportunity. We suggest regularly updating states on recruitment progress and enrollment for the Online Panel.

Earned Media

A successful media relations effort creates a steady drumbeat of media placements over time. Beginning this effort in May will pay off later on as we build goodwill with the media and other entities that are eager for more information on Common Core-aligned assessments. In early May, we will begin to selectively push earned media outreach to help drive awareness about the Online Panel. This will include announcements to trade media and a launch op-ed to help increase awareness about the achievement level setting activities and drive traffic to the registration site.

This will lay the groundwork for maintaining or developing strong relationships between the Consortium's core partners and supporters and the news media. We can work with Smarter Balanced staff to evaluate and develop a series of media opportunities at the national, state, and local levels. The primary goals of our summer earned media campaign will be to:

- Promote awareness of and participation in the Online Panel and In-Person Panel.
- Educate state and national education writers and influencers on the achievement level setting process.
- Ensure that this process rolls out in a transparent, inclusive manner that will build greater acceptance of the final cut scores and assessment results.

The first phase of media outreach will focus on building an initial understanding of the achievement level setting process with key education writers. After securing approval from Smarter Balanced for targeted pitches, Hager Sharp will focus on major national and regional publications with dedicated education coverage. As needed, we can provide news release templates and media lists to Smarter Balanced states who wish to engage in additional outreach to support their recruitment activities for the Online Panel.

Content Calendar

An overarching content calendar will help create a continuous flow of communications and outbound messages from the Consortium. We will begin the rollout of the earned media push by providing Smarter Balanced with a content calendar for summer months, highlighting internal and external events and milestones throughout the achievement level setting process. We will provide a content calendar in early May (or once the plan is approved) to help plan ahead for internal and external events, to streamline content from Smarter Balanced and partner organizations, and to provide structure and time-scale organization to outreach/recruitment tactics.

This content calendar will provide a roadmap for linking an organization's work with issues in the news and earned media opportunities. We can work with the Consortium team to identify major milestones and develop a content calendar around report releases, conferences, knowledge leadership, and news items of interest to target audiences, aligned with a larger strategy of helping to reassert the value of common achievement standards. We will produce recommendations on how to integrate those activities and events into outreach activities and recommend new additions, as appropriate. As part of both the recruitment effort and our outreach to partners, we will also craft corresponding information for your website, blog posts, and social media channels.

Media Announcement/Webinar

Following the launch of public recruitment for the Online Panel, we propose making an announcement to key trade outlets and education writers regarding the achievement level setting process. We suggest holding at least one webinar, possibly in conjunction with the Education Writers Association, to convene reporters and other education stakeholders and begin a dialogue about the technical process for achievement level setting.

Priority Media Targets for Recruitment and Stakeholder Outreach

- American Educator
- Associated Press
- Bay Area News Group
- *The Chronicle of Higher Education*
- *Detroit Free Press*
- *District Administrator*
- *Education Week*
- Huffington Post
- *Inside Higher Ed*
- *Los Angeles Times*
- POLITICO Pro
- NPR
- *San Francisco Chronicle*
- *The Detroit News*
- *The New York Times*
- *The Wall Street Journal*
- *The Washington Post*
- *USA Today*

Op-Eds and Letters to the Editor

A launch op-ed in one or two prominent education news outlets/blogs (*Inside Higher Ed*, *The Chronicle of Higher Education*, and *Education Week*) would provide significant visibility to teachers and administrators and help drive traffic toward the Online Panel registration site. We could work to place an op-ed from Joe Willhoft or a leader from a governing state. Education-specific outlets such as *Education Week* and industry blogs offer opportunities to place op-ed articles that articulate the value of the Common Core State Standards and assessments. We can lend our expertise in writing thoughtful guest columns for news outlets; help you choose pertinent topics; and write and place guest columns, blog posts, or other entries to reach audiences.

In many newspapers, the letters to the editor/community conversation section is still one of the most widely read and a valuable conduit for your messages. Letters need to be short (usually about 250 words) and direct to convey a message or position. If there are high-profile problems that develop related to the Online Panel or there is a need to counter misconceptions, we can develop a letter to the editor (LTE) to help clarify the process for achievement level setting.

Social Media

Over the course of the summer, we will monitor developments in social media, specifically with an eye toward traffic to the registration page. Digital outreach and social media accounts are vital dissemination tools, and Hager Sharp will prepare content and response strategy for these platforms, including websites and blog, if appropriate. Our goal would be to provide the audience with an interactive and engaging experience and use these tools to build awareness and support for the assessments and the achievement level setting process.

Social media channels should work in concert to develop campaigns, build a campaign's audience, establish knowledge leadership, drive web traffic, and, ultimately, help achieve strategic goals. We can craft a strategy and content for you to use to reach education stakeholders and other key audiences, train employees in using social media channels effectively, and counsel you on managing various accounts. Each social media network reaches different audiences, broken down by any number of traits, and we can ensure that the campaign has an effective, comprehensive presence based on which platforms each audience uses.

Digital tactics could include targeted SEO (search engine optimization) efforts to improve site traffic and increase engagement.

Additional tactics include:

- **Social Media Campaigns:** We can improve the existing social media presence (Facebook, Twitter, Google+, LinkedIn, and YouTube) to expand influence among target audience and push content that supports key messages.
- **Social Media Commentary:** Guest-blogging/bylined articles on key websites will help drive traffic to the registration pages.
- **Promoted Social Media Posts:** Depending on the initial success of the recruitment efforts, we can use digital advertising tactics to be introduced on a pilot basis, auditing results to understand which calls-to-action are most effective in engaging target audiences. Facebook is a paid media platform for brands, and promoted posts on Twitter reach a much larger audience than organic tweets.
- **Recruitment Blitz:** The crowd-sourced nature of the Online Panel suggests that we consider a crowdsourcing method for encouraging recruitment. As we approach the deadline, we

suggest investigating low cost crowd-sourced social media tactics to help drive participation in the Online Panel and increase awareness about the achievement level setting process among target audiences.

- **Twibbon Campaign:** A Twibbon Campaign allows brands to create a microsite where users can support a particular cause through a variety of actions, including social media posts and adding a picture to their social media account.
- **Thunderclap:** Thunderclap is the first “crowd-speaking” platform that leverages the power of major social networks by allowing campaign supporters to link their accounts and blast out a Twitter, Facebook, or Tumblr at a strategic time. Organizers have to first set a goal for number of supporters; once the goal is reached, Thunderclap can disseminate a uniform message from all your supporters, creating a wave of attention at a strategic point in time.

Partner Outreach

We will share customizable email invitations and newsletter content for all of our key audience segments (parents, teachers, higher education, business), allowing them to disseminate information about the Online Panel with their membership and allies. For each audience segment, we propose building out a customized pitch and developing a specific “ask” or call to action. The goal of these touches will be to drive participation in the Online Panel, but they will also work toward the overarching goal of building credibility, awareness, and support for Smarter Balanced and the achievement level score recommendations.

Educator/Administrator Outreach

For the Online Panel recruitment and outreach effort, our campaign must begin by winning the trust and support of teachers. We can help you explore partnerships with organizations such as the national teacher organizations for each academic field, statewide teachers’ organizations, and institutions of higher education and university presidents in member states, the American Association of School Administrators, both national school principals’ organizations, the National School Boards Association, and state-specific education policy organizations.

The Online Panel provides a critical opportunity for educators from Smarter Balanced states to participate in the achievement level setting process in numbers far greater than is traditional. This democratic process has both operational and strategic value to the Consortium; widespread participation from teachers creates a much larger sample than an In-Person Panel alone and will add greatly to our ability to defend the validity of the cut scores. As a result, engaging educators and the associations that represent them is an important part of our process. Although some of the key associations have already held their annual conferences, we propose leveraging the conferences that remain as an opportunity to inform educators about the Online Panel opportunity. Pending approval by Smarter Balanced staff, we will also explore the possibility of coordinating outreach through the State Network of Educators.

Priority Targets for Educator/Administrator Outreach¹

- American Association of School Administrators
- American Federation of Teachers
- ASCD

¹ See Appendix I for master list of partnership targets.

- Council of the Great City Schools
- Council of Chief State School Officers
- Education Commission of the States
- International Reading Association
- National Association of State Boards of Education
- National Association of Secondary School Principals
- National Council of Teachers of English
- National Council of Teachers of Mathematics
- National Education Association
- National School Boards Association

Business Community Outreach

Outreach to business and workforce groups serves two purposes in our larger goal defending the Smarter Balanced assessments and achievement level scores. Partnerships with the business community will help drive awareness of and participation in the Online Panel. Secondly, reaching out early to business groups will also serve as a long-term strategy for strengthening the community of Common Core supporters. The highest priority target will be conducting outreach to the U.S. Chamber of Commerce Foundation, which leads Education and Workforce activities for the Chamber and played an important role in the successful “Business for Core” and “Ready Kentucky” campaigns. These campaigns, and similar efforts in others states, built community and political will for Core-aligned assessments through an echo chamber of voices in support of a “new baseline” based on higher standards.

The Business Roundtable, who is partnering with the Chamber on the Higher State Standards Partnership, has also indicated its desire to be a persuasive force and conduct a campaign to advance the Common Core. As part of the recruitment effort for the Online Panel, we will deliver turnkey materials to Smarter Balanced states and conduct outreach on the national level to groups that represent business interests as well as individual firms that promote preparedness and competitiveness. We can provide materials that states can use to do the same on the state and local levels. We propose seeking active partnerships with the national and state-level Business Roundtable groups, the national and state chambers of commerce, and others.

Priority Targets for Business Outreach

- U.S. Chamber of Commerce Foundation (Business for Core)
- Business Roundtable
- Business Higher Education Forum
- California Economic Summit
- Change the Equation
- GE Foundation
- National Skills Coalition
- Society for Human Resource Management

Higher Education

Recruitment and outreach to higher education influencers will build awareness and credibility for the assessments and encourage institutions to recognize Smarter Balanced scores as evidence that students are ready for credit-bearing courses and may be exempted from developmental education. Our engagement with the higher education sector will begin with a focus on disciplinary and leadership associations with the most immediate stake in the grade 11 exam (developmental English and math faculty, community colleges, State Higher Education Executive Officers [SHEEO]). Our outreach will gradually expand to include other presidential, disciplinary, and membership associations who we can engage in a discussion about the importance of common assessments and common college- and career-ready standards.

Priority Targets for Higher Education Outreach

- SHEEO
- Conference Board of the Mathematical Sciences
- American Association of Colleges for Teacher Education
- American Association of University Women
- American Association of Community Colleges
- American Association of State Colleges and Universities
- American Association of University Professors
- American Council on Education
- Association for Institutional Researchers
- Association of Public Land Grant Universities

Phase 2 (September–November)

Strategic Goals

- Develop compelling messages on the performance standards to present to K-12 and higher education decision makers (e.g., Chief State School Officers and State Higher Education Executive Officers) in advance of states voting to approve the performance standards.
- Identify and deliver compelling messages on the Common Core State Standards, the Smarter Balanced Assessment System, the standard setting design, and the performance standards to key Smarter Balanced constituents.
- Respond to contingency situations arising at the Online Panel and In-Person Panel.
- Respond to potential crises as a result of new achievement level recommendations and state consideration of those recommendations.
- Promulgate info on the new performance standards to state audiences.

Communications Packet: Turnkey Materials for States Pt. 2.

- Talking points, Q&As for states explaining cut scores and need for high standards
- Fact sheets for policymakers and the general public explaining the achievement levels
- Grassroots toolkit
 - Generic op-ed/letter to the editor for general public

- Generic talking points for use by parents/community members, business leaders, higher education, teachers
- Op-ed/letter to the editor for grasstops leaders (e.g., business leaders/owners, school leaders/board members, labor, civil rights leaders, faith and clergy, military leaders and families, charter school supporters, religious affiliated groups, current and retired elected officials)
- Op-eds explaining the value of the standards and Smarter Balanced assessment which would be customized for placement in local media
- Exemplar items from the test to demonstrate alignment with CCSS
- Ready-to-use PowerPoint presentations on CCSS and the Smarter Balanced assessment that could be customized for use when testifying before a state legislature, a state board of education, a local school board, and Congress
- A video package on the process for setting achievement levels to be released after In-Person and Online Panels conclude

Strategies and Tactics for Promoting the Achievement Levels

A key component of our work would be providing communications counsel and turnkey materials to Governing States as they conduct outreach and prepare to release data at the state level. These materials will convey to parents and other stakeholders in the states the necessity of setting a new baseline for test scores, aligned to higher standards, so that all students are on track for success after high school. Before the close of registration for the Online Panel in September, we will work to provide turnkey materials to Smarter Balanced states, explaining the achievement level setting process and expected outcomes from that process.

One of the primary goals of this campaign will be to provide materials and strategies for states. In order to help states explain and promote the new achievement levels and test, the bulk of our work will focus on developing state capacity for communications through a train-the-trainer model. By sharing customizable content and materials with states and using new and existing meetings and forums to train and collaborate with state leads, we will create a much larger platform for reaching public audiences than we would through only individual engagement with the states.

In addition to explaining the new achievement levels, we will be working to coordinate media inquiries and participation at the In-Person Panel in Dallas and the Online Panel.

Owned Media

Smarter Balanced should be cautious in the use of owned media channels during the second phase of the communications scope. We will need to have a very careful and deliberate approach as well as a unified voice in our dialogue with the media. Common Core issues could be under intense scrutiny by the time the election arrives, and we must exercise discipline in our communications with external parties during what could potentially be a much more volatile environment. Working with Smarter Balanced, we will prepare vetted messages and materials for a variety of scenarios that may arise after the achievement level setting activities conclude.

Earned Media

In preparation for what will likely be a challenging rollout of the new achievement levels and cut scores, we will work concurrently with our general communications planning to develop crisis response protocols and topline messages for the most likely contingency scenarios.

When the media calls, it is important to be ready. Going “on the record” can be as simple as issuing a statement to targeted contacts or making direct contact with the media, but a rapid response is crucial. The ability to monitor and quickly respond to ongoing news developments and breaking news stories is a key component of a proactive public relations campaign. Having the ability to quickly develop and then receive internal sign-off on position statements or news releases is a must. Our crisis communications plan will include protocols to ensure a timely media response in fast-moving situations. Hager Sharp will focus our crisis planning on potential crises related to the standard setting process and the “score cliff.”

These scenarios include:

- Technical or operational problems at the In-Person Panel or involving the administration of the Online Panel;
- A disagreement or negative exposure following the In-Person Panel;
- A disagreement or divided vote at the fall collaboration conference when states vote to endorse the new achievement levels and cut scores;
- Concerns about the achievement level setting process from the Technical Advisory Committee;
- Exposure of test items via a variety of media (e.g., students tweeting photos of test questions);
- The release of a third-party review questioning the validity of the standard setting process (e.g., the National Academy of Sciences report criticizing NAEP achievement levels); and
- Intense media exposure or political mobilization on the issue of the “score cliff.”

The crux of a rapid-response media strategy is to anticipate, monitor, and quickly respond to a moving news story. Having a chain of command in place is critical to being able to respond on short notice to an issue that impacts the Consortium as well as key stakeholders.

We will work with Smarter Balanced spokespersons to ensure they are prepared for media interviews, testimony before federal and state bodies, and other presentations, including development and dissemination of talking points and position statements for using during on-the-record comment. We can provide media training to Smarter Balanced spokespersons and work to develop and coordinate a network of educator and parent leaders who can speak about why the Common Core State Standards are important to them.

In addition to the development and completion of a crisis communications plan, we will look for opportunities to ensure that Smarter Balanced spokespeople and stakeholders communicate effectively over the course of the summer. At any time during the engagement, we can provide additional media training for your executive team, member states, and allies.

Social Media

There will likely be a continuous stream of social media “chatter” mentioning Smarter Balanced during the Online Panel. During the period when panelists are actively participating in the Online Panel, it is important that we continue proactive outreach through owned media and social media in order to communicate key messages to participants. During the Online and In-Person Panels, we will be closely monitoring social media activity and reviewing comments and posts for their source, tonality, and audience type. If social media behavior begins to show a particular pattern or trend, we will provide recommendations for responding and correcting the record so that misinformation does not take root.

We also recommend taking steps to build a sense of community around participants in the Online Panel. The concept of awarding a badge to show a skill or completion of a task has been in use since the Middle Ages. Only recently, however, with increased access to online technology, have digital badges become prevalent in education. Social media users are displaying digital badges on their social media profiles, websites, blogs, and online dashboards and backpacks. We suggest giving participants in the Online Panel a “digital badge” to show that they took part in this unique opportunity to shape assessment levels for our students. The goal of a social media badge would be to build pride and community among participants. We will also want to provide teachers with a certificate of participation that can be used for professional development credit.

Grassroots Engagement

To support this integrated public affairs campaign, we suggest looking to an old expression: All politics are local. In addition to the reactive media outreach, we can seek out partnerships with authentic grassroots supporters whose voices resonate at state and local levels.

We can work to strategically identify grassroots supporters in key states and markets, educate them on the value of the Common Core Standards and the Smarter Balanced assessment, dispel myths about the standards and assessments, and activate them during key time periods. We can investigate messaging and partnership opportunities with a variety of other stakeholder groups such as parents (including charter school parents), military families and spouses, religious-affiliated groups, and current and retired elected officials of both political parties. As part of the turnkey materials delivered to states, we will include a grassroots toolkit with op-eds and other materials that will be easily customizable to grassroots and grassroots supporters in the states. We will focus our outreach on partner organizations such as Stand for Children and the PIE Network.

Outreach for Spanish Speaking Audiences

Spanish-Language Messaging

Our initial inquiries suggest that messages for Spanish speakers should focus on advancing educational equality, high standards for all students, and greater quality of life and job prospects. Depending on the Consortium’s desire to reach Spanish-speaking audiences, it may be necessary to conduct informal focus groups with bilingual and Spanish-speaking parents to gauge their involvement, interest, and reactions to Smarter Balanced. This will be an opportune time to identify reading levels and word preferences. When developing messages for Hispanic groups, their education attainment and language preferences should be taken into consideration. Some materials will only need to be translated while others will need to be adapted so as to engage the reader with information that is specific to them and avoid generalization of student experiences.

At minimum, we will need to produce fact sheets and other written materials that can be made available inside and outside school settings. Experience shows that English language learners (ELLs) are less likely to attend an after-school program or parent-teacher conference, where most parents would receive this information.

During the recruitment period for qualified Hispanic educators and policy specialists, Smarter Balanced should consider contacting national organizations that work closely with the community. They could also serve in the long-term for community outreach during the messaging and media outreach phase. A network of support groups throughout the country should have Spanish and English materials available for parents. These groups can be reached through the national organizations contacted for the achievement level setting panels. They will play an important role by ensuring that information about the assessments and college- and career-ready standards reaches parents, especially those in low-income communities.

It should be noted that participants in the Online and In-Person Panels will be required to be English speakers, because the test items and discussions will be conducted in English. We should prepare FAQs in advance about ELL considerations in the development of the assessment. For the English language Online Panel, we will reach out to national organizations to encourage their members to participate, and we will help them understand the process.

In order to ensure success and relevance with Spanish-speaking audiences, it is important that a Hispanic, Spanish-speaking spokesperson or representative be available for media interviews. Experience also suggests that Hispanic audiences, especially parents, will be more receptive to a message that is coming from an experienced and relatable professional. The absence of a Spanish-speaking spokesperson from Smarter Balanced could make media outreach difficult. Media outreach should be primarily in Spanish via major radio and television networks and print publications such as *La Opinion*.

The team should be prepared to produce Hispanic-specific press releases and talking points in English for national, online media channels, such as NBC Latino, Yahoo en Español, MSN Latino, Fox News Latino, etc.

Priority Targets For Spanish-Language Partnerships Include:

- National Latino Education Network
- Excelencia in Education
- National Council of La Raza
- Committee for Hispanic Children and Families
- Hispanic Association of Colleges and Universities
- National Association of Latino Elected and Appointed Officials

Project Conclusion

At the conclusion of the project in November, we suggest scheduling an inclusive debrief discussion involving key personnel from the campaign. In addition to regular feedback we will provide throughout the course of the engagement, we intend to submit an after-action report that will provide insights acquired over the course of engagement on specific messages, tactics, and strategies that succeeded and those that failed, and a retrospective on issues we encountered from kickoff to execution.

The report will contain candid insights on the campaign's success from various perspectives, feedback and suggested next steps for communications priorities, detailed evaluation metrics, (suggested by Hager Sharp and approved by Smarter Balanced) and an overarching analysis of success mapped to the contract's original objectives. Our report will also contain media coverage highlights from the engagement along with final drafts of all key documents and materials.

Appendix I: Partnership Targets

<p>Educators</p> <ul style="list-style-type: none"> ○ Alliance for Excellent Education ○ American Association of School Administrators ○ American Federation of Teachers ○ Association for Supervision and Curriculum Development) ○ Council of the Great City Schools ○ Council for Exceptional Education ○ Education Commission of the States ○ International Reading Association ○ International Society for Technology in Education ○ International Association of Special Education ○ National Association of State Boards of Education ○ National Association of State Directors of Special Education ○ National Association of Secondary School Principals ○ National Association for Bilingual Education ○ National Council for the Social Studies ○ National Council of Teachers of English ○ National Council of Teachers of Mathematics ○ National Education Association ○ National Rural Education Association ○ National School Boards Association ○ Teach for America ○ Teachers of English to Speakers of Other Languages ○ Teach Plus 	<p>Higher Education</p> <ul style="list-style-type: none"> ○ American Association of Collegiate Registrars and Admissions Officers ○ American Association of Community Colleges ○ American Association of State Colleges and Universities ○ American Association of University Professors ○ American Council on Education ○ American Association of Colleges for Teacher Education ○ Association of Public and Land Grant Universities ○ Conference Board of the Mathematical Sciences ○ Midwestern Higher Education Compact ○ National Association for College Admission Counseling ○ New England Board of Higher Education ○ SHEEO ○ Southern Regional Education Board ○ Western Interstate Commission for Higher Education
<p>Parents</p> <ul style="list-style-type: none"> ○ National Parent Teachers Association (PTA) ○ School districts ○ State and local PTA affiliates ○ Democrats for Education Reform ○ Students First 	<p>Business Community</p> <ul style="list-style-type: none"> ○ America's Edge/Ready Nation ○ America's Promise Alliance ○ Business Roundtable ○ Business-Higher Education Forum ○ California Economic Summit ○ Change the Equation

	<ul style="list-style-type: none">○ Foundation for Excellence in Education○ GE Foundation○ National Skills Coalition○ National Governor’s Association○ State Chamber affiliates○ Society for Human Resource Management○ Thomas B. Fordham Institute○ U.S. Chamber of Commerce Foundation
--	---



Communications Plan Achievement Level Setting

The Smarter Balanced Assessment Consortium brings together states to create a common, innovative assessment system for Mathematics and English Language Arts that is aligned with the Common Core State Standards and helps prepare students for college and careers. The Consortium involves educators, researchers, policymakers, and community groups in a transparent and consensus-driven process to help all students thrive in a knowledge-driven global economy. The Consortium's projects are funded through a four-year, \$175 million grant from the U.S. Department of Education, comprising 99% of activity resources, with the remaining support provided through generous contributions of charitable foundations. Membership is open to any interested U.S. state. For more information, please visit <http://www.smarterbalanced.org/>.

Copyright 2013 Smarter Balanced Assessment Consortium



Smarter Balanced Assessment Consortium: Focus Groups with Teachers, School Principals, and Parents in California and Michigan

A Report Prepared by Hager Sharp

May 22, 2014

CONFIDENTIAL



Table of Contents

Table of Contents	130
Background	131
Methods	131
Insights and Recommendations.....	132
Educators	132
Statement Testing	135
Phrase Testing.....	137
Parents.....	138
Statement Testing	140
Phrase Testing.....	141
Word Choice.....	142
Detailed Findings: Teachers and School Principals.....	142
Awareness, Knowledge, and Attitudes Regarding Common Core, Common Core Assessment, and Standard Setting.....	142
Standards and Assessments	142
Smarter Balanced Assessment Consortium	143
Common Core Assessment.....	143
Standard Setting.....	146
Message Testing	146
Educator Invitation Language.....	146
Specific Statement Testing.....	148
Phrase Testing.....	150
Trusted Leaders in Public Education.....	150
Other Feedback.....	151
Detailed Findings: Parents	151
Awareness, Knowledge, and Attitudes Regarding Common Core, Common Core Assessment, and Standard Setting.....	151
Standards and Assessments	151
Smarter Balanced Assessment Consortium	152
Common Core Assessment.....	152
Standard Setting.....	154
Message Testing	154
Parent Invitation Language.....	154
Specific Statement Testing.....	155
Phrase Testing.....	157
Word Choice.....	158
Trusted Leaders in Public Education.....	158
Other Feedback.....	159
Appendix A: Moderator’s Guide - Educators	160
Appendix B: Moderator’s Guide - Parents.....	165

Background

The Smarter Balanced Assessment Consortium (SBAC) is a consortium working collaboratively with its 22 member states to develop a student assessment system aligned with a common core of academic content standards for English language arts/literacy and mathematics. A field test of the assessment system is being conducted this spring, and the assessment system will be finalized later this year.

Smarter Balanced will be conducting both in-person standard-setting workshops and an online panel to obtain input on achievement levels for the assessment. Smarter Balanced is developing materials to invite and encourage educators to participate in the online panel. Smarter Balanced is also developing educational materials to inform educators and the general public about the assessment system that is being developed.

To gather feedback from key audience segments on messages and materials for the upcoming Smarter Balanced standard setting activities, Hager Sharp conducted online focus groups among parents, teachers, and school principals from two Smarter Balanced governing states, California and Michigan. These states were chosen because they offer political, geographic, and cultural diversity, and they represent significant numbers of the Consortium's target audiences. For this round of research, the team focused on parents, teachers, and school principals, as feedback from other audiences, including the higher education and business communities, will be obtained through other channels such as professional associations and existing contacts. Feedback from the focus groups provides valuable insights into the messages and information audiences find most appealing and relevant. These insights will inform the team's outreach efforts for the standard-setting activities May through October 2014. They will also help Smarter Balanced prepare for various contingencies leading up to and following the standard-setting workshops, including changes in student performance standards and test scores.

Methods

Hager Sharp conducted eight 90-minute online focus groups during May 1-8, 2014, using WebEx as a platform. The focus groups were configured as follows:

In California:

- Eight lower-income (household income \$40,000 or less) parents of public school students in grades 3-12
- Eleven middle-higher-income (household income \$50,000 or more) parents of public school students in grades 3-12
- Eleven public school teachers of grades 3-12
- Nine public school principals of grades 3-12

In Michigan:

- Six lower-income (household income \$40,000 or less) parents of public school students in grades 3-12
- Eleven middle-higher-income (household income \$50,000 or more) parents of public school students in grades 3-12
- Ten public school teachers of grades 3-12
- Eight public school principals and one superintendent of grades 3-12

Focus groups explored the following content areas:

- Awareness, knowledge and attitudes regarding standards and assessments
 - Awareness of the Smarter Balanced Assessment Consortium
 - Familiarity and experience with Common Core State Standards
 - Familiarity with assessment system Smarter Balanced is developing
 - Concerns and/or questions about the new assessment being developed
 - Questions about the field testing

- Attitudes about a potential drop in scores with the new assessment compared to previous assessments
- Perceived benefits and concerns regarding standards and assessments generally
- People and sources they trust for information about educational assessments and standards
- Perceptions of Smarter Balanced online standard setting opportunity
 - Reactions to invitation to participate
 - Questions and concerns about the opportunity
- Reactions to specific messages associated with Smarter Balanced standard setting and assessment
 - Reactions to statements about Smarter Balanced and Common Core objectives and mission
 - Reactions to statements about overall objectives and mission of schools
- Discussion of information needs and communication preferences

To prepare this report, moderator notes were reviewed and analyzed with the goal of identifying common themes, areas of consensus and differences of opinion among participants. The moderator guides for educator and parent sessions are included as appendices A and B in this report.

Insights and Recommendations

Key insights and recommendations from the focus group sessions are outlined in this section. A detailed summary of the discussions among educators and parents are included in subsequent sections of this report.

Educators

Key insights and resulting recommendations from the educator focus groups are outlined in the table below.

TEACHERS AND PRINCIPALS	
Insight	Recommendation
Assessments and Standards in General	
<p>Teachers and principals want assessments to provide ‘actionable data’ they can use for the benefit of individual students. They have several specific concerns and recommendations:</p> <ol style="list-style-type: none"> 1. The current assessment measures student performance on a single day, and many do not believe this is a fair indicator of student ability. More frequent but shorter tests may be more useful in measuring student progress over time. 2. The assessment is conducted too late in the school year to be of interventional value before students move to the next grade. A baseline test in the fall and follow-up test in the spring may provide more actionable data. 3. The results of the assessment arrive too late to be useful for student intervention. Educators hope an online assessment will facilitate a quicker turnaround in results reporting. 	<p>At a policy level, states may consider adjusting the timing and frequency of assessments to enable results reporting that teachers can use to help improve individual student performance.</p> <p>To the extent possible, educators would appreciate a shorter turnaround time in results reporting of assessments.</p>

Assessments and Standards in General (continued)	
<p>Educators are skeptical of the credibility of those who are developing the new assessment. They fear assessment developers are “too far removed from the classroom.”</p> <p>They also want more details on “what was learned in the field testing” – e.g., they want to know what changes will be made as a result of the field test findings and what activities will occur between now and the launch of the new assessment. Several educators said they want to hear from other educators who participated in the field testing to understand what they thought of the new assessment and how stakeholders are responding to challenges in content, navigation, format, estimated time for administration, and technological difficulties.</p>	<p>Communication efforts should describe the credentials of assessment developers and the credibility of the development process. This should be communicated through trusted sources, particularly teachers with current classroom experience who participated in field tests.</p>
<p>Educators want more specific details about certain aspects of the assessment, including:</p> <ol style="list-style-type: none"> 1. How writing skills will be assessed and scored; 2. How ‘critical thinking’ skills will be assessed and scored; 3. If students will receive partial credit on questions with multiple response options that are correct; 4. Specific content that will be included on the assessment; and 5. An accurate estimate of how long the assessment will take to administer (they are hearing a variety of estimates from three to 15 hours). 	<p>Communication efforts should address these requests for detail with demonstrations of questions and scoring as well as an FAQ containing details of content areas and time requirements.</p>
<p>Educators have significant technology concerns in transitioning to a computer-based assessment, including:</p> <ol style="list-style-type: none"> 1. They are concerned schools do not have enough computers and tech resources to conduct the new assessments. 2. They have concerns about broadband and wireless capacity at the school level. 3. They believe students, especially younger students, lack the keyboard skills to perform well on the assessment. 4. They are concerned the assessment system will crash if too many students take the test simultaneously. 	<p>At a policy level, states may consider conducting an audit of technology capacity for the new assessment to make certain schools have adequate resources.</p> <p>Communication efforts may address some of these concerns by providing insights into best practices from the field testing experience.</p>
<p>Educators want reassurance that states will commit to the new assessment for a significant amount of time. They want states to “stop moving the target.”</p>	<p>States may consider communication opportunities to reassure educators of their commitment to the new standards and assessment.</p>
<p>Educators want to know what modifications will be made to the test to accommodate students with special needs, learning disabilities and/or physical disabilities.</p>	<p>Communication efforts should provide this information in an FAQ about the new assessment.</p>
<p>Educators are generally not concerned about a potential decline in scores compared to previous assessments, but they stress the importance of communicating the context of the new assessments to parents early in the process.</p> <p>Educators stated the importance of talking to parents early to help them</p>	<p>Communication efforts should describe to parents the differences in the test compared to previous tests and ensure they understand the first year should be treated as a baseline rather than a point of comparison to</p>

understand “what the new test looks like” and what their students are being tested on.	previous tests.
Assessments and Standards in General (continued)	
<p>Teachers specifically want to know how the new scores will be interpreted (e.g., percent proficient or high/medium/low), how they will be used, and what type of impact they will have on their performance evaluations.</p> <p>Educators generally want to know if state performance goals will be adjusted based on implementation of the Common Core standards and the new assessment. Some suggested the first year of the assessment should be regarded as a “nationwide pilot,” given potential technological problems and other expected challenges.</p>	States may consider communication opportunities to clarify expectations and implications of the new assessment.
A majority of educators want to be helpful resources for parents regarding the new assessment, but many stated they would need more information and a better understanding of the new assessment before they could attempt to help parents understand it.	States should undertake an intensive effort to educate teachers and principals about the new assessment before engaging in outreach with parents and the public. Educators can then serve as resources and guides in the transition.
Online Standard Setting Opportunity	
Educators were completely unaware of both the in-person and online standard setting opportunities.	A more concentrated communication effort is needed to make educators aware of these opportunities.
Educators appreciate the opportunity to provide input through the online panel, but they want reassurance that their input will be used and details about how it will be used.	Communication efforts should provide specific details about how the input will be used for the final recommendations for the standard setting.
<p>Educators expressed considerable concern about participation among the general public in the online panel for the following reasons:</p> <ul style="list-style-type: none"> • The open invitation is vulnerable to sabotage by special interest groups and others with a political agenda. • Higher education stakeholders, representatives from the business community, and other non-educators may skew the process, as they do not have practical experience teaching in the classroom. • “The public” likely does not have the experience, education, and/or expertise to provide valuable input and may be counterproductive to the rating efforts. Educators were somewhat offended that their expertise was not valued more highly, as they felt they were being “lumped in with” the general public in the online panel. • The open invitation process makes it difficult to ensure participation among a true cross-section of the population, leading to a potentially biased sample. • Students should also have an opportunity to provide feedback. 	Communication efforts should clarify to educators how input from the general public will be used in standard setting.

Online Standard Setting Opportunity (continued)	
<p>Regarding the draft invitation for the online standard setting, educators made the following recommendations:</p> <ol style="list-style-type: none"> 1. The language should be clear that the focus is not limited to students who are college-bound. 2. The online panel—the real purpose of the invitation—should be mentioned in the opening paragraph. 3. Educators are skeptical of the terms, “higher standards” and “academic checkup,” as they are vague and subjective. 4. Teachers questioned the use of the words “fair” and “accurate” for all students, stating there has never been a test that equally measures the range from special education students through gifted students. 5. Educators stated concern with the phrase “college and career readiness,” and questioned how this differs from “grade-level readiness.” 	<p>Communication efforts should incorporate these recommendations in materials for educators.</p>
<p>Questions asked by the educators in response to the invitation language include the following:</p> <ul style="list-style-type: none"> • Will field test results, including sample answers, be included in the panel process? • How are “Achievement Levels” defined, and will the scores be categorized (such as basic, below basic, etc.) or numeric? • Providing input in October seems like a short time frame in which to finalize the assessment for this next school year. How will the input provided during this panel be incorporated into the assessment? • Will there be periodic assessments to give students an opportunity to practice between now and the official administration of the new assessment? • With an emphasis on ensuring high school students are prepared for “success after graduation” as indicated by the exam given at the end of the 11th grade, what is supposed to be taught/focused on in 12th grade? 	<p>Communication efforts should include an FAQ to address these questions.</p>
<p>Educators wanted clarity on whether or not the three-hour window could be split into multiple sessions or had to be done in one sitting, as a single three-hour block may be harder to schedule.</p>	<p>Communication efforts should include an FAQ to address these questions.</p>

Statement Testing

Smarter Balanced developed sets of statements about the importance of the effort to inform educators and the general public about the assessment system that is being developed. The following table outlines the statements presented to educators and summarizes participants’ level of agreement with the statements and other specific feedback. The feedback should be used to refine messaging for communication efforts.

Statement	Educator Feedback
-----------	-------------------

<p><i>The Common Core State Standards will have a positive effect on students' ability to think critically and use reasoning skills.</i></p>	<p>General agreement Educators generally agreed with this statement, saying emphasis on critical thinking is a positive shift from existing state assessments. However, some stated that while the philosophical intent is good, it may not align with reality. To address skepticism, educators need evidence to support the claims in this message.</p>
--	--

Statement	Educator Feedback
<p><i>Consistent standards will provide clear guidelines for teachers and principals and improve efficiency, especially when integrating students from other states.</i></p>	<p>General agreement Educators generally agreed consistency is desirable, as long as the standards are not changed frequently, as has been the case in the past. Some suggested the potential for improved efficiency is there if assessment results can be returned in a timely manner. Educators want actionable data from the assessments that they can use to help students learn more effectively.</p>
<p><i>The Smarter Balanced assessment is a test by states and for states, developed with ongoing input from educators.</i></p>	<p>General disagreement Educators conveyed the following reasons for disagreeing with this statement:</p> <ul style="list-style-type: none"> • It implies each state is involved, which is not the case. • The development process for the assessment is generally not well understood, including the claim of “ongoing input” from educators, of which many of the participants were skeptical. They wondered who the educators are who have provided input and if their input is valid and representative. • There is no proof, evidence of credibility, or frame of reference to back up the statement.
<p><i>The Smarter Balanced assessments are a key part of implementing the Common Core standards and preparing all students for success in college and careers. The assessments will replace existing tests and provide an “academic checkup” for students, parents, and teachers by measuring real-world skills such as critical thinking and problem solving. In addition, they will provide information that will give teachers and parents a better picture of where students are succeeding and where they need help.</i></p>	<p>General disagreement Educators generally disagreed with this statement, saying that “a key part” is an overstatement of the role and impact of the assessment, which they believe will likely provide less value than more regular assessments of student learning (such as weekly or biweekly tests). Other educators agreed the results are useful for guiding future instruction but do not offer benefits to the individual students taking the tests. In other words, teachers and principals want information that “can help current students where they need help, which is not likely with these assessments.” Educators again expressed concern about the phrase “success in college and careers.”</p>
<p><i>Participating in the Consortium will help educators share and learn from best practices through collaboration with other participating states.</i></p>	<p>General agreement Educators were generally encouraged by this statement and agreed with the concept, although some stated skepticism about its reality. Several educators specifically highlighted the benefit of multiple states’ involvement. Educators want more information about what is meant by “best practices” and who is participating in the Consortium.</p>
<p><i>Educators will have widespread opportunity for input through the Online Panel, in addition to an in-depth</i></p>	<p>General agreement Educators were generally encouraged and interested to learn more about these opportunities. Several were skeptical as to whether the</p>

<i>opportunity for input through the In-Person Panel.</i>	opportunities were real and they wanted more details.
<i>Educators have the opportunity to see actual test items and experience the test as their students will experience it.</i>	Strong agreement Several educators conveyed strong agreement with this statement, suggesting the biggest benefit to participation would be in experiencing the test as students will. A few principals suggested this statement would be particularly compelling to their teachers in encouraging their participation.

Phrase Testing

Educators assessed seven potential objectives for schools to pursue, indicating the relative importance of each in terms of benefits to students, schools and society more broadly. The feedback should be used to refine messaging for communication efforts.

<i>How important is it for schools to...</i>	Educator Feedback
<i>Prepare students for success in college, careers, and life.</i>	Medium to high importance The majority of educators agreed this is of high importance for schools. Elementary and middle school principals indicated medium, stating their students are generally too young to be focused on these long-term aims.
<i>Teach students to think and communicate.</i>	High importance Educators agreed this is of high importance, as “it is impossible to teach all of the content, and students need to be able to apply it and think.” One principal recommended rewording this to: “Teach students to problem solve, reason and communicate effectively.”
<i>Give families access to effective education options.</i>	Undecided (statement seen as too vague) Most educators were confused by the statement, not understanding what is meant by “options.” Educators questioned if this includes factors like charter schools, vouchers, alternatives to public education, online education, etc.
<i>Deliver on consistent, high expectations.</i>	Medium to high importance Educators agreed standards must be consistent across schools and states and apply to all students, but with the recognition that “what’s high for one student isn’t necessarily high for another.”
<i>Equip students for the real world.</i>	High importance Educators agreed this is of high importance but questioned the meaning of “real world.” One principal suggested rewording to “equip students to be better problem solvers.”
<i>Deliver real, measurable results.</i>	High importance Educators agreed this is of high importance but questioned the meaning of “real, measurable results.” One suggested replacing “real” with “actionable.” A few principals also mentioned the value of more frequent testing with timely results.
<i>Reward effective schools.*</i>	Low importance Educators expressed concerns about rewarding effective schools and questioned how “effective” would be measured. They suggested too many factors out of schools’ control (such as student background and socioeconomic status) could impact “effectiveness.” Several also felt “effective” should be expected without warranting additional rewards.
<i>Hold failing schools accountable.*</i>	Low importance Educators were generally offended by this statement, suggesting success is defined differently for different students. Similar to the last statement, they believe there are too many factors schools cannot control that may

	contribute to overall “success” or “failure.”
--	---

Parents

Key insights and resulting recommendations from the parent focus groups are outlined in the table below.

PARENTS	
Insight	Recommendation
Assessments and Standards in General	
<p>Several parents specifically stated that standardized testing provides the following benefits:</p> <ol style="list-style-type: none"> 1. Understanding where the school or state compares on average to other districts, schools, states, or countries 2. Feedback on specific students’ strengths and areas for improvement 3. Opportunities for scholarships based on the results of assessments 4. The concept of using assessments to hold teachers and schools accountable 	<p>These perceived benefits should be emphasized in communication efforts with parents.</p>
<p>Despite these benefits, parents expressed many concerns about academic assessments, including:</p> <ol style="list-style-type: none"> 1. The time taken away from instructional learning time and resources (time and money) put into testing 2. An emphasis on memorization and “teaching to the test” and a perceived lack of emphasis on critical thinking skills or creativity 3. An emphasis on using the results to determine school and teacher recognition and funding/compensation rather than using the results to benefit students 4. The lack of flexibility in standardized assessments to cater to students with different backgrounds, resources, learning styles and abilities, and comfort levels with academic testing 5. A failure of testing to provide valuable feedback of student progress in a timely manner that can be acted upon 6. The potential for inaccurate measures of student achievement as a result of shortcomings in the assessments 7. The current over-testing of students and the pressure testing puts on students to perform 8. An overemphasis on the roles schools and teachers should play in student achievement and an under-emphasis on the roles parents and communities should play 	<p>FAQs and other materials should be developed to address these concerns in communications efforts with parents.</p>
<p>Parents had low awareness of the new assessment Smarter Balanced is developing. When presented with a brief overview, parents asked the following questions:</p> <ol style="list-style-type: none"> 1. What is the format of the new assessment? 2. How will it be administered? 3. How is this assessment different from the existing standardized test in the state? 4. Will it replace the existing test in the coming year? 5. Is the new assessment proven to be better than previous assessments, and if so, how was it proven? 6. What will the new assessment measure specifically—especially compared to what previous tests measured? 7. Is it worth the money spent to transition from existing tests? 8. How long has CCSS been in place, and shouldn’t the schools teach CCSS for a 	<p>A more concentrated communication effort is needed to make parents aware of the new assessment and what it means for their children.</p> <p>FAQs and other materials should be developed to</p>

<p>few years before testing?</p> <p>9. What is Smarter Balanced hoping to achieve in field testing, and what information will be used to determine if it was a success or not?</p> <p>10. What is the value of the new assessment to the students themselves? Will it improve future chances of success?</p>	<p>address these questions in communications efforts with parents.</p>
Assessments and Standards in General (Continued)	
<p>11. What will the results of the new assessment look like, and how will we know if the test is measuring what it is supposed to measure?</p> <p>12. What is the long-term goal? Will the state commit to this assessment for the long term?</p> <p>13. What will be done to prepare teachers and students for the new assessment to alleviate stress from the transition?</p> <p>14. If the goal is to apply these standards and assessments to all states, why are only 22 states currently on board?</p> <p>15. On the other hand, if 22 states are already involved, will this become a federal program? [Several parents had to be reminded that a state consortium is developing the new assessment, not the federal government]</p> <p>16. How are the standards and assessments supposed to be consistent across states if I move to a non-Smarter Balanced state or a PARCC state?</p> <p>17. Will results of the testing affect school funding at the state level?</p>	
<p>The groups discussed the possibility that states may see lower test scores compared to previous assessments after implementing the new Smarter Balanced assessment. When asked what their reaction would be, lower-income parents generally conveyed slightly more concern about this potential drop in scores compared to higher-income parents.</p>	<p>Communication efforts should describe to parents the differences in the test compared to previous tests, and ensure they understood the first year should be treated as a baseline rather than a point of comparison to previous tests.</p>
Online Standard Setting Opportunity	
<p>No parents from either state or income group were aware of a standard-setting process or the opportunity to get involved through in-person workshops or an online panel to support the development of the new assessment. Despite not having heard of this opportunity, about three-quarters of parents were interested in learning more and said it would be an interesting opportunity.</p>	<p>If increased participation of parents in the online panel is a goal, a more concentrated communication effort is needed to make parents aware of the opportunity.</p>
<p>In response to the invitation to the online panel, several parents expressed interest, but were concerned they lacked the expertise for meaningful input. Many preferred to leave the input to “true education experts,” as they were concerned they would bias the results of the standard-setting exercise. Without a background or expertise in education, some felt they could participate only if they could be prepared with</p>	<p>Parents should be made aware of various options for engagement, including options in</p>

<p>information and resources ahead of time.</p> <p>Most parents were interested in the opportunity as a way to experience what their students experience in taking the assessment. While some parents are interested in taking a sample test, most prefer to trust teachers to define the standards and measures for each grade level.</p>	<p>addition to participating in the online panel, as many have concerns about their qualifications.</p>
--	---

Statement Testing

Smarter Balanced developed sets of statements about the importance of the effort to inform educators and the general public about the assessment system that is being developed. The following table outlines the statements presented to parents and summarizes their level of agreement and other specific feedback. The feedback should be used to refine messaging for communication efforts.

Statement	Parent Feedback
<p><i>Nothing is more important than giving all students an education that prepares them for success. Yet independent research shows America’s schools are failing our children. Though many assume that poverty in America is pulling down education scores, research shows that middle-class U.S. students are far behind their peers in other countries. We need to ensure that our public education systems prepare students for success after high school.</i></p>	<p>General agreement Parents said it is important for the education system to ensure all students are prepared for life after high school, not just those who are college-bound, by equipping them with knowledge and skills necessary to enter the workforce. Schools should prepare students for “whatever opportunities are appropriate for them.”</p>
<p><i>The Common Core standards represent an effort to ensure that every student in every state has the opportunity to learn the same topics and master the same skills through a common national standard, so that no student is denied basic learning.</i></p>	<p>General agreement Despite general agreement that standards should be consistent across states and communities, parents said standards need to be flexible to accommodate both “high and low achievers.”</p>
<p><i>High standards will help more children succeed. Assessments are a tool for ensuring we reach those standards.</i></p>	<p>Mixed response About half of parents agreed assessments help measure progress toward meeting standards and place accountability on the teacher and student to perform well. The other half stated the following concerns or points of disagreement:</p> <ul style="list-style-type: none"> • There is too much pressure on students to perform, which may lead them to fail. • Standards should be more inclusive of job-related skills rather than just English and math. • Standards need to be “reasonable and attainable”
<p><i>The Smarter Balanced assessments are a key part of implementing the Common Core standards and preparing all students for success in college and careers. The assessments will replace existing tests and provide an “academic checkup” for students, parents, and teachers by measuring real-world skills such as critical thinking and problem solving. In addition, they will provide information that will give teachers and parents a better picture of where students are succeeding and where they need help.</i></p>	<p>General agreement This statement was seen as wordy by many, but most parents agreed preparing students for success after high school—not just college and careers—is important. Several parents liked the emphasis on “real-world skills and critical thinking,” but they want proof that “real-world skills” will be measured. Parents are skeptical that a test can accurately assess these skills.</p>
<p><i>It is better for all states to have the same tests at each grade level in math and English so test scores can be compared across states.</i></p>	<p>General agreement Parents generally agreed with this statement, while several asked the question of what modifications</p>

	would be made to accommodate children with special needs or those with different learning styles.
<i>Assessments need to be more effective, accurately measuring true knowledge and skills—not just memorized information.</i>	General agreement Parents agreed the movement away from memorization and to greater emphasis on critical thinking skills was a positive change. However, they want reassurances of the assessment’s accuracy.
<i>The Common Core State Standards and assessments will help reduce inequality and ensure we hold ourselves accountable for improving education.</i>	Mixed response About half of parents agreed with this statement, but a few wanted assurance that teachers would have the resources necessary to teach to these standards and not be penalized for reasons out of their control.

Phrase Testing

Parents assessed seven potential objectives for schools to pursue, indicating the relative importance of each in terms of benefits to students, schools and society more broadly.

<i>How important is it for schools to...</i>	Parent Feedback
<i>Prepare students for success in college, careers, and life.</i>	High importance Parents agreed this is of high importance, but want to emphasize the role parents, church, and/or the larger community also play in a child’s success.
<i>Teach students to think and communicate.</i>	High importance Several parents suggested more emphasis should be placed on critical thinking and problem solving rather than “think and communicate.” With the abundance of social media and other online communication, parents said they want students to communicate more <i>effectively</i> , rather than simply communicating <i>more</i> .
<i>Give families access to effective education options.</i>	Undecided Similar to educators, the majority of parents found this statement too vague. One lower-income parent found this to be of high importance, suggesting parents should always have access to better options for their children.
<i>Deliver on consistent, high expectations.</i>	Medium to high importance One parent stating this was of high importance interpreted this to mean that schools are doing the best they can, while another questioned, “whose expectations and what serves as the benchmark?” Two parents suggested expectations be realistic and not just high, as they cannot be the same for everyone.
<i>Equip students for the real world.</i>	Medium to high importance Parents agree education needs to be updated to help students live productive lives after high school, but “real world” should be defined.
<i>Deliver real, measurable results.</i>	Undecided One parent labeled this of high importance, but several others found it too vague or subjective. They wanted to know what the results would be measured against.
<i>Reward effective schools.</i>	Low importance Parents stated there is too much competition among teachers, principals, and schools, suggesting more of the focus should be placed on students. Some parents worried that rewarding effective schools may lead to greater inequalities, and suggested expectations be adjusted for schools with different levels of resources.
<i>Hold failing schools accountable.</i>	Mixed response Parents suggested schools need to be evaluated within the context of the resources they are provided and the setting in which they operate. Holding failing schools accountable will not be fair unless these factors can be controlled. Further, withholding resources from failing schools may cause them to fail further. However, lower-income parents in particular agreed this is very important and suggested that if schools are not held

accountable, they will continue to perform poorly.

Word Choice

Parents were asked to share their preference between different words being used to describe components of the new assessment. The table below outlines feedback from parents on specific word choices.

Word Choice	Parent Feedback
<i>“Assessment” or “Test”</i>	Assessment The majority of parents preferred “assessment,” stating it implies “more of a gauge of where students <i>are</i> versus placing the emphasis on right or wrong.” “Assessment” puts less pressure on students, and has less of a negative connotation. One parent preferred “test,” stating that it means the same thing but is “more honest about calling it what it is.”
<i>“Math and English” or “mathematics and English Language Arts”</i>	Math and English Parents preferred the simpler “math and English,” this is generally what students call it. However, one parent highlighted that it does get confusing when schools use different terms for English, language arts, etc.
<i>“Preparing students for college and career” or “Preparing students for success after graduation”</i>	Preparing students for success after graduation Parents unanimously preferred this “more inclusive” phrase for all students, especially those who are not college-bound.

Detailed Findings: Teachers and School Principals

Responses to each section of the moderator’s guide for educators are summarized as follows. For the most part, teachers and principals in both California and Michigan felt similarly about assessments, the Common Core State Standards, and Smarter Balance Assessment Consortium. It is explicitly stated in the report when feedback differed among participants or was mentioned by only one segment or another. Direct quotations from participants are indicated by quotation marks.

Awareness, Knowledge, and Attitudes Regarding Common Core, Common Core Assessment, and Standard Setting

Standards and Assessments

When asked about the importance of assessments, educators gave mixed responses. Most teachers and principals agreed it is important to gauge student progress and use results to address student needs, but they believe assessments are often unfairly used to judge teacher performance. One teacher described assessments as “micromanaging” teachers to “see if they are doing everything they are supposed to.” In contrast, several principals said they find assessments to be a sound measure of school and teacher performance, while two principals stated specifically they like using the results to identify areas of intervention for students but prefer not to hold teachers accountable for student scores. A few others expressed concern that assessments only capture a snapshot of one day in time, and that the “one test, one day” model does not help teachers monitor progress over time for students. According to three educators, more frequent assessments would provide “more valid measures with more utility,” and many participants agreed with this.

Several educators mentioned challenges with the timing of the test within the school year (generally toward the end of the spring), expressing concern that it is administered too late to be of interventional value before students go on to the next grade. They also expressed concern about the turnaround time in receiving results from assessments. Several educators stated that by the time they receive results for a given assessment, they have a whole new set of students, and two others described the results as a “reflection piece” or an “autopsy” rather than actionable data with which to make changes to help students. Having the scores earlier would help

educators “determine class placements and set master schedules around the data rather than the other way around,” according to one principal. Educators want assessment results they can act on for the benefit of students, and they believe they need either more frequent testing or testing that is more strategically timed. One teacher suggested administering a baseline test in the fall and a follow-up in the spring, and others expressed support for this concept.

Three teachers stated there is a “disconnect” between what is asked on tests and what is expected in the classroom. One further suggested that the external parties charged with creating the assessments are too far removed from the classroom, and the expectation feels very “top-down” rather than inclusive of the teacher perspective. Teachers often feel pressure to change the curriculum continuously to meet the standards measured by the tests, according to two educators, but frequent changes to the standards can put stress on both the teachers and the students in an already challenging and time-limited school year. Another concern, stated by one teacher, is that standardized tests do not often address writing skills, ability to think critically, or communicate an issue fluently, and that is a limitation. Further, if tests do include writing that is computer-graded, students—and teachers—quickly pick up on how to “cut corners” and write to the test in a formulaic way that lacks creativity. Along these lines, two educators mentioned teachers are “smart at adapting,” and they learn how to teach to the testing system, especially when they perceive their job performance ratings are associated with outcomes. Another teacher also mentioned a limitation of standardized tests is that students do not receive partial credit for their responses, which led some to question the accuracy of assessment results.

Further, standardized testing in general is seen among many educators as disruptive to regular class time, diminishing a teacher’s ability to teach important content for each grade level.

Finally, a number of educators expressed concern about the lack of flexibility in standardized testing for students with different learning abilities or disabilities, college or career aspirations, comfort levels with testing and/or technology, and socioeconomic or cultural backgrounds. Further, some students do not take standardized testing as seriously as others, particularly if grades are not associated with them or if particular students are less vested in school-based learning. Educators argue all of these factors impact student performance, which in turn impact teachers’ wages, bonuses, and career progression. Two educators expressed concern about “the push” for students younger than 5th grade to take standardized tests, arguing that these younger students are less inclined to understand the importance of the assessments or take them seriously. These educators believe other state or country benchmarks are more appropriate for lower grade levels.

Smarter Balanced Assessment Consortium

Most, if not all, educators had heard of Smarter Balanced in the context of the new assessments and the current field testing.

Common Core Assessment

General Awareness and Attitudes

Educators were generally aware of—and many were in the middle of—field testing. Despite being exposed to the testing through the piloting and general updates, one principal still felt the entire process and the developing assessment were “shrouded in mystery.” This educator wanted to provide input on the process, be reassured his input would be used, and receive a greater understanding of *how* his input would be used. Another stated there has been a lot of conflicting information shared with principals and that one estimate was that students would spend 15 hours of class time on this testing, which was a significant source of concern.

One elementary school teacher shared her experience with kindergartners and 3rd graders coming in to take the test and lacking the computer and keyboard skills necessary to log in, type full sentences, and complete the test. However, another 5th grade teacher had a positive experience with the computer-based format and found this shift in format to be more appropriate for this grade level. With regard to younger students taking the test, one teacher stated that kindergarten and 3rd grade is too young for such a rigorous testing environment, and it could result in taking away students’ love for learning at such a young age.

Principals in particular highlighted the benefit of having a single, consistent assessment across states to be able to compare progress. However, several teachers and principals want assurance and commitment that the standards will not continue to change as they have many times over recent years. As one principal put it, consistent standards will be a good thing, “just stop moving the target!”

Finally, two Michigan educators discussed political issues they believe may result in the delayed adoption of the new assessments in their state. One other science teacher asked specifically whether and when science is going to appear in new assessments, as science is his area of teaching and his main concern with his students.

Perceived Barriers and Limitations

One limitation of the assessment mentioned by several educators, both teachers and principals, is that their schools simply are not technologically or otherwise equipped to test every student in the impacted grade levels. Schools have too few computers, too many students, not enough staff available to proctor, and inconsistencies with internet availability or capacity of wireless access points, among other issues, and districts or states are not providing the necessary resources to get schools adequately prepared to administer the test successfully. These issues were expressed by educators from both high- and low-resource schools. Further, several educators expressed concern that the assessment system itself may not be built for the needed capacity, with too many users potentially crashing the system.

Several educators expressed concern about the accuracy of the new assessment, generally out of a lack of familiarity with what will be tested and the specific content for which students will be held accountable. One teacher described the assessment development process as “an outside company evaluating students and teachers they know nothing about.” He likened it to asking someone else to test drive a car for him that he was going to purchase, allowing someone else to evaluate his students at a national level without taking into account regional or school-level differences. One teacher took the field test that is available online and felt there were a number of questions she was unsure of as an educated adult. She expressed concern that 4th or 5th graders may not have the critical thinking skills needed to perform well on this test. In contrast, one teacher stated he has been getting regular and consistent information from his principal about the Smarter Balanced assessments and is on board with the change. He sees a lot of benefits in the shift in content being tested and the update in format. Regarding the higher grade levels, one principal expressed concern about the lack of a Smarter Balanced assessment between 8th and 11th grade, which potentially negatively impacts how students will be able to demonstrate progression appropriately and accurately.

With regard to the content, while many educators appreciate the intention of shifting the content to measure more critical thinking skills, many are worried about the achievement levels and if students will be able to grasp the new approach. Many see it as a significant change from the previous memorization-based assessments. The new approach has the potential to be a positive shift, but educators and parents both want the assessment to be fair and reasonable for the students taking it.

A number of educators questioned if modifications would be made to the test to accommodate special education students or others with learning or physical disabilities. Particularly for a test that is designed to measure critical thinking, some educators believe some students with special needs may have an unfair disadvantage for not having been taught to think this way.

While most educators expressed some degree of concern about adequate access to technology and bandwidth for online testing, most participants stated the computer-based format would not be a significant challenge for most students. They generally believe the computer-based format is, in theory, a step in the right direction, although some expressed concern that students may have varying comfort levels in working on computers—as opposed to smartphones or tablets—which could impact scores. Two educators expressed that the computer-based format would be more engaging than a paper-pencil format and that the use of technology plays to most students’ strengths.

Reactions to Potential Drop in Scores

The moderator introduced the possibility that states may see lower test scores compared to previous assessments after implementing the new Smarter Balanced assessment. When asked what their reactions would be, all of the educators stated they would not be surprised and they would expect a drop in scores resulting from a change in format and content. Most felt comparing the Smarter Balanced assessment to previous assessments would be like “comparing apples to oranges,” as they are essentially different tests. Teachers specifically wanted to know how the new scores would be interpreted (e.g., percent proficient or high/medium/low), how they would be used, and what type of impact they would have on their performance evaluations. Ultimately, the educators expressed the hope that teachers and students would adjust to the new assessment over time, and improvement would be seen in scores in subsequent years. Several educators suggested the states allow for a “transition period” to enable teachers to adjust to the new test before it impacts their performance evaluations. One principal felt assured that a drop in scores would not impact any high school student’s college prospects, as higher education admissions officers are more concerned with scores from ACT, SAT, and AP exams.

To reassure parents about a potential drop in scores, educators suggested they would describe to parents the differences in the test and ensure they understood the first year should be treated as a baseline rather than a point of comparison to previous tests. With this, several educators stated the importance of talking to parents early to help them understand “what the new test looks like” and what their students are being tested on, so they can be “vested in their students’ lifelong learning.” One teacher suggested—and several agreed—that ‘Gallery Walks’ to show how students are being assessed would be a helpful immersion experience for parents. Two educators expressed concern about parents asking them why they are testing students if the assessment does not count; they suggested consistency in what is told to parents will be critical. Another teacher suggested parents at her school really do not care about standardized test results and will not think it is “a big deal” if they see a drop in scores. A possible exception to this could be parents of the highest achieving students, who are generally more involved and aware than the average parent. Two teachers questioned the type of details parents would see from assessment reporting in terms of benchmarks or subject-level scores. They suggested parents might find this information useful as they support their children in the learning process. A majority of educators wanted to be helpful resources for parents regarding the new assessment, but many stated they would need more information and a better understanding of the new assessment before they could attempt to help parents understand it.

Several educators also suggested the students themselves may be confused or concerned by a drop in scores, and this will also need to be addressed early and with clear explanations and expectations to reduce anxiety.

One Michigan principal spoke specifically about a state-level mandate to get all schools to 85% math and English proficient by 2022 and questioned whether this expectation would be adjusted based on implementation of the Common Core standards and the new assessment. He said expectations are generally good to have in place, but the changes in curriculum and assessment may warrant some flexibility in the timing of the transition.

Results of Field Testing

Coming out of the field testing, educators want “transparency” – e.g., they want to know what changes will be made as a result of the field test findings and what activities will occur between now and the launch of the new assessment. Several educators said they want to hear from other educators who participated in the field testing to understand what they thought of the new assessment and how stakeholders are responding to challenges in content, navigation, format, estimated time for administration, and technological difficulties. Educators want to be able to plan for these challenges to manage them as smoothly as possible in their schools. Three teachers stated they simply want to see more sample questions, as they had heard so little about specific content on the assessment. One principal suggested having the first year of the new assessment count as a “nationwide pilot,” given the anticipated technological difficulties, the potential drop in scores, and other expected challenges.

Standard Setting

No educators from either state were aware of a standard-setting process to support the development of the new assessment, and none had heard of an opportunity to participate in the in-person workshops or online panel. Despite not having heard of this opportunity, nearly all educators were interested in learning more and stated it would be an interesting opportunity. However, in reaction to this opportunity to get involved, a few educators questioned why the standard-setting was happening only at this point in the development and had not happened already among experts in the field, stating, “That’s not how standardized tests should be developed.” Further, two principals reacted adversely, saying they should not be the ones participating; rather, it should be the teachers, who are more knowledgeable and better equipped to provide valuable feedback to the process.

One teacher, who had previously expressed concern about the questions on the test after trying to answer them herself, stated she would want to be involved in the assessment’s development simply to ensure the test did not include “trick questions” or questions that are “so diluted that the students would not be able to sift through what is actually being asked.” Another teacher, after taking the test himself, claimed the development of the new assessment “has left large audiences out.” He suggested the development process needs to ensure various perspectives are represented appropriately. One principal wanted to know specifically if the assessment is criteria-referenced or norm-referenced, and wanted to know more about the bookmarking process in standard setting.

Message Testing

Educator Invitation Language

When presented with the following invitation to participate in the online standard setting process, educators were generally interested in the opportunity but had specific feedback about aspects of the language in the invitation.

Nothing is more important than giving all students an education that prepares them for success. That is the goal of the Common Core State Standards. The standards help ensure that all students leave school prepared for success after graduation and also provide clear, consistent guidelines for educators and school principals.

The Smarter Balanced Assessment Consortium is developing new assessments aligned to these higher standards. The tests will provide a meaningful and fair ‘academic checkup’ and help educators chart progress toward the goal of improving student achievement.

To ensure that the tests are rigorous, fair, and accurate for all students, Smarter Balanced is inviting educators to participate in an Online Panel for recommending Achievement Levels on their assessments. Participants will recommend an Achievement Level score that demonstrates how much students should know or be able to do in order to meet the grade-level standards and to be on track for eventual college and career readiness. You can participate from the comfort of your own home or office, at a time that works best for you.

Educators will have the opportunity to preview and analyze the test’s items and structure and provide input into the process of recommending Achievement Level scores. Participation is free and open to the public. This opportunity will allow a large group of educators and principals to help establish consistent measures of progress for the Smarter Balanced interim and summative assessments.

The Online Panel to recommend Achievement Levels will occur October 6–17, 2014. It will take participants up to 3 hours to complete the orientation process, review test questions, and recommend a score. Each participant will have a two-day window to complete the entire process.

In the first paragraph, several educators stated “success after graduation” can mean something different for each student. They suggested the language should be clear that the focus is not limited to students who are college-bound. Similarly, both teachers and principals expressed concern that the CCSS and the assessment itself are too focused on the skills students need to succeed in the college-bound path and do not place enough emphasis on other or alternative skills students may need to enter the workforce. Another teacher stated there was simply too much information provided in the first paragraph, making it “hard to digest.” Further, one teacher suggested the online panel—the real purpose of the invitation—should be mentioned in the opening paragraph for clarity.

Similar to concerns with the first paragraph, educators again stated the terms “higher standards” and “meaningful and fair” are difficult to apply to all students of every level and every background. Further, several principals stated that to be “meaningful,” testing should be conducted at regular, ongoing intervals to gauge individual student progress. A few educators disliked the term “higher standards,” saying it has always been their intention to teach to a higher standard. They believe this new curriculum and assessment are just “applying a new format,” and not necessarily helping to achieve higher standards. One teacher suggested language should be added to convey, “this is the goal [to be achieved] in a couple years, when we work out all the kinks,” indicating that the first year or two will have challenges and may not live up to the ideal described in the language of the invitation. Three principals disliked “academic checkup,” saying it seemed too informal a phrase when the results impact evaluations, salaries, and school resources. They felt the language should better reflect the gravity of the assessment and how it is used. They also commented the term “sounds more like a medical exam or an autopsy than a system in place to provide meaningful data.”

Regarding the third paragraph, educators generally agreed the language clearly stated the purpose of the online panel and they appreciated the opportunity to be involved in the process. Two teachers questioned the use of the words “fair” and “accurate” for all students, stating there has never been a test that equally measures the range from special education or English as a Second Language (ESL) students through gifted students. Several other teachers and principals stated concern with the phrase “college and career readiness,” and questioned how this differs from “grade-level readiness.” One principal agreed with the concept of involving a wide range of stakeholders in the panel but was hesitant about the benefit and practicality of a score being dictated by such a wide range of people. Finally, one high school English teacher was critical of the grammar used in this paragraph and insisted the writing should be better if the intention is to support Common Core skills.

In the fourth paragraph, participants focused on the sentence, “Participation is free and open to the public.” Educators expressed considerable concern about this sentence for the following reasons:

- The open invitation is vulnerable to sabotage by special interest groups and others with a political agenda.
- Higher education stakeholders, representatives from the business community and other non-educators may skew the process, as they do not have practical experience teaching in the classroom.
- “The public” likely does not have the experience, education and/or expertise to provide valuable input and may be counterproductive to the rating efforts. Educators were somewhat offended that their expertise was not valued more highly, as they felt they were being “lumped in with” the general public in the online panel.
- The open invitation process makes it difficult to ensure participation among a true cross-section of the population, leading to a potentially biased sample.
- Students should also have an opportunity to provide feedback.

One principal pointed out the fourth paragraph is really asking for three different activities related to the assessment: (1) previewing, (2) analyzing the items and structure, and (3) providing input into the process. He suggested different segments of the population may have valuable contributions to each of these three distinct activities. One educator was confused by the paragraph’s switching between the educator’s role and the role of the public and wanted this to be clearer. Finally, one educator wanted a definition of “Achievement Levels,” which is a “nebulous concept” in the context of the invitation.

For the final paragraph, a few educators suggested an incentive (such as compensation or professional development credit) for participation may increase participation rates and the likelihood that a broad range of participants is represented, especially given teachers' already heavy workload. Several educators were concerned about the time commitment and the narrow window for participation, especially given typically busy schedules in the fall with sports events, conferences, etc. Several others wanted clarity on whether or not the three-hour window could be split into multiple sessions or had to be done in one sitting, as a single three-hour block may be harder to schedule. However, most educators agreed the time commitment seems feasible, and many agreed that despite it being difficult to find the time, many interested educators would make it work because of the importance of the effort. In contrast, one principal conveyed skepticism about the whole process, saying her teachers remember going through this process with the Michigan Educational Assessment Program (MEAP), and suggesting it would be difficult to engage them again for a new assessment.

Questions asked by the educators in response to the invitation language include the following:

- Will field test results, including sample answers, be included in the panel process?
- How are "Achievement Levels" defined, and will the scores be categorized (such as basic, below basic, etc.) or numeric?
- Providing input in October seems like a short time frame in which to finalize the assessment for this next school year. How will the input provided during this panel be incorporated into the assessment?
- Will there be periodic assessments to give students an opportunity to practice between now and the official administration of the new assessment?
- With an emphasis on ensuring high school students are prepared for "success after graduation" as indicated by the exam given at the end of the 11th grade, what is supposed to be taught/focused on in 12th grade?

Specific Statement Testing

Smarter Balanced developed sets of statements about the importance of the effort to inform educators and the general public about the assessment system that is being developed. The following table outlines the statements presented to educators and summarizes participants' level of agreement with the statements and other specific feedback.

Statement	Educator Feedback
<i>The Common Core State Standards will have a positive effect on students' ability to think critically and use reasoning skills.</i>	General agreement Educators generally agreed with this statement, saying emphasis on critical thinking is a positive shift from existing state assessments. However, a few stated that while the philosophical intent is good, it may not align with reality. One educator disliked the words "positive effect," while another—a fourth grade teacher—stated she is already seeing positive effects with the field testing, and her students have expressed pride in applying their critical thinking skills. One high school teacher disagreed with the statement, stating the English standards require limited higher level thinking skills. Another teacher also disagreed with the statement, claiming teachers have been aiming to teach critical thinking for a long time, and "the CCSS are not what will bring this change about," as "these standards are just words on paper with the teachers and students being the real stakeholders for change." Another educator commented that national media outlets are "making fun" of CCSS, and this is negatively coloring his perspective.
<i>Consistent standards will provide clear guidelines for teachers and principals and improve efficiency, especially when integrating students from other states.</i>	General agreement Educators generally agreed consistency is desirable, as long as the standards are not changed frequently, as has been the case in the past. However, one high school science teacher expressed concern that the current CCSS are too vague to lead to consistency. A few teachers questioned the phrase "improve efficiency" and wanted clarity on what this means. Another said the potential for improved efficiency is there if

	assessment results can be returned in a timely manner. Educators want actionable data from the assessments that they can use to help students learn more effectively.
<i>The Smarter Balanced assessment is a test by states and for states, developed with ongoing input from educators.</i>	<p>General disagreement Educators conveyed the following reasons for disagreeing with this statement:</p> <ul style="list-style-type: none"> • “By the states” implies each state is involved, which is not the case. • The development process for the assessment is generally not well understood, including the claim of “ongoing input” from educators, of which many of the participants were skeptical. They wondered who the educators are who have provided input and if their input is valid and representative. • There is no proof, evidence of credibility, or frame of reference to back up the statement.
<i>The Smarter Balanced assessments are a key part of implementing the Common Core standards and preparing all students for success in college and careers. The assessments will replace existing tests and provide an “academic checkup” for students, parents, and teachers by measuring real-world skills such as critical thinking and problem solving. In addition, they will provide information that will give teachers and parents a better picture of where students are succeeding and where they need help.</i>	<p>General disagreement Educators generally disagreed with this statement, saying that “a key part” is an overstatement of the role and impact of the assessment, which they believe will likely provide less value than more regular assessments of student learning (such as weekly or biweekly tests). Other educators agreed the results are useful for guiding future instruction but do not offer benefits to the individual students taking the tests. In other words, teachers and principals want information that “can help current students where they need help, which is not likely with these assessments.” Educators again expressed concern about the phrase “success in college and careers.”</p>
<i>Participating in the Consortium will help educators share and learn from best practices through collaboration with other participating states.</i>	<p>General agreement Educators were generally encouraged by this statement and agreed with the concept, although some stated skepticism about its reality. Several educators specifically highlighted the benefit of multiple states’ involvement. They suggested educators are simply scouring the Internet for available resources, and a consortium may aid this process of getting teachers the materials they need. A few educators wanted more information about what is meant by “best practices” and who is participating in the Consortium.</p>
<i>Educators will have widespread opportunity for input through the Online Panel, in addition to an in-depth opportunity for input through the In-Person Panel.</i>	<p>General agreement Educators were generally encouraged and interested to learn more about these opportunities. A few were skeptical as to whether the opportunities were real and they wanted more details. Specifically, two educators wanted to know how their input would be used. Several principals suggested providing some sort of incentive (compensatory time, monetary incentives, or professional development credits) to support their teachers to participate.</p>
<i>Educators have the opportunity to see actual test items and experience the test as their students will experience it.</i>	<p>Strong agreement Several educators conveyed strong agreement with this statement, suggesting the biggest benefit to participation would be in experiencing the test as students will. A few principals suggested this statement would be more compelling to their teachers than any of the other statements to encourage teacher participation in the process.</p>

Phrase Testing

Educators assessed seven potential objectives for schools to pursue, indicating the relative importance of each in terms of benefits to students, schools, and society more broadly.

<i>How important is it for schools to...</i>	Educator Feedback
<i>Prepare students for success in college, careers, and life.</i>	Medium to high importance The majority of educators agreed this is of high importance for schools. Two elementary and middle school principals indicated medium, stating their students are generally too young to be focused on these long-term aims.
<i>Teach students to think and communicate.</i>	High importance Educators agreed this is of high importance, as “it is impossible to teach all of the content, and students need to be able to apply it and think.” One principal recommended rewording this to: “Teach students to problem solve, reason and communicate effectively.”
<i>Give families access to effective education options.</i>	Undecided (statement seen as too vague) While two elementary teachers listed this as low importance, most educators were confused by the statement, not understanding what is meant by “options.” Educators questioned if this includes factors like charter schools, vouchers, alternatives to public education, online education, etc., and two stated that options are good, as long as each is held to consistent standards.
<i>Deliver on consistent, high expectations.</i>	Medium to high importance Educators agreed standards must be consistent across schools and states and apply to all students, but with the recognition that “what’s high for one student isn’t necessarily high for another.” A few principals perceived this to mean that educators were not already doing this, which they found offensive, or that higher expectations will solve problems within public education, to which they disagreed.
<i>Equip students for the real world.</i>	High importance Educators agreed this is of high importance but questioned the meaning of “real world.” One principal suggested rewording to “Equip students to be better problem solvers.” Another questioned where this statement fits in with the CCSS.
<i>Deliver real, measurable results.</i>	High importance Educators agreed this is of high importance but questioned the meaning of “real, measurable results.” One suggested replacing “real” with “actionable.” A few principals also mentioned the value of more frequent testing with timely results.
<i>Reward effective schools.*</i>	Low importance Educators expressed concerns about rewarding effective schools and questioned how “effective” would be measured. They suggested too many factors out of schools’ control (such as student background and socioeconomic status) could impact “effectiveness.” Several also felt “effective” should be expected without warranting additional rewards.
<i>Hold failing schools accountable.*</i>	Low importance Educators were generally offended by this statement, suggesting success is defined differently for different students. Similar to the last statement, they believe there are too many factors schools cannot control that may contribute to overall “success” or “failure.”

Trusted Leaders in Public Education

Educators were asked whom they would trust and rely on for more information about the new assessment. Teachers generally indicated they prefer to get information from their principals, department heads, and other teachers. Several teachers mentioned they have less trust in and agreement with the superintendent or other leaders in public education who are further removed from the classroom. They trust teachers who have practical classroom experience. Other responses include:

- Former or master teachers who have risen to positions in the districts, but “only those who are not so far removed from classroom experience” and with “broad enough experience to apply to all students, rather than a specialized group like gifted education”;
- Professional development experts in the areas of English and math;
- District-level representatives from each school; and
- Union representatives.

Principals preferred to hear from experts who have been involved in the development of the curriculum and the assessments, such as a contact from the Silicon Valley Mathematics Initiative, as mentioned by one administrator. They also wanted to hear from other teachers and principals who have gone through the field testing and have first-hand experience with the new assessments. Most principals valued the opinion and credibility of their teachers, expressing their respect for their teachers’ expertise and classroom experience.

Educators generally did *not* want to hear from higher education/university experts, not knowing how well-versed they are in the school- and district-level processes and challenges and highlighting the difference between theoretical approaches to education and “what happens in real life.” One principal mentioned these leaders and researchers are often too pressured by school associations to present unbiased information.

In terms of format, educators preferred to see additional information in videos—possibly delivered by the principals with testimonials from real teachers,

Other Feedback

Following the focus groups, most participants agreed they had learned from the conversation more about the new assessments coming to their schools, but none of them had significantly changed their opinions on the topic. One principal was reassured that other people in his position are “in the same boat” and jokingly said, “But that doesn’t mean the boat isn’t sinking.” He appreciated others are facing the same challenges and have the same concerns. Another educator simply wanted assurance that the new standards and the associated assessment are bipartisan, vetted, voted for, rolled out correctly, funded, and supported completely.

Detailed Findings: Parents

Responses to each section of the moderator’s guide for parents are summarized as follows. Discussions reflected some similarities among parents from both income segments and in both states regarding assessments and the standard-setting process, but a number of differences were expressed in reaction to the potential drop in scores as well as the value of standards and assessments in general. It is explicitly stated in the report when feedback differed among participants or was mentioned by only one segment or another. Direct quotations from participants are indicated by quotation marks.

Awareness, Knowledge, and Attitudes Regarding Common Core, Common Core Assessment, and Standard Setting

Standards and Assessments

When asked in general how they felt about academic testing in schools, parents from both income groups responded with a mix of positive and negative feelings. Several parents specifically stated that standardized testing provides the following benefits:

- Understanding where the school or state compares on average to other districts, schools, states, or countries
- Feedback on specific students’ strengths and areas for improvement
- Opportunities for scholarships based on the results of assessments
- The concept of using assessments to hold teachers and schools accountable

Despite these benefits, parents expressed many concerns about academic assessments, including:

- The time taken away from instructional learning time and resources (time and money) put into testing
- An emphasis on memorization and “teaching to the test,” and a perceived lack of emphasis on critical thinking skills or creativity
- An emphasis on using the results to determine school and teacher recognition and funding/compensation rather than using the results to benefit students
- The lack of flexibility in standardized assessments to cater to students with different backgrounds, resources, learning styles and abilities, and comfort levels with academic testing
- A failure of testing to provide valuable feedback of student progress in a timely manner that can be acted upon
- The potential for inaccurate measures of student achievement as a result of shortcomings in the assessments
- The current over-testing of students and the pressure testing puts on students to perform
- An overemphasis on the roles schools and teachers should play in student achievement and an under-emphasis on the roles parents and communities should play
- Some students’ perceived lack of importance of assessments and their resulting poor performance

One parent of a 4th grade charter school (Waldorf) student stated her school and the school’s parents care very little about standardized testing and “do not make a big deal of it given their philosophies.” She suggested testing is not essential for student achievement.

Smarter Balanced Assessment Consortium

While none of the parent participants had heard of the Smarter Balanced Assessment Consortium, about half had heard of a test being developed to replace the current standardized test in their states, and about three-quarters were familiar with the Common Core State Standards (CCSS)—with higher levels of familiarity among the middle and high-income parents. Of those aware of the CCSS, several parents supported the concept of having consistent standards from state to state, but one parent expressed a distrust of standards coming from the federal government (despite being told these standards are being developed by a consortium of states). One parent expressed concern that the CCSS cannot possibly be appropriate for all levels of students and thus will be unfair to certain student populations.

Common Core Assessment

General Awareness and Attitudes

When presented with a brief overview of the new assessment Smarter Balanced is developing, parents asked the following questions:

- What is the format of the new assessment?
- How will it be administered?
- How is this assessment different from the existing standardized test in the state – i.e., in Michigan, the Educational Assessment Program (MEAP) and in California, the Standardized Testing and Reporting (STAR)?
- Will the MEAP or STAR be given simultaneously with the new assessment, or will it replace the existing test in the coming year?
- Is the new assessment proven to be better than previous assessments, and if so, how was it proven?
- What will the new assessment measure specifically—especially compared to what previous tests measured?
- Is it worth the money spent to transition from existing tests?
- How long has CCSS been in place, and shouldn’t the schools teach CCSS for a few years before testing?
- What is Smarter Balanced hoping to garner from the field testing, and what information will be used to determine if it was a success or not?
- What is the value of the new assessment to the students themselves? Will it improve future chances of success?

- What will the results of the new assessment look like, and how will we know if the test is measuring what it is supposed to measure?
- What is the long-term goal? Will the state commit to this assessment for the long term?
- What will be done to prepare teachers and students for the new assessment to alleviate stress from the transition?
- If the goal is to apply these standards and assessments to all states, why are only 22 states currently on board?
- On the other hand, if 22 states are already involved, will this become a federal program? [Several parents had to be reminded that a state consortium is developing the new assessment, not the federal government]
- How are the standards and assessments supposed to be consistent across states if I move to a non-Smarter Balanced state or a PARCC state?
- Will results of the testing affect school funding at the state level?

Several parents reiterated the importance of the standards and the benefits of having consistent assessments across states; although one expressed skepticism saying, “I’m sure [consistency] would never happen.”

Two parents suggested their children would not have a hard time “picking up on the new format,” as “they are constantly exposed to technology” and it would be like “mastering a new video game.”

Reactions to Potential Drop in Scores

The moderator introduced the possibility that states may see lower test scores compared to previous assessments after implementing the new Smarter Balanced assessment. When asked what their reaction would be, lower-income parents generally conveyed slightly more concern about this potential drop in scores compared to higher-income parents.

Several lower-income parents suggested a drop may indicate teachers are not teaching what they should in schools or being held accountable, resulting in their students not knowing what they need to know. Other lower-income parents said a drop in scores would prompt her to question why they had switched to the new format if students had performed better on the previous tests. When provided with additional information about the tests being more focused on critical thinking skills, she said a drop in scores would be more acceptable, but she would want to see how her son’s scores and those of his school compared to scores across the state to ensure the drop was not out of line with what other students were showing.

In contrast, about half the parents across income groups said they would feel comfortable with a drop in scores as long as their school communicated to them ahead of time the reason for the drop and additional details about the changing assessments. These parents generally understood the new assessment scores should be seen as a baseline rather than be compared to previous standardized tests, as the new assessment incorporates new content in a new format and requires higher level problem solving skills. This understanding led to their acceptance that the new assessment will provide a valid measure of student ability in critical thinking skills. They also agreed the new assessment could help schools, teachers, and parents understand critical areas for improvement. One parent suggested a drop in scores would be more indicative of school and teacher performance than student performance, especially when compared to other schools or districts. However, several higher-income parents suggested a drop in scores would unfairly put the spotlight on lower-income schools with fewer resources and more challenging student backgrounds, placing too much of an emphasis on how teachers are doing and not enough on the societal factors that are out of a teacher’s control.

Two parents shared their optimism that the new assessments would be more reflective of a child’s achievement level because teachers would be unable to “teach to the test” or “fool the testing system.” One said the current California STAR test is outdated and too focused on memorization, so the drop in scores would not be concerning, as it would likely be a more accurate measure. Another parent suggested the new assessments would better enable schools and teachers to focus on areas for improvement. With the possibility of declining scores, one higher-income parent said she would definitely *not* be concerned, and instead, would want to use the opportunity to determine:

- If the assessment is measuring what it is supposed to measure,

- Where and how it is working,
- How culture and community may be impacting scores, and
- How “we” can support communities in need of information and resources to help students fulfill their potential.

This same parent wanted to know how long it would be before an improvement was seen and how long stakeholders would be committed to the process before “throwing it out and trying something new... [because] education seems to do this a lot.”

Several middle-higher-income parents, who earlier expressed disapproval of standardized testing, said they would not be concerned about a drop in scores because they do not “put very much stock in the new assessment any more than the older ones.” One parent said she would not want her son to suffer academic consequences as a result of lower scores, while another said he would not be concerned because the standardized testing scores no longer impact scholarship opportunities like MEAP has in the past.

Standard Setting

No parents from either state or income group were aware of a standard-setting process or the opportunity to get involved through in-person workshops or an online panel to support the development of the new assessment. Despite not having heard of this opportunity, about three-quarters of parents were interested in learning more and said it would be an interesting opportunity.

Message Testing

Parent Invitation Language

When presented with the following opportunities to contribute or participate in the development of the new assessment system, the majority of parents said they would be interested:

- Read more about the results of the spring field test, and follow announcements about the new assessment system.
- Take a sample test to get a better understanding of what will be on the assessment.
- Participate along with teachers and educational experts in an online activity to help determine and set a passing score for the new assessment.

With regard to reading more about the field test, one parent emphasized she would be interested in reading about results across different districts in varying geographic settings and patient populations. In response to the second opportunity, one parent jokingly asked, “Do I have to?” and another hoped it would not be a timed, proctored test. However most parents thought this would be an interesting way to experience what their children will experience. One parent said the hands-on nature of the second opportunity would “better absorb into her brain and help her understand” what her son would be tested on, and another parent stated it would help to see changes and improvements over the previous assessment. One parent was interested in taking a sample test, but prefers to trust the teachers’ judgment on it.

The following invitation was presented to parents regarding the opportunity to participate in the online panel.

The Smarter Balanced Assessment Consortium is inviting parents of school-aged children to participate in an online panel that will inform the assessment’s achievement levels. The consortium is developing a new testing system for English language arts and mathematics aligned to the Common Core State Standards, which will be administered in the coming year. To ensure that the assessments achievement levels accurately reflect what students need to know to be college- and career-ready, Smarter Balanced is asking parents to weigh in on the discussion. Your participation can help ensure these tests are rigorous and fair for all students. The online panel will take up to three hours over a 2-day window that you select between October 6 and October 17, 2014.

In response to the invitation, several parents expressed interest, but were concerned they lacked the expertise for meaningful input. Many preferred to leave the input to “true education experts,” as they were concerned they would bias the results of the standard-setting exercise. Without a background or expertise in education, some felt they could participate only if they could be prepared with information and resources ahead of time.

Among the two-thirds of parents wanting to participate, one parent needed assurance that her participation would not be “lost” and that her input would actually be used. Another wanted assurance that his input would be used to fix issues with the CCSS and the assessment, rather than used to try to “fix students.” One participant suggested parents should weigh in as much as educators, politicians, and others. Another said he would want to provide his input primarily to ensure the test was not designed to “deliberately knock a child back” and ensure it is “a fair opportunity for all students.” One lower-income parent especially liked that the assessments would help his child be “college- and career-ready.”

Among participants preferring to leave this activity to the experts, one parent of a 3rd and 6th grader asked, “How am I supposed to know whether a particular concept or content belongs in 2nd or 3rd grade?” Others agreed with this sentiment.

One higher-income parent commented the time commitment for the online panel seems long, and suggested this could impact who would be able to participate, potentially skewing the results to include a greater proportion of those who do not work full-time. Other parents stated the time commitment seems reasonable and feasible. One parent followed this up by suggesting there be other avenues of participation to ensure the broadest possible representation of stakeholders.

Specific Statement Testing

Smarter Balanced developed sets of statements about the importance of the effort to inform educators and the general public about the assessment system that is being developed. The following table outlines the statements presented to parents and summarizes participants’ level of agreement with the statements and other specific feedback.

Statement	Parent Feedback
<i>Nothing is more important than giving all students an education that prepares them for success. Yet independent research shows America’s schools are failing our children. Though many assume that poverty in America is pulling down education scores, research shows that middle-class U.S. students are far behind their peers in other countries. We need to ensure that our public education systems prepare students for success after high school.</i>	General agreement One parent said it is important for the education system to ensure all students are prepared for life after high school, not just those who are college-bound, by equipping them with knowledge and skills necessary to enter the workforce. She said education should prepare students for “whatever opportunities are appropriate for them.” A few parents suggested the competitiveness of the U.S. compared to other countries goes beyond what students are learning in school and is a larger societal issue. Several parents suggested high school students are given too many options (such as adult education) that allow them to leave the education system, to the detriment of their success. Two disagreeing parents respectively said 1. U.S. scores are being compared to only the wealthiest students in other countries, and it is not a fair comparison and 2. The U.S. performs better in the workforce and academic assessments do not accurately reflect this type of success.
<i>The Common Core standards represent an effort to ensure that every student in every state has the opportunity to learn the same topics and master the same skills through a common national standard, so that no student is denied basic learning.</i>	General agreement Despite general agreement that standards should be consistent across states and communities, one parent said standards like CCSS have been around for years and the larger problem is that “districts are not being held accountable.” Further, three parents said standards need to be flexible to accommodate both high and low achievers, with one stating a more preferable approach might be minimum standards that allow the states to customize based on their given population. One strongly disagreeing higher-income parent said these standards were a government intrusion on educational decisions that

	<p>should be made at the local level. Another parent suggested a 22-state consortium is not nationally representative, and therefore not yielding a national standard.</p>
<p><i>High standards will help more children succeed. Assessments are a tool for ensuring we reach those standards.</i></p>	<p>Mixed response About half of responding parents agreed assessments help measure progress toward meeting standards and place accountability on the teacher and student to perform well. The other half stated the following concerns or points of disagreement:</p> <ul style="list-style-type: none"> • Two parents said there is too much pressure on students to perform, and they may give up or fall short of their full potential. • One parent said standards should be more inclusive of job-related skills rather than just English and math, as these may not be relevant to their long-term success. • Several parents suggested the current focus on testing emphasizes school/teacher funding rather than student achievement. • One parent suggested that for every student succeeding in reaching high standards, there are equally as many falling short and becoming frustrated and discouraged. He suggested standards need to be “reasonable and attainable” and maintained that “high standards alone will not make the U.S. globally competitive.”
<p><i>The Smarter Balanced assessments are a key part of implementing the Common Core standards and preparing all students for success in college and careers. The assessments will replace existing tests and provide an “academic checkup” for students, parents, and teachers by measuring real-world skills such as critical thinking and problem solving. In addition, they will provide information that will give teachers and parents a better picture of where students are succeeding and where they need help.</i></p>	<p>General agreement While this statement was seen as wordy by many, most parents agreed that preparing students for success after high school—not just college and careers—is important. One parent liked the term “academic checkup” as something that would be valuable, and another agreed the opportunity to identify where students are struggling is a definite benefit. Several parents liked the emphasis on “real-world skills and critical thinking,” while one parent wanted to know how these standards compare to those being tested with previous standardized tests. Two others wanted proof that “real-world skills” would be measured. One of these was skeptical that a test could accurately assess these skills, as “they cannot be gained without hands-on experience and applying the skills, along with learning from mistakes.” Further, two parents suggested there is “no such thing as one size fits all” in testing, and this test may not be a fair assessment of all students. Another questioned how this test would be administered to students with special needs to ensure fairness.</p>
<p><i>It is better for all states to have the same tests at each grade level in math and English so test scores can be compared across states.</i></p>	<p>General agreement Parents generally agreed with this statement, while several again asked the question of what modifications would be made to accommodate children with special needs or those with different learning styles. However, one parent stated, “not every kid is going to be at the same level even if they’re taught the same information,” insisting comparisons across states still may not be accurate.</p>
<p><i>Assessments need to be more effective, accurately measuring true knowledge and skills—not just memorized information.</i></p>	<p>General agreement Parents agreed that the movement away from memorization and to greater emphasis on critical thinking skills was a positive change. However, a few higher-income parents questioned the accuracy of measures of students taking a computer-based test if they have limited exposure to computers. Another parent asked what determines “effective” and would want to see success of students in career paths later in life, suggesting it would take time</p>

	to evaluate effectiveness.
<i>The Common Core State Standards and assessments will help reduce inequality and ensure we hold ourselves accountable for improving education.</i>	<p>Mixed response</p> <p>About half of parents questioned agreed with this statement, but a few wanted assurance that teachers would have the resources necessary to teach to these standards and not be penalized for reasons out of their control. Several parents expressed uncertainty about how the assessments would reduce inequality. One parent disagreed, stating he is not convinced the new way of thinking with its “convoluted processes” is an improvement over the old approach to teaching. Another parent “just does not know how the standards will play out and if these outcomes will be achieved.”</p>

Phrase Testing

Parents assessed seven potential objectives for schools to pursue, indicating the relative importance of each in terms of benefits to students, schools, and society more broadly.

<i>How important is it for schools to...</i>	Parent Feedback
<i>Prepare students for success in college, careers, and life.</i>	<p>High importance</p> <p>Parents agreed this is of high importance, but two parents wanted to emphasize the role parents, church, and/or the larger community also play in a child’s success.</p>
<i>Teach students to think and communicate.</i>	<p>High importance</p> <p>Several parents suggested more emphasis should be placed on critical thinking and problem solving rather than “think and communicate.” Another struggled with prioritizing critical thinking versus communicating, insisting that good communication skills will help students become better problem solvers and critical thinkers. With the abundance of social media and other online communication, two parents said they want students to communicate more <i>effectively</i>, rather than simply communicating <i>more</i>.</p>
<i>Give families access to effective education options.</i>	<p>Undecided</p> <p>Similar to educators, the majority of parents found this statement too vague. One lower-income parent found this to be of high importance, suggesting parents should always have access to better options for their children.</p>
<i>Deliver on consistent, high expectations.</i>	<p>Medium to high importance</p> <p>One parent stating this was of high importance interpreted this to mean that schools are doing the best they can, while another questioned, “whose expectations and what serves as the benchmark?” Two parents suggested expectations be realistic and not just high, as they cannot be the same for everyone. Expectations “need to be challenging, but not frustrating,” according to one parent. Another parent assigning medium importance to this statement said education is “more than just about high grades but also about the learning experience.” Another suggested high expectations “put a lot of pressure on kids who are already going through a lot, especially in middle and high school.”</p>
<i>Equip students for the real world.</i>	<p>Medium to high importance</p> <p>Parents agree education needs to be updated to help students live productive lives after high school, but “real world” should be defined. One parent suggested this statement is more important for middle and high school students than for those in elementary school. One parent emphasized this responsibility belongs to more than just schools, falling on parents and the greater society as well.</p>
<i>Deliver real, measurable results.</i>	<p>Undecided</p> <p>One parent labeled this of high importance, but several others found it too vague or subjective. They wanted to know what the results would be measured against.</p>
<i>Reward effective schools.</i>	<p>Low importance</p> <p>Parents stated there is too much competition among teachers, principals, and schools, suggesting more of the focus should be placed on students. Some parents worried that rewarding effective schools may lead to greater inequalities, and suggested expectations be adjusted for schools with different levels of resources. Further, one parent said being</p>

	effective should be the norm, not something that is rewarded.
<i>Hold failing schools accountable.</i>	Mixed response Parents suggested that schools need to be evaluated within the context of the resources they are provided and the setting in which they operate, and holding failing schools accountable will not be fair unless these factors can be controlled for. Further, withholding resources from failing schools may cause them to fall further, according to a few parents. However, a few lower-income parents agreed that this is very important and suggested that if schools are not held accountable, they will continue to perform poorly.

Word Choice

Parents were asked to share their preference between different words being used to describe components of the new assessment. The table below outlines feedback from parents on specific word choices.

Word Choice	Parent Feedback
<i>“Assessment” or “Test”</i>	Assessment The majority of parents preferred “assessment,” stating it implies “more of a gauge of where students are versus placing the emphasis on right or wrong.” “Assessment” puts less pressure on students, and has less of a negative connotation. One parent preferred “test,” stating that it means the same thing but is “more honest about calling it what it is.”
<i>“Math and English” or “mathematics and English Language Arts”</i>	Math and English Parents preferred the simpler “math and English,” this is generally what students call it. However, one parent highlighted that it does get confusing when schools use different terms for English, language arts, etc.
<i>“Preparing students for college and career” or “Preparing students for success after graduation”</i>	Preparing students for success after graduation Parents unanimously preferred this “more inclusive” phrase for all students, especially those who are not college-bound.

Trusted Leaders in Public Education

Parents were asked whom they would trust and rely on for more information about the new assessment. The majority of parents across the groups wanted to hear from teachers, suggesting they are “generally honest and comfortable sharing information with parents,” they are trained, and “they are the ones with hands-on classroom credibility and know the students best.” Parents placed special emphasis on teachers who are currently participating in field testing, as they are better equipped to share information, having experienced the new assessment.

Other potential sources for information, according to one or two parents, include:

- Other parents and students, especially those currently participating in the field testing;
- Experts in the education field, especially those with classroom experience, but not those with “just theoretical expertise” or those who were “not involved in the development of the test”;
- Reading materials, emails, or news announcements from the education department or local district or on their websites;
- U.S. Department of Education, because it lacks the potential bias that a local, city, or state department may have;
- Newspapers; and
- Other Internet sources like reviews or blogs.

In terms of format, a few parents preferred to get information in a group setting such as the school auditorium, delivered in a presentation from the school’s principal. Another mentioned she would not want the burden of sharing information to fall on the teacher, as they already have so much to do.

Other Feedback

Parents were appreciative of being brought into the conversation and given an opportunity to share their opinions. A few parents stated their opinions on assessments, the CCSS, and other topics discussed in the focus groups had not changed since the beginning of the conversation, but many expressed interest in learning more and being involved in the standard setting process.

Appendix A: Moderator's Guide - Educators

WELCOME AND INTRODUCTIONS (10 MINUTES)

- Welcome to our focus group discussion, and thank you for coming.
 - My name is [Christina Nicols / Roshni Devchand], and I have been hired to facilitate our discussion today. I am an independent moderator - not an expert on the topic we are going to discuss today, and I am primarily interested in *your* views on this topic. During our discussion, you may have questions that I can't answer, but I will refer you to resources that will have more information for you.
- PURPOSE OF THE GROUP:
 - We're here to discuss educational testing in schools on behalf of the Smarter Balanced Assessment Consortium. The Smarter Balanced Assessment Consortium is a state-led consortium working to develop next-generation school assessments that accurately measure student progress toward college- and career-readiness.
 - We will be looking at some statements and materials on your screens throughout the group today; please make sure you have reading glasses if you need them.
 - I did not create any of the materials we are going to talk about today, so please feel free to be open and honest in your comments.
- FOCUS GROUP PROCESS AND GENERAL RULES:
 - This is meant to be an open discussion. It is important to hear what everyone thinks, so please speak up and please let others do the same.
 - You don't have to wait for me to call on you, but please speak one at a time.
 - You can respond to each other as well as to me.
 - We are recording the discussion so I can write a report. However, no names will be included in the report and all comments mentioned in the report will be anonymous.
 - There are some interested colleagues listening to our discussion.
 - There are no right or wrong answers. All comments, positive and negative, what you know and don't know, are important.
 - Respect each other and personal opinions; you may not always agree, but we will be respectful of each other.
 - If I ask a question you're not comfortable answering, feel free not to answer.
 - My job is to keep us on the right track and moving along so that we're done in an hour and a half. So, I may have to jump in from time to time to keep our conversation moving.
 - If you need to go to the restroom or need a break, feel free to step away, but please come back as soon as possible because all opinions are important.
 - Please turn off or silence all of your electronic devices to avoid disturbing others. Please refrain from multi-tasking on your computers or other devices during this focus group discussion.
 - Do you have any questions before we get started?
- PARTICIPANT INTRODUCTIONS:
 - Let's get started with introductions - Please tell us your first name ONLY, and what grade and subject you teach. Additionally, if you have subject matter expertise, such as instruction of English Language Learners or Students with Disabilities, please let us know.

AWARENESS / KNOWLEDGE / ATTITUDES REGARDING COMMON CORE, COMMON CORE ASSESSMENT, AND STANDARD SETTING (20 MINUTES)

- Generally speaking, how do you feel about assessments in schools?
 - How important are they to you as an educator?
 - Do you think they are necessary to achieving high standards?
 - What are some of the benefits they provide?
 - Do you have any concerns about them?
 - What concerns do you have?

- Earlier I mentioned we are having this discussion on behalf of the Smarter Balanced Assessment Consortium. Had you heard of this organization before I mentioned it earlier?
 - What have you heard?
 - How familiar are you with the assessment system Smarter Balanced is working to develop?
 - What have you heard about this assessment?
 - Do you have any questions or concerns about the new assessment system as it is being developed?
 - What questions or concerns do you have?
 - What additional information would be helpful to you?
- [As you are probably aware] The Smarter Balanced Assessment Consortium is a consortium working collaboratively with its member states to develop a student assessment system aligned with a common core of academic content standards for English language arts/literacy and mathematics. A field test of the assessment system is being conducted this spring, and the assessment system will be finalized later this year.
 - Have you been following developments of the field test?
 - Did your students participate in the Smarter Balanced field test?
 - Do you feel like you have enough information about the field test? What more would you like to know?
- I'd like to get your opinion about a potential scenario regarding the new assessment system. After your state has conducted the new assessment in schools for the first time, we may see lower scores than we've seen on assessments that were used prior to the implementation of the Common Core standards. If your students scored lower on the new assessment compared to previous assessments, what do you think your reaction would be?
 - How concerned would you be?
 - What specific questions or concerns would you have?
 - With the implementation of more challenging standards and assessments, do you think it is reasonable that the scores could be lower – at least initially – compared to previous assessments measuring less rigorous standards?
 - If the more challenging standards and assessments better prepare students for success after high school, is the initial drop in scores worth it?
 - Why or why not?
 - What questions or concerns do you have about this?
 - What do you think would be the best way to communicate a drop in scores to parents?
 - What do you think they will need to hear to be reassured that the new assessment is moving in the right direction?
- You may have recently seen or heard about a nomination process for an in-person standard setting to support the development of the new assessment. Smarter Balanced will also be conducting an online panel to obtain input on achievement levels for the assessment. Have you heard anything about this?
 - What have you heard?
 - Does it sound like an interesting opportunity?

MESSAGE AND MATERIALS TESTING (55 MINUTES)

- Smarter Balanced is developing materials to invite and encourage educators to participate in the online panel. I'd like to get your perceptions of an invitation. I'll show it on two slides, and I'd like you to take a few minutes to read it and then answer a few questions.

Nothing is more important than giving all students an education that prepares them for success. That is the goal of the Common Core State Standards. The standards help ensure that all students leave school prepared for success after graduation and also provide clear, consistent guidelines for educators and school administrators.

The Smarter Balanced Assessment Consortium is developing new assessments aligned to these higher standards. The tests will provide a meaningful and fair 'academic checkup' and help educators chart progress toward the goal of improving student achievement.

To ensure that the tests are rigorous, fair, and accurate for all students, Smarter Balanced is inviting educators to participate in an Online Panel for recommending Achievement Levels on their assessments. Participants will recommend an Achievement Level score that demonstrates how much students should know or be able to do in order to meet the grade-level standards and to be on track for eventual college and career readiness. You can participate from the comfort of your own home or office, at a time that works best for you.

Educators will have the opportunity to preview and analyze the test's items and structure and provide input into the process of recommending Achievement Level scores. Participation is free and open to the public. This opportunity will allow a large group of educators and administrators to help establish consistent measures of progress for the Smarter Balanced interim and summative assessments.

The Online Panel to recommend Achievement Levels will occur October 6–17, 2014. It will take participants up to 3 hours to complete the orientation process, review test questions, and recommend a score. Each participant will have a two-day window to complete the entire process.

- Let's start with the first paragraph in this invitation:
 - Does this seem credible to you?
 - What do you like about it? What do you dislike?
 - Are there any specific words or phrases that stand out?
- Second paragraph:
 - Does this seem credible to you?
 - What do you like about it? What do you dislike?
 - Are there any specific words or phrases that stand out?
- Third paragraph:
 - Does this seem credible to you?
 - What do you like about it? What do you dislike?
 - Are there any specific words or phrases that stand out?
- Fourth paragraph:
 - Does this seem credible to you?
 - What do you like about it? What do you dislike?
 - Are there any specific words or phrases that stand out?
- Fifth paragraph:
 - Does this seem credible to you?
 - What do you like about it? What do you dislike?
 - Are there any specific words or phrases that stand out?
- Are you likely to participate in the online panel?
 - Why or why not?
- Is there anything you would suggest to make it more compelling or motivating?
- Smarter Balanced is also developing educational materials to inform educators and the general public about the assessment system that is being developed. I'd like to get your opinion on some language that may be used for these materials.
- First let's start with some statements about the importance of this effort. I'll show you a series of brief statements on the screen. I'd like for you to read them, and then I'll ask a few questions about each one.

Statement 1: The Common Core State Standards will have a positive effect on students' ability to think critically and use reasoning skills.

- Do you agree with this statement?
- Why or why not? What do you like about it? What do you dislike?

Statement 2: Consistent standards will provide clear guidelines for teachers and administrators and improve efficiency, especially when integrating students from other states.

- Do you agree with this statement?
- Why or why not? What do you like about it? What do you dislike?

Statement set 3: The Smarter Balanced assessments are a key part of implementing the Common Core standards and preparing all students for success in college and careers. The assessments will replace existing tests and provide an “academic checkup” for students, parents, and teachers by measuring real-world skills such as critical thinking and problem solving. In addition, they will provide information that will give teachers and parents a better picture of where students are succeeding and where they need help.

- Do you agree with this set of statements?
- Why or why not? What do you like about it? What do you dislike?

Statement 4: The Smarter Balanced assessment is a test by states and for states, developed with ongoing input from educators.

- Do you agree with this statement?
- Why or why not? What do you like about it? What do you dislike?

Statement 5: Participating in the Consortium will help educators share and learn from best practices through collaboration with other participating states.

- Do you agree with this statement?
- Why or why not? What do you like about it? What do you dislike?

Statement 6: Educators will have widespread opportunity for input through the Online Panel, in addition to an in-depth opportunity for input through the In-Person Panel.

- Do you agree with this statement?
- Why or why not? What do you like about it? What do you dislike?

Statement 7: Educators have the opportunity to see actual test items and experience the test as their students will experience it.

- Do you agree with this statement?
- Why or why not? What do you like about it? What do you dislike?

Now I'd like to read a series of phrases, and I'd like you to tell me *how important* you think it is for schools to do each of the following [high, medium, low importance]:

- Preparing students for success in college, careers, and life.
- Teaching students to think and communicate.
- Giving families access to effective education options.
- Delivering on consistent, high expectations.
- Equipping students for the real world.
- Delivering real, measurable results.
- Rewarding effective schools.
- Holding failing schools accountable.
- By raising standards for all, we can challenge students to do better.

CONCLUSION (5 MINUTES)

- We've now spent an hour or so talking about the Common Core standards and assessments. Have your perceptions or feelings about this topic changed from before this discussion? If yes, how have they changed?
- Is there more information you would like to know on this topic? What would be of interest to you?
- How would you like to receive this information? Through your school? Through education-related media? Do you have any preferred communication methods?
- I'd like to read a list of types of leaders who often provide opinions on public education. I'd like you to tell me *how likely* you are to trust and agree with taken by these leaders when you think about public education:
 - Master teachers
 - State and local school superintendents

- Higher education leaders (e.g. college presidents)
 - Leaders/owners of successful local businesses
 - Labor/teacher's union leaders
 - Elected officials
 - Education research experts
 - Other parents
 - Reporters/news commentators
 - Religious leaders
 - Are there other leaders who I did not list here from you would like to hear more opinions on education?
- That's all of my questions. Do you have any final comments on anything we've talked about?
 - More information and updates can be found at: <http://www.smarterbalanced.org/>
 - Thank you so much; your help has been invaluable. You will receive your thank you check in the mail.

Appendix B: Moderator's Guide - Parents

WELCOME AND INTRODUCTIONS (10 MINUTES)

- Welcome to our group discussion, and thank you for coming.
- My name is [Christina Nicols / Roshni Devchand], and I have been hired to facilitate our discussion today. I am an independent moderator - not an expert on the topic we are going to discuss today - and I am primarily interested in your views on this topic. During our discussion, you may have questions that I can't answer, but I will refer you to resources that will have more information for you.

PURPOSE OF THE GROUP:

- We're here to discuss educational testing in schools on behalf of the Smarter Balanced Assessment Consortium. The Smarter Balanced Assessment Consortium is a state-led consortium working to develop new school assessments that accurately measure student progress toward college- and career-readiness.
- We will be looking at some statements and materials on your screens throughout the group today; please make sure you have reading glasses if you need them.
- I did not create any of the materials we are going to talk about today, so please feel free to be open and honest in your comments.

FOCUS GROUP PROCESS AND GENERAL RULES:

- This is meant to be an open discussion. It is important to hear what everyone thinks, so please speak up and please let others do the same.
- You don't have to wait for me to call on you, but please speak one at a time.
- You can respond to each other as well as to me.
- We are recording the discussion so I can write a report. However, no names will be included in the report and all comments mentioned in the report will be anonymous.
- There are some interested colleagues listening to our discussion.
- There are no right or wrong answers. All comments, positive and negative, what you know and don't know, are important.
- Respect each other and personal opinions; you may not always agree, but we will be respectful of each other.
- If I ask a question you're not comfortable answering, feel free not to answer.
- My job is to keep us on the right track and moving along so that we're done in an hour and a half. So, I may have to jump in from time to time to keep our conversation moving.
- If you need to go to the restroom or need a break, feel free to step away, but please come back as soon as possible because all opinions are important.
- Please turn off or silence all of your electronic devices to avoid disturbing others. Please refrain from multi-tasking on your computers or other devices during this focus group discussion.
- Do you have any questions before we get started?

PARTICIPANT INTRODUCTIONS:

Let's get started with introductions - Please tell us your first name ONLY, and what grade your child is in at his or her school.

AWARENESS / KNOWLEDGE / ATTITUDES REGARDING COMMON CORE AND COMMON CORE ASSESSMENT (25 MINUTES)

- Earlier I mentioned we are having this discussion on behalf of the Smarter Balanced Assessment Consortium. Had you heard of this organization before I mentioned it earlier?
 - What have you heard?
 - How familiar are you with the school assessment system Smarter Balanced is working to develop?
 - What have you heard about this assessment system?

- What more would you like to know about it?
- The Smarter Balanced Assessment Consortium is a consortium working collaboratively with its member states to develop a student assessment system aligned with a common core of academic content standards for English language arts and mathematics. A field test of the assessment system is being conducted this spring, and the assessment system will be finalized later this year.
 - Are you familiar with the Common Core State Standards?
 - What is your experience with them?
 - What do you know about them?
 - What more would you like to know?
- Generally speaking, how do you feel about academic testing in schools?
 - How important are they?
 - What are some of the benefits they provide?
 - Do you have any concerns about them?
 - What concerns do you have?
- Do you have any concerns or questions specifically about the new assessment system that is being developed?
 - What concerns or questions do you have?
- I'd like to get your opinion about a potential scenario regarding the new assessment system. After your state has conducted the new assessment in schools for the first time, we may see lower scores than we've seen on previous assessments. If your child scored lower on the new assessment compared to previous assessments, what do you think your reaction would be?
 - How concerned would you be?
 - What questions would you have?
 - With the implementation of more challenging standards and assessments, do you think it is reasonable that the scores could be lower – at least initially – compared to previous assessments?
 - If the more challenging standards and assessments better prepare students for success after high school, would the initial drop in scores be worth it?
 - Why or why not?
 - What questions or concerns do you have about this?
 - What would you need to hear from teachers, principals and other school administrators if this happened?
- I'd like to read a list of types of leaders who often provide opinions on public education. Can you tell me *how likely* you are to trust and agree with positions taken by these leaders when you think about *your* child's education:
 - Master teachers
 - State and local school superintendents
 - Higher education leaders (e.g. college presidents)
 - Leaders/owners of successful local businesses
 - Labor/teacher's union leaders
 - Elected officials
 - Education research experts
 - Other parents

- Reporters/news commentators
- Religious leaders
 - Are there other types of leaders whom I did not list here whose opinions on education you would trust?
- If you had an opportunity to contribute to or participate in the development of the new assessment system, would you be interested? Which of the following activities would you be willing to do?
 - Read more about the results of the spring field test, and follow announcements about the new assessment system.
 - Take a sample test to get a better understanding of what will be on the assessment.
 - Participate along with teachers and educational experts in an online activity to help determine and set a passing score for the new assessment.
 - I'm going to show you more information about this opportunity, and I'd like you to read it and answer a few questions about it:

The Smarter Balanced Assessment Consortium is inviting parents of school-aged children to participate in an online panel that will inform the assessment's achievement levels. The consortium is developing a new testing system for English language arts and mathematics aligned to the Common Core State Standards, which will be administered in the coming year. To ensure that the assessments achievement levels accurately reflect what students need to know to be college- and career-ready, Smarter Balanced is asking parents to weigh in on the discussion. Your participation can help ensure these tests are rigorous and fair for all students. The online panel will take up to three hours over a 2-day window that you select between October 6 and October 17, 2014.

- Are you likely to sign up for this and participate?
 - Why or why not?
- What would you want to know about this activity before deciding to register?

MESSAGE AND MATERIALS TESTING (55 MINUTES)

- Smarter Balanced is currently developing educational materials to inform parents and others about the assessment system that is being developed. I'd like to get your opinion on some language that may be used for these materials.
- First let's start with some statements about the importance of this effort. I'll show you a series of brief statements on the screen. I'd like for you to read them, and then I'll ask a few questions about each one.

Statement set 1: Nothing is more important than giving all students an education that prepares them for success. Yet independent research shows America's schools are failing our children. Though many assume that poverty in America is pulling down education scores, research shows that middle-class U.S. students are far behind their peers in other countries. We need to ensure that our public education systems prepare students for success after high school.

- Do you agree with this set of statements?
- Why or why not? What do you like about it? What do you dislike?

Statement 2: The Common Core standards represent an effort to ensure that every student in every state has the opportunity to learn the same topics and master the same skills through a common national standard, so that no student is denied basic learning.

- Do you agree with this statement?
- Why or why not? What do you like about it? What do you dislike?

Statement 3: High standards will help more children succeed. Assessments are a tool for ensuring we reach those standards.

- Do you agree with this statement?
- Why or why not? What do you like about it? What do you dislike?

Statement set 4: The Smarter Balanced assessments are a key part of implementing the Common Core standards and preparing all students for success in college and careers. The assessments will replace existing tests and provide an “academic checkup” for students, parents, and teachers by measuring real-world skills such as critical thinking and problem solving. In addition, they will provide information that will give teachers and parents a better picture of where students are succeeding and where they need help.

- Do you agree with this set of statements?
- Why or why not? What do you like about it? What do you dislike?

Statement 5: It is better for all states to have the same tests at each grade level in math and English so test scores can be compared across states.

- Do you agree with this statement?
- Why or why not? What do you like about it? What do you dislike?

Statement 6: Assessments need to be more effective, accurately measuring true knowledge and skills—not just memorized information.

- Do you agree with this statement?
- Why or why not? What do you like about it? What do you dislike?

Statement 7: The Common Core State Standards and assessments will help reduce inequality and ensure we hold ourselves accountable for improving education.

- Do you agree with this statement?
- Why or why not? What do you like about it? What do you dislike?

- Now I'd like to read a series of phrases, and I'd like you to tell me *how important* you think it is for schools to do each of the following, particularly when you think about *your child's* education:
 - Preparing students for success in college, careers, and life.
 - Teaching students to think and communicate.
 - Giving families access to effective education options.
 - Delivering on consistent, high expectations.
 - Equipping students for the real world.
 - Delivering real, measurable results.
 - Rewarding effective schools.
 - Holding failing schools accountable.
 - By raising standards for all, we can challenge them to do better.
- I'd also like to get your opinion on specific preferences in word choice:
 - When you are reading or hearing about this topic, do you prefer the term 'assessment' or 'test'?
 - Do you prefer 'math and English' or 'mathematics and English Language Arts' when you are hearing about the content of the assessments?
 - When we are talking about the ultimate goal of the standards and assessments, is it better to say we are 'preparing students for college and career' or 'preparing students for success after graduation'?

CONCLUSION (5 MINUTES)

- We've now spent an hour or so talking about common core standards and assessments. Have your perceptions or feelings about this topic changed from before this discussion? If yes, how have they changed?
- Is there more information you would like to know on this topic? What would be of interest to you?
- How would you like to receive this information?
 - Through your child's school?
 - Through media? Are there particular channels, magazines, web sites or other media outlets you pay attention to? What are these?
- More information and updates can be found at: <http://www.smarterbalanced.org/>
- That's all of my questions. Do you have any final comments on anything we've talked about?

- Thank you so much; your help has been invaluable. You will receive your thank you check in the mail.

The Smarter Balanced Assessment Consortium brings together states to create a common, innovative assessment system for Mathematics and English Language Arts that is aligned with the Common Core State Standards and helps prepare students for college and careers. The Consortium involves educators, researchers, policymakers, and community groups in a transparent and consensus-driven process to help all students thrive in a knowledge-driven global economy. The Consortium's projects are funded through a four-year, \$175 million grant from the U.S. Department of Education, comprising 99% of activity resources, with the remaining support provided through generous contributions of charitable foundations. Membership is open to any interested U.S. state. For more information, please visit <http://www.smarterbalanced.org/>.

Achievement Level Setting Overview for Educators

[PRESENTER NAME]



PRESENTATION DATE



At-a-Glance

WHAT:

- Achievement level setting for the Smarter Balanced assessments

WHO:

- Educators and other interested members of the public

WHEN:

- October 2014

WHERE:

- Online and in-person activities



What is achievement level setting?

- Commonly-known as “standard-setting”
- Process of establishing one or more scores for proficiency on a test
- Allows education agencies to create categories of performance for students



Goals

- Obtain proficiency scores for mid-year and summative assessments
- Process ~~driven by~~ involves a large group of educators and practitioners in Smarter Balanced states
- Collect input from other participants including members of the general public
- Make sense of proficiency scores across grades
- States approve recommended scores based on existing processes



Approach

- Collaborative and transparent
- Inclusive and consensus-driven process
- Scientifically-rigorous design
- Multiple reviews, technical certification

Achievement Level Setting Timeline

April 30

- States Approve Achievement Level Setting Plan

October 6-17

- Online Panel open to thousands of educators and other constituents

October 13-19

- In-Person Panel

October 22-23

- Vertical Articulation Committee: Subgroup of In-Person

October 30

- Technical Advisory Committee Review

November 6

- State Vote on Achievement Level Recommendations

2014-15

- States Adopt Achievement Levels

Major Activities

- **Online Panel**
 - Crowd sourcing – thousands of participants
 - Recommendations broken down by subgroup (teachers, higher education, etc.)
- **In-Person Panel**
 - 500 participants in grade/subject panels
 - Bookmark procedure
- **Vertical Articulation**
 - 60 panelists from In-Person Panel
 - Review Online Panel and In-Person Panel results
 - Recommend changes



Achievement Level Setting Questions

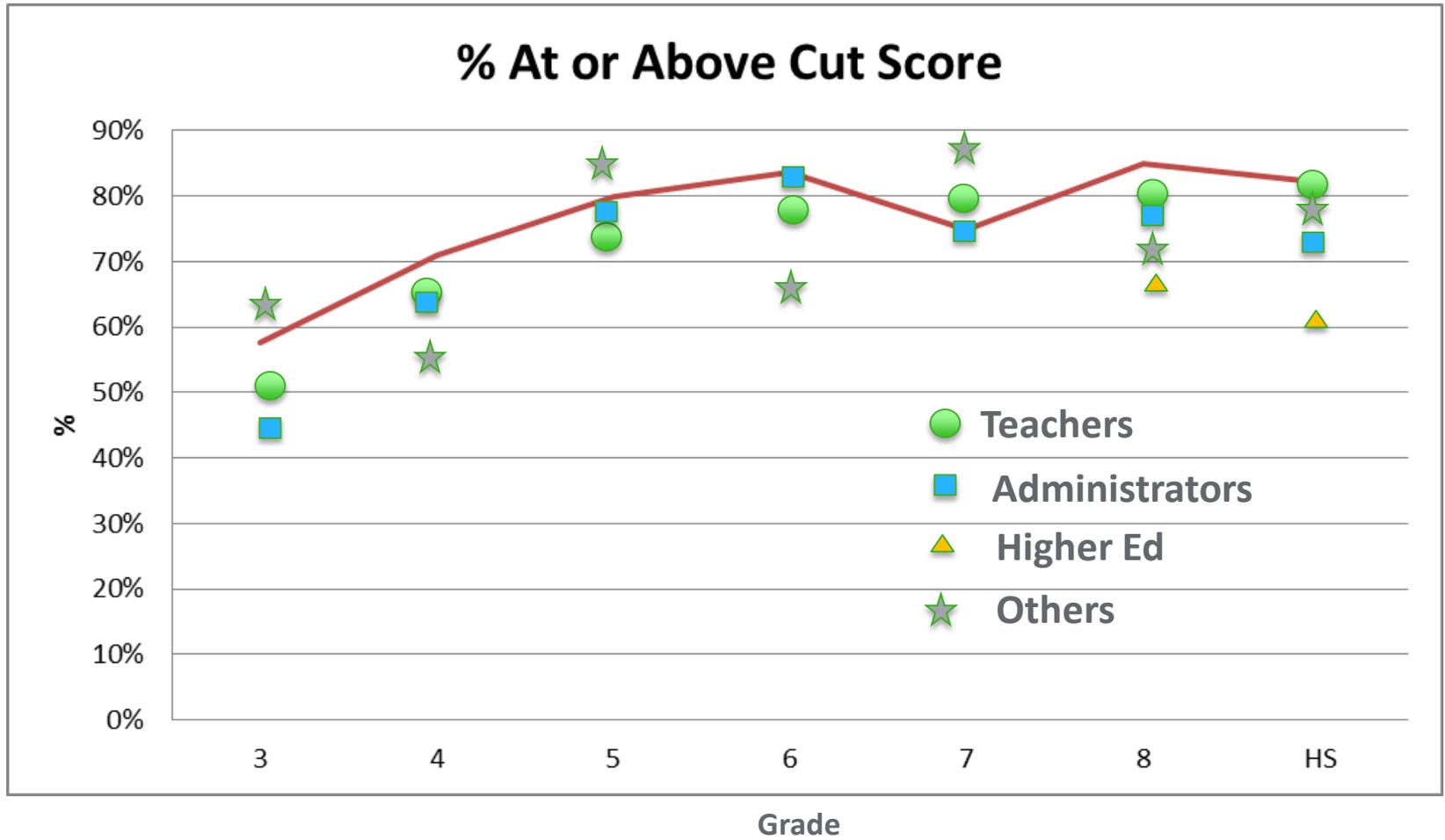
Panelists will be prompted with questions, such as:

- What knowledge or skill is required to earn this point?
- What makes later items more difficult than earlier items?
- Think of a large group of students at the cut score for this level. Would ~~at least 2/3~~ most of them earn this point?

The Bookmark Procedure

- One of the most widely used procedures for developing achievement level scores
- Created by staff of CTB/McGraw-Hill
- Well documented
- Test items placed in a special booklet in difficulty order
- Panelists mark places in the booklet where difficulty shifts
- Bookmarks are translated into cut scores

How Input Will Be Used





Benefits to Educators

- Experience the Smarter Balanced test as your students will experience it
- Capacity for thousands of educators to participate
- Opportunity for continued input into the assessments
- Review actual test items and recommend scores for proficiency
- Correlate student performance to achievement level descriptors

FAQ about the Online Panel

- **Cost:** Open to the public. Participation is free.
- **Privacy:** Personally identifiable information of individual participants will not be released to the general public.
- **Accessibility:** Accessible through virtually any computing device that connects to the Internet, including tablets (Smart Phones not recommended due to screen size).
- **Reminder:** Panelists will receive reminders before their two-day window opens.
- **Time Commitment:** Up to 3 hours over a two-day window.
- **Log in/Log Out:** Ability to save work and log out, and then log back in to complete session within the two-day window.

Get Involved

Make your voice heard by registering for the Online Panel for Achievement Level Setting.

Go to SmarterBalanced.org/OnlinePanel to get started.

- ✓ Step 1 - Select a content area and grade level
- ✓ Step 2 – Enter your email address and personal information
- ✓ Step 3 – Check your email for a message and confirm your email address
- ✓ Step 4 – Select a two-day window between October 6 and 17, 2014

Setting a New Baseline

- **New standards:** provide clear guidelines for educators and ensure more students are on track for college and career
- **New assessments:** challenging and engaging for students, provide teachers and school leaders with better information to help students
- **New baseline:** proficiency scores developed by educators that reflect high expectations for students

Key Takeaways

- Achievement level setting: recommending scores for proficiency based on achievement level descriptors
- Inclusive, consensus-based and scientifically rigorous design
- Thousands of educators will be able to participate
- Information and resources are available – visit SmarterBalanced.org

Item	Description	Target Audience	Format	Status	
Sample Communications Plan	This summary planning document will provide recommendations for developing and launching a sustained communications campaign in support of the new performance standards. The document will suggest rollout dates, targets and tactics for the announcement and will summarize key message points for explaining the new performance standards and the process by which they were developed.	<ul style="list-style-type: none"> • Internal only 	<ul style="list-style-type: none"> • Word Document 		
Q&A Document for States	This Q&A document will help answer anticipated questions about the achievement level setting and new performance standards for the Smarter Balanced assessments. These can be included in student packets for parents at the beginning of the school year, posted to a school district's website, posted on a school's physical bulletin board, or any other relevant method of distribution. This will also provide content for the FAQ section on state department of education websites to help explain achievement level setting and the new performance standards.	<ul style="list-style-type: none"> • School Districts • Parents • Educators 	<ul style="list-style-type: none"> • Word Document 		
Fact Sheet for Policymakers/ Stakeholders	These fact sheets will help states explain the Smarter Balanced achievement levels to policymakers and education stakeholders.	<ul style="list-style-type: none"> • Policymakers • Parents • Education stakeholders 	<ul style="list-style-type: none"> • PDF 		
Op-ed/letter to the editor template	These customizable tools will explain the new performance standards on the Smarter Balanced assessments and how they were developed. They will also provide messages and supporting points to reinforce the value of the assessment and college and career ready standards. The template will include generic submission instructions to	<ul style="list-style-type: none"> • Policymakers • Parents • Education stakeholders • Higher education 	<ul style="list-style-type: none"> • Word Document 		

	expedite placement in local news outlets.	institutions			
Fact Sheets for Higher Ed	This fact sheet will address questions from institutions of higher education and support the claim that Smarter Balanced scores are evidence that students are ready for credit-bearing courses.	<ul style="list-style-type: none"> Higher education institutions 	<ul style="list-style-type: none"> PDF 		
Exemplar Test Items	Items from the Smarter Balanced exam that states can use to demonstrate alignment between the standards, the assessments and the new cut scores.	<ul style="list-style-type: none"> Policymakers Education stakeholders Higher education institutions Parents 	<ul style="list-style-type: none"> PDF 		
PowerPoint Presentation	This Ready-to-use PowerPoint presentations on achievement level setting and the new Smarter Balanced performance standards will be customized for use when testifying or presenting before a state legislature, a state board of education and, a local school board.	<ul style="list-style-type: none"> Higher education institutions Policymakers State education leaders School districts 	<ul style="list-style-type: none"> PowerPoint 		
Video Package	This video of achievement level setting process will explain the purpose and structure of the process through interviews with multiple panelists. The overarching goal will be preview the cut scores and the impact with states and communicate broader messages about the value of the assessments and to. The video will be used to document the process around achievement level setting for purposes of peer review.	<ul style="list-style-type: none"> Policymakers State education leaders Education stakeholders 	<ul style="list-style-type: none"> Videos 		

<p>Messaging for Grassroots Leaders</p>	<p>These message points for grassroots leaders can be used to provide content for op-eds, news release quotes, speeches, and other communications materials. The grassroots messaging will provide content for stakeholder groups and grassroots leaders, including:</p> <ul style="list-style-type: none"> • Teachers • Business leaders/owners • Principals and administrators • Higher education leaders • School board members • Labor • Civil rights leaders • Faith and clergy • Military families and military leaders • PTA/Parent leaders • Teachers with English Language Learner/Students with Disabilities focus 	<ul style="list-style-type: none"> • Policymakers • Education stakeholders • Higher education institutions • Parents 	<ul style="list-style-type: none"> • Word Document 		
<p>Spanish Outreach Materials</p>	<p>Outreach materials to Spanish-speaking parents to explain the value of the standards and the Smarter Balanced assessments and how they will help promote student success and access to college and career options. Communications materials for Spanish-speaking parents and communities will include:</p> <ul style="list-style-type: none"> • A FAQ sheet to provide standardized message points for states to use; • A parent notification letter with information about the assessment and new performance standards. • A news release for states and districts to use for Spanish language media. • A short news script for broadcast outlets. 	<ul style="list-style-type: none"> • Spanish speaking parents • Spanish-language media • Education stakeholders 	<ul style="list-style-type: none"> • Word Document 		



Smarter Balanced Assessment Consortium: Contingency Plan for Contract 21

Prepared by: Hager Sharp

July 7, 2014



Table of Contents

Background	191
Messaging to use during rapid response	191
Detailed messages	192
Contingency Communications Objectives	193
The Contingency Communications Plan	194
1. Define the Situation.....	194
Situation analysis.....	195
Steps to define the situation.....	196
2. Establish Response Team, Enforce Internal Communications Protocols, and Identify Spokespersons.....	197
Response Team.	198
3. Develop and Deliver Messages	201
Tips for message development.....	202
Developing message maps.	202
4. Review Course of Action.....	204
5. Establish Timeline and Implement	204
Steps for Day One of a Contingency Situation.	204
Steps for Day Two.	205
6. Evaluate Results and Recalibrate	207
Scenarios.....	208
Priority Scenarios	208
Divided Vote on Achievement Levels.....	208
Scenarios: Acceptance of the Achievement Levels	210
Highly-visible criticism by an external commentator or panelist.....	212
Additional Scenarios: In-Person Panel.....	214
a. Ejection of a rogue panelist.....	214
b. Disturbance at the plenary session.	216
c. Operational problems at the In-Person Panel.	217
d. Repeal of Common Core State Standards/loss of Governing State status prior to panel, but after panelist selection.	218
e. Leaked information and negative exposure during or after the In-Person Panel.....	218
Additional Scenarios: Online Panel.....	219
Appendix 1: Internal and External Audience Stakeholders.....	220
Appendix 2: Message Map for Achievement Level Setting	221

Background

This contingency communications plan is intended to enhance understanding of operational vulnerabilities; maintain credibility and reputation; involve appropriate cross-divisional staff through effective preparation and tasks; and protect participating states, employees, and a range of Smarter Balanced stakeholders.

This plan was developed to help implement best practices during a contingency situation as well as incidents that may occur before, during, and after the achievement level setting process. This document proposes protocols and messages for responding to key audiences through a variety of channels.

In order to assess the different types of external threats and internal weaknesses, Hager Sharp conducted research on Smarter Balanced operations and planned achievement level setting activities using several primary methods, including:

- A kickoff meeting and subsequent planning meetings with Smarter Balanced and representatives from WestEd (PMP), Measurement Incorporated, CTB/McGraw-Hill, and GMMB.
- In-depth phone interviews with key employees and contractors including leadership staff, the lead facilitators for achievement level setting and project leads from PMP.

The plan will help Smarter Balanced respond rapidly to opposition and potential threats, especially those related to the achievement level setting process and acceptance of the new cut scores. In preparation for what will likely be a challenging rollout of the new achievement levels and cut scores, we will work concurrently with our general communications planning to develop response protocols and topline messages for the most likely contingency scenarios. The contingency communications plan includes protocols to ensure a timely media response in fast-moving scenarios, such as:

- Technical or operational problems at the In-Person Panel or involving the administration of the Online Panel;
- A disagreement or negative exposure following the In-Person Panel;
- A disagreement or divided vote at the fall Collaboration Conference when states vote on a motion to endorse the new achievement levels and cut scores;
- Concerns about the achievement level setting process from Smarter Balanced internal groups or stakeholders; and
- The release of a third-party review questioning the validity of the achievement level setting process (e.g., the National Academy of Sciences report criticizing National Assessment of Education Progress [NAEP] achievement levels) or validity of cut scores in relation to external measures such as NAEP, PISA, SAT, or ACT.

Messaging to use during rapid response

It is important to establish some fundamentals for responding to a contingency situation. During a time of crisis, an organization must fall back on its brand. A positive brand identification helps when an organization faces a contingency situation. The work of Smarter Balanced is guided by a set of core principles. The recommendations that follow are designed to help Smarter Balanced navigate a challenging communications environment without compromising these core values and messages. Regardless of the specific type of contingency situation or critical incident that Smarter Balanced and its Governing States encounter, the underlying message to education constituents and the public must be the same.

The work of Smarter Balanced is guided by a set of core principles. In addition to providing additional factual information about a situation, Smarter Balanced should incorporate messages that reinforce its commitment to this set of principles. In the case of achievement level setting, Smarter Balanced should strive to communicate four key messages about the process.

1. Inclusion, respect for diverse views from all states
2. Voluntary, state-led consortium
3. Collaborative and transparent process
4. Scientifically rigorous design

Detailed messages

- **The process is collaborative and transparent:** educators and others from all Smarter Balanced states working together using a process that has been thoroughly documented and approved by the Governing States in an open meeting.
- **The process is inclusive:** The final recommendations for achievement level scores will be based on input from thousands of interested stakeholders and community members. Multiple rounds of test item review and discussion conducted by classroom teachers, higher education faculty, parent leaders and community members. States have extensive representation through the In-Person Panel and an Online Panel with capacity for thousands of educators to participate.
- **The process is scientifically rigorous:** panelists will use a procedure (Bookmark) that has been researched and used in more states than any other method. Multiple reviews of the achievement level score results will occur. The integrity of this process is essential. Member states have been involved extensively in the design of this process and continuously monitoring progress.
- **Educator involvement is key.** Our achievement levels are being set primarily by practitioners. Higher education faculty will help recommend achievement level scores for high school, high school teachers will help recommend scores for middle school, middle school teachers for elementary school, along with on-grade-level teachers.
- **Setting a new baseline for assessment and raising the bar:** These activities will allow member states to work toward higher standards, which use performance standards for students recommended by educators. The scores will be based on input from a diverse group of educators and stakeholders. Smarter Balanced is committed to conducting this work in a transparent and open way.

Contingency Communications Objectives

- Maintain the credibility of Smarter Balanced and respond in a manner that supports its mission, core values, and established messages of transparency and collaborative work with governing states.
- Assure all relevant stakeholders, including employees, of commitment to resolving the crisis.
- Help ensure accuracy and balance in news coverage and internal and external information dissemination.
- Minimize rumors and limit misinformation.

The primary principle in a risky or volatile communications environment is simple: **FIRST DO NO HARM**. When faced with a challenging reactive media/communications situation, ask yourself the following questions:

- What information is crucial to convey in initial messages in order to prompt appropriate responses?
- What are the messages to be delivered prior to, during, and after the event?
- What are the obstacles to effectively communicate your messages and how can they be overcome?
- What are the opportunities for effective communications and how can they be maximized?
- What questions can we anticipate from the public? Media? Education leaders? Critics/competitors? Partners? Government/agency representatives? Staff?
- Under what time constraints are you operating? Media deadlines? Federal reporting deadlines?

When dealing with such an environment, it is important to give thought to what you will say before making any public comment—be it a written statement on a Smarter Balanced Web page, a letter or notice to states, a 30-second sound bite or interview with a reporter, or comments to staff. When preparing to respond:

1. **Assess the environment.** Is it a heated debate or a genial discussion? What are the circumstances into which you will be providing information? Who is your audience? What are their concerns? What are their likely perceptions and biases? Will they be receptive or resistant to your messages?
2. **Review your remarks to gauge their impact.** Be sure to use existing Smarter Balanced messages (in combination with any situation-specific messages developed for the situation) and to try to avoid language that may be inflammatory to critics.
3. **Understand your audience.** If it is a reporter, respect his or her deadlines. If the concern comes from a state employee or another stakeholder, understand how his or her audiences may be impacted and the consequences they may be facing. If it is a government official, understand that he or she needs specifics and will need to hear both sides of the story—and will pay close attention to constituent voices first. If it is staff, understand that personal relationships may be involved and professional and reputational risk may be an issue.
4. **Do not speak unless you are comfortable doing so.** Poorly prepared spokespeople make most communications mistakes.
5. **Monitor the resulting media coverage/stakeholder feedback.** Analyze this for accuracy and any needed follow-up activities. Also, pay particular attention to the messages and delivery of those on the other side of the debate. There is much to learn from monitoring all aspects of a story or issue. Additionally, being familiar with an opposing viewpoint assists with the anticipatory needs of communicating in an unstable environment.

The Contingency Communications Plan

No individual or organization is immune to crisis. With proper preparation, a sense of perspective, and communications discipline, a critical incident or contingency situation can be managed by engaging in the communications process instead of trying to avoid it. Contingency situations cause confusion and create an unstable communications environment. When they arise, stakeholders, Smarter Balanced member states, staff, media, etc., will want answers to three basic questions:

- What is the story here—what happened exactly and who is impacted?
- Who is to blame? What/who is the culprit?
- What is Smarter Balanced doing to fix the problem/minimize or limit the damage?

For our purposes, a contingency situation is defined as any emergency or disruption that has the potential to damage the reputation and/or disrupt the operations of Smarter Balanced. Resulting public scrutiny may have legal, political, financial, and/or governmental implications.

Communication experts define a crisis in simple terms, noting there are three main stages: before, during, and after a crisis.¹ Ongoing Smarter Balanced public relations and outreach efforts are very much a part of the “before stage.” Smarter Balanced is working to communicate the value of the assessment, as well as communicate its approach to delivering those services in a way that allows for control of the message, builds a record of its success, and informs audiences of the ongoing value of membership in the Consortium.

In the event of a contingency situation, Smarter Balanced will need to ease concern among state leads, advisers, and other stakeholders about the specifics of the situation. What follows are guidelines for communicating with various and diverse audiences during a contingency situation or unstable communications environment. These tactics and recommendations can be used as strategies and procedures to assist in planning and communicating in such situations.

Overriding principles for communication during a contingency situation are:

- Tell the truth.
- Remember that stakeholder concerns are your concerns.
- Establish Smarter Balanced as the best initial and ongoing source for information.
- Provide accurate and factual information in a timely manner.
- Reinforce key messages.
- Exercise communications discipline in a challenging environment.

It is also essential to look at “360 degrees” of audiences (see Appendix 1) who may be impacted, be involved, or otherwise have an interest in a critical incident. The more Smarter Balanced can connect with affected parties and related organizations, the more it can minimize misperception and inaccurate or incomplete information dissemination and, more importantly, maintain the Smarter Balanced brand and reputation.

1. Define the Situation

When a contingency situation arises, the focus needs to be on working with staff and all related parties to understand the situation and assess the implications for Smarter Balanced. This section discusses the general parameters that govern these situations and presents a list of questions that should be answered to better understand the situation.

¹ Marconi, J. (2005). *Public Relations: The Complete Guide*, American Marketing Association/Thomson, 2004.

Situation analysis.

It is critical to define the appropriate level of response to contingency situation. All potential crises demand an individualized level of response. To help determine the significance of a potential situation, Smarter Balanced should address the following questions:

- What happened, exactly?
 - Provide the most complete picture of what exactly happened—build a complete, factual record of what is involved in the incident.
- Who is impacted?
- How serious is the situation?
 - To Smarter Balanced staff, advisors, and states
 - To Smarter Balanced reputation
- What is involved in returning the situation to normal?
 - In some cases, the solution will be readily apparent. In others, it may take some time before you will have a clear idea of what will be needed. In either case, try to explain what will be involved to all stakeholders, both internal and external audiences, throughout the situation. Both audiences will need to be identified and informed of what steps will be taken to rectify the situation.
- Who will be involved in returning the situation to normal?
 - Will Smarter Balanced staff be able to execute the remediation internally or will it require external expertise? What, if any, role will impacted parties need to play in returning the situation to normal?
- How long will that take?
 - This is a question that may not have an immediate answer. Some situations can be resolved in a relatively short period of time. Others may take longer. **Avoid speculation. Stick to the facts. If it will take time, say so.**

When collecting facts, consider the following:

- Do you have *all* the facts (to the best of your knowledge)?
- What *other information* do you need to put the event into perspective?
- Has the situation been *confirmed*?
- Was your information source(s) *credible*?
- Is information *consistent* from several sources?

In some cases, the media may be alerted to the situation before all of these facts can be determined. Even if you do not have all of the information yet, it is important to notify the Response Team and, if appropriate, provide the media or other internal or external stakeholders with a statement indicating that the situation is under investigation and that more information will be provided as soon as it is available.

In assessing the situation, it is vital to determine, in a timely manner, the seriousness and potential impact of the situation. Any assessment should always include the potential for escalation to a higher level. The following table is designed to provide an understanding of common descriptors and related characteristics. It should serve as a guide to the seriousness of the situation, along with providing an understanding of the appropriate level of response.

Steps to define the situation.

1. Identify the scope—Collect all relevant facts.

- a. Define the situation and what exactly has happened.
- b. How was the matter brought to your attention?
 - i. By internal staff—who and at what level are they operating?
 - ii. By external parties—a reporter or stakeholder?
If so, who and what is his/her contact information and relationship to Smarter Balanced?
- c. How long has the situation been unfolding?
- d. What are the implications for Smarter Balanced in terms of reputation?
- e. What, if any, potential is there for negative exposure?
- f. What is the scope?
 - i. How are member states and partners impacted?
 - ii. How many are impacted? How widespread is the problem?
 - iii. Are staff impacted or in danger?
 - iv. Is there a disruption of service?

Once an issue has been identified, senior management needs to bring together all relevant staff who can help gather information pertaining to the situation. Staff members should catalog and qualify all information sources, including incoming calls, incoming feedback or reporting, and any news stories originating from the media.

In doing so, staff should pay attention to the following:

2. Collect and centralize all the details related to what happened.

In a contingency situation, knowing every detail about an incident contributes to your ability to serve as a credible source of facts. Reporters, impacted parties, the public, and others will want to know “who, what, when, where, and why.” The more information can be provided in an accurate and proactive manner, the more likely Smarter Balanced will be to convey credibility and keep internal and external audiences informed.

Suggested action step: Appoint one person to manage all information intakes and serve as primary record keeper from the start of the situation to its conclusion. This individual should track incoming calls, internal and external meetings with all stakeholders, and media inquiries, as well as perform ongoing auditing of media coverage of the situation.

- a. *Maintain a chronology of events.* It is critical that the exact timeline of the incident be established—when it occurred, first response and related actions, and when other parties

became involved. Then, identify how Smarter Balanced will continue to provide help and assistance. By understanding the chronological sequence of events, you remain in control of the “knowable” parts of the evolving situation.

⇒ *Suggested action step:* Create a timeline/flowchart of events and update it frequently as the situation unfolds.

Agreed-upon briefing sessions should be established on a daily basis, with all necessary staff to review information and event timelines as information is uncovered. This will help ensure that everyone is on the same page as information is learned and allow the team to review any necessary changes in outreach based on what is uncovered.

3. Understand how it happened—examine the role of systems.

- a. Review Smarter Balanced policies pertaining to the specific issue.
- b. Did the systems break down?
- c. Review with management the role of policies in the situation. Were Smarter Balanced policies adhered to, but the outcome is in dispute? Or were policies violated?
- d. Has Smarter Balanced adhered to its guidelines in relations with member states?

4. Assess the implications—is there any possible effect on ability to deliver services?

- a. Will the situation cause local, state, or federal authorities, media or other oversight entities to investigate Smarter Balanced?
- b. What is the potential impact with on states, students, families, policymakers, board members, etc.?
- c. What is the potential impact on states?
- d. To what extent will activity impact ongoing operations with states?
- e. Will services be affected?

5. What is the solution—how will it be fixed?

- a. Identify the ideal scenario to fix the problem.
- b. Summarize the opinions of all involved stakeholders in defining what a solution looks like.
- c. Assess internal capabilities and experience—is outside counsel needed?
- d. Present the worst-case scenario.
- e. Given all the inputs, present the realistic outcome for a resolution and timeframe.

2. Establish Response Team, Enforce Internal Communications Protocols, and Identify Spokespersons.

Set up a chain of command and assign specific tasks to individuals to ensure that all areas are covered and that no one is duplicating efforts. In doing so, make clear who is responsible for communicating facts to internal and external audiences (the spokesperson), and make sure that another individual is in charge of coordinating all incoming and external calls (usually a media contact).

The composition of this Response Team, especially the choice of spokesperson, should be determined after an assessment of the situation. Smarter Balanced will need to determine, as quickly as possible, the seriousness of the incident and its potential impact. The challenge will be to qualify and act accordingly. Any assessment should always include the potential for escalation to a higher level. **However, keep these three roles in mind: operations, communications, and subject matter expertise.**

Response Team.

The core operations and communications personnel who form the Response Team for Smarter Balanced consists of:

- Joe Willhoft, Ph.D., Executive Director
- Jacqueline E. King, Ph.D., Director, Higher Education Collaboration
- Nicole Siegel, Communications Associate

Contract 21 Contractors

- Mike Bunch, Senior Vice President for Measurement Incorporated
- Barb Scherich, Project Manager for Measurement Incorporated
- Nancy Arnold, Contract Manager for Project 21
- Debra Silimeo, Executive Vice President for Hager Sharp
- Ted Eismeier, Account Supervisor for Hager Sharp
- Lisa Matthews, James Elias, Hager Sharp and others as needed

For the purposes of achievement level setting, key contractors should also be involved:

Depending on the nature of the situation and severity of its impact, Smarter Balanced would add **subject matter experts** through other Executive Team members, including:

- Brandt Redd, Chief Technology Officer (technology/systems)
- Magda Y. Chia, Ph.D., Director, Support for Under-Represented Students (accommodation issues)
- Marty McCall, Ph.D., Lead Psychometrician (psychometrics expert)

The Smarter Balanced response team will need to strike a balance between acting quickly with lean decision-making, collecting information from relevant divisions within the organization, and consulting with and communicating to internal audiences. Note that current Smarter Balanced personnel may not be sufficient to handle the demands of the situation (whether it is an “all-hands-on-deck” PR and communications need or “fixing” the problem, should it be technical in nature or require some sort of outside expertise, such as facility or hotel staff at the In-Person Panel). As part of the planning process for creating a contingency response team, outside resources that could supplement or fill gaps in the team should be considered and relationships developed, as appropriate.

While not part of the Executive Team, communications staff should always be active members of the Response Team, along with relevant project managers. The organizational grid below is one way to think through the members of a Response Team:

Response Team		
Role	Responsibilities	Name
External Spokesperson	<p>Phase 1: Act as the face of the organization, communicating to the media and other audiences all facts and positions agreed by response team.</p> <p>Phase 2: Continue to serve as designated spokesperson for all organizational issues to maintain consistency.</p>	Joe Willhoft Or Jaci King,
Executive Briefer	<p>Phase 1: Inform impacted parties of the incident and all relevant facts and anticipated outcomes.</p> <p>Phase 2: Serve as main contact between the board members, leadership team, and response team.</p>	Joe Willhoft
Stakeholder Contact	<p>Phase 1: Immediately contact all potentially affected stakeholders to inform them of the incident and promise to update with more facts ASAP.</p> <p>Phase 2: Link to ensure critical information coming in from the field flows to the communications team.</p>	Jaci King,
Critical Media Monitor	<p>Phase 1: Concentrate on news breaking in social media, online, and broadcast space, as well as feedback from stakeholders, and inform response team of issues and tone.</p> <p>Phase 2: Continue online and broadcast monitoring and expand to include traditional and print.</p>	Nicole Siegel

Once the situation is defined based on the information available, it must be presented to the Response Team with the goal of identifying the single spokesperson who will respond on the part of Smarter Balanced. This assessment will be based on risk against the national profile and reputation of Smarter Balanced.

The designation of the appropriate spokesperson should be determined after a thorough and candid assessment of the situation analysis, with a discussion of the public profile of the issue. (NOTE: “public” could refer to the “general” public and/or Smarter Balanced publics, such as partners, staff, media, etc.) Management must be flexible with time during a contingency situation. Interested

parties often want to hear from the executive director or senior management rather than a representative of the communications team.

Once the spokesperson is designated, daily reporting should include all members of the Response Team. Prior to any public comment, all messages, statements, and related Q&A should be reviewed and signed off on by the Response Team.

In the event of a potential or definite contingency situation, all primary members of the Response Team are contacted. Sample team roles and responsibilities are outlined in the chart below

Team Responsibilities	
Response Team Role	Name and Support Staff
<p>Team Leader</p> <ul style="list-style-type: none"> Oversee development of situation analysis and review of key timelines and consequences. Initiate priority response procedures and meetings. Review and approve all messages and correspondence to publics to ensure one voice. Designate responsibilities to team members. Convene final evaluation and debrief. 	<p>Chief Executive Officer</p> <p>Joe Willhoft, Executive Director</p>
<p>Message Development</p> <ul style="list-style-type: none"> Oversee development of statement/messaging and all related communications materials. Track material development in chronological file. Cross reference and notes message changes and modifications throughout. Maintain record of all inputs, including sources and dates of input. Record ongoing internal conversations supporting adjustments to (or lack thereof) core messages. Oversee development of final evaluation and debrief. Audit media/stakeholders to assess feedback and next steps (if any). 	<p>Communications</p> <p>Jaci King, Interim Director of Communications</p> <p>Nicole Siegel, Communications Associate</p> <p>Hager Sharp</p> <p>Debra Silimeo, Executive Vice President</p> <p>Ted Eismeier, Account Supervisor</p>
<p>Logistics Coordinator and Record Keeper</p> <ul style="list-style-type: none"> Maintain record of incoming media and stakeholder inquiries and responses. Provide daily or as needed briefing on deadlines for each day of the response and immediate action steps as needed. Maintain chronological file of stakeholder outreach. Develop weekly summary and next actions for the following day/week/month as needed. Identify unresolved issues for the week. 	<p>Communications</p> <p>Nicole Siegel, Communications Associate</p> <p>Hager Sharp staff</p>

Team Responsibilities	
Response Team Role	Name and Support Staff
<ul style="list-style-type: none"> Provide task-driven assistance such as scheduling, logistics, and research. 	
Stakeholder Outreach <ul style="list-style-type: none"> Maintain relationship with states/stakeholders. Record outreach and media response plans of partners. Capture and records messaging of states/stakeholders. Conduct ongoing liaison to states/stakeholders throughout event. 	State Services/Higher Education Collaboration Jaci King, Director of Higher Education Collaboration, Director of Communications
Executive and Internal Communications <ul style="list-style-type: none"> Inform Executive Committee of developments once risk assessment and response strategy are completed. Conduct outreach to Technical Advisory Committee members, senior advisors and relevant consultants. Provide internal updates to staff to explain the situation and organizational response. Designate individuals who are in charge of response. Develop messaging that will be simple for everyone to understand and remember. 	Chief Executive Officer Joe Willhoft, Executive Director

3. Develop and Deliver Messages

With the situation and key deadlines identified, a message track should be developed for dealing with external and internal inquiries relating to the situation. Elements of a typical messaging approach include:

- Three messages is the recommended number to provide the foundation for all public communications: talking points, internal and forward-facing correspondence, media releases, news alerts, website and social media content, and standby statements. The need for consensus is essential with this task.
- Once final messaging is developed, it will be date- and time-stamped and circulated to appropriate Smarter Balanced senior staff, along with the appropriate confidentiality warnings. Then, the team will determine the need for supporting materials based on stakeholder and media inquiries (e.g., statement, interviews, media releases, website and social media content, external correspondence, etc.).
- As part of this task, a Q&A document will be developed and maintained on a chronological basis to address how the story breaks in the media (if it does). It will be based on inquiries from stakeholders.

- On a regular basis (depending on the situation this could be hourly, daily, weekly, etc.), current messaging will be reviewed and revised as needed to stay up to date with stakeholder inputs and the news cycles. These materials will be date- and time-stamped throughout to ensure that all team members are working with current messaging.

Using only one spokesperson will help Smarter Balanced ensure message discipline and consistency. In fast-moving scenarios and media environments, there is potential for conflicting messages from staff and states/stakeholders (online, via Twitter, Facebook, over the phone, etc.) and there must be a single source for information throughout to eliminate the potential for confusion.

However, state agencies also have their own spokespeople, and they may need to respond to media inquiries depending on the nature of the situation. This could come either as the result of a situation involving one state or an issue that applies to all member states. As part of the response, Smarter Balanced should notify and coordinate its response with member states as appropriate and provide pertinent messaging and information to help them effectively communicate at the state level.

Tips for message development.

1. ***Resist the urge to become combative, and stay disciplined.*** Stress will run high, and reporters and other interested parties will want answers within a fixed period of time. You may be asked to comment on a number of issues that are directly related—or not related—to the situation. Regardless of what you know or surmise, it is critical that you stay focused on providing information about your operation.
2. ***Use a human touch.*** Spokespeople have a dual role in speaking out during a contingency situation: disseminating information and serving as an ambassador to targeted audiences. Both roles carry equal importance in achieving communication objectives. Whenever possible, temper statements of fact with expressions of concern for those involved in the situation.
3. ***Speculation is a dangerous game.*** The media, the public, or targeted stakeholders may want immediate answers to complex questions. Transparency is critical, but resist any temptation to comment on things that are uncertain.

Developing message maps.

Message maps structure information essential for responding to public concerns. There are two categories of messages to include:

Messages specific to the situation.

- Current efforts to address the situation;
- Specifics about how Smarter Balanced is working with those impacted;
- A sense of what it will take to fix the situation and when; and
- The systems in place to ensure proper operation.

Spokespersons and expert sources should be:

- Knowledgeable about Smarter Balanced and/or the specific crisis.
- Able to communicate complex information clearly and simply.
- Someone who will be viewed as a trusted source by Smarter Balanced audiences.

Overall messages in support of Smarter Balanced. During a situation, an organization must stand on its brand reputation. Examples of these goals include:

- Dedication to helping all students achieve—regardless of disability, language, or subgroup;
- Delivering high-quality assessment systems to member states;
- Providing tools for teachers and schools to improve instruction and help students succeed;
- Soliciting the input of experienced educators, researchers, state and local policymakers, and community groups; and
- Collaborating through a transparent and consensus-driven process.

Here are seven steps to follow when developing message maps:

- 1. *Identify stakeholders.*** Stakeholders are interested, affected, or influential parties that would be or are currently affected by the situation. See page 20 for list of potential Smarter Balanced stakeholders.
- 2. *Identify concerns.*** Develop a complete list of specific concerns for each important stakeholder group.
- 3. *Identify underlying general concerns.*** Analyze all concerns to identify common sets of underlying general concerns. NOTE: This should be done as time allows—initial messages or holding statements may need to be made before this step can take place.
- 4. *Develop key messages.*** Messages should be crafted in response to each stakeholder question, concern, or perception. Initial messages should address top of mind concerns, e.g., information compromised, employees' safety/well-being (if a workplace- or weather-related incident), specifics and timing of an outage, what is being done to combat and/or alleviate the situation, etc.
- 5. *Develop supporting facts and proof points for each key message.*** Supporting facts provide the continuity and details needed to support the key message. Key messages should have no more than three supporting facts.
- 6. *Conduct systematic message testing.*** To validate the accuracy of technical information, subject matter experts not directly involved in the original message development process should test messages. This will need to be done very quickly. Sharing and testing messages internally with board members, subject matter experts, etc. ensures message consistency and coordination.
- 7. *Plan for delivery.*** Prepare for delivery of the message maps by a trained spokesperson or through the appropriate communication channels.

Once communication objectives have been established and key messages have been finalized, Smarter Balanced must address how to deliver those messages. This step involves identification of mechanisms for updating key stakeholders, interested parties, and the media with information. This could include holding a media briefing, sending out regular updates via statements to the media and/or social media, updating certain Smarter Balanced web pages, disseminating a communication to all impacted parties and/or ALL Smarter Balanced contacts and/or a subset of contacts, or reaching out to key stakeholders individually.

As you reach out, please keep in mind the following communications guidelines:

- Stay on message – bridge back to your key points during questioning.

- Assure stakeholders and reporters that a Smarter Balanced representative will always be available to answer questions and will follow up with more information as soon as it is available.
- Be responsive. Respond to queries quickly, even if there is no new information to share.
- Readily admit if you do not have an answer, but be prompt in getting back with the answer or additional information if it is available. Communicate only verified information.
- Keep accurate records of all stakeholder inquiries, news coverage,

4. Review Course of Action

The final step before launching the response is to review that all procedures have been adhered to and that the systems are in place to respond to and monitor the situation.

5. Establish Timeline and Implement

It is essential to document the flow of events and communication starting with the first time Smarter Balanced learns about a potential situation. All communications with the media, stakeholders, and interested parties must be documented and tracked to:

1. Ensure an ongoing record.
2. Document communications flow and revisions.
3. Ensure that decisions are executed within a well-established and realistic timeline, and to eliminate any false sense of urgency.

Steps for Day One of a Contingency Situation.

To help provide an understanding of issues that should be addressed and managed during the situation, the following checklist has been developed to manage Day One actions and ensure that ongoing actions are consistent throughout.

Response team leader.

- Identify and catalogue the most critical problems facing Smarter Balanced. This will give you a chance to independently gauge the situation and reflect on the next steps.
- Gather the facts as quickly as possible. Be thorough, but do not speculate about anything. You will likely not have all the answers to all the questions right away.
- Do not overreact.
- Once the fact-gathering process has begun and an initial assessment of the situation has been made, members of the Response Team should be notified of the situation. If the situation is fast moving, establish a time as soon as possible for the Response Team to meet in person or via conference/video call. The team should begin to review the Contingency Plan while waiting to receive all details of the situation from those closest to it.

Response team.

- All responders must understand their individual roles and responsibilities. This will facilitate effective communication and decision-making.
- All incoming communications (from the media, stakeholders, partners, and staff) should be identified, logged, and systematically passed on to a designated member of the Response Team.

- The team member designated to monitor media activity should begin all facets of an up-to-the-minute media audit. Determine who is covering the story. This will need to be done with an understanding that Smarter Balanced may not necessarily want to alert the media to its concerns. The same should occur with other members of staff and contractors who receive incoming communication from external parties to monitor/gauge communications from stakeholders.
- Contact relevant counsel as well as Legal, Public Relations/Communications, and Human Resources, as appropriate.
- Begin process for confirming and prepping appropriate spokesperson.
- Make sure all departments and relevant players know who is speaking publicly (includes internal and external communications) for Smarter Balanced and that all communications must be coordinated through that person.
- The executive management liaison should notify senior management of the situation and what steps are currently underway.
- Receptionists, executive assistants, and other support staff, as well as any employee who handles incoming calls and who will be the first point of contact for any external phone calls, emails, or communications, should be advised of the situation and provided messages and talking points. They should be reminded, however, that they are not spokespersons (unless told otherwise).
- Those staff members not on the Response Team should be reminded they are not authorized to speak with the media and should not communicate with external audiences unless given the authority to do so on a case-by-case basis.
- If media coverage is anticipated, begin monitoring the news media and record stories about the situation. Assess and report on media awareness and interest to this point. Prepare a statement for any media queries. Work with the spokesperson on draft media responses.

Steps for Day Two.

In Day Two, the situation should be evaluated by the Team Leader to determine if Response Team members should be added (or subtracted) to provide further assistance and/or expertise to the situation.

- Evaluate the situation. Discuss with the Response Team if additional support is needed. This needs to be done to determine if the appropriate support group members are all in play.
- Discuss with the group whether all the basics have been covered. Has Smarter Balanced done everything that should be done at the early stage?
- Evaluate current external reaction (from customers, partners, media, etc.) and tone of coverage/communication. Are any rumors—true or false—being spread? If so, develop a plan for dealing with them quickly.
- Evaluate the need for media contact, whether by statement or spokesperson response. If Day One events have not driven the need for media implementation, begin draft of brief media standby statement based on what is known.
- Engage Legal in review of statements and documents.

- Review how well you are executing the plan. What is not being done that should be? What needs to be done that is not covered in the plan? Make notes for the future.
- Establish an order of priority for communicating with the various audiences.
- Reassess whether you have identified all appropriate audiences.
- Consider outreach to potential allies.
- Identify who, if anyone, is missing from the team and what gaps need to be filled.
- Distribute the draft media statement (if appropriate) and messages (if updated from Day One) to every member of the team who will be communicating with any constituency, and underscore the point that no one is to speculate beyond the information provided.
- Revisit the worst-case scenario. Is Smarter Balanced prepared for this?
- Keep all members of the Response Team in the loop through regular teleconferences. Report on current understanding of the situation and any decisions that have been made by executive management.

If the situation continues or escalates, the following steps should be considered:

Team leader.

- Continue to assess how the story is being communicated internally and externally. Are the messages resonating? Are additional steps necessary to alter the way audiences are viewing and/or communicating about the situation or Smarter Balance's handling of it?
- Assess whether the facts and the landscape have changed substantially. If so, are we applying the right solutions or do we need to adjust?
- Assess what needs to be done to keep the regular operations running smoothly. Determine whether makeup of the Response Team remains appropriate or if other individuals need to be involved. If the latter, designate responsibilities.
- Assess your own role. Does the situation demand your continuing involvement as the leader, or should you be focusing on other Smarter Balanced activities?

Team.

- What feedback are you getting from the various audiences with whom you are in contact? Are any special steps necessary to improve their understanding of Smarter Balance's messages?
- Continue monitoring of media and social media channels. Work with Team Leader to assess and reassess media and communications strategy going forward.
- Overall, how is it going? Should you stay the course, or are strategic changes required? Consider the perceptions of the problem, especially the news media, stakeholders, and employees. What is the worst-case scenario? What can be done to reduce the likelihood of the worst case?
- Have additional audiences been identified that need attention?

6. Evaluate Results and Recalibrate

All too often, once a situation is resolved, there is a tendency to get back to work. However, the contingency situation should be catalogued along with all related call logs, statements, media coverage, and stakeholder inputs in order to establish a permanent record of what happened.

As part of this, a final briefing should be held with the Response Team to identify what worked and what did not. These inputs can then be pulled together into a one- to two-page briefing sheet that can be used to help inform future situations. During the briefing, share any input you believe would help improve this guide.

As part of the risk management assessment process, the team should also commit to reconvening on a quarterly basis to revisit the contingency communications plan and review the steps to managing a coordinated response. This will ensure that when a situation arises, while the particulars of the situation may be new and beyond the control of Smarter Balanced, the way in which the organization processes the situation and responds to it is not.

Once time has passed to enable the team to reflect on the situation, the team leader should review lessons learned in partnership with the Executive Team and other key Smarter Balanced leadership and stakeholders.

When evaluating the response, consider the following questions:

- Did the Response Team respond appropriately? Did everyone have the facts needed to make good decisions?
- Which approaches worked well during the response?
- What did not work well?
- Were all the appropriate activities covered within the first hour of a crisis? Within the second hour? Within Day One?
- What responses would we delete or add in what to do in the first hour of a crisis?
- What would we do next time to improve on those things?
- Do we feel that any Smarter Balanced policies need to be revisited? Why?
- Would you change any of the procedures?
- How did the spokesperson do?
- If the media was involved, how were they managed throughout the crisis? What should have been done differently in working with the media? Would you have done anything the same? Why?
- Were all the stakeholders communicated to directly and in a timely manner?
- Does Smarter Balanced need to follow up with any stakeholders?
- Knowing what it knows at the end of the crisis, would Smarter Balanced do anything differently?

Scenarios

There are a number of potential contingency scenarios to which Smarter Balanced must be prepared to respond in a way that minimizes the impact to get out the truth and helps the public understand the situation. The information that follows should be used for planning purposes only. Every critical incident requires a response based on the facts of the situation. Each occurrence should be dealt with individually based on the facts of the situation at the time it was made known to the Consortium. This plan cannot anticipate all the specifics of such a situation should one occur. Please use this as a guide for building skills and internal capacity for rapid response.

Priority Scenarios

Divided Vote on Achievement Levels

Failure of the Governing States to unanimously endorse the cut scores and achievement levels at the fall collaboration conference is a scenario that would require a rapid response from Smarter Balanced. Smarter Balanced always seeks a unanimous vote on motions considered by the Governing States. If the states fail to reach a unanimous vote, Smarter Balanced would hold additional votes on the motion.

While member states reviewed the RFP and proposals for Contract 21 and approved the achievement level setting design, the final decision to endorse or not endorse the achievement level scores rests with the chiefs. However, it's possible that one or more states would object to the cut score recommendations.

For example, a state could claim that the expectations would create unfair expectations for their students. States could also have technical or process-related concerns, such as a lack of representation in the process or questions.

During the session, states could voice concerns about the achievement levels and cut scores. States may have concerns about the process being too heavily weighted toward one set of recommendations. **Key messages for a divided voted include:**

- Smarter Balanced is a voluntary, state-led consortium. We trust in the decision-making of our member states. The majority of states in the Consortium have chosen to endorse the achievement level scores and move forward with the recommendations from educators, higher education faculty and community members so that the assessment system will be fully-operational for the 2014-2015 school year.
- The process for developing achievement level scores included extensive opportunities for input from our member states. Throughout this process, we have been committed to ensuring that educators and other stakeholders in our states took the lead in the final achievement levels and cut scores.
- Our process has relied on the most thoroughly-researched and well-documented processes for achievement level setting. However, our member states are also taking a pioneering step by agreeing to use a set of common performance standards. Several states have chosen not to endorse the achievement level and cut score recommendations.
- Member states had extensive representation in this process through panelists nominated by state employees at the in-person event, the innovative Online Panel with capacity for thousands of educators to participate, as well as final oversight of recommendations from those two groups.

- Moving forward, Smarter Balanced will continue to listen to the needs of our member states and provide the highest level of service to all of our member states and territories. We are committed to working with all of our states and partners to implement an assessment system based on fair expectations for all students.

Goal: Communicate key messages about the achievement level setting process and relay factual information about the state vote.

Response actions:

- Assemble the Response Team immediately.
- Determine the intensity level of the incident.
- Inform and coordinate with stakeholders, consistent with the seriousness of the incident, to make them aware of the event and agree on next steps.
- Identify a spokesperson.
- Draft and review all statements and messages with the executive team.

Media/audience response and message development:

Develop a statement and/or talking points, incorporating key messages, which may include the following:

- Factual statement(s) about the vote and the desire of many states to move forward with the recommendations.
- Key details about that state's opportunities to contribute to the process, including higher education involvement
- The details as we know them at this point in time (be sure to keep this updated)
- Details on state concerns as needed—and what the rejection of the scores means for states involvement in Smarter Balanced
- Details on whether Smarter Balanced is working with the member state to reevaluate
- Information, designed for a lay audience, about the achievement level setting process—including an outline of the state's involvement in the process—and addressing the state board of education's concerns.
- Reiterate the Consortium's overarching organizational messages, including its status as a state-led initiative, and any history dealing with a similar situation, if appropriate.
- Develop a question and answer document that is specific to the state in question, and ensure that the Response Team—and especially the identified spokesperson, should there be a need for one—is conversant on and understands the talking points.

The Response Team will need to determine how to communicate with the Consortium's various audiences, including those specifically impacted by the state board's decision, and what messages need to be disseminated.

Additional activities:

- Inform Executive Committee of the response plan and provide them with a statement in advance of public communications. Ask them to route all inquiries to the executive director or the designated Response Team representative.
- Continually monitor social media and other news media and feedback from staff, state leadership, and other stakeholders, and external reaction to the event as it unfolds.

Prepare an after-action report that would summarize lessons learned from the contingency situation and review how the incident was managed and whether other actions were needed as events unfolded

Scenarios: Acceptance of the Achievement Levels

The success of achievement level setting does not begin or end with the In-Person and Online Panels. In most Consortium member states, the state board of education holds final authority over the performance levels for the assessment. The Consortium's achievement level setting process is also a multiple-grade process, making it inherently more complex. The typical process for adopting performance levels in these states would be as follows:

- a. Smarter Balanced follows a process for achievement level setting approved by Governing States.
- b. The scores would be on the agenda for a state board of education meeting that would take place in the weeks following the conclusion of the achievement level setting activities.
- c. The standards would go undergo a review by the state education agency before review by the board.
 - i. The standards would be presented to the board following this review, and the board would have the option to accept; reject; or modify.

While not common, state boards in the past have made extensive modifications to standards. As an example, during AIR's standard setting with South Carolina, the board made sweeping changes to the cut scores. This was a standard setting activity involving standards for multiple grades, and the board went as far as three standard deviations from each cut score.

Extensive modification or rejection of the cut scores in at least one member state would seem the likeliest of these scenarios. An exacerbating factor is the high rate of turnover on state boards of education during the process; typically, standards are set in a much tighter window.

A potentially more likely scenario would be one or more state boards determining that they are simply not going to abide by the recommended achievement level scores. In other words, if a board were to make its own cut scores very different from the scores recommended. Context for communication around a change in scores is important; parents and educators agreed that a potential change in scores would not worry them, as long as they received adequate information in advance.

Goal: Communicate key messages about the achievement level setting process and commitment to the Common Core State Standards and aligned assessments and reinforce.

Response actions:

- Assemble the Response Team immediately.
- Determine the intensity level of the incident.

- Inform and coordinate with stakeholders, consistent with the seriousness of the incident, to make them aware of the event and agree on next steps.
- Identify a spokesperson.
- Draft and review all statements and messages with the executive team.

Media/audience response and message development:

Develop a statement and/or talking points, incorporating key messages, which may include the following:

- Factual statement about what occurred
- The details as we know them at this point in time (be sure to keep this updated)
- Details on what was done to the cut scores—how greatly they have been changed, or what the rejection of the standards means for the states involvement in the Consortium
- Details on whether Smarter Balanced is working with the member state to reevaluate its commitment to the cut scores
- Details on whether modification or rejection of the cut scores is due to a state’s legislative decision-making or other external events largely out of the Consortium’s control
- Details on next steps
- Information, designed for a lay audience, about the achievement level setting process—including an outline of the state’s involvement in the process—and addressing the state board of education’s concerns.
- Reiterate the Consortium’s overarching organizational messages, including its status as a state-led initiative, and any history dealing with a similar situation, if appropriate. (For example, the South Carolina state board of education’s decision to set its own cut scores.)
- Develop a question and answer document that is specific to the state in question, and ensure that the Response Team—and especially the identified spokesperson, should there be a need for one—is conversant on and understands the talking points.

The Response Team will need to determine how to communicate with the Consortium’s various audiences, including those specifically impacted by the state board’s decision, and what messages need to be disseminated. It is possible, if not likely, that different audiences will require different levels of detail.

Additional activities:

- Inform key leadership of the response plan and provide them with a statement in advance of public communications. Ask them to route all inquiries to the executive director or the designated Response Team representative.
- Continually monitor the news media and feedback from staff, state leadership, and other stakeholders, and external reaction to the event as it unfolds.
- Prepare an after-action report that would summarize lessons learned from the contingency situation and review how the incident was managed and whether other actions were needed as events unfolded.

Highly-visible criticism by an external commentator or panelist

Another likely scenario would be the publication of a damaging article/blog either by an external commentator, scholar or a panelist who participated in the In-Person Panel or the Online Panel. Smarter Balanced faces opposition from prominent and vocal critics on the political right and left, who will be ready to pounce on any issues at the panels as a reason to criticize the Common Core State Standards and/or Smarter Balanced. There have been cases of participants in standard-setting activities publishing criticisms of the process in newspapers and blogs.

Panelists have been nominated by states and have been chosen for their expertise and commitment to public education. They have received detailed information from Measurement Incorporated/CTB and Smarter Balanced about the goals of the achievement level setting activities. They will receive training and instruction in Dallas. Despite all the work that has gone into creating a thoughtful and consensus-based process, it is based on voluntary participation and the free expression of the individuals who attend.

It is critical to establish and maintain the reputation of Smarter Balanced as a trusted source of facts about the standards and assessments to counter opponents' claims. If or when a criticism of the process is published, Smarter Balanced will need to be prepared to immediately respond with factual information about the achievement level setting process, reinforcing its broadly-inclusive design and its grounding in well-documented and researched methods (see *Messaging for use during rapid response*).

Goal: Communicate key messages about the achievement level setting process and relay factual information about the state vote.

Response actions:

- Assemble the Response Team immediately.
- Determine the intensity level of the incident.
- Inform and coordinate with stakeholders, consistent with the seriousness of the incident, to make them aware of the event and agree on next steps.
- Identify a spokesperson.
- Draft and review all statements and messages with the executive team.
- Seek outside validators: states, education leaders, experts.

Media/audience response and message development:

Develop a statement and/or talking points, incorporating key messages, which may include the following:

- Information, designed for a lay audience, about the achievement level setting process—including an outline of the state's involvement in the process—and addressing the state board of education's concerns.
- Reiterate the Consortium's overarching organizational messages, including its status as a state-led initiative, and any history dealing with a similar situation, if appropriate.
- Communicate relevant technical details about methodology to counteract misinformation about the process

- Develop a question and answer document that is specific to the state in question, and ensure that the Response Team—and especially the identified spokesperson, should there be a need for one—is conversant on and understands the talking points.

The Response Team will need to determine how to communicate with the Consortium's various audiences, including those specifically impacted by the state board's decision, and what messages need to be disseminated.

Additional activities:

- Inform Executive Committee of the response plan and provide them with a statement in advance of public communications. Ask them to route all inquiries to the executive director or the designated Response Team representative.
- Continually monitor the news media and feedback from staff, state leadership, and other stakeholders, and external reaction to the event as it unfolds.
- Prepare an after-action report that would summarize lessons learned from the contingency situation and review how the incident was managed and whether other actions were needed as events unfolded.

Additional Scenarios: In-Person Panel

Smarter Balanced selected an achievement level setting process using the best elements of many accepted methods contained in *The Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999) and best practices based on pilot and field test results. The achievement level setting design for Smarter Balanced includes elements of the:

- NAEP standard-setting process (including bookmark, modified Angoff, and some form of body-of-work for the writing achievement levels);
- National Assessment Governing Board approach for recruitment, training, and procedures;
- American Diploma Project (Achieve’s project led by Sheila Byrd Carmichael); and
- State of Texas Assessments of Academic Readiness (STAAR).

Extensive planning and training has gone into preparing facilitators and staff for the In-Person Panel. Mitigation steps have been taken as part of contingency planning for the In-Person Panel. But the unexpected can always occur. There are still risk factors that could disrupt or adversely affect the outcomes of the Panel achievement level setting process. Broadly speaking, the greatest risk factor for the In-Person Panel is a situation that causes the dynamic to become difficult to control—either as a result of a technical or operational issue or some type of real-time disturbance involving the panelists. This could either prevent facilitators and panelists from completing the task of standard setting or skew the results in such a way that would call into question their validity. A technical failure could cause delays that exceed the time constraints of the planned in-person activities.

Based on our interviews with experts and facilitators from Measurement Incorporated and CTB/McGraw-Hill, the scenarios below outline potential contingency situations that could develop before, during and after the In-Person Panel activities.

a. Ejection of a rogue panelist.

For the In-Person Panel, one of the worst-case scenarios would be if an individual or a particular set of circumstances disrupts the event. If such an individual slips through the nomination and vetting and gets invited to the In-Person Panel, they could disrupt an entire room and conceivably disrupt an entire three-day panel. To provide realistic scale for this example, a single outspoken participant could significantly disrupt a single room with 29 or 30 other panelists. In this scenario, this person is referred to as a “rogue.”

Typical behavior for a person like this would include placing an extreme cut score. This is generally not a problem since there are enough other participants in the room to balance out a clear outlier. It is possible that a high-level or well-known participant could engage in this type of rogue behavior, hijacking conversation in the room and influencing others to do the same. Occurrences can cause large deviations in cut scores. However, in the experience of experts interviewed for the purposes of this plan, such deviations have offset each other in the past.

Outspoken or plaintive commentary would not pose a major problem. In fact, this should be expected as part of a discussion and large-scale activity involving hundreds of diverse participants. However, a panelist could “filibuster” on a cut score decision, spend excessive time speaking negatively of the test or Smarter Balanced, or repeatedly bring the discussion to questions unrelated to the task at hand. If the individual has been asked to return to task repeatedly, then escalation is considered.

As an example, if someone begins ranting, using offensive language, or becomes physically or verbally abusive, he or she would be removed. In the experience of key personnel managing the

Smarter Balanced achievement level setting, it is very unlikely that a panelist would have to be removed, but it is definitely a possibility.

Mitigation strategy: If necessary, Measurement Incorporated will use an escalation procedure approach to mitigate an issue. The facilitator—a trained staff member from Measurement Incorporated and the first line of defense in the room—would attempt to moderate a rogue panelist’s behavior. The facilitator has received training and a guide for managing the panel, and, in this particular case, each has previous experience in facilitating standard setting. Smarter Balanced should communicate expectations for behavior and decorum to states and panelists in advance. The escalation procedure for disruptive panelists is as follows:

1. Facilitator attempts to resolve.
2. Failing resolution, facilitator escalates issue to lead facilitator (Michael Bunch or Dan Lewis).
3. Lead facilitator attempts to resolve.
4. Failing resolution, lead facilitator escalates issue to executive director.
5. Executive director attempts to resolve.
6. Failing resolution, executive director asks the panelist to leave and contacts representatives from the Governing State.

Goal: Communicate to impacted panelists and Governing States that a panelist has been asked to leave the In-Person Panel, refer to the training and orientation provided to panelists, and reinforce the importance of professional courtesy and civility in the context of this activity.

Key messages:

- Facilitators will take this action only as a last resort as part of an established escalation procedure and after thorough discussion with the involved panelist(s).
- In order to ensure that the In-Person Panel can achieve its goal on behalf of Smarter Balanced states, educators, students and the general public, it is critical that facilitators have the ability to maintain an orderly environment in which all panelists have their voices heard.
- Panelists are invited and encouraged to share their views. However, as a condition of participation, Smarter Balanced also asks that they remain polite, courteous, and respectful to other panelists and facilitators and allow others a chance to share their views. There are important “ground rules” for participation: thoughtful and civil discussion and mutual respect.
- Panelists at the In-Person Panel for Achievement Level Setting have been selected by Governing States to represent stakeholders, students, and the general public. We appreciate their commitment to public education and willingness to go “above and beyond” through their participation in this event.

Response actions:

- Assemble the Response Team immediately.
- Determine the intensity level of the situation.
- Inform and coordinate with stakeholders, consistent with the seriousness of the incident, to make them aware of the event and agree on next steps.
- Identify a stakeholder contact (executive director or deputy).
- Develop messaging and statements for dissemination.

Message frequency:

- Notify impacted panelist of request to leave the In-Person Panel, only as a last resort after other options in the escalation procedure are exhausted.
- Notify state lead(s) by phone upon a panelist's ejection.

b. Disturbance at the plenary session.

The most severe scenario at the In-Person Panel would be if a “rogue panelist” chooses to go rogue during the plenary session. The sensitivity behind the plenary sessions stems from the fact that it will be difficult for facilitators get back on track if a disruptive situation arises.

Depending on the circumstances, a plenary session can be an appropriate venue to address grievances or hear legitimate concerns. In contrast with a smaller forum of delegates from Smarter Balanced Governing States, handling a large number of questions or concerns from standard setting panelists could open a floodgate of questions that Smarter Balanced is not prepared to handle. Due to the large size, multi-state nature of this standard setting, it is important to keep the conversation focused on reviewing test items and setting achievement levels.

Mitigation strategy: It is important to establish trust during the plenary session in order to maintain calm, order, and stability of the activities. When Smarter Balanced leaders and facilitators have the floor, use the “airtime” to:

- Clearly explain the goals for the activities;
- Thank panelists for their participation;
- Set the tone of the meeting up front by reminding all panelists of the purpose of the session and its importance for students;
- Issue a polite reminder about professional courtesy; and
- Encourage participants to engage in a thorough, thoughtful discussion and make their voices heard.

The escalation procedure for a disturbance at the plenary session is as follows:

1. Facilitator attempts to resolve.
2. Failing resolution, facilitator escalates and attempts to resolve the issue.
3. Failing resolution, facilitator escalates issue to executive director.
4. Executive director attempts to resolve the issue following the plenary session.
5. Failing resolution, executive director asks the panelist to leave and contacts representatives from the Governing State.

Goal: Communicate to impacted panelists and Governing States that a panelist has been asked to leave the In-Person Panel, refer to the training and orientation provided to panelists, and reinforce the importance of professional courtesy and civility in the context of this activity.

Key messages: See messages from Scenario 1.a.

Audiences: Panelists at the In-Person Panel, Smarter Balanced state leads, contractors and staff, impacted panelist(s)

Message Frequency:

- Impacted panelist(s) should be notified of request to leave the In-Person Panel, but only as a last resort after other options in the escalation procedure have been exhausted. Notify state lead(s) by phone upon a panelist's ejection.

c. Operational problems at the In-Person Panel.

A number of events beyond our control during the week could occur at the In-Person Panel (for example, a fire that leads to a fire evacuation at a hotel and causes a loss of 45 minutes). Scenarios could arise—travel cancellations, emergency weather conditions, local protests, etc.—that cause panelists who were expected to attend and participate to drop out of the In-Person Panel. It is also possible for an event with nationwide impact and significance to occur that would necessitate cancellation of the In-Person Panel. Computer systems, including servers or individual machines, could go down, or facilitators could experience some other type of failure.

A power failure at the venue could impact success factors for the In-Person Panel. The most damaging consequence would be a major time delay for the panels. The backend data processing that needs to occur on-site is very intense, with data elements related the operation of the standard setting panels happening on the fly.

The hotel also has far more guest rooms and meeting rooms than we have booked, providing some buffer and backup options for physical space.

Additionally, it will be critical to maintain physical security and privacy at the event. Business meetings involving Smarter Balanced states tend to have open access in public venues such as hotels, without a host checking to see who is there and whether they are supposed to be there. Despite the fact that the event will not be announced publicly aside from the event manifest for the hotel, it may be possible that members of the media, parents, local educators, or political activists will attempt to gain entry into the proceedings. Given the proprietary nature of the In-Person Panel, Smarter Balanced should take steps to credential each participant. Smarter Balanced should instruct hotel staff that registration for the event is required and that participants will be asked to show identification.

Mitigation strategy: Measurement Incorporated has mitigated against this risk by moving from laptops and wireless connections to servers with hardwired connections. On its own, a power failure or loss of Internet connectivity is unlikely to cause a lengthy disruption in the activities. However, short of reserving and paying for a second location, there is no backup plan for a complete operational failure at the In-Person Panel. Through the conference contract, the Hilton has indemnified Measurement Incorporated and Smarter Balanced against negligence, gross negligence, and misconduct of their employees. This does not completely protect against the possibility of an A/V failure or a problem with rooms, but it does provide some protection against catastrophic failure.

6. *Response actions:*

- Assemble the Response Team immediately.
- Determine the intensity level of the issue by undertaking an assessment of the situation. Can the In-Person Panel move forward or will there be significant delays? Is there an imminent risk to the building itself? Can technical/operational issues be resolved in a timely fashion?
- Establish the facts and rapidly develop an event timeline. The timeline should include when the problem was discovered; who discovered it; whether impacted parties (e.g., member states) have been notified and, if so, what they were told; and the current situation status.
- Inform and coordinate with stakeholders, consistent with the seriousness of the incident, to make them aware of the event and agree on next steps.
- Identify a spokesperson.

- Develop messaging and statements for dissemination.

Media/audience response and message development:

Develop a statement and/or talking points, incorporating key messages, which may include the following:

- Factual statement about what occurred
- What the Consortium is currently doing to preserve the integrity of the In-Person Panel
- The details as we know them at this point in time (be sure to keep this updated)

d. Repeal of Common Core State Standards/loss of Governing State status prior to panel, but after panelist selection.

In the lead up to the standard setting activities in November, it is possible that a state could repeal or otherwise change its state content standards prior to the In-Person Panel, or another condition could occur in which it would no longer be a member of Smarter Balanced. A number of states are considering repealing all or parts of the Common Core State Standards or revoking their memberships in the Consortium. In the months leading up to the achievement level setting activities, state legislatures will be winding down their 2014 legislative sessions. This reduces the likelihood that additional states will repeal or modify their state standards in advance of the In-Person Panel.

One of the core tenets of the Consortium is that it is a voluntary, state-led initiative. Therefore, Smarter Balanced is committed to recognizing that while there are important benefits of participation to member states, the Consortium is respectful of the primacy of state decision-making and public oversight in the management of public education. Should this contingency occur, Smarter Balanced will need to reconsider the status of panelists from that member state, as they would be forfeiting their right to participate in the process for developing achievement level scores. If a state drops out in September, there will be a small number of alternates, but there may not be enough to cover an entire state delegation.

Mitigation strategy: In this scenario, Measurement Incorporated and CTB/McGraw-Hill will go over the roster of current panelists. Should a Governing State pull out in the next few months, Measurement Incorporated and CTB/McGraw-Hill will approach the remaining states, explain the situation, and ask that they nominate additional panelists for consideration at the In-Person Panel.

e. Leaked information and negative exposure during or after the In-Person Panel.

The problem of item security is an issue that Smarter Balanced has dealt with extensively in the past. On a number of occasions, students participating in the field test or other advance activities involving secure test items have used their social media account to post images of testing materials. There is a process in place for contacting school districts and notifying school leaders and administrators about such incidents. Similarly, it is possible that a panelist at the In-Person Panel could leak images of the activities or reproductions of specific materials to news organizations or social media. Measurement Incorporated and CTB/McGraw-Hill will ask confirmed panelists to sign a Non-Disclosure Agreement; however, panelists could disregard the agreement or, perhaps unknowingly, breach the terms of the agreement.

Additional Scenarios: Online Panel

The Online Panel provides a great opportunity for Smarter Balanced to involve educators and other stakeholders in achievement level setting on a large—potentially massive—scale. For the same reason, the Online Panel also makes an appealing target for entities that are hostile to assessments or to the Common Core.

It seems unlikely that individuals would attempt to manipulate the results of the Online Panel or “game the system” due in part to its length. However, increasingly, anti-testing advocates are organizing via social media and the Internet. These activists are committed and, typically, at odds with testing generally and the goals of the Common Core assessment consortia in particular. Systematic infiltration of the Online Panel by opposition groups could significantly disrupt the achievement level setting process.

Checks are in place to sort data and reject outliers. There are additional logistical obstacles, too, such as the amount of time required to complete the Online Panel, and the fact that registration will occur far in advance. However, this scenario remains a possibility. Opposition groups could aggressively publicize the fact that the Consortium has rejected certain participants, creating a perception that the Consortium’s commitment to transparency and an open, democratic process was less than sincere.

Goal: Clearly communicate the current situation by providing a public response that addresses what the Consortium knows, the nature of the attack or disruptive activity, and the steps being taken to rectify the situation and confirms that the integrity of the process has been preserved.

Response actions:

- Assemble the Response Team immediately.
- Determine the severity of the attack; end the attack, if not already done.
- Determine the intensity level of the situation.
- Establish the facts and rapidly develop an event timeline. The timeline should include when the problem was discovered; who discovered it; whether impacted parties (e.g., member states) have been notified and, if so, what they were told; and the current situation status.
- Inform and coordinate with stakeholders, consistent with the seriousness of the incident, to make them aware of the event and agree on next steps.
- Identify a spokesperson.
- Develop messaging and statements for dissemination.

Media/audience response and message development:

Develop a statement and/or talking points, incorporating key messages, which may include the following:

- Factual statement about what occurred
- How the Consortium discovered the attack
- What the Consortium is currently doing to preserve the integrity of the Online Panel
- The details as we know them at this point in time (be sure to keep this updated)

- Information about the nature of the Consortium’s screening procedure for achievement level setting, including security and auditing procedures to prevent/detect tampering during the

Type	Stakeholder
------	-------------

process—Will Smarter Balanced review those procedures and resources in light of this situation? What are immediate next steps for the Online Panel? How will the Consortium determine which cut scores to reject?

- Reiterate the systems that are in place to ensure the security of the Online Panel and data.
- Reiterate the Consortium’s overarching organizational messages, including its commitment to pioneering the most open and inclusive achievement level setting in the history of assessment, and its commitment to protecting the integrity of the achievement level setting process through the combination of the In-person and Online Panel.
- Develop a question and answer document that is specific to the event and ensure that the Response Team—and especially the identified spokesperson, should there be a need for one—is conversant on and understands the talking points.

Additional activities:

- Inform key leadership of the response plan and provide them with a statement in advance of public communications. Ask them to route all inquiries to the executive director or the designated Response Team representative.
- Continually monitor the news media, feedback from staff and other stakeholders, and social media reaction to the event as it unfolds.
- Prepare a follow-up report reviewing how the situation was managed and whether other actions were needed as events unfolded.

Appendix 1: Internal and External Audience Stakeholders

Internal	Executive Team
	Smarter Balanced Staff
	Executive Committee
	Governing States (i.e. chief state school officers, assessment directors, higher education leads)
	Regional Advisers
	Technical Advisory Committee
External	School districts (i.e. teachers and school leaders)
	Media (national, trade, state and local)
	Funders and strategic partners
	State and federal policymakers (i.e. state boards of education, legislators, U.S. Department of Education)
	K-12 and higher education associations and advocates
	Parents of public school students
	Influencers (i.e. foundations, think tanks, researchers, opinion leaders)
	Contractors/vendors
	Business Community

Appendix 2: Message Map for Achievement Level Setting

**WORKSHEET:
Intake Form for Panelist Interviews**

Instructions: This form is to be used by panel facilitators and staff from Measurement Incorporated, CTB/McGraw-Hill and Hager Sharp for purposes of identifying qualified panelists and spokespeople to participate in video interviews.

Interview Subject	
Prepared by (Facilitator Name) :	_____ Date: _____ Time: _____
Panelist Name:	_____
Panelist's State:	_____
Panelist Type (check one)	K-12 Educator () Higher Education () Parent/Community Member () Business Leader ()
Panelist Type II (check all that apply)	K-12: SWD Experience () ELL Experience () Administrator () Higher Ed: Faculty Member () Faculty Member ()
Title:	_____
Content Area (check one)	English language arts/literacy () Math ()
Grade-Level (K-12 only)	_____

Additional Notes

Share Your Feedback: Participate in a Video Interview

Instructions for Panelists: You are invited to participate in an interview to share your experience at the In-Person Panel and the achievement level setting process. If you are interested, please fill out the sign-up form and follow the instructions at the bottom of the page. As there will be a limited number of opportunities for interviews, please be aware that there are no guarantees that you will be interviewed.

Background Information: As part of the In-Person Panel for Achievement Level Setting, Smarter Balanced will be developing a video compilation to document the process for member states, policy makers and other interested parties. Panelists interested in providing feedback and comments on the process will have the opportunity to participate in an optional video interview for use in multimedia materials.

Interview Signup	
Prepared by (For Facilitator Use Only)	Date: Time:
Panelist Name	
Panelist State:	
Panelist Role	K-12 Educator (<input type="checkbox"/>) Higher Education (<input type="checkbox"/>) Parent/Community Member (<input type="checkbox"/>)
Content Area	English language arts/literacy (<input type="checkbox"/>) Math (<input type="checkbox"/>)

To complete your interview:

- Visit the Wyeth Room located in the same hallway as your assigned meeting room. The screen outside the interview room will include a display, "Video Interview Room," and should be easy to find.
- Deliver this form to the camera crew.
- For questions, contact Ted Eismeier, 315-335-9222, or see the facilitator.



Interview and Multimedia Permission and Release Form

Dear [Panelist Name]:

Congratulations on your selection to participate in the In-Person Panel for Achievement Level Setting for the Smarter Balanced Assessment Consortium.

The In-Person Panel will be held in Dallas, Texas, on October 13–15, 2014. As part of this in-person panel, the Smarter Balanced Assessment Consortium will be developing a short video to document achievement level setting process for member states and other interested parties. The video compilation will explain the achievement level setting process to members of the general public the process for the achievement level setting.

Panelists interested in providing feedback and comments on the process will have the opportunity to participate in an optional video interview for use in multimedia materials. The attached Interview Multimedia Permission and Release form will allow the Smarter Balanced Assessment Consortium to document the In-Person panel and use your likeness in future multimedia and print materials.

As part of participation, panelist are asked to complete the attached permission and release form and return it electronically.

If you have any questions, please contact Mandy Hunter at MI_smarterbalanced21@measinc.com.

Thank you for your participation. We look forward to working with you.



Interview and Multimedia Permission and Release Form

Interview and Multimedia Permission and Release Form

In consideration for the opportunity to be interviewed and/or photographed in connection with the Smarter Balanced Assessment Consortium, I agree to the following:

- The interview or materials obtained through this release may be used by Smarter Balanced Assessment Consortium and its licensees for any purpose in any medium in the perpetuity, including without limitation scholarly, educational, promotional and/or commercial purposes at Smarter Balanced Assessment Consortium's discretion.
- I grant to Smarter Balanced Assessment Consortium and its licensees the irrevocable and perpetual right to record, reproduce, distribute, transmit, display, publish, or otherwise use my name, photograph, voice, likeness, comments, in any manner and medium, in whole or in part, individually, or in conjunction with other materials, for any purpose in perpetuity, including the right to edit the Materials. I understand that I will not receive any compensation for or have any right of approval over any use of the Materials. I release and hold Smarter Balanced Assessment Consortium and its licensees harmless from all claims and demands arising out of or in any way connected with the Materials, including without limitation any and all claims for invasion of privacy, infringement of my right publicity, defamation, libel, and any other personal and/or property rights.
- All statements that I make during the interview and/or all other materials related to the Materials will reflect my honest opinions, beliefs, and/or experiences, and any representations I make shall be true and accurate to the best of my knowledge.

Signature _____

Please Print Name _____

E-mail _____

Phone # _____

Date _____

City, State _____



Interview and Multimedia Permission and Release Form

For Smarter Balanced Use Only

Event _____

Distinguishing Features _____

Interviewer Name _____

Notes _____

The Smarter Balanced Assessment Consortium brings together states to create a common, innovative assessment system for mathematics and English language arts that is aligned with the [Common Core State Standards](#) and helps prepare students for college and careers. The Consortium involves educators, researchers, policymakers, and community groups in a transparent and consensus-driven process to help all students thrive in a knowledge-driven global economy. The Consortium's projects are funded through a four-year, \$175 million grant from the U.S. Department of Education and the generous contributions of charitable foundations. Membership is open to any interested U.S. state, territory, or commonwealth. For more information, please visit: www.SmarterBalanced.org.

Smarter Balanced Achievement Levels Surveys among Parents and Students

Contents

Methodology	227
Executive Summary of Highlights	228
Conclusions.....	229
Characteristics of the Parent Sample	230
Highlights of Parent Responses.....	230
Annotated Questionnaire: Parents of Children in Grades 3-12.....	232
Characteristics of the Student Sample	238
Highlights of Student Responses	238
Annotated Questionnaire: Students in Grades 8-12.....	239

Methodology

Smarter Balanced commissioned two surveys to test potential descriptors for student achievement levels on the assessment currently in development. The surveys were fielded online by Ipsos, a leading market research firm, on October 17-24, 2014. The sample for the survey among parents consists of 615 parents of children 18 years old or younger, with 271 men and 344 women. The sample for the survey among students consists of 505 children ages 12 to 17, with 258 boys and 247 girls. Additional information about respondent characteristics are included in this report.

Respondents were asked for their opinions on various aspects of the following four sets of descriptors:

<p>Option 1 Level 4: Advanced Level 3: Proficient Level 2: Basic Level 1: Minimal</p>	<p>Option 2 Level 4: Advanced Level 3: On track Level 2: Needs improvement Level 1: At risk</p>
<p>Option 3 Level 4: Exceeding expectations Level 3: Meeting expectations Level 2: Partially meeting expectations Level 1: Not meeting expectations</p>	<p>Option 4 Level 4: Exceeding grade level standards Level 3: Meeting grade level standards Level 2: Partially meeting grade level standards Level 1: Below grade level standards</p>

In a final survey question, respondents were asked to select the single option they found most effective in describing student performance on the new assessment.

Executive Summary of Highlights

Parent Survey

When asked to select the single most effective set of descriptors from among the four options, parents selected the following by a substantial margin over the other three:

Level 4: Exceeding grade level standards
Level 3: Meeting grade level standards
Level 2: Partially meeting grade level standards
Level 1: Below grade level standards

Feedback on specific terms:

- Option 1: Parents agreed ‘advanced’ and ‘proficient’ are effective descriptors (73% and 61% agreed respectively), but found ‘minimal’ less effective (49% agreed).
 - Women were more likely than men to object to the use of ‘minimal’ – 33% of women somewhat or strongly *disagreed* that it is a good way to describe the lowest level of student performance, compared with 24% of men.
- Option 2: A majority of parents agreed ‘on track’, ‘needs improvement’ and ‘at risk’ are effective descriptors (73%, 71% and 67% agreed respectively).
- Option 3: While a majority agreed ‘meeting expectations’ and ‘not meeting expectations’ are effective descriptors (63% and 62% agreed respectively), only half (55%) agreed ‘partially meeting expectations’ is effective.
- Option 4: Three-quarters (75%) agreed ‘exceeding grade level standards’ is “a good way to describe the top level of student performance.”
 - A majority agreed ‘meeting grade level standards’ and ‘below grade level standards’ are effective descriptors (66% and 67% agreed respectively).
 - Women were more likely than men to agree with the effectiveness of the terms in this set of descriptors.

Student Survey

When asked to select the single most effective set of descriptors from among the four options, students selected the following by a substantial margin over the other three:

Level 4: Exceeding grade level standards
Level 3: Meeting grade level standards
Level 2: Partially meeting grade level standards
Level 1: Below grade level standards

Feedback on specific terms:

- Option 1: Most students agreed (80%) that ‘advanced’ and ‘proficient’ are good words to use to tell them if they did well on the test, but fewer agreed (58.5%) that ‘basic’ and ‘minimal’ are good words to use to tell them if they need extra help.
 - Most students agreed (82%) ‘advanced’ is a good way to describe the top level of student performance.
 - A majority agreed (69%) ‘proficient’ is a good way to describe a student performance that is on track for college and career readiness.

- Only slightly more than half (57%) agreed ‘minimal’ is a good way to describe the lowest level of student performance.
- Option 2: A majority of students agreed ‘on track’, ‘needs improvement’ and ‘at risk’ are effective descriptors (73%, 75% and 70% agreed respectively).
- Option 3: While a majority agreed ‘meeting expectations’ and ‘not meeting expectations’ are effective descriptors (66% and 70% agreed respectively), fewer students (59%) agreed ‘partially meeting expectations’ is effective.
- Option 4: Most students (82%) agreed ‘exceeding grade level standards’ is “a good way to describe the top level of student performance.”
 - A majority agreed ‘meeting grade level standards’ and ‘below grade level standards’ are effective descriptors (67% and 71% agreed respectively).

Conclusions

Parents had a strong preference for the following set of descriptors:

Level 4: Exceeding grade level standards
Level 3: Meeting grade level standards
Level 2: Partially meeting grade level standards
Level 1: Below grade level standards

The weakest descriptor in this set is the Level 2 descriptor. An alternate descriptor to consider for this level may be ‘Needs improvement in grade level standards,’ as ‘needs improvement’ tested well in Option 2 and the addition of ‘grade level standards’ makes it consistent with the other descriptors in this set.

Students also preferred the following set of descriptors:

Level 4: Exceeding grade level standards
Level 3: Meeting grade level standards
Level 2: Partially meeting grade level standards
Level 1: Below grade level standards

As with parents, the weakest descriptor in this set is the Level 2 descriptor. An alternate descriptor to consider for this level may be ‘Needs improvement in grade level standards,’ as ‘needs improvement’ tested well in Option 2 among students and the addition of ‘grade level standards’ makes it consistent with the other descriptors in this set.

Characteristics of the Parent Sample

The sample consists of 615 parents of children 18 years old or younger, with 271 men and 344 women. When asked what grade(s) in school their child or children are in, the parents responded with the following distribution (*note: percentages total more than 100%, accounting for multiple children in multiple grades*):

Grade	%	Grade	%	Grade	%
First	1%	Fifth	12%	Ninth	14%
Second	4%	Sixth	17%	Tenth	12%
Third	13%	Seventh	12%	Eleventh	14%
Fourth	16%	Eighth	15%	Twelfth	10%

Other demographic characteristics of parents are outlined as follows:

- Age of parents:
 - 355 (58%) ages 18-44
 - 260 (42%) ages 45+
- Education of parents:
 - 316 (51%) with no college degree
 - 299 (49%) with college degree
- Parent employment status:
 - 372 (60%) employed full-time
 - 75 (12%) employed part-time
 - 157 (26%) not employed
 - 11 (2%) retired
- Race of parents:
 - 522 (85%) Caucasian
 - 93 (15%) other race
- Household income (HHI):
 - 198 (32%) HHI under \$50,000
 - 417 (68%) HHI \$50,000+
- Marital status of parents:
 - 463 (75%) married
 - 152 (25%) single/divorced/widowed
- Geographic region:
 - 106 (17%) Northeast
 - 131 (21%) Midwest
 - 232 (38%) South
 - 146 (24%) West

Highlights of Parent Responses

When asked to select the single most effective set of descriptors from among the four options, parents selected the following by a substantial margin over the other three: Level 4: Exceeding grade level standards, Level 3: Meeting grade level standards, Level 2: Partially meeting grade level standards, Level 1: Below grade level standards. The responses to the question are summarized in the table below.

Of the four options you have reviewed, please indicate what you think is the single best option for describing student performance levels on the new assessment.

Option	%	Option	%
Option 1 Level 4: Advanced Level 3: Proficient Level 2: Basic Level 1: Minimal	12%	Option 2 Level 4: Advanced Level 3: On track Level 2: Needs improvement Level 1: At risk	25%
Option 3 Level 4: Exceeding expectations Level 3: Meeting expectations Level 2: Partially meeting expectations Level 1: Not meeting expectations	18%	Option 4 Level 4: Exceeding grade level standards Level 3: Meeting grade level standards Level 2: Partially meeting grade level standards Level 1: Below grade level standards	41%

In addition, 4% of parents indicated none of these represented the best option for describing the achievement levels.

Parents also provided feedback on specific aspects of the four sets of descriptors. The highlights are summarized as follows, and additional details are included in the annotated questionnaire below.

- **Option 1: Advanced, Proficient, Basic, Minimal**
 - Overall, only about half of the parents (55%) agreed this set provides a clear indication of a student’s performance on the assessment.
 - Men were more likely to agree than women on this point.
 - Parents agreed ‘advanced’ and ‘proficient’ are effective descriptors (73% and 61% agreed respectively), but found ‘minimal’ less effective (49% agreed).
 - Women were more likely than men to object to the use of ‘minimal’ – 33% of women somewhat or strongly *disagreed* that it is a good way to describe the lowest level of student performance, compared with 24% of men.
 - Similarly, only about half of parents (53%) agreed this set of descriptors “tell me if my child needs extra help.”
 - Women (30%) were more likely than men (18%) to *disagree* with this statement.
 - Overall, this option was least preferred among parents.

- **Option 2: Advanced, On Track, Needs Improvement, At Risk**
 - A majority of the parents (65%) agreed this set provides a clear indication of a student’s performance on the assessment.
 - A majority of parents agreed ‘on track’, ‘needs improvement’ and ‘at risk’ are effective descriptors (73%, 71% and 67% agreed respectively).
 - Nearly three-quarters (73%) agreed this set tells them if their child “needs extra help.”
 - Overall, this option was the second choice among parents.

- **Option 3: Exceeding Expectations, Meeting Expectations, Partially Meeting Expectations, Not Meeting Expectations**
 - Overall, slightly more than half of the parents (57%) agreed this set provides a clear indication of a student’s performance on the assessment.

- While a majority agreed ‘meeting expectations’ and ‘not meeting expectations’ are effective descriptors (63% and 62% agreed respectively), only half (55%) agreed ‘partially meeting expectations’ is effective.
- Overall, this option was the third choice among parents.
- ***Option 4: Exceeding Grade Level Standards, Meeting Grade Level Standards, Partially Meeting Grade Level Standards, Below Grade Level Standards***
 - A majority of the parents (65%) agreed this set provides a clear indication of a student’s performance on the assessment.
 - Three-quarters (75%) agreed ‘exceeding grade level standards’ is “a good way to describe the top level of student performance.”
 - A majority agreed ‘meeting grade level standards’ and ‘below grade level standards’ are effective descriptors (66% and 67% agreed respectively).
 - A majority agreed this set tells them if their child has performed well (67%) and if their child needs extra help (66%).
 - Women were more likely than men to agree with the effectiveness of the terms in this set of descriptors.
 - Overall, this option was the first choice among parents.

Annotated Questionnaire: Parents of Children in Grades 3-12

This spring, some states in the U.S. will start using new school assessments to replace their current tests. In some of those states, these new assessments may be used for student graduation requirements or teacher evaluation. The score reports for the new assessments will show four levels of student performance. For example, some current assessments describe the four levels as “Below Basic,” “Basic,” “Proficient” and “Advanced.” We are asking parents for their opinions about the most effective descriptors for each of the four performance levels. Please review the following sets of descriptors and answer questions for each set. In all sets, the Level 3 descriptor indicates a student who is on track to be ready for college or career after high school.

Option 1

Level 4: Advanced

Level 3: Proficient

Level 2: Basic

Level 1: Minimal

To what extent do you agree or disagree with the following statements about this set of descriptors?

If I saw these descriptors on a results report, I would have a clear indication of how my child did on the assessment.

- 55% somewhat or strongly agree
 - Men (61%) are more likely than women (51%) to agree
- 21% neutral
- 24% somewhat or strongly disagree

These descriptors tell me if my child performed well.

- 58% somewhat or strongly agree
- 23% neutral

- 19% somewhat or strongly disagree
 - Women (23%) are more likely than men (13%) to disagree

These descriptors tell me if my child needs extra help.

- 53% somewhat or strongly agree
 - Men (59%) are more likely than women (49%) to agree
- 22% neutral
- 25% somewhat or strongly disagree
 - Women (30%) are more likely than men (18%) to disagree

“Proficient” is a good way to describe a student who has demonstrated that she/he is on track for college- and career-readiness.

- 61% somewhat or strongly agree
- 22% neutral
- 17% somewhat or strongly disagree

“Advanced” is a good way to describe the top level of student performance on the assessment.

- 73% somewhat or strongly agree
- 15% neutral
- 12% somewhat or strongly disagree

“Minimal” is a good way to describe the lowest level of student performance on the assessment.

- 49% somewhat or strongly agree
- 22% neutral
- 29% somewhat or strongly disagree
 - Women (33%) are more likely than men (24%) to disagree

Option 2

Level 4: Advanced

Level 3: On track

Level 2: Needs improvement

Level 1: At risk

To what extent do you agree or disagree with the following statements about this set of descriptors?

If I saw these descriptors on a results report, I would have a clear indication of how my child did on the assessment.

- 65% somewhat or strongly agree
 - No college degree (70%) are more likely than college degree (60%) to agree
- 20% neutral
 - Men (24%) are more likely than women (17%) to be neutral
 - College degree (24%) are more likely than no college degree (16%) to be neutral
- 15% somewhat or strongly disagree

These descriptors tell me if my child performed well.

- 67% somewhat or strongly agree
- 21% neutral

- 12% somewhat or strongly disagree
 - No college degree (16%) are more likely than college degree (9%) to disagree

These descriptors tell me if my child needs extra help.

- 71% somewhat or strongly agree
 - Women (75%) are more likely than men (67%) to agree
- 17% neutral
- 11% somewhat or strongly disagree

“On track” is a good way to describe a student who has demonstrated that she/he is on track for college- and career-readiness.

- 73% somewhat or strongly agree
- 14% neutral
- 12% somewhat or strongly disagree

“Needs improvement” is a good way to communicate that a student is not yet on track for college- and career-readiness.

- 71% somewhat or strongly agree
- 16% neutral
- 13% somewhat or strongly disagree

“At risk” is a good way to describe the lowest level of student performance on the assessment.

- 67% somewhat or strongly agree
- 17% neutral
- 16% somewhat or strongly disagree

Option 3

Level 4: Exceeding expectations

Level 3: Meeting expectations

Level 2: Partially meeting expectations

Level 1: Not meeting expectations

To what extent do you agree or disagree with the following statements about this set of descriptors?

If I saw these descriptors on a results report, I would have a clear indication of how my child did on the assessment.

- 57% somewhat or strongly agree
- 23% neutral
 - Men (28%) are more likely than women (19%) to be neutral
- 20% somewhat or strongly disagree
 - Women (23%) are more likely than men (16%) to disagree

These descriptors tell me if my child performed well.

- 60% somewhat or strongly agree
- 22% neutral
- 18% somewhat or strongly disagree

These descriptors tell me if my child needs extra help.

- 63% somewhat or strongly agree
- 20% neutral
- 17% somewhat or strongly disagree

“Meeting expectations” is a good way to describe a student who has demonstrated that she/he is on track for college- and career-readiness.

- 63% somewhat or strongly agree
- 20% neutral
- 17% somewhat or strongly disagree

“Partially meeting expectations” is a good way to communicate that a student is not yet on track for college- and career-readiness.

- 55% somewhat or strongly agree
- 37% neutral
 - Men (43%) are more likely than women (32%) to be neutral
- 23% somewhat or strongly disagree

“Not meeting expectations” is a good way to describe the lowest level of student performance on the assessment.

- 62% somewhat or strongly agree
- 21% neutral
- 17% somewhat or strongly disagree

Option 4

Level 4: Exceeding grade level standards

Level 3: Meeting grade level standards

Level 2: Partially meeting grade level standards

Level 1: Below grade level standards

To what extent do you agree or disagree with the following statements about this set of descriptors?

If I saw these descriptors on a results report, I would have a clear indication of how my child did on the assessment.

- 65% somewhat or strongly agree
- 19% neutral
 - Men (24%) are more likely than women (15%) to be neutral
- 16% somewhat or strongly disagree

These descriptors tell me if my child performed well.

- 67% somewhat or strongly agree
- 21% neutral
 - Men (25%) are more likely than women (17%) to be neutral
- 12% somewhat or strongly disagree

These descriptors tell me if my child needs extra help.

- 66% somewhat or strongly agree

- Women (70%) are more likely than men (61%) to agree
- 20% neutral
 - Men (27%) are more likely than women (15%) to be neutral
- 14% somewhat or strongly disagree

“Meeting grade level standards” is a good way to describe a student who has demonstrated that she/he is on track for college- and career-readiness.

- 66% somewhat or strongly agree
- 18% neutral
- 16% somewhat or strongly disagree

“Exceeding grade level standards” is a good way to describe the top level of student performance on the assessment.

- 75% somewhat or strongly agree
 - Women (78%) are more likely than men (70%) to agree
- 14% neutral
- 11% somewhat or strongly disagree

“Below grade level standards” is a good way to describe the lowest level of student performance on the assessment.

- 67% somewhat or strongly agree
- 19% neutral
- 15% somewhat or strongly disagree

Of the four options you have reviewed, please indicate what you think is the single best option for describing student performance levels on the new assessment.

Option 1

Level 4: Advanced
Level 3: Proficient
Level 2: Basic
Level 1: Minimal

12% of parents selected this as the best option

Option 2

Level 4: Advanced
Level 3: On track
Level 2: Needs improvement
Level 1: At risk

25% of parents selected this as the best option

- No college degree (29%) are more likely than college degree (20%) to prefer this option

Option 3

Level 4: Exceeding expectations
Level 3: Meeting expectations
Level 2: Partially meeting expectations
Level 1: Not meeting expectations

18% of parents selected this as the best option

- College degree (22%) are more likely than no college degree (14%) to prefer this option

Option 4

Level 4: Exceeding grade level standards

Level 3: Meeting grade level standards

Level 2: Partially meeting grade level standards

Level 1: Below grade level standards

41% of parents selected this as the best option

In addition, 4% of parents indicated none of these represented the best option for describing the achievement levels.

DRAFT

Characteristics of the Student Sample

The sample consists of 505 children ages 12 to 17, with 258 boys and 247 girls. The demographic characteristics of students are outlined as follows:

- Grade:
 - 251 (50%) grades 8-9
 - 254 (50%) grades 10-12
- Race:
 - 522 (85%) Caucasian
 - 93 (15%) other race
- Household income (HHI):
 - 194 (38%) HHI under \$50,000
 - 311 (62%) HHI \$50,000+
- Geographic region:
 - 86 (17%) Northeast
 - 109 (21%) Midwest
 - 190 (38%) South
 - 120 (24%) West

Highlights of Student Responses

When asked to select the single most effective set of descriptors from among the four options, students selected the following by a substantial margin over the other three: Level 4: Exceeding grade level standards, Level 3: Meeting grade level standards, Level 2: Partially meeting grade level standards, Level 1: Below grade level standards. The responses to the question are summarized in the table below.

Of the four options you have reviewed, please indicate what you think is the single best option for describing student performance levels on the new assessment.

Option	%	Option	%
Option 1 Level 4: Advanced Level 3: Proficient Level 2: Basic Level 1: Minimal	12%	Option 2 Level 4: Advanced Level 3: On track Level 2: Needs improvement Level 1: At risk	29%
Option 3 Level 4: Exceeding expectations Level 3: Meeting expectations Level 2: Partially meeting expectations Level 1: Not meeting expectations	16%	Option 4 Level 4: Exceeding grade level standards Level 3: Meeting grade level standards Level 2: Partially meeting grade level standards Level 1: Below grade level standards	38%

In addition, 5% of students indicated none of these represented the best option for describing the achievement levels.

Students also provided feedback on specific aspects of the four sets of descriptors. The highlights are summarized as follows, and additional details are included in the annotated questionnaire below.

- **Option 1: Advanced, Proficient, Basic, Minimal**
 - A majority of students (64%) agreed this set provides a clear indication of a student's performance on the assessment.
 - Most students agreed (80%) that 'advanced' and 'proficient' are good words to use to tell them if they did well on the test, but fewer agreed (58.5%) that 'basic' and 'minimal' are good words to use to tell them if they need extra help.
 - Most students agreed (82%) 'advanced' is a good way to describe the top level of student performance.
 - A majority agreed (69%) 'proficient' is a good way to describe a student performance that is on track for college and career readiness.
 - Only slightly more than half (57%) agreed 'minimal' is a good way to describe the lowest level of student performance.
 - Overall, this option was least preferred among students.
- **Option 2: Advanced, On Track, Needs Improvement, At Risk**
 - A majority of the students (68%) agreed this set provides a clear indication of a student's performance on the assessment.
 - A majority of students agreed 'on track', 'needs improvement' and 'at risk' are effective descriptors (73%, 75% and 70% agreed respectively).
 - Overall, this option was a strong second choice among students.
- **Option 3: Exceeding Expectations, Meeting Expectations, Partially Meeting Expectations, Not Meeting Expectations**
 - A majority of students (65%) agreed this set provides a clear indication of a student's performance on the assessment.
 - While a majority agreed 'meeting expectations' and 'not meeting expectations' are effective descriptors (66% and 70% agreed respectively), fewer students (59%) agreed 'partially meeting expectations' is effective.
 - Overall, this option was the third choice among students.
- **Option 4: Exceeding Grade Level Standards, Meeting Grade Level Standards, Partially Meeting Grade Level Standards, Below Grade Level Standards**
 - Nearly three-quarters of students (71%) agreed this set provides a clear indication of a student's performance on the assessment.
 - Most students (82%) agreed 'exceeding grade level standards' is "a good way to describe the top level of student performance."
 - A majority agreed 'meeting grade level standards' and 'below grade level standards' are effective descriptors (67% and 71% agreed respectively).
 - Overall, this option was the first choice among students.

Annotated Questionnaire: Students in Grades 8-12

This spring, some states in the U.S. will start using new statewide English and math tests to replace their current tests. In some of those states, these new tests may be used for student graduation requirements or teacher evaluation. The score reports for the new tests will show four levels of student performance. For example, some current tests describe the four levels as "Below Basic," "Basic," "Proficient" and "Advanced." We are asking students for their opinions about the best way to describe each of the four performance levels. Please review the following sets of descriptions and answer questions for each set. In all sets, the Level 3 description indicates a student who is on track to be ready for college or a career after high school.

Option 1

Level 4: Advanced

Level 3: Proficient

Level 2: Basic

Level 1: Minimal

How much do you agree or disagree with the following statements about this set of descriptions?

If I saw these descriptions on a results report, I would have a clear understanding of how I did on the test.

- 64% somewhat or strongly agree
- 21% neutral
- 15% somewhat or strongly disagree

“Advanced” and “proficient” are good words to use to tell me if I did well on the test.

- 80% somewhat or strongly agree
- 13% neutral
- 7% somewhat or strongly disagree

“Basic” and “minimal” are good words to use to tell me if I need extra help.

- 58.5% somewhat or strongly agree
- 22% neutral
- 19.5% somewhat or strongly disagree

“Proficient” is a good way to describe a student who has demonstrated that she/he is on track for college- and career-readiness.

- 69% somewhat or strongly agree
- 23% neutral
- 8% somewhat or strongly disagree

“Advanced” is a good way to describe the top level of student performance on the assessment.

- 82% somewhat or strongly agree
- 13% neutral
- 5% somewhat or strongly disagree

“Minimal” is a good way to describe the lowest level of student performance on the assessment.

- 57% somewhat or strongly agree
- 22% neutral
- 21% somewhat or strongly disagree

Option 2

Level 4: Advanced

Level 3: On track

Level 2: Needs improvement

Level 1: At risk

To what extent do you agree or disagree with the following statements about this set of descriptors?

If I saw these descriptions on a results report, I would have a clear understanding of how I did on the test.

- 68% somewhat or strongly agree
- 22% neutral
- 10% somewhat or strongly disagree

“Advanced” is a good word to use to tell me if I did well on the test.

- 81% somewhat or strongly agree
- 14% neutral
- 5% somewhat or strongly disagree

“On track” is a good way to describe a student who has demonstrated that she/he is on track for college- and career-readiness.

- 73% somewhat or strongly agree
- 20% neutral
- 7% somewhat or strongly disagree

“Needs improvement” is a good way to communicate that a student is not yet on track for college- and career-readiness.

- 75% somewhat or strongly agree
- 19% neutral
- 6% somewhat or strongly disagree

“At risk” is a good way to describe the lowest level of student performance on the assessment.

- 70% somewhat or strongly agree
- 19% neutral
- 11% somewhat or strongly disagree

Option 3

Level 4: Exceeding expectations

Level 3: Meeting expectations

Level 2: Partially meeting expectations

Level 1: Not meeting expectations

To what extent do you agree or disagree with the following statements about this set of descriptors?

If I saw these descriptions on a results report, I would have a clear understanding of how I did on the test.

- 65% somewhat or strongly agree
- 22% neutral
- 13% somewhat or strongly disagree

“Meeting expectations” is a good way to describe a student who has demonstrated that she/he is on track for college- and career-readiness.

- 66% somewhat or strongly agree
- 24% neutral
- 10% somewhat or strongly disagree

“Partially meeting expectations” is a good way to communicate that a student is not yet on track for college- and career-readiness.

- 59% somewhat or strongly agree
- 26% neutral
- 15% somewhat or strongly disagree

“Not meeting expectations” is a good way to describe the lowest level of student performance on the assessment.

- 70% somewhat or strongly agree
- 19% neutral
- 11% somewhat or strongly disagree

Option 4

Level 4: Exceeding grade level standards

Level 3: Meeting grade level standards

Level 2: Partially meeting grade level standards

Level 1: Below grade level standards

To what extent do you agree or disagree with the following statements about this set of descriptors?

If I saw these descriptions on a results report, I would have a clear understanding of how I did on the test.

- 71% somewhat or strongly agree
- 18% neutral
- 11% somewhat or strongly disagree

“Meeting grade level standards” is a good way to describe a student who has demonstrated that she/he is on track for college- and career-readiness.

- 67% somewhat or strongly agree
- 23% neutral
- 10% somewhat or strongly disagree

“Exceeding grade level standards” is a good way to describe the top level of student performance on the assessment.

- 82% somewhat or strongly agree
- 13% neutral
- 5% somewhat or strongly disagree

“Below grade level standards” is a good way to describe the lowest level of student performance on the assessment.

- 71% somewhat or strongly agree
- 17% neutral

- 12% somewhat or strongly disagree

Of the four options you have reviewed, please indicate what you think is the single best option for describing how students did on the test.

Option 1

Level 4: Advanced
Level 3: Proficient
Level 2: Basic
Level 1: Minimal

12% of students selected this as the best option

Option 2

Level 4: Advanced
Level 3: On track
Level 2: Needs improvement
Level 1: At risk

29% of students selected this as the best option

Option 3

Level 4: Exceeding expectations
Level 3: Meeting expectations
Level 2: Partially meeting expectations
Level 1: Not meeting expectations

16% of students selected this as the best option

Option 4

Level 4: Exceeding grade level standards
Level 3: Meeting grade level standards
Level 2: Partially meeting grade level standards
Level 1: Below grade level standards

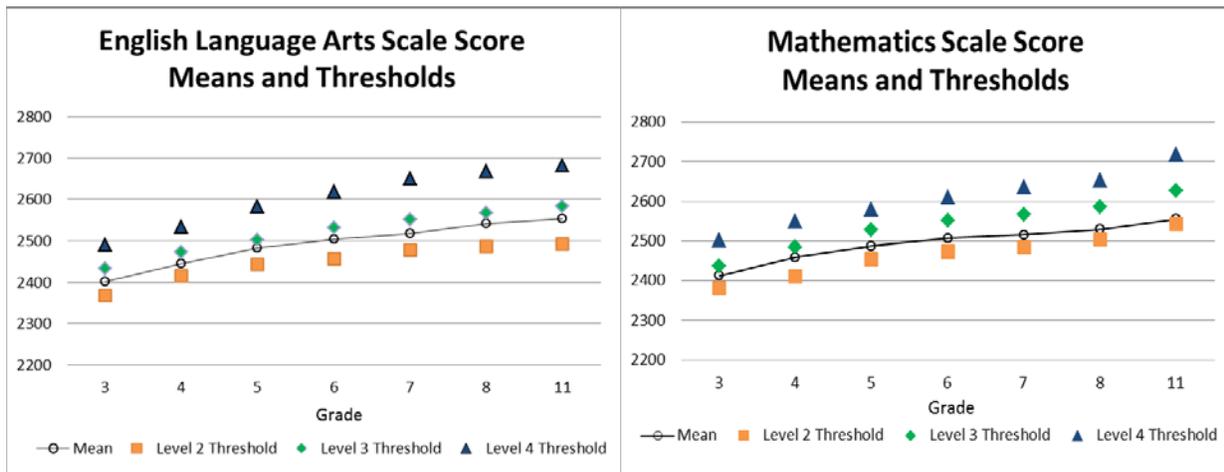
38% of students selected this as the best option

In addition, 5% of students said none of the above options are effective descriptors.

Smarter Balanced Scale Scores

Smarter Balanced scale scores form genuine growth scales that can be used to follow student performance from grade 3 through grade 11. The scales were developed to span grades through a procedure that permits direct comparison of scores in one grade to those in adjacent grades.

The range of the scale scores, 2000 to 3000, was chosen specifically to allow growth to be shown in whole numbers and to distinguish Smarter Balanced scores from other assessments such as those of the National Assessment of Educational Progress (NAEP) or the College Board (SAT). The figures below show the mean scale scores for each grade, as well as the threshold scores for Levels 2, 3 and 4.



When teachers, parents, higher education faculty, and community leaders met to recommend threshold scores for the achievement levels, they focused on the actual test questions students would need to answer correctly to be performing at Level 2, 3, or 4. Their content recommendations were translated into scale scores through a process known as the Bookmark procedure. This process assigns a scale value to each test question as well as to each student so that a student who has the same scale score as a particular question will have a 50 percent chance of answering that question correctly. For example, by identifying the hardest question a student just entering Level 3 would have a 50 percent chance of answering correctly, achievement level setting panels were assigning the scale score of that question to that student. Since that student had been defined as just entering Level 3, that scale score became the threshold score for Level 3.

A New Baseline: Results from Achievement Level Setting

Joe Willhoft, Ph.D., Smarter Balanced



Media Briefing
November 6, 2014

Agenda

- Achievement Level Setting Process
- Overview of New Achievement Levels
- Q&A

Achievement Level Setting Process



Goals

- Engage a large group of educators and practitioners from Smarter Balanced states
- Collect input from interested community members
- Develop achievement level score recommendations for interim comprehensive and summative assessments
- Align achievement levels appropriately across grades
- Recommend scores to states for final adoption

Achievement Level Setting Timeline

April 30	<ul style="list-style-type: none">• States approved achievement level setting plan
October 6-18	<ul style="list-style-type: none">• Online Panel for thousands of educators and other constituents
October 13-19	<ul style="list-style-type: none">• In-Person Panel (500+ educators, others)
October 20	<ul style="list-style-type: none">• Cross-Grade Review: Subgroup of In-Person
October 30	<ul style="list-style-type: none">• Technical Advisory Committee review
November 6	<ul style="list-style-type: none">• State chiefs vote to approve achievement level recommendations
Fall/Winter	<ul style="list-style-type: none">• States adopt achievement levels
Spring 2015	<ul style="list-style-type: none">• First administration of new end-of-year assessments

Who participated?

Educators, higher education faculty, parents, community leaders



Astrid
8th Grade Mathematics Teacher
U.S. Virgin Islands



Eric
Mathematics Faculty Member
Hawaii



Kim
PTA Leader
Wisconsin



Kimberly
Middle School ELA Teacher
Idaho



David
High School Mathematics Teacher
Washington



Erika
ELL Specialist
Maine



Princess
High School Mathematics Teacher
California



Richard
Mathematics Teacher
Connecticut

What are participants saying?

I was very excited when I found out that I was going to be able to participate in the Smarter Balanced process. I was excited to hear that additional stakeholders besides just teachers and administrators were involved.

– Parent, Wisconsin

I think it's extremely important that the voice of the educator is involved in the process. I'm thrilled that we were allowed the chance to do that and I think that it provides a lot of buy-in from teachers and that they feel confident--more confident about the scores.

– Gifted Student Specialist, Maine

I am very grateful for the opportunity I was given to represent our state at this event. I was surprised and impressed by the thoughtfulness that has gone into this Consortium and the detailed methodology that has brought it through test design, rubric descriptors, and field test.

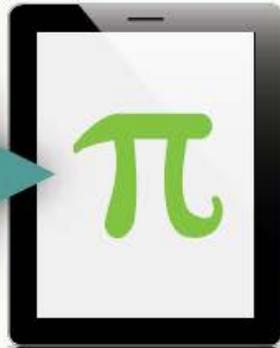
– Higher Ed Faculty, Michigan

Step 1: Online Panel



Participants:

- Teachers and administrators
- Higher education faculty
- Business and community leaders
- Parents



Description:

Thousands of teachers and other interested parties will review test questions online and recommend the level of performance required for students to be considered on-track to college and career readiness.

Purpose:

The online panel allows educators, parents, and other concerned citizens to have unprecedented input on achievement levels.

Step 2: In-Person Panel

Description:

States have nominated educators, higher education faculty, and community members to participate in an in-person panel to deliberate and make recommendations on all four achievement levels. A diverse cross-section of teachers will participate, representing grades 3-8 and high school in English language arts and math. Educators with experience teaching students with disabilities and English language learners also will participate, helping ensure the levels accurately reflect achievement for all students.



Participants:

- Teachers and administrators
- Higher education faculty
- Business and community leaders
- Parents

Purpose:

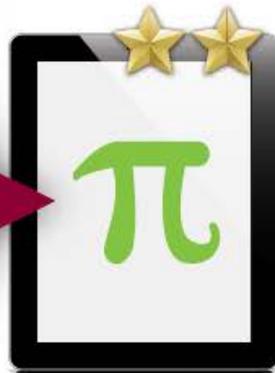
The in-person panel allows teams of educators and other stakeholders nominated by states to deliberate and recommend threshold scores for all four achievement levels for each grade and content area.

Step 3: Cross-Grade Review



Participants:

- A subset of representatives from the in-person panel



Description:

A cross-grade review committee (a subset of primarily educators from the in-person panel) will meet to review the suggested achievement levels and recommend any changes needed to ensure that the levels are aligned appropriately from grade to grade. In this way, teachers and parents can understand how students have progressed from year to year.

Purpose:

This step ensures that the achievement levels are appropriately aligned across grades and accurately reflect student progress from year to year.

Step 4: State Approval

Description:

Smarter Balanced member state and territory leaders will meet to review and endorse the achievement level recommendations. Higher education leaders will participate in decisions regarding the Grade 11 achievement levels to ensure that they reflect the expectations of colleges and universities. An independent Technical Advisory Committee and expert auditor will certify that the process used to arrive at these recommendations followed the plan approved by members. The achievement levels will then be subject to existing approval processes within individual states and territories.



Participants:

- Smarter Balanced member states and territories

Purpose:

This step provides all students in Smarter Balanced member states and territories a clear, consistent measure of progress toward college and career readiness.

Participation Summary

	Online Panel	In-Person Panel	Cross-Grade Review Committee
Teachers	65%	64%	66%
Non-Teacher Educators & Administrators	18%	18%	19%
Higher Education	9%	12%	11%
General Public	7%	6%	5%
Total	2,660	482	64

Figures reflect the number of individuals who submitted a bookmark during the Online Panel and who attended the In-person Panel.

Results



Achievement Levels

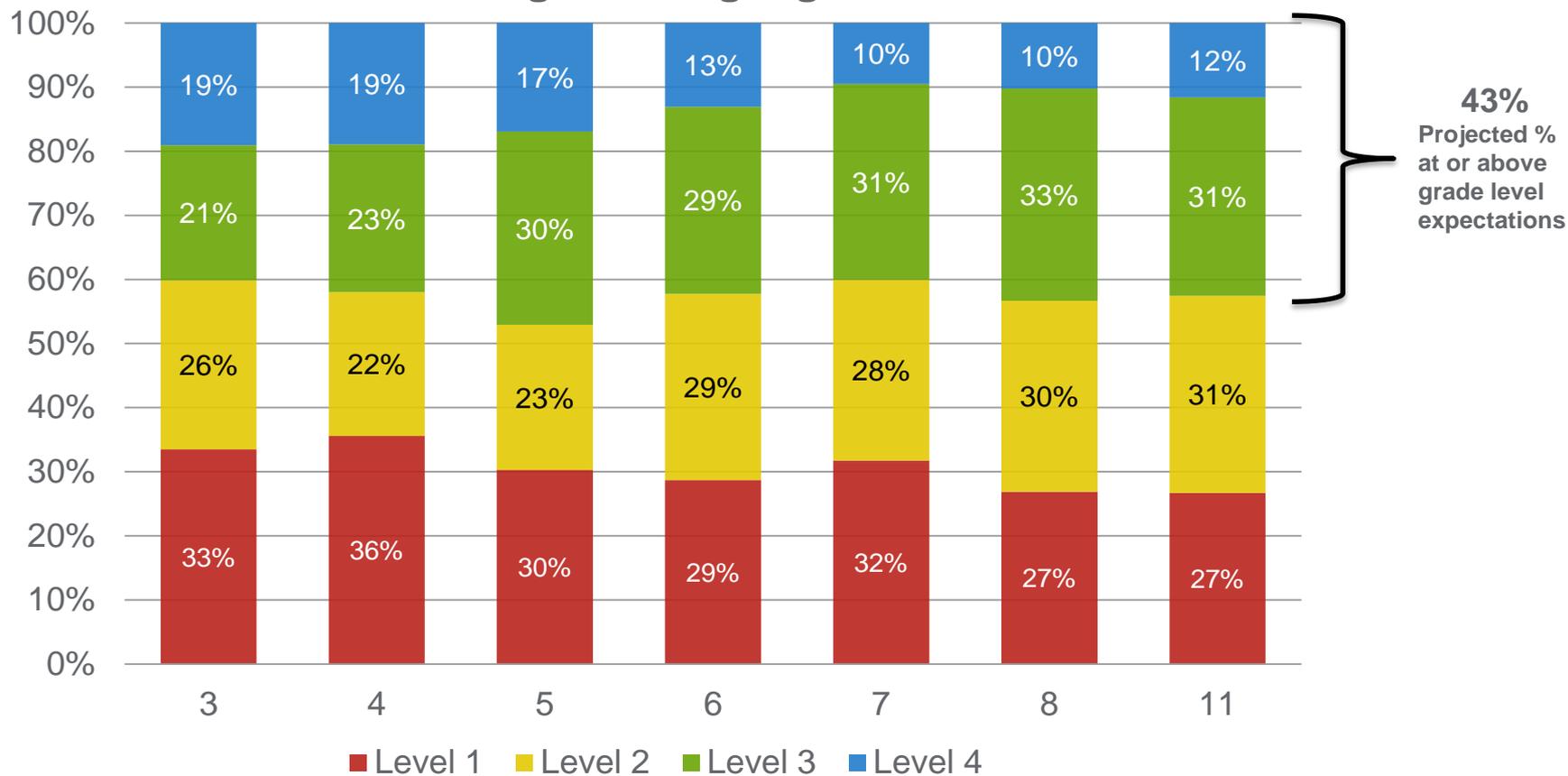
- Level 4
 - Above grade level expectations
- Level 3
 - Meets grade level expectations
- Level 2
 - Partially meets grade level expectations
- Level 1
 - Does not meet grade level expectations

← Threshold



Projected Impact of the New Levels

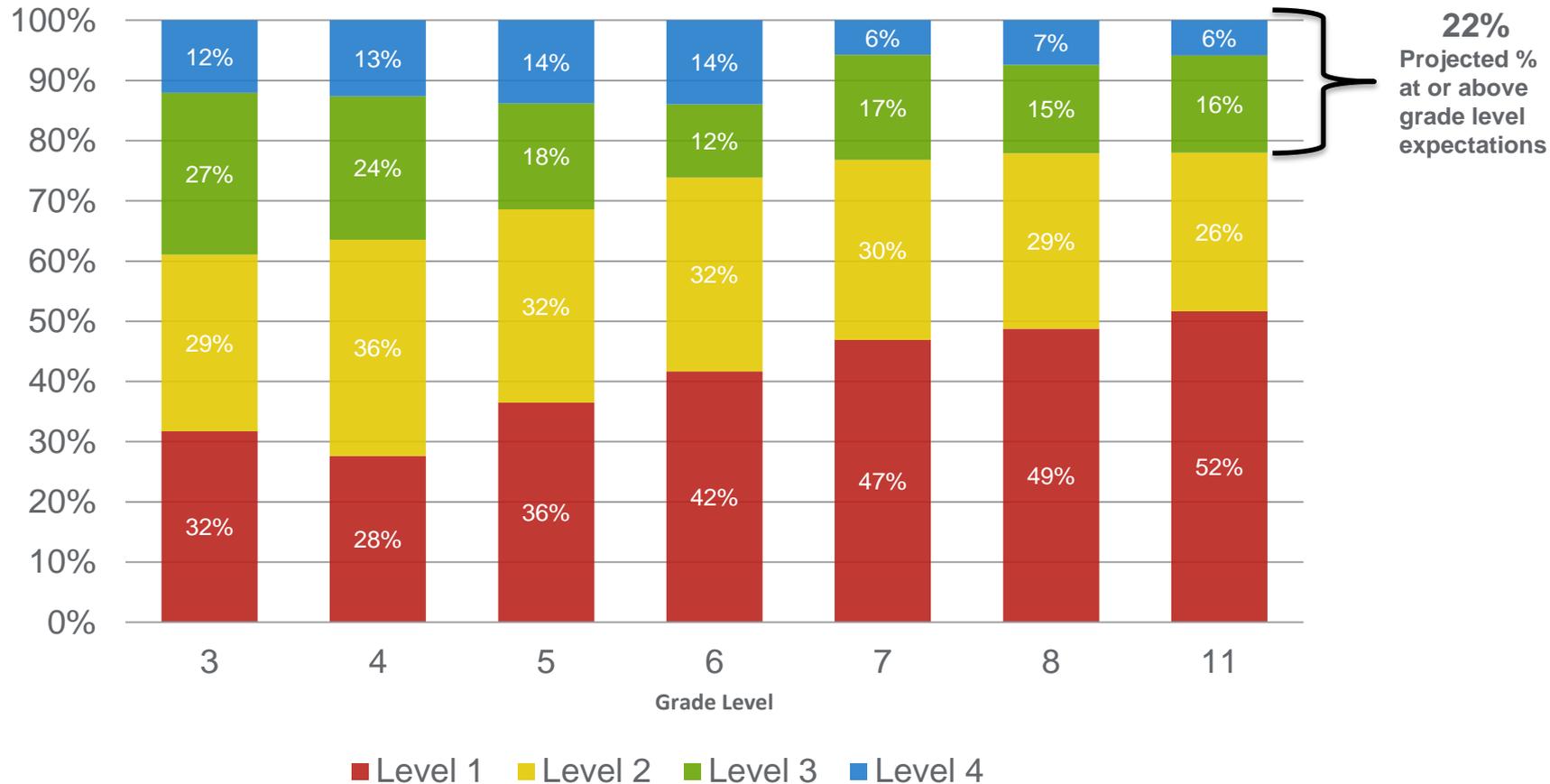
Projected Percentage of Students at Each Level:
English Language Arts



*Projections based on student performance on Smarter Balanced Field Test in 21 states and the U.S. Virgin Islands (March 25 – June 6, 2014).

Projected Impact of the New Levels

Projected Percentage of Students at Each Level: Mathematics



*Projections based on student performance on Smarter Balanced Field Test in 21 states and the U.S. Virgin Islands (March 25 – June 6, 2014).

Wrap-Up



New Assessments and a New Baseline

- **New standards:** provide clear expectations to help keep students on track for college and/or career
- **New assessments to replace existing tests:** challenging and engaging for students, provide teachers and school leaders with better information
- **New baseline:** achievement standards developed by educators reflect high expectations for students

Final Thought

“These recommendations are crucial to aligning exactly what we are supposed to be teaching with what they are actually testing to get really good quality information on the achievement of our students.”

-Washington State teacher



Questions?



A New Baseline: Results from Achievement Level Setting

Joe Willhoft, Ph.D., Smarter Balanced



Media Briefing
November 6, 2014

Agenda

- Achievement Level Setting Process
- Overview of New Achievement Levels
- Q&A

Achievement Level Setting Process



Goals

- Engage a large group of educators and practitioners from Smarter Balanced states
- Collect input from interested community members
- Develop achievement level score recommendations for interim comprehensive and summative assessments
- Align achievement levels appropriately across grades
- Recommend scores to states for final adoption

Achievement Level Setting Timeline

April 30

- States approved achievement level setting plan

October 6-18

- Online Panel for thousands of educators and other constituents

October 13-19

- In-Person Panel (500+ educators, others)

October 20

- Cross-Grade Review: Subgroup of In-Person

October 30

- Technical Advisory Committee review

November 6

- State chiefs vote to approve achievement level recommendations

Fall/Winter

- States adopt achievement levels

Spring 2015

- First administration of new end-of-year assessments

Who participated?

Educators, higher education faculty, parents, community leaders



Astrid
8th Grade Mathematics Teacher
U.S. Virgin Islands



Eric
Mathematics Faculty Member
Hawaii



Kim
PTA Leader
Wisconsin



Kimberly
Middle School ELA Teacher
Idaho



David
High School Mathematics Teacher
Washington



Erika
ELL Specialist
Maine



Princess
High School Mathematics Teacher
California



Richard
Mathematics Teacher
Connecticut

What are participants saying?

I was very excited when I found out that I was going to be able to participate in the Smarter Balanced process. I was excited to hear that additional stakeholders besides just teachers and administrators were involved.

– Parent, Wisconsin

I think it's extremely important that the voice of the educator is involved in the process. I'm thrilled that we were allowed the chance to do that and I think that it provides a lot of buy-in from teachers and that they feel confident--more confident about the scores.

– Gifted Student Specialist, Maine

I am very grateful for the opportunity I was given to represent our state at this event. I was surprised and impressed by the thoughtfulness that has gone into this Consortium and the detailed methodology that has brought it through test design, rubric descriptors, and field test.

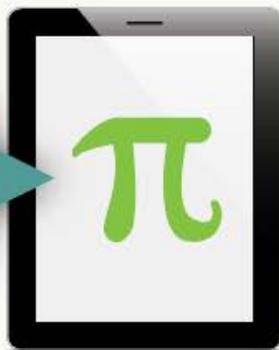
– Higher Ed Faculty, Michigan

Step 1: Online Panel



Participants:

- Teachers and administrators
- Higher education faculty
- Business and community leaders
- Parents



Description:

Thousands of teachers and other interested parties will review test questions online and recommend the level of performance required for students to be considered on-track to college and career readiness.

Purpose:

The online panel allows educators, parents, and other concerned citizens to have unprecedented input on achievement levels.

Step 2: In-Person Panel

Description:

States have nominated educators, higher education faculty, and community members to participate in an in-person panel to deliberate and make recommendations on all four achievement levels. A diverse cross-section of teachers will participate, representing grades 3-8 and high school in English language arts and math. Educators with experience teaching students with disabilities and English language learners also will participate, helping ensure the levels accurately reflect achievement for all students.



Participants:

- Teachers and administrators
- Higher education faculty
- Business and community leaders
- Parents

Purpose:

The in-person panel allows teams of educators and other stakeholders nominated by states to deliberate and recommend threshold scores for all four achievement levels for each grade and content area.

Step 3: Cross-Grade Review



Participants:

- A subset of representatives from the in-person panel



Description:

A cross-grade review committee (a subset of primarily educators from the in-person panel) will meet to review the suggested achievement levels and recommend any changes needed to ensure that the levels are aligned appropriately from grade to grade. In this way, teachers and parents can understand how students have progressed from year to year.

Purpose:

This step ensures that the achievement levels are appropriately aligned across grades and accurately reflect student progress from year to year.

Step 4: State Approval

Description:

Smarter Balanced member state and territory leaders will meet to review and endorse the achievement level recommendations. Higher education leaders will participate in decisions regarding the Grade 11 achievement levels to ensure that they reflect the expectations of colleges and universities. An independent Technical Advisory Committee and expert auditor will certify that the process used to arrive at these recommendations followed the plan approved by members. The achievement levels will then be subject to existing approval processes within individual states and territories.



Participants:

- Smarter Balanced member states and territories

Purpose:

This step provides all students in Smarter Balanced member states and territories a clear, consistent measure of progress toward college and career readiness.

Participation Summary

	Online Panel	In-Person Panel	Cross-Grade Review Committee
Teachers	65%	64%	66%
Non-Teacher Educators & Administrators	18%	18%	19%
Higher Education	9%	12%	11%
General Public	7%	6%	5%
Total	2,660	482	64

Figures reflect the number of individuals who submitted a bookmark during the Online Panel and who attended the In-person Panel.

Results



Achievement Levels

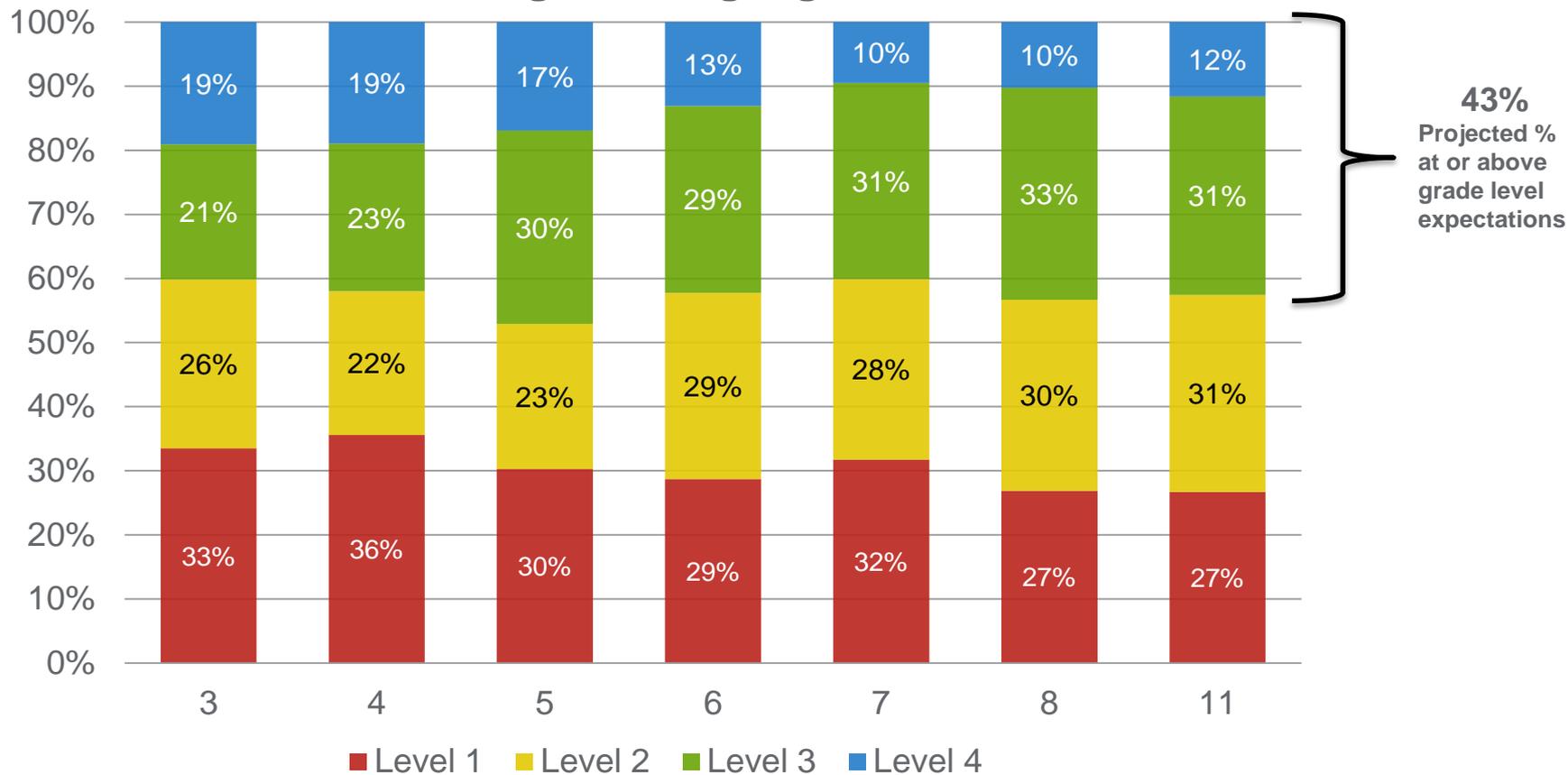
- Level 4
 - Above grade level expectations
- Level 3
 - Meets grade level expectations
- Level 2
 - Partially meets grade level expectations
- Level 1
 - Does not meet grade level expectations

← Threshold



Projected Impact of the New Levels

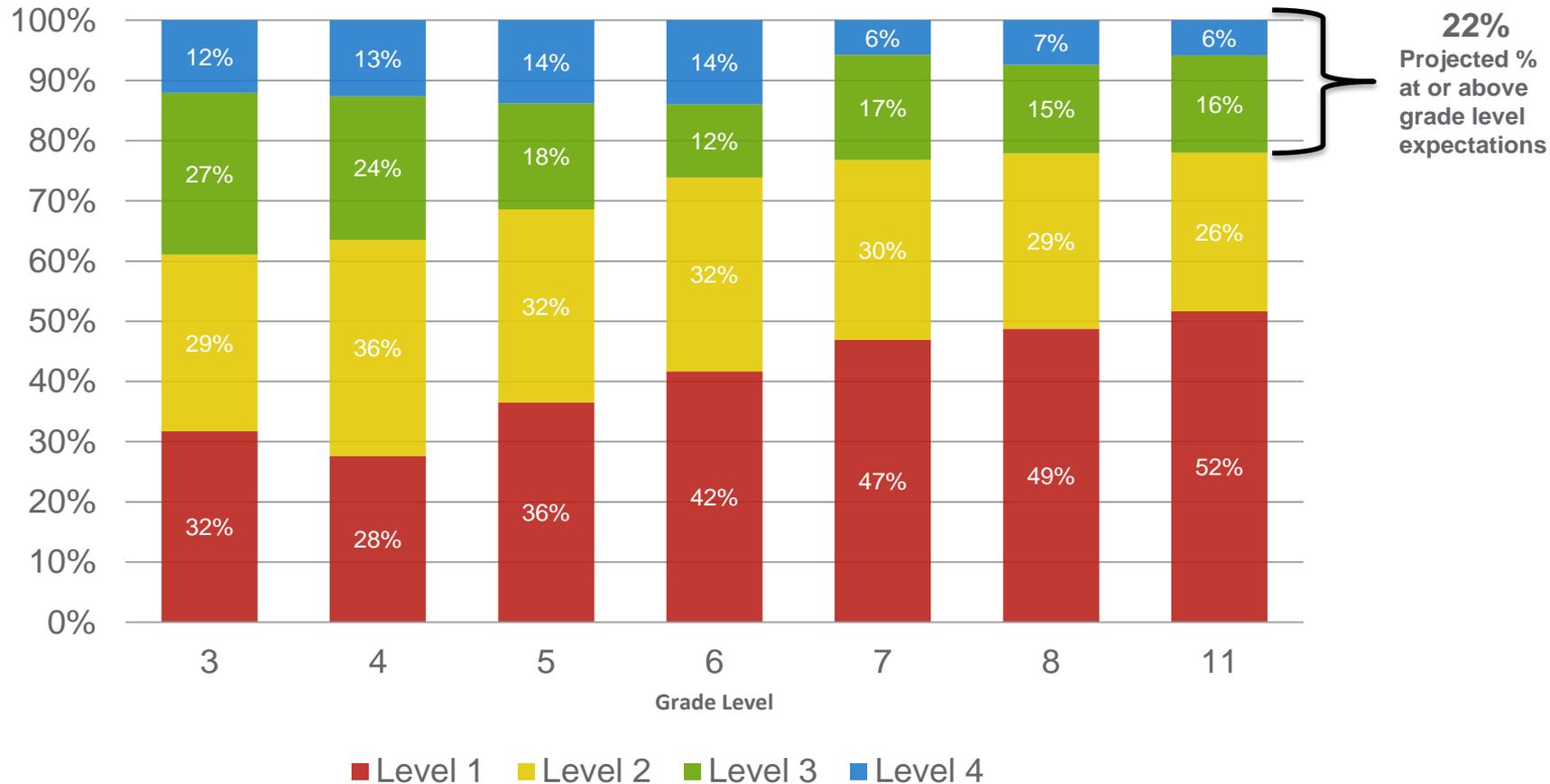
Projected Percentage of Students at Each Level:
English Language Arts



*Projections based on student performance on Smarter Balanced Field Test in 21 states and the U.S. Virgin Islands (March 25 – June 6, 2014).

Projected Impact of the New Levels

Projected Percentage of Students at Each Level: Mathematics



*Projections based on student performance on Smarter Balanced Field Test in 21 states and the U.S. Virgin Islands (March 25 – June 6, 2014).

Wrap-Up



New Assessments and a New Baseline

- **New standards:** provide clear expectations to help keep students on track for college and/or career
- **New assessments to replace existing tests:** challenging and engaging for students, provide teachers and school leaders with better information
- **New baseline:** achievement standards developed by educators reflect high expectations for students

Final Thought

“These recommendations are crucial to aligning exactly what we are supposed to be teaching with what they are actually testing to get really good quality information on the achievement of our students.”

-Washington State teacher



Questions?



Raising the Bar: Achievement Levels on the Smarter Balanced Assessments

Presenter



[Presentation Name]
[Date]

Agenda

- Smarter Balanced assessments update
- Overview of the achievement level setting process
 - Teacher and constituent involvement
- Results of achievement level setting
- New baseline for student achievement
- Questions



Changes This School Year

- Smarter Balanced replaces existing state assessments
- Students take new computer-adaptive assessment measuring progress toward Common Core
- Universal access for all students, including students with disabilities and English language learners
- New technology and performance tasks measure critical thinking and problem-solving
- Higher bar for grade-level performance

Rollout of Full Assessment System in 2014-15

Online assessments
that measure progress
toward readiness for
college and careers



Digital Library

Available Now

Formative assessment resources and practices that teachers can use as needed throughout the year



Interim Assessments

Available Beginning Winter 2014-15

Optional assessments that allow educators to check student progress and provide information to inform instruction during the year



Summative Assessments

Available Spring 2015

Year-end assessments for grades 3-8 and 11 with a computer adaptive test and performance tasks in math and English

**Smarter Balanced gives educators information
and tools to improve teaching and learning**



Setting a New Baseline

Adjusting our definition for grade-level performance

New standards: provide clear expectations to help keep students on track for college and/or career

New baseline: achievement levels developed by educators reflect high expectations for students



New assessments: challenging and engaging for students, provide teachers and school leaders with better information

Achievement Level Setting



Major Points

- A large group of educators and practitioners from Smarter Balanced states participated
- Interested community members provided input
- Panelists developed achievement level score recommendations for interim comprehensive and summative assessments
- Achievement level scores align appropriately across grades
- Levels are recommendations to states, who consider for final adoption

Achievement Level Setting Timeline

April 30

- States approved achievement level setting plan

October 6-18

- Online Panel for thousands of educators and other constituents

October 13-19

- In-Person Panel (500+ educators, others)

October 20

- Cross-Grade Review: Subgroup of In-Person

October 30

- Technical Advisory Committee review

November 6

- State chiefs vote to approve achievement level recommendations

Fall/Winter

- States adopt achievement levels

Spring 2015

- First administration of new end-of-year assessments

How This Process Works



Overview



Educators and practitioners recommended achievement level scores for mid-year and end-of-year assessments.



Interested community members and parents provided input.



Achievement levels and scores are aligned appropriately across grades.



States approved final scores based on an existing process.



Who participated?

Educators, higher education faculty, parents, community leaders



Astrid
8th Grade Mathematics Teacher
U.S. Virgin Islands



Eric
Mathematics Faculty Member
Hawaii



Kim
PTA Leader
Wisconsin



Kimberly
Middle School ELA Teacher
Idaho



David
High School Mathematics Teacher
Washington



Erika
ELL Specialist
Maine



Princess
High School Mathematics Teacher
California



Richard
Mathematics Teacher
Connecticut

Participation Summary

	Online Panel	In-Person Panel	Cross-Grade Review Committee	Representatives From [INSERT STATE]
Teachers	65%	64%	66%	Insert state specific information
Non-Teacher Educators & Administrators	18%	18%	19%	
Higher Education	9%	12%	11%	
General Public	7%	6%	5%	
Total	2,660	482	64	

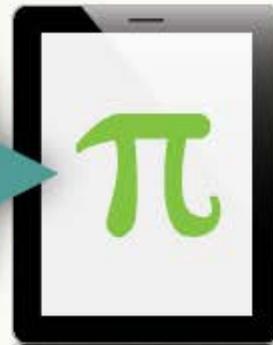
Figures reflect the number of individuals who submitted a bookmark during the Online Panel and who attended the In-person Panel.

Step 1: Online Panel



Participants:

- Teachers and administrators
- Higher education faculty
- Business and community leaders
- Parents



Description:

Thousands of teachers and other interested parties reviewed test questions online and recommended the level of performance required for students to be considered on-track to college and career readiness.

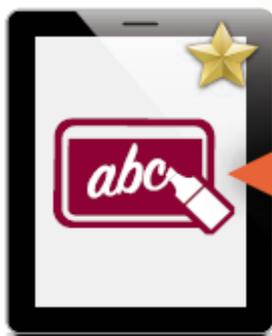
Purpose:

The online panel allowed educators, parents, and other concerned citizens to have unprecedented input on achievement levels.

Step 2: In-Person Panel

Description:

States nominated educators, higher education faculty, and community members to participate in an in-person panel to deliberate and make recommendations on all four achievement levels. A diverse cross-section of teachers participated, representing grades 3-8 and high school in English language arts and math. Educators with experience teaching students with disabilities and English language learners also participated, helping ensure the levels accurately reflect achievement for all students.



Participants:

- Teachers and administrators
- Higher education faculty
- Business and community leaders
- Parents

Purpose:

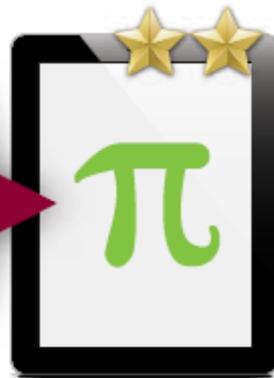
The in-person panel allowed teams of educators and other stakeholders nominated by states to deliberate and recommend threshold scores for all four achievement levels for each grade and content area.

Step 3: Cross-Grade Review



Participants:

- A subset of representatives from the in-person panel



Description:

A cross-grade review committee (a subset of primarily educators from the in-person panel) met to review the suggested achievement levels and recommended any changes needed to ensure that the levels were aligned appropriately from grade to grade. In this way, teachers and parents can understand how students have progressed from year to year.

Purpose:

This step ensured that the achievement levels were appropriately aligned across grades and accurately reflect student progress from year to year.

Step 4: State Approval

Description:

Smarter Balanced member states and territory leaders will meet to review and endorse the achievement level recommendations. Higher education leaders will participate in decisions regarding the Grade 11 achievement levels to ensure that they reflect the expectations of colleges and universities. An independent Technical Advisory Committee and expert auditor will certify that the process used to arrive at these recommendations followed the plan approved by members. The achievement levels will then be subject to existing approval processes within individual states and territories.



Participants:

- Smarter Balanced member states and territories

Purpose:

This step provides all students in Smarter Balanced member states and territories a clear, consistent measure of progress toward college and career readiness.

Step 5: Final Achievement Levels



In the Classroom:

The final product of this inclusive, collaborative process will be a set of new achievement levels on the Smarter Balanced assessments. Because Smarter Balanced members have set high expectations for students—and the new tests are designed to assess student performance against these expectations—the definition of grade level performance will be higher than it used to be in many states. Initial test results will represent a new baseline that provides a more accurate indicator for educators, students, and parents as they work to meet the rigorous demands of college and career readiness.



Major Points

- Collaborative and transparent
- Inclusive and consensus-driven process
- Scientifically rigorous design
- Using best practices
- Multiple reviews, technical certification

Results





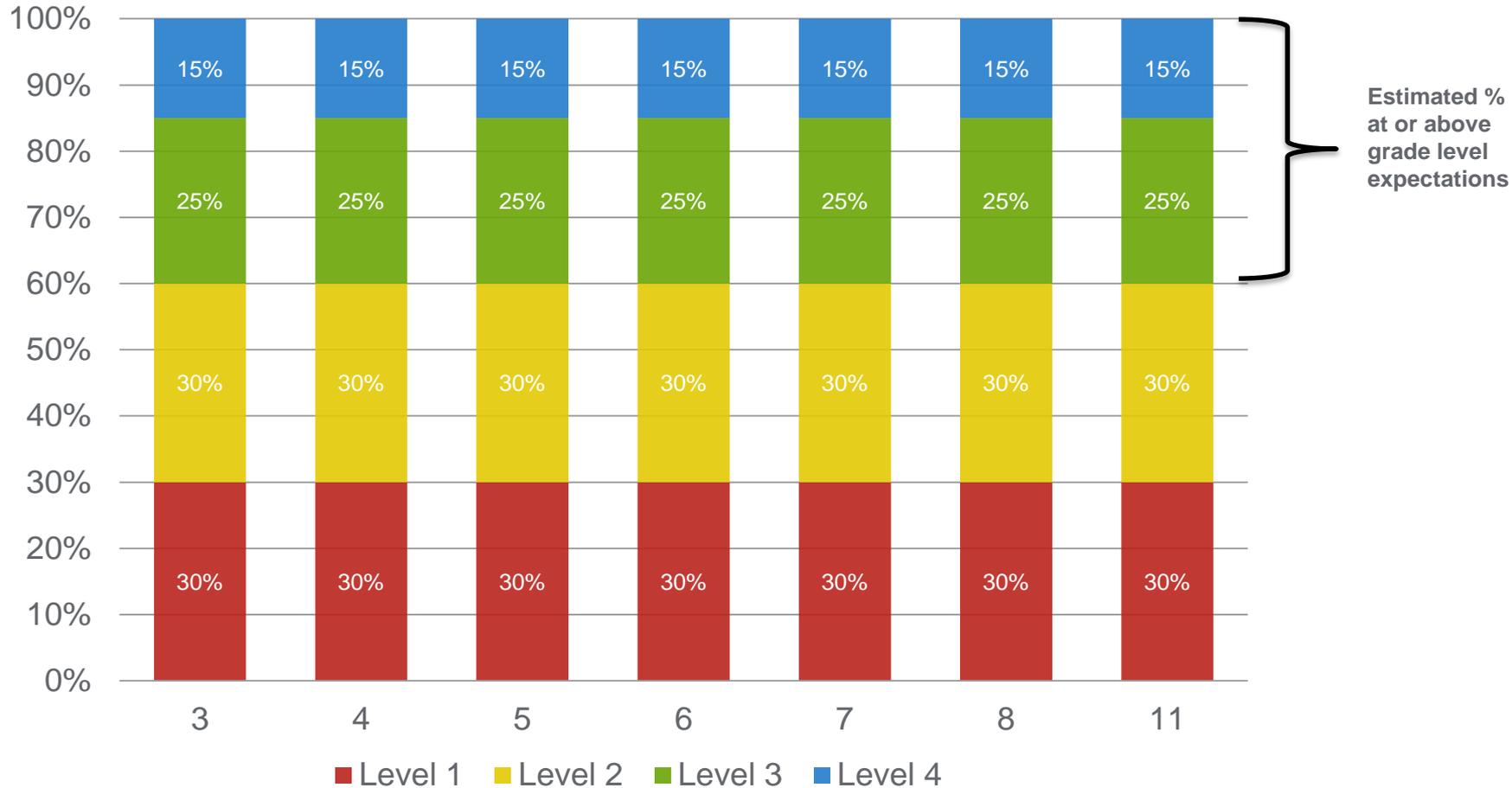
Achievement Levels

- Level 4
 - Above grade level expectations
- Level 3
 - Meets grade level expectations
- Level 2
 - Partially meets grade level expectations
- Level 1
 - Does not meet grade level expectations

← Threshold

Impact of the New Levels

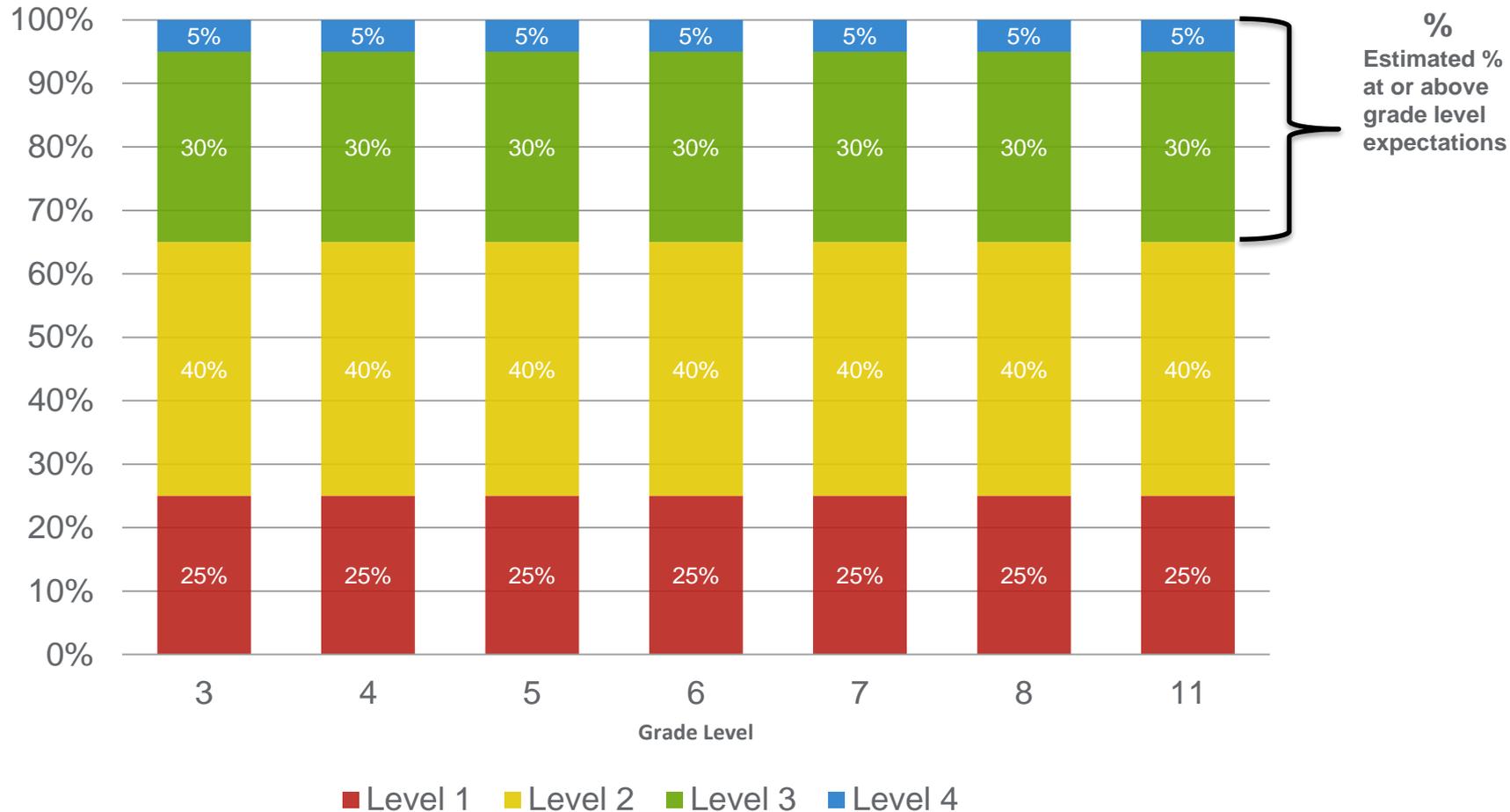
Estimated % of Students at Each Level: ELA



*Based on student performance on Smarter Balanced field test in 21 states and the U.S. Virgin Islands (March 25₃₀₅ June 6, 2014)

Impact of the New Levels

Estimated % of Students at Each Level: Math



*Based on student performance on Smarter Balanced field test in 21 states and the U.S. Virgin Islands (March 25³⁰⁶ June 6, 2014)

Score Recommendations

- Four achievement levels at each grade level and content area
- Level 3: threshold score for students at grade level
- Varies with the difficulty at each grade level test
- Adjusted in future years depending on accuracy



Wrap-Up





Key Points

- **Consistent expectations** give all students the opportunity to succeed.
- The achievement levels are based on recommendations from **teachers, parents, and community leaders** representing the Smarter Balanced member states.
- The achievement level scores will establish **a new baseline for student achievement** in line with higher standards used in the classroom.
- Moving forward, the achievement levels will give teachers and parents **reliable and accurate information** about students who are succeeding and those who need support.



Moving Forward

2014-2015

New baseline:

Achievement level scores developed by educators that reflect high expectations for students



Student success:

All students regardless of who they are on track to succeed after high school.



Student progress: Teachers, students and parents work together to address areas for improvement and make sure student needs are met.



Questions?



Appendix B – Logistics Documentation

Logistics Plan

- B01_Logistics Plan 313

Kickoff Meeting

- B02_Kickoff Meeting Agenda 333
- B03_Kickoff Meeting Participants. 335
- B04_Kickoff Meeting Travel Information 336

Advisory Panel Meetings

- B05_Advisory Panel Agenda for July 338
- B06_Advisory Panel Agenda for September 339
- B07_Advisory Panel Agenda for October 340
- B08_Advisory Panel Members 341
- B09_Advisory Panel Travel Information 342
- B10_October 28 Advisory Panel Presentation 343

In-Person Workshop/Cross-Grade Panel Meetings

- B11_In-Person Cover E-mail 366
- B12_Smarter Balanced Resource Link 369
- B13_In-Person Cover E-mail ALTs.. 371
- B14_Final Information Cover E-mail 374
- B15_Last-Minute Cover E-mail 376
- B16_VAC Reminder 377
- B17_Evening Shuttle Service 378
- B18_IPP Thank You Certificate 379
- B19_State Lead Nominated Observer Cover E-mail 380
- B20_Nominated Observer Cover E-mail 381
- B21_Observer Cover E-mail 382
- B22_Guidelines for Observers 383
- B23_Smarter Balanced Staff Cover E-mail 384
- B24_E-mail to Staff 386
- B25_Non-Disclosure Agreement.... 387

- B26_Stipend and Honorarium Claim Voucher 388
- B27_Substitute Teacher Reimbursement Form..... 389
- B28_Travel Reimbursement Form 390
- B29_Guest Rooms 391
- B30_Hilton Anatole Atrium I Map.. 392
- B31_In-Person Materials List 394
- B32_NetOps Diagram 395
- B33_NetOps Schedule 396
- B34_Breakout Room Assignments 397
- B35_In-Person Panelists 401
- B36_Panelist Travel Dates..... 412
- B37_Panelist Registration Envelope..... 413
- B38_Color Print Badge Template .. 414
- B39_Name Badge Instructions 415
- B40_Table Tent Template..... 417
- B41_Super Shuttle Voucher 418

Chiefs Meetings

- B42_NCSA Standard Setting Session 419
- B43_November 6 Overview 429
- B44_RSVP Email..... 463
- B45_Dear Chiefs and State Leads 464
- B46_November 14 Overview 465
- B47_November 14 Motions 487

Smarter Balanced Standard Setting Logistics Plan

The successful completion of the three phases of standard setting (Distributed, In-Person, and Vertical Articulation) rests on careful planning and faithful execution of the plan. Logistics—recruitment, communication, coordination, support, record keeping, and tracking—will play a key role in the success of each phase. Measurement Incorporated (MI) appreciates the significance of effective logistics management for proceedings of this size and scope. We will ensure that all meetings associated with all phases of standard setting will be carried out effectively and efficiently.

This plan addresses seven aspects of logistics:

- Kick-Off Meeting
- In-Person Standard Setting
- Vertical Articulation Committee
- Online Meeting of the TAC
- Other Meetings
- Smarter Balanced Task Force
- Security of the Standard Setting Results

Kick-Off Meeting

On February 24–25, 2014, staff of Measurement Incorporated and CTB/McGraw-Hill met with representatives of Smarter Balanced in Minneapolis to review the initial plan outlined in the proposal and solidify plans for the life of the project. The agenda for that meeting is included as Appendix A. Deliverables associated with that meeting are listed below and have been forwarded to Smarter Balanced under separate cover:

- Updated master schedule
- Updated Program Work Plan including the management plan addressing communication, meetings, and management reporting
- Updated critical program success metrics and quality standards that will drive the program
- Meeting Notes/Decision Log/Open Action List

Mandy Hunter (MI's Assistant Project Manager for Meetings and Logistics) made arrangements for participation of Smarter Balanced leadership (Joe Willhoft, Christyan Mitchell, Nancy Arnold, and Michael Middleton) plus eight Consortium members. Subsequent to the meeting, Ms. Donna Merritt¹ (MI's project manager) provided the minutes.

In-Person Workshop

In October 2014, we will assemble 504 panelists, alternates, up to 13 Smarter Balanced staff and work group members, and observers and auditors to recommend cut scores on 14 tests. We will make all the necessary arrangements to transport all participants to and from a suitable location and provide for their lodging, meals, and related needs for the time that they are on site.

Meeting site. MI has entered into an agreement with the Hilton Anatole in Dallas, Texas, for in-person standard setting October 13–19, 2014. Ms. Hunter will make all the necessary arrangements to transport all participants to and from the meeting site and provide for their lodging, meals, and related needs for the time that they are on site. MI, CTB, and Hager Sharp staff are currently working with Smarter Balanced leadership to identify all panelists for the in-person standard setting. Once

¹ On March 21, Ms. Barbara Scherich assumed the role of project manager.

panelists are identified, Ms. Hunter will work directly with them to make their flight arrangements and room assignments. Panelists will be able to reach our travel agent through a toll-free number, and all associated costs will be billed directly to MI.

MI will underwrite all costs associated with the meeting, including meeting and sleeping rooms, meals, snacks, coffee, juice, water, and all equipment necessary to conduct standard setting. Each training room and breakout room will have a beverage station for coffee, tea, and water. While we will provide full breakfast and lunch for all participants each day, we will also provide a meal allowance for dinner as well as meals participants consume en route to and from the meeting. Although each group is scheduled to leave mid-day on their third day, we have arranged for lunch to be served to all departing panelists as well as those staying for the afternoon rounds. We will secure and have budgeted for sleeping rooms and meal service for the numbers of panelists ranging from 167 to 392 people per day.

Meeting rooms at the Hilton Anatole are quite adequate for our purposes. Appendix B shows a complete day-by-day listing of general session and breakout rooms for the week.

In our proposal, we indicated that we would provide wireless Internet access in each breakout room for panelists to use their own laptops to complete the standard-setting activities. Subsequent negotiations with Smarter Balanced leadership resulted in MI providing computers for all panelists. Therefore, we will provide computers with all required software preloaded in order to have more control over access to content, negating the need for constant Internet access. Members of our IT staff will work with hotel staff in the weeks leading up to the October meeting to make sure all equipment is in place, tested, and ready to function as required.

Identification and recruitment of panelists. Recruitment of in-person panelists will begin with a presentation to State Leads on April 22 and continue through May 23. MI, CTB, and Smarter Balanced staff will work together to identify 504 panelists and 10 alternates by June 9. MI/CTB will notify all selected panelists and alternates by June 30. Panelist notification packets will include the following:

- Security/Nondisclosure Agreement *
- Travel/Dietary/Housing restrictions *
- Housing and Travel policy and guidelines
- Confirmed dates of travel and on-site participation *
- Committee Assignment (grade/subject or Alternate)
- Links to CCSS, ALDs, and Smarter Balanced Practice Tests
- Sample expense voucher and guidelines
- Sample substitute teacher reimbursement form and guidelines
- Receipt Acknowledgment form

These materials will spell out in detail what we will and will not pay for and how panelists will be reimbursed for any out-of-pocket expenses. We will ask each panelist to acknowledge receipt of all materials and return the forms marked with an asterisk (*) above.

From this point, Measurement Incorporated staff will work with the panelists to secure hotel reservations and make travel arrangements. We have booked a block of rooms at the Hilton Anatole in Dallas, Texas, and are working with major airlines to set up accounts so that no panelist has to pay for airline or hotel reservations.

Support and conditions. We have budgeted for substitute teacher pay at the rate of \$150 per day for teachers who serve as panelists and require substitutes, and stipends of \$150 for teachers

who do not require substitutes. We plan to extend the daily stipend to all in-person standard setting panelists within the constraints of their individual state laws and regulations.

In addition to underwriting the costs of participation of all panelists, MI will assume costs of attendance of up to 13 Smarter Balanced staff and state representatives. These costs include travel, lodging, and meals.

Prior to attending in-person standard setting, all panelists and observers will complete and sign a non-disclosure agreement (see above). Ms. Hunter will keep all non-disclosure agreements and scan them into a file for permanent storage.

We are committed to inclusive meetings and will make every effort to ensure that the in-person standard setting is inclusive. In selecting meeting sites, we will make sure there are accessible meeting rooms, sleeping rooms, restaurants, and restrooms. During the recruiting phase, we will learn of any dietary restrictions of potential panelists and make sure we are able to accommodate them. For each large-group presentation, we will hire at least one ASL interpreter. If, through the recruiting and vetting process, we learn that we will have two or more panelists requiring ASL translation services on a given day, we will hire enough ASL interpreters to meet the need.

Program. We have asked Dr. Joe Willhoft, Executive Director of the Smarter Balanced Assessment Consortium, to serve as the keynote speaker for each of the plenary sessions (Monday, Wednesday, and Friday mornings), and he has agreed to serve. In the event that Dr. Willhoft is unable to be present to address the other grade-level groups, we will record his presentation and share it with Grades 6–8 on Wednesday morning and Grades 3–5 on Friday morning. We believe it will be important for panelists to hear his vision of the ultimate use and impact of the Smarter Balanced assessments as they consider their role in carrying out that vision. MI staff will work closely with Dr. Willhoft to prepare for these presentations.

Flow. Figure 1 illustrates the flow of groups through the week of October 13-19. The grade 11 panelists arrive Sunday, October 12, and start Monday morning, October 13. They finish at noon on October 15. The grades 6-8 panelists arrive Tuesday, October 14, and start Wednesday morning, October 15. They will be receiving orientation while the Grade 11 panelists are completing Round 3. The grades 3-5 panelists arrive Thursday, October 16, and start Friday morning, October 17, as the Grades 6-8 panelists are completing Round 3.

Figure 1. Group Timetable for In-Person Workshop

	October 13	October 14	October 15	October 16	October 17	October 18	October 19
Grade 11	■	■	■				
Grades 6-8			■	■	■		
Grades 3-5					■	■	■

Table 1 shows the event-by-event flow of activities for the In-Person Workshop. Panelists spend most of the first day receiving training in the content of the tests, the Achievement Level Descriptors, the Common Core State Standards, and the Bookmark procedure. They will begin using their training before the first day is over.

Table 1. Agenda For In-Person Workshop

Agenda	
Day 1	
Time	Event
7:30 a.m.	Registration opens (Atrium I Lobby) <ul style="list-style-type: none"> Badges, place cards, and room and table assignments Continental breakfast
8:30 a.m.	Opening session (Carpenter Ballroom) <ul style="list-style-type: none"> Keynote address and charge Overview of schedule of events for the In-Person Panel Housekeeping matters
9:15 a.m.	Review of CCSS and ALDs (Carpenter Ballroom) <ul style="list-style-type: none"> Review of materials send in advance Discussion of CCSS and ALDs Completion of Readiness Form Part 1
10:15 a.m.	Break (adjourn to breakout rooms)
10:30 a.m.	Introduction to the test <ul style="list-style-type: none"> Overview of test development and scoring Practice Test administration Table wide discussion of tests Closing comments on test construction Completion of Readiness Form Part 2
12:30 p.m.	Lunch (Atrium I Lobby)
1:30 p.m.	Introduction to the bookmark procedure (Carpenter Ballroom) <ul style="list-style-type: none"> PowerPoint presentation Practice round Questions and answers
3:15 p.m.	Completion of Readiness Form Part 3
3:30 p.m.	Adjourn to breakout rooms
3:45 p.m.	Begin Round 1 <ul style="list-style-type: none"> Panelists work in 6-person groups to review OIB and item map, discuss possible bookmarks, and enter bookmarks individually
4:45 p.m.	Wrap-up for Day 1 <ul style="list-style-type: none"> Panelists log out; computers secured

Agenda	
5:00 p.m.	Adjourn
5:15 p.m.	Smarter Balanced, MI, CTB, and auditor debriefing
6:00 p.m.	Secure meeting rooms and depart
Day 2	
7:30 a.m.	Registration opens (Atrium I Lobby) <ul style="list-style-type: none"> Badges, place cards, and room and table assignments Continental breakfast
8:30 a.m.	Convene in Breakout rooms <ul style="list-style-type: none"> Review Round 1 process Questions and answers on work thus far for Round 1
9:15 a.m.	Resume Round 1 as on previous afternoon
11:30 a.m.	Lunch (Atrium I Lobby) <ul style="list-style-type: none"> MI/CTB staff analyze data and prepare reports while panelists eat lunch. Confer with Smarter Balanced staff to approve Round 1 analyses and reports Identify Vertical Articulation Committee panelists
1:00 p.m.	Review of Round 1 <ul style="list-style-type: none"> Grade group facilitators review results of Round 1; lead discussion of bookmark placements, impact data, and rationales for bookmark placements
2:30 p.m.	Completion of Readiness Form Part 4
2:45 p.m.	Begin Round 2 <ul style="list-style-type: none"> Panelists work in 6-person groups as in Round 1
4:45 p.m.	Wrap-up for Day 2 <ul style="list-style-type: none"> Panelists log out; computers secured
5:00 p.m.	Adjourn <ul style="list-style-type: none"> MI/CTB staff analyze data and prepare reports. Meet with Vertical Articulation Committee panelists.
5:30 p.m.	Smarter Balanced, MI, CTB, and auditor debriefing
6:00 p.m.	Secure meeting rooms and depart
Day 3	
7:30 a.m.	Registration opens (Atrium I Lobby)

Agenda	
	<ul style="list-style-type: none"> Badges, place cards, and room and table assignments Continental breakfast
8:30 a.m.	Convene in Breakout rooms <ul style="list-style-type: none"> Review Round 2 results; grade group facilitators review results of Round 1; lead discussion of bookmark placements, impact data, and rationales for bookmark placements
9:45 a.m.	Completion of Readiness Form Part 5
10:00 a.m.	Begin Round 3 <ul style="list-style-type: none"> Panelists work in 6-person groups as in Rounds 1 and 2.
11:30 a.m.	Wrap-up for Day 3 <ul style="list-style-type: none"> MI/CTB staff process and share Round 3 results for this group plus others that have completed previously (not applicable to Grade 11; Grade 6–8 groups will see only their results and those for Grade 11; Grade 3–5 groups may see all results).
11:45 a.m.	Completion of Evaluation Forms; log out; all computers secured
12:00 p.m.	Adjourn for lunch (Atrium I Lobby) <ul style="list-style-type: none"> Panelists may stay for lunch or leave if they have successfully logged out. MI/CTB staff process Round 3 data confer with Smarter Balanced staff and auditor, update standard setting report.

Vertical Articulation Committee (VAC)

MI and CTB/McGraw-Hill will conduct two simultaneous vertical articulation activities following the in-person standard setting in October. The VAC meetings are scheduled for October 20–21 and will be held as webinars. Because we will have just completed in-person standard setting at the Dallas Hilton Anatole, we plan to remain on site to conduct these webinars from the hotel rather than run the risk of mishap between Dallas and our home locations. We also believe that it will be helpful to be able to confer with one another after the first day of each VAC.

Table 2 shows the agenda for the two-day vertical articulation. Both sessions begin at 1:00 p.m. Eastern Time and conclude at 5:00 p.m. Eastern Time, thereby allowing panelists in multiple time zones the opportunity to participate during a typical work day.

Table 2. VAC Agenda

Agenda	
October 20	[All times are Eastern.]
1:00 p.m.	Welcome and introductions
1:15 p.m.	Orientation to VAC (PowerPoint presentation)

Agenda	
2:00 p.m.	Questions and Answers
2:45 p.m.	Completion of VAC Readiness Form Part 1
3:00 p.m.	Presentation of Round 3 cut scores and impact data
3:15 a.m.	Discussion of Level 3 (CCR) cut across grades with Online Panel data (using voting feature of WebEx)
3:45 p.m.	<ul style="list-style-type: none"> • Formal voting on cut score changes for Level 3 • Panelist recommends change; change receives a second (message to whole group) • Facilitator opens floor for comments • Facilitator calls for the vote • Panelists enter Yes/No vote on specific motion on the floor • Facilitator announces results of vote • Facilitator opens the floor for motions for more changes as above • Facilitator calls for motion to approve all changes for Level 3 • Panelist moves for approval; another seconds • Facilitator calls for discussion, then vote on approval of changes
5:00 p.m.	Adjourn Day 1
October 21	
1:00 p.m.	Review of Day 1; completion of VAC Readiness Form Round 2
1:15 p.m.	Presentation of Round 3 cut scores and impact data updated from Day 1
2:00 p.m.	Discussion of Level 2 cut scores across grades
2:30 p.m.	Formal voting on cut score changes for Level 2 as above
3:30 p.m.	Completion of VAC Readiness Form Round 3
3:45 p.m.	Discussion of Level 4 cut scores across grades
4:00 p.m.	Formal voting on cut score changes for Level 4 as above
5:00 p.m.	Adjourn

To obtain vertical articulation feedback in a virtual environment, we will take the following steps:

1. *Identify, recruit, and select panelists.* VAC panelists will be selected from the in-person standard setting panelists. We will identify each of the 30 VAC panelists in advance of in-person standard setting in order that they can be attuned to relevant information from their individual subject/grade panels to bring forward to the VAC.
2. *Schedule the VAC meeting.* We plan to conduct 2 four-hour sessions for each VAC, beginning at 1:00 p.m. and ending at 5:00 p.m. EDT each day. This schedule allows panelists from multiple time zones to participate in the course of a normal work day (e.g., 10:00 a.m. to 1:00 p.m. PDT or 8:00 a.m. to noon in Hawaii). Ms. Hunter will arrange for the webinar, notify all panelists well in advance, confirm participation during the in-person standard setting, and send reminders the day before the webinar.
3. *Verify access.* Some of the panelists may never have participated in a webinar before. We will send an invitation to a pre-webinar to all VAC panelists in advance of the In-Person Workshop so that they can practice logging in and dialing in to the webinar, using the device they plan to use for the VAC, and become comfortable with navigation. Prior to coming to the In-Person Workshop, they will log in to the practice site and answer some basic questions to verify their participation.
4. *Provide a secure log-in.* MI will provide secure log-in information via email and confirm receipt during the in-person standard setting. We will open the site 30 minutes prior to the webinar so that all 30 panelists will be able to log in, dial in, and be in place so that we can start on time each day.
5. *Provide backup.* MI's IT staff will be on site in Dallas, working with hotel IT staff to make sure the Internet service is uninterrupted and will have backup sources standing by.
6. *Train panelists.* MI staff will prepare web-based training materials that reinforce the in-person training and introduce VAC panelists to the task before them. The Standard Setting Plan contains details of this training.
7. *Conduct a secure online VAC meeting.* We will follow Smarter Balanced security protocols and augment those with lessons learned during the SBAC 16/17 contract. Each panelist will sign a second non-disclosure agreement applicable just to the VAC during the in-person standard setting.
8. *Provide visual feedback.* Using software field tested with the DSS, we will present tabular and graphical data. We provide additional detail about this feedback in the Standard Setting Plan.

Online Support. We will conduct two simultaneous online meetings, one for English language arts/literacy and another for mathematics. Each webinar will be on a secure website such as the WebEx site we currently use for conferences. Dr. Bunch and Dr. Lewis are scheduled to lead these two webinars. Each webinar will begin with roll call. During the call, the presenters will have available all ordered item booklets from the In-Person Workshop, Achievement Level Descriptors, Common Core State Standards, and Round 3 results from the In-Person Workshop as well as results of the Online Panel. They will share these documents at appropriate times, open the floor for motions and seconds (to change a cut score), and then lead discussions via chat and voice. Panelists will then vote for or against the motion using a voting slide.

Online Meetings of the Technical Advisory Committee (TAC)

On March 6, MI hosted an online meeting of the TAC, and Dr. Bunch assisted Dr. Willhoft in the presentation of the standard setting plan. In advance of that meeting, MI staff prepared presentation materials in draft form for Smarter Balanced review, revised those materials as necessary, and made

them ready for online presentation. MI staff also solicited questions and comments from TAC members and incorporated those into the presentation.

Working with Smarter Balanced, Ms. Hunter and Dr. Bunch will identify the specific requirements for each online TAC meeting. Dr. Bunch's role will be to respond to technical questions and concerns regarding methodology for both standard setting and vertical articulation, working in concert with Dr. Daniel Lewis of CTB/McGraw-Hill. Ms. Hunter will set up the webinars and send invitations to the TAC members and other required and optional attendees.

For the post-VAC meeting, Dr. Bunch will prepare a briefing package in consultation with Dr. Lewis during the in-person standard setting and refine it immediately after the DSS. The briefing package will document the conduct of the DSS, in-person standard setting, and VAC. Its purpose will be to provide background for the final virtual TAC meeting.

As with other webinars, Ms. Hunter will schedule the meeting, obtain log-in information, share that information with TAC members in advance, send out reminders, and be available from one hour prior to the meeting to one hour after the meeting to assist with technical problems. The webinar itself will permit but not require the use of webcams by the TAC members. Control of the screen, while normally reserved for the host, can be passed from the host to members in order that they may more readily share information.

The log-in process will capture identifying information about each participant and make it available both to the host and to other participants. During conversations, a drop-down box will show who is speaking. Simultaneously, participants may use the chat feature of the webinar to share comments silently with specific other participants or the entire committee. While Dr. Bunch will serve as host, the Meeting Manager will also be online to record recommendations and decisions made, queue up files to display, and assist with technical problems or questions TAC members may have.

Other Meetings

MI will provide logistical support for other meetings such as virtual TAC meetings and face-to-face Collaboration Conferences. We will provide logistical support for additional virtual TAC meetings in accordance with the procedures described above. We will provide for the attendance of Dr. Bunch and Ms. Scherich at two Collaboration Conferences.

Over the life of the project, as part of the public outreach and communication plan, Ms. Scherich, Dr. Bunch, and other contractor staff will attend targeted meetings. We have made allowance for 2 four-day trips to the June 2014 NCSA meeting in New Orleans.

In addition to the meetings described above, we have budgeted for 4 two-day meetings for Smarter Balanced staff and members to meet with Ms. Scherich and Dr. Bunch. We will underwrite the costs for each of these meetings, including airfare, parking, ground transportation, lodging, meals, snacks, meeting sites, and support. Ms. Hunter will work out dates for those meetings with Smarter Balanced staff.

Smarter Balanced Task Force

The Smarter Balanced Task Force (Task Force) is another stakeholder group requiring special attention. This group will include high-level educational administrators as well as prominent technical experts and policy advisors. It will differ from the TAC in that its charge will be broader and more policy oriented. Table 3 shows the names and affiliations of Task Force Nominees. From this list of 18, we will select 12 to serve on the Task Force.

Table 3. Task Force Nominees.

First Name	Last Name	Affiliation
Susan	Loomis	Consultant
Bill	Schmidt	Michigan State University
Mark	Reckase	Michigan State University
Karla	Egan	NCIEA
Scott	Marion	NCIEA
Joseph	Martineau	Smarter Balanced
Joe	Willhoft	Smarter Balanced
Marty	McCall	Smarter Balanced
Jacqueline	King	Smarter Balanced
Greg	Cizek	Smarter Balanced TAC
Brian	Gong	Smarter Balanced TAC
David	Conley	Smarter Balanced TAC
Joe	Ryan	Smarter Balanced TAC
Sandra	Alberti	Student Achievement Partners – ELA
Luz	Bay	The College Board
Stephen	Sireci	University of Massachusetts–Amherst
Barbara	Plake	University of Nebraska–Lincoln
Jim	Wollack	University of Wisconsin

MI has drafted an invitation email for Smarter Balanced approval. Upon approval of the email, we will contact nominees and form the Task Force from those responding positively.

Over the life of the contract, we will meet two to three times with the Task Force, tentatively in May or June and again in October. Ms. Hunter will arrange for each meeting at a site that is mutually acceptable or via webinar. MI will underwrite all costs for each of the meetings, including lodging, meeting space, meals, snacks, per diem, and flights and additional travel expenses. Ms. Scherich will prepare an agenda for each meeting in consultation with Consortium staff, assign a scribe to take notes during each meeting, and provide a written account of discussions and decisions reached. Ms. Scherich and Dr. Bunch will be present at each meeting, fully prepared to discuss any aspect of standard setting or to answer questions about the design, implementation, or reporting and follow-up activities.

Security of the Standard Setting Results

Security of standard setting results encompasses online, in-person, and vertical articulation activities, as well as the events leading up to them.

Security overview. MI has an off-site, high-tech, Tier III data center to house critical computer and network operations. Extensive planning went into its design and implementation to provide maximum reliability, redundancy, and security to protect internal and external client data and systems. The building is constructed of solid brick, ballistic-resistant walls that are reinforced to withstand sustained Category 3 hurricane-force winds. In addition, a sophisticated grounding system protects the building and computer systems from lightning strikes and surge damage.

Inside, the building is equipped with an automated, non-damaging chemical fire suppression system with fire alarms for automatic notification and dispatch. The inside temperature and humidity are precisely controlled and monitored by redundant air handlers and chillers to maintain system reliability. The exterior of the data center is protected by barbed wire fencing, security lighting, video surveillance, a security monitoring system, and keycard-authorized entry. In the event of a minor or catastrophic power failure, the data center is equipped with two fully redundant generators that are supplied by a natural gas line and tested weekly.

MI employs on-site state-of-the-art Distributed Denial of Service (DDoS) protection in its Tier III data center as a complement to our Internet Service Provider (ISP) based upstream DDoS services. Behavioral protection and advanced challenge/response techniques allow us to mitigate modern DDoS attacks that use new methods to exploit areas that traditional security solutions, such as firewalls, are not equipped to handle. We are able to automatically defend against network flood and application layer attacks while maintaining low latency and high performance.

At the heart of the MI information technology system is a firewall implementation that allows us to block, audit, and respond to both internal and external threats. MI currently employs 19 separate firewalls to provide layered and redundant protection. These firewalls utilize state-of-the-art deep packet inspection, port blocking, proxying, address translations, heuristics, and trend analysis to provide security. In addition, this multi-vendor solution limits exposure to potential weaknesses associated with each implementation that might be exploited in the future. In recent years new attack vectors have emerged that largely bypass many of the port based security protections that traditional firewalls provide. In response, MI has implemented state-of-the-art enterprise class HTTP attack and other application layer protection appliances. Our firewalls, both virtual and physical, are from industry-recognized leaders such as Juniper, Radware, and Checkpoint. MI immediately updates all defenses as soon as emerging threats and countermeasures are identified.

MI also employs an Intrusion Detection System that allows rule sets to be updated automatically to block unwanted traffic in real time, whether the source is internal or external. To further complement these capabilities, MI has deployed software that detects, removes, and destroys viruses, spyware, and other forms of malicious software. This software is updated at least daily through automated means backed by constant monitoring by Network Operations staff. We also routinely deploy security patches and updates for operating systems and commercial software using a central update management server.

MI deploys Web services in an untrusted domain separate from the main corporate network and with additional layers of firewall protection. We employ compartmentalization to divide internal users on a “need to know” basis with controlled access between client and project-related resources and discrete subnets within the network. Our Active Directory-based Identity Management Services provide the foundation for this capability. The internal network users are monitored in a “trust but verify” approach that audits and records the actions of users and prevents the pass-through of undesired traffic. This audit allows us to monitor for irregular and illegal acts and provides reinforcement of MI standards.

In order to verify the effectiveness of all our security measures MI periodically engages an independent expert Cyber Security company to audit our operations. The audits include but are not limited to: penetration testing, web application testing, network infrastructure security testing, best security practices reviews, wireless audits, and social engineering security testing. If at any time critical deficiencies are identified, immediate remedial action is taken. Other less critical security

deficiencies are scheduled for remediation upon the next release or maintenance window as appropriate.

Backup and recovery. In addition to the redundancy and fault tolerance of our data storage outlined in Storage Architecture above, MI employs time-proven methods for data backups and disaster recovery. Several Ultrium tape autoloader solutions, backup to disk solutions, and SAN solutions provide the mechanisms for primary backups along with the latest snap-shot and off-site replication technologies. We back up all data, including our servers, nightly from our primary Network Operating Center in Durham, NC, to other facilities (our Durham Data Center, our Greensboro office, and the Brodie building in Durham) for immediate diversification. If a disaster occurs to one building, we are able to begin recovery immediately from another location.

Security of online standard setting. With the opening of the online standard setting to the public, security of test items is effectively compromised. Smarter Balanced has therefore decided to treat the online panel activity as an open activity and use items that have either already been released (e.g., the Practice Tests) or use limited numbers of items that will be released once used in the online activity.

Beyond item integrity, however, is data integrity. In order to make sure that individuals or groups do not intentionally skew results of the online panel, MI has introduced security measures specific to the online panel activity beyond those described above. The registration page, for example, contains characters that the registrant must recognize and type into a box. These characters are not readable by malware designed to invade websites. In addition, individuals registering for the online panel must go through a two-stage process and must use a consistent email address throughout the process.

Security during in-person standard setting. Prior to coming to in-person standard setting, all panelists will have completed and signed a non-disclosure agreement. All panelists will also have been thoroughly vetted by Teacher Involvement Coordinators (TICs), Higher Education Leads, or other known individuals in their states. Once panelists arrive, we will provide explicit instructions about security on the first day of training. We have found that instructions such as the following are quite helpful:²

- You can't take any of this with you when you leave.
- You can't talk to anyone about any test content, now or ever.
- You can talk to others about the process but not the data or test items.
- You can't call another person (e.g., your supervisor, colleague, friend, spouse) to ask for advice.
- You can't take this back to your room to study tonight or to lunch to discuss with colleagues.

MI staff will arrive at the site the week prior to the in-person standard setting to set up computer and server equipment. Working with hotel IT staff, MI's IT staff will secure all servers and computers and test all software to make sure not only that it works but that it works securely. The computers we will assign to panelists will be disabled except to run the programs we preload. Thus, it will not be possible for panelists to use email or other outward facing programs unless we authorize such activity.

² Taken from Bunch, M. B. (2012). Practical issues in standard setting. In G. J. Cizek (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations* (2nd Ed.). New York: Routledge.

Security during vertical articulation. In advance of the vertical articulation, panelists will receive a secure log-in code that they will confirm during in-person standard setting. MI and CTB staff leading the VACs will conduct them from the Dallas Hilton Anatole rather than from MI or CTB offices in order to minimize the risk of mishap between the last day of in-person standard setting (October 19) and the first day of VAC (October 20).

Appendix A Kick-Off Meeting Agenda

Smarter Balanced 21: Standard Setting & Communications

Meeting Title:

Smarter Balanced Project 21 Kick-Off
Radisson Plaza Hotel Minneapolis
35 S 7th Street
Minneapolis, MN 55402

Type of Meeting:

Project Kick-off Meeting

Date of Meeting:

2/24 – 2/25, 2014

Facilitator:

Measurement Incorporated, Project Contractor

Expected Attendees: PRELIMINARY LIST – to be confirmed

- **Smarter Balanced:** Jackie King, Juan D'Brot, Luci Willits, Dacia Hopfensberger, Marty McCall, Steve Slater, Patty Reiss, Joseph Martineau, Joe Willhoft, Nancy Arnold
- **OSPI:** Mike Middleton
- **PMP:** Christyan Mitchell
- **Contractor: MI:** Hank Scherich, Mike Bunch, Nelson Androes, Mandy Hunter, Donna Merritt
- **CTB:** Antonia "Toni" Deoudes, David Breen, Rick Mercado, David Cosio
- **ETS:** Dirk Mattson, Bihua Xiang

Meeting Access Information:

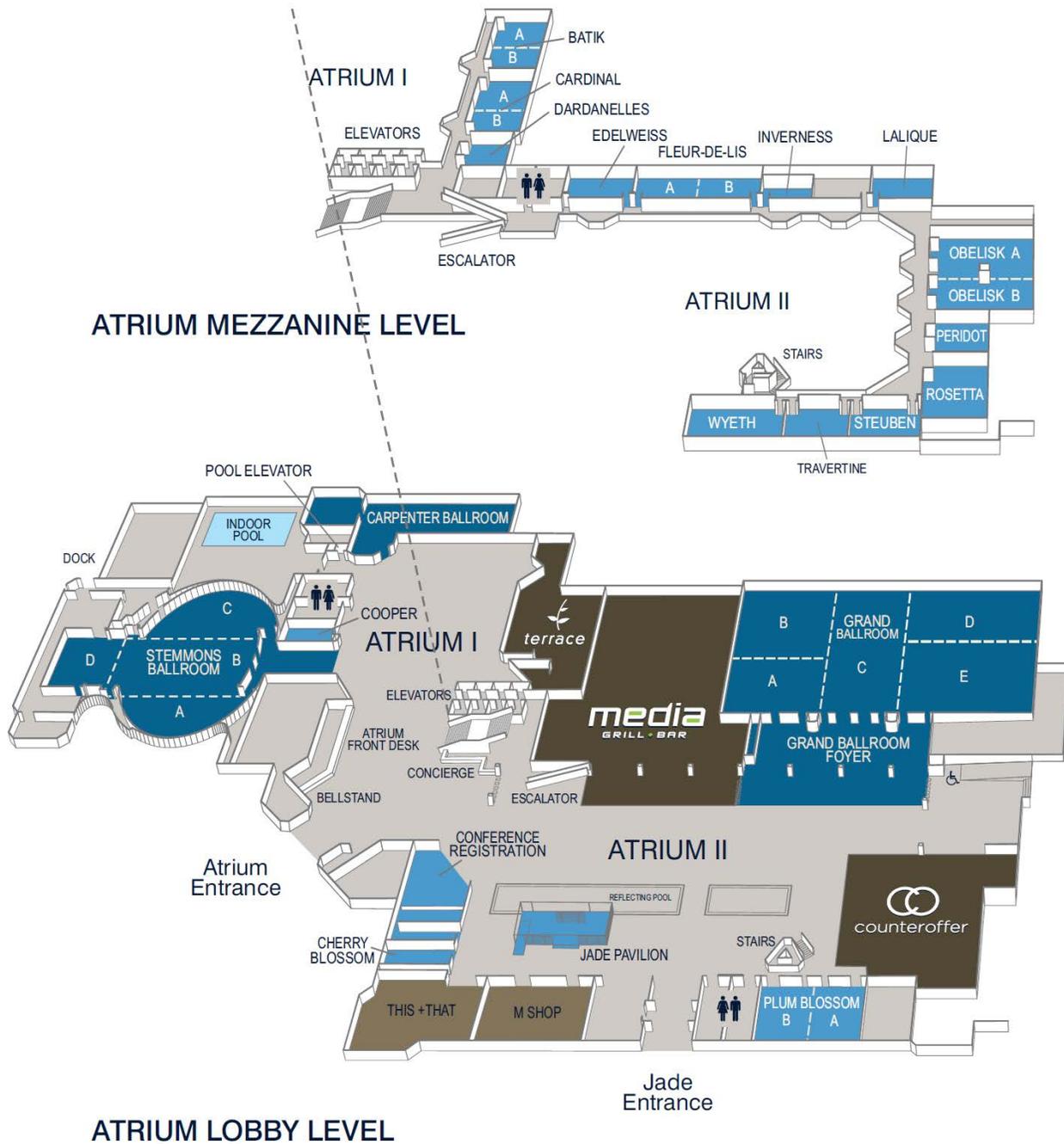
- **Meeting Conference Call-in Number: 1 (877) 820-7831**
- **Access Code: 843853**

Agenda Topic : Day 1		Presenter
9:30 – 10:15	Welcome and Introductions Confirm Attendees Hotel Information	Donna Merritt
10:15 – 11:45	Task Cluster 7: Project Schedule & Tasks	
11:45 – 12:00	Break	
12:00 – 1:15	Lunch	
1:15 – 2:00	Task Cluster 1: Communication Plan Discussion	Donna Merritt
2:00 – 3:00	Overview of Functional Groups and Tasks Discussion	

Agenda Topic : Day 2		Presenter
9:30 – 11:00	Task Cluster 5: Standard Setting - Cross-Contract Project Schedule & Task Discussion	MI, CTB, ETS
11:00 – 11:30	Task Cluster 4: Participant Recruitment Activities Discussion	CTB
11:30 – 12:00	Upcoming Meeting Updates	
12:00 – 1:00	Lunch	
1:00 – 1:30	Project Schedule & Task Discussion Wrap-up	

Appendix B

Dallas Hilton Anatole Meeting Space



Meeting Room Configuration

Date	Start Time	End Time	Function	Room	Setup	Agr	
Fri, 10/10/14	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Batik A/B	Classroom	40	
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Cardinal A/B	Classroom	40	
	8:00 a.m.	11:00 p.m.	Office	Cooper			
Sat, 10/11/14	8:00 a.m.	11:00 p.m.	Office	Cooper			
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Batik A/B	Classroom	40	
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Cardinal A/B	Classroom	40	
Sun, 10/12/14	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Batik A/B	Classroom	40	
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Cardinal A/B	Classroom	40	
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Dardanelles	Classroom	40	
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Edelweiss	Classroom	40	
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Fleur de Lis A/B	Classroom	40	
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Lalique	Classroom	40	
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Rosetta	Classroom	40	
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Steuben	Classroom	40	
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Wyeth	Classroom	40	
	8:00 a.m.	11:00 p.m.	Office	Cooper	Office	10	
	Mon, 10/13/14	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Batik A/B	Classroom	40
8:00 a.m.		11:00 p.m.	Breakout/Syndicate	Cardinal A/B	Classroom	40	
8:00 a.m.		11:00 p.m.	General Session/Congress	Carpenter Ballroom	Theater	200	
8:00 a.m.		11:00 p.m.	Breakout/Syndicate	Dardanelles	Classroom	40	
8:00 a.m.		11:00 p.m.	Breakout/Syndicate	Edelweiss	Classroom	40	
8:00 a.m.		11:00 p.m.	Breakout/Syndicate	Fleur de Lis A/B	Classroom	40	
8:00 a.m.		11:00 p.m.	Breakout/Syndicate	Lalique	Classroom	40	
8:00 a.m.		11:00 p.m.	Breakout/Syndicate	Rosetta	Classroom	40	
8:00 a.m.		11:00 p.m.	Breakout/Syndicate	Steuben	Classroom	40	
8:00 a.m.		11:00 p.m.	Breakout/Syndicate	Wyeth	Classroom	40	
7:00 a.m.		8:00 a.m.	Continental Breakfast	Atrium I Lobby	Lounge	167	
10:00 a.m.		10:30 a.m.	Coffee Break	Atrium I Lobby	Lounge	167	
Noon	2:00 p.m.	Lunch	Atrium I Lobby	Round Tables of 10	167		
3:00 p.m.	3:30 p.m.	Coffee Break	Atrium I Lobby	Lounge	167		
8:00 a.m.	11:00 p.m.	Office	Cooper	Office	10		
8:00 a.m.	11:00 p.m.	Meeting	Inverness	Hollow Square	30		
Tue, 10/14/14	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Batik A/B	Classroom	40	

Date	Start Time	End Time	Function	Room	Setup	Agr
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Cardinal A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Dardanelles	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Edelweiss	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Fleur de Lis A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Lalique	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Rosetta	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Steuben	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Wyeth	Classroom	40
	7:00 a.m.	8:00 a.m.	Continental Breakfast	Atrium I Lobby	Lounge	167
	10:00 a.m.	10:30 a.m.	Coffee Break	Atrium I Lobby	Lounge	167
	Noon	2:00 p.m.	Lunch	Atrium I Lobby	Round Tables of 10	167
	3:00 p.m.	3:30 p.m.	Coffee Break	Atrium I Lobby	Lounge	167
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Steuben	Classroom	40
	8:00 a.m.	11:00 p.m.	Office	Cooper	Office	10
	8:00 a.m.	11:00 p.m.	Meeting	Inverness	Hollow Square	30
Wed, 10/15/14	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Batik A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Cardinal A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	General Session/Congress	Carpenter Ballroom	Theater	200
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Dardanelles	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Edelweiss	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Fleur de Lis A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Lalique	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Rosetta	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Steuben	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Wyeth	Classroom	40
	7:00 a.m.	8:00 a.m.	Continental Breakfast	Atrium I Lobby	Lounge	354
	10:00 a.m.	10:30 a.m.	Coffee Break	Atrium I Lobby	Lounge	354
	Noon	2:00 p.m.	Lunch	Atrium I Lobby	Round Tables of 10	354
	3:00 p.m.	3:30 p.m.	Coffee Break	Atrium I Lobby	Lounge	354
	8:00 a.m.	11:00 p.m.	Office	Cooper	Office	10
	8:00 a.m.	11:00 p.m.	Meeting	Inverness	Hollow Square	30
Thu, 10/16/14	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Batik A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Cardinal A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Dardanelles	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Edelweiss	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Fleur de Lis A/B	Classroom	40

Date	Start Time	End Time	Function	Room	Setup	Agr
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Lalique	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Rosetta	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Steuben	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Wyeth	Classroom	40
	7:00 a.m.	8:00 a.m.	Continental Breakfast	Atrium I Lobby	Lounge	206
	10:00 a.m.	10:30 a.m.	Coffee Break	Atrium I Lobby	Lounge	206
	Noon	2:00 p.m.	Lunch	Atrium I Lobby	Round Tables of 10	206
	3:00 p.m.	3:30 p.m.	Coffee Break	Atrium I Lobby	Lounge	206
	8:00 a.m.	11:00 p.m.	Office	Cooper	Office	10
	8:00 a.m.	11:00 p.m.	Meeting	Inverness	Hollow Square	30
Fri, 10/17/14	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Batik A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Cardinal A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	General Session/Congress	Carpenter Ballroom	Theater	200
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Dardanelles	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Edelweiss	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Fleur de Lis A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Lalique	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Rosetta	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Steuben	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Wyeth	Classroom	40
	7:00 a.m.	8:00 a.m.	Continental Breakfast	Atrium I Lobby	Lounge	392
	10:00 a.m.	10:30 a.m.	Coffee Break	Atrium I Lobby	Lounge	392
	Noon	2:00 p.m.	Lunch	Atrium I Lobby	Round Tables of 10	392
	3:00 p.m.	3:30 p.m.	Coffee Break	Atrium I Lobby	Lounge	392
	8:00 a.m.	11:00 p.m.	Office	Cooper	Office	10
	8:00 a.m.	11:00 p.m.	Meeting	Inverness	Hollow Square	30
Sat, 10/18/14	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Batik A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Cardinal A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Dardanelles	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Edelweiss	Classroom	40
	8:00 a.m.	11:00 p.m.	Teardown	Fleur de Lis A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	Teardown	Lalique	Classroom	40
	8:00 a.m.	11:00 p.m.	Teardown	Rosetta	Classroom	40
	8:00 a.m.	11:00 p.m.	Teardown	Steuben	Classroom	40
	8:00 a.m.	11:00 p.m.	Teardown	Wyeth	Classroom	40
	7:00 a.m.	8:00 a.m.	Continental Breakfast	Atrium I Lobby	Lounge	206

Date	Start Time	End Time	Function	Room	Setup	Agr
	10:00 a.m.	10:30 a.m.	Coffee Break	Atrium I Lobby	Lounge	206
	Noon	2:00 p.m.	Lunch	Atrium I Lobby	Round Tables of 10	206
	3:00 p.m.	3:30 p.m.	Coffee Break	Atrium I Lobby	Lounge	206
	8:00 a.m.	11:00 p.m.	Office	Cooper	Office	10
	8:00 a.m.	11:00 p.m.	Meeting	Inverness	Hollow Square	30
Sun, 10/19/14	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Batik A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Cardinal A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Dardanelles	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Edelweiss	Classroom	40
	8:00 a.m.	11:00 p.m.	Breakout/Syndicate	Fleur de Lis A/B	Classroom	40
	8:00 a.m.	11:00 p.m.	Teardown	Lalique	Classroom	40
	8:00 a.m.	11:00 p.m.	Teardown	Rosetta	Classroom	40
	8:00 a.m.	11:00 p.m.	Teardown	Steuben	Classroom	40
	8:00 a.m.	11:00 p.m.	Teardown	Wyeth	Classroom	40
	7:00 a.m.	8:00 a.m.	Continental Breakfast	Atrium I Lobby	Lounge	206
	10:00 a.m.	10:30 a.m.	Coffee Break	Atrium I Lobby	Lounge	206
	Noon	2:00 p.m.	Lunch	Atrium I Lobby	Round Tables of 10	206
	3:00 p.m.	3:30 p.m.	Coffee Break	Atrium I Lobby	Lounge	206
	8:00 a.m.	11:00 p.m.	Office	Cooper	Office	10
	8:00 a.m.	11:00 p.m.	Meeting	Inverness	Hollow Square	30
Mon, 10/20/14	8:00 a.m.	5:00 p.m.	Teardown	Batik A/B	Classroom	40
	8:00 a.m.	5:00 p.m.	Teardown	Cardinal A/B	Classroom	40
	8:00 a.m.	5:00 p.m.	Teardown	Cooper	Office	
	8:00 a.m.	5:00 p.m.	Teardown	Dardanelles	Classroom	40
	8:00 a.m.	5:00 p.m.	Teardown	Fleur de Lis A/B	Classroom	40
	8:00 a.m.	5:00 p.m.	Teardown	Inverness	Hollow Square	30

Smarter Balanced 21: Standard Setting & Communications

Meeting Title:

Smarter Balanced Project 21 Kick-Off
 Radisson Plaza Hotel Minneapolis
 35 S 7th Street
 Minneapolis, MN 55402

Type of Meeting:

Project Kick-off Meeting

Date of Meeting:

2/24 – 2/25, 2014

Facilitator:

Measurement Incorporated, Project Contractor

Expected Attendees:

- **Smarter Balanced:** Jackie King, Juan D’Brot, Luci Willits, Dacia Hopfensberger, Marty McCall, Steve Slater, Patty Reiss, Joseph Martineau, Joe Willhoft, Nancy Arnold
- **OSPI:** Mike Middleton
- **PMP:** Christyan Mitchell
- **Contractor MI:** Hank Scherich, Mike Bunch, Nelson Androes, Mandy Hunter, Donna Merritt
- **CTB:** Antonia “Toni” Deoudes, David Breen, Rick Mercado, David Cosio
- **ETS:** Dirk Mattson, Bihua Xiang

Meeting Access Information:

- **Meeting Conference Call-in Number: 1 (877) 820-7831**
- **Access Code: 843853**

Agenda Topic : Day 1		Presenter
9:30 – 10:15	Welcome and Introductions Confirm Attendees Hotel Information	Donna Merritt
10:15 – 11:45	Task Cluster 7: Project Schedule & Tasks	
11:45 – 12:00	Break	
12:00 – 1:15	Lunch	
1:15 – 2:00	Task Cluster 1: Communication Plan Discussion	Donna Merritt
2:00 – 3:00	Overview of Functional Groups and Tasks Discussion	

Agenda Topic : Day 2		Presenter
9:30 – 11:00	Task Cluster 5: Standard Setting - Cross-Contract Project Schedule & Task Discussion	MI, CTB, ETS
11:00 – 11:30	Task Cluster 4: Participant Recruitment Activities Discussion	CTB
11:30 – 12:00	Upcoming Meeting Updates	
12:00 – 1:00	Lunch	
1:00 – 1:30	Project Schedule & Task Discussion Wrap-up	

Dates: February 24-25, 2014
 Location: Radisson Plaza Hotel Minneapolis
 Contractor pays: Up to 15 Consortium representatives

Kickoff Meeting Participants

First Name	Last Name	Meeting Component	State / Affiliation	Role	Email	Availability	Notes
Jaci	King	Communications	Executive Staff	Director of Higher Education Collaboration	jacqueline.king@smarterbalanced.org	1	
Juan	D'Brot	Communications	WV	Executive Committee member	jdbrot@access.k12.wv.us	1	
Luci	Willits	Communications	ID	Executive Committee member	lbwillits@sde.idaho.gov>	1	
Dacia	Hopfensberger	Communications	Executive Staff	Director, State Services	dacia.hopfensperger@smarterbalanced.org	1	
Marty	McCall	Technical	Executive Staff	Lead Psychometrician	marty.mccall@smarterbalanced.org	1	
Steve	Slater	Technical	OR	Co-Chair: Test Design / Validation and Psychometrics Work Group	steve.slater@state.or.us	1	
Patty	Reiss	Technical	HI	Co-Chair: Test Design / Validation and Psychometrics Work Group	patricia_reiss@notes.k12.hi.us	1	
Joseph	Martineau	Technical	MI	EC Liaison: Test Design / Validation and Psychometrics Work Group	martineauj@michigan.gov	1	
Joe	Willhoft	Other	Executive Staff	Executive Director	joe.willhoft@smarterbalanced.org	1	
Tony	Alpert	Other	Executive Staff	Chief Operating Officer	tony.alpert@smarterbalanced.org		
Mike	Middleton	Other	OSPI	Fiscal Agent	michael.middleton@k12.wa.us	1	
Christyan	Mitchell	Other	PMP	Project Management	cmitch@wested.org	1	
Magda	Chia	Communications, Technical	Executive Staff	Director of Support for Under-represented Students	magda.chia@smarterbalanced.org	TBD	Unsure of availability
Nancy	Arnold	Other	Contract Manager	Contract-21 Contract Manager	nancy.arnold@comcast.net>	TBD	Contract initiation in progress
Jane	Wellman	Communications	CA	Higher Education Regional Senior Advisor	janevwellman@gmail.com	TBD; Dial-in, if available	1/28: Jaci will inquire about availability

Smarter Balanced 21

Travel Information

Hotel Accommodations

Radisson Plaza Hotel Minneapolis
35 South 7th St.
Minneapolis, MN 55402
612.339.4900

http://www.radisson.com/minneapolis-hotel-mn-55402/mpls_dt

- Accommodations have been made for **all** attendees (including other contractors) for Sunday and Monday, February 23 and 24. Please let us know if you require accommodations for Tuesday, February 25 as soon as possible.
- Accommodations for Smarter Balanced and PMP attendees have been direct billed to Measurement Incorporated; other contractors will be responsible for room costs.
- Confirmation numbers will be provided once available.

Meeting Rooms

- All meeting rooms are located on the second floor
 - Kickoff Meeting, Fjords 2
 - Breakout Conference Room, Fjords 1
 - Lunch, Fjords 3
- The kickoff meeting room and the breakout conference room will each be equipped with a conference phone and secure conference line. Dial-in numbers and access codes will be provided once available.

Meals

Please note that the hotel restaurant is currently closed due to renovations.

- The hotel serves a daily breakfast buffet from 6-10 AM; the cost is \$15 per person. Due to the on-going renovations the buffet will be served on the second floor, adjacent to the meeting rooms.
- Coffee and tea service will be available all day in the kickoff meeting room.
- We will break for lunch at noon each day.
- In-room dining is available in the evenings.
- Lunch will be direct billed to Measurement Incorporated; all other dining expenses (breakfast buffet and dinner) will be reimbursed for Smarter Balanced and PMP attendees.

Airfare

Smarter Balanced and PMP attendees should Contact **Connie Garrett** at **Cardinal Travel Service** to arrange flights. All airfare will be direct billed to Measurement Incorporated.

- E-mail: connie.g@cardinaltravelservice.com
- During office hours (8 AM – 5 PM EST) dial: 919.433.0021
- After hours dial: 800.424.2479

Ground Transportation

Transportation from Minneapolis–St. Paul International Airport to Radisson Plaza Hotel

- Metro Transit Rail
 - 23 minute ride, runs every 15 minutes
 - Take Blue Line towards downtown Minneapolis to the Nicollet Mall Station, Radisson Plaza Hotel is two blocks away
 - <https://www.metrotransit.org/maps-schedules>
- Taxi Cab
 - 12 miles, 25 minute ride

Reimbursement

- Measurement Incorporated will provide payment vouchers for Smarter Balanced and PMP attendees for any additional expenses that are incurred during the kick off meeting. Please save receipts.

Agenda for July Task Force Meeting

1. Achievement Level Setting Plan and Supplement
 - a. Policy implications
 - b. Further refinements
2. Purpose and Design of Field Tests of Online, In-Person, and VAC
 - a. Purpose – to test viability of training materials, timing, procedures
 - b. Design – use MI facilities and staff
 - c. Follow-up – present findings to Task Force in September and revise as necessary
3. Timing of Chiefs Review of Cut Scores
 - a. Issue – variability in authority of chiefs
 - b. Possibility of advance involvement to set boundaries (e.g., Hofstee approach)
 - c. Security and communication issues
 - d. Guidance sought
4. Plans for Work Beyond Contract End
 - a. Review of proposal/plans
 - b. Policy issues
 - c. Technical issues
 - d. Guidance sought

Contract - 21 Standard Setting

Meeting Title:

Achievement Level Setting Advisory Panel
(in person)

Type of Meeting:

Contract Management

Date of Meeting:

September 22, 2014
8:30 a.m. – 4:00 p.m. Central

Facilitator:

Michael Bunch, Measurement Incorporated
Wireless Access:

Scribe:

Barbara Scherich

Expected Attendees:

- **Smarter Balanced:** Joe Willhoft, Marty McCall, Jaci King, Joseph Martineau
- **Contract Manager:** Nancy Arnold
- **MI:** Michael Bunch, Barbara Scherich
- **Task Force Members:** Luz Bay, Mike Briscoe, Chad Buckendahl, Brian Gong, Bob Lissitz, Susan Loomis, Susan Phillips, Joe Ryan, Bill Schmidt

Advance Documents: July meeting minutes; ALS Field Test Report; Final Report Outline; ALS Follow-Up Plans Proposal

Agenda Topic & Presenter	Approximate Times	Discussion	Action Items
Hot Breakfast Buffet	7:30 a.m.		
Welcome and Introductions: Joe Willhoft	8:30 a.m.		
Approval of July meeting minutes (Advance Copy): Barbara Scherich	8:45 a.m.		
Review of ALS Field Test and subsequent changes (Advance Copy): Mike Bunch	9:15 a.m.		
Update on reporting to Chiefs: Mike Bunch/Joe Willhoft	10:45 a.m.		
Final Report Outline (Advance Copy): Mike Bunch	11:15 a.m.		
Lunch	11:45 a.m.		
Review of ALS Follow-Up Plans Proposal (Advance Copy): Mike	12:30 p.m.		
Plans for October: Mike Bunch	2:15 p.m.		
Topics for Next Meeting: All	3:15 p.m.		
Adjourn	4:00 p.m.		

Contract - 21 Standard Setting

Meeting Title: Achievement Level Setting Advisory Panel (webinar)	Type of Meeting: Contract Management	Date of Meeting: October 28, 2014 9:00 a.m. – 1:00 p.m. Pacific
---	--	--

Facilitator: Michael Bunch, Measurement Incorporated	Scribe: Barbara Scherich
--	------------------------------------

Expected Attendees:

- **Smarter Balanced:** Joe Willhoft, Marty McCall, Jaci King, Joseph Martineau
- **Contract Manager:** Nancy Arnold
- **MI:** Michael Bunch, Barbara Scherich
- **Task Force Members:** Luz Bay, Mike Briscoe, Chad Buckendahl, David Conley, Bob Lissitz, Susan Loomis, Susan Phillips, Joe Ryan, Bill Schmidt, Stephen Sireci

Advance Documents: ALS Advisory Panel Agenda Recap 9-22-14; SBAC Standard Setting Report (2014-10-23 FINAL); Online Panelist Submissions FINAL; DemographicDataWithNationalComparison8-25; In-Person Panelists_No Names; In-PersonPanelistCountByStateandRolexls; Facilitator Script Overview Final; Chiefs Agenda 2014_11_6_DRAFT 10 24 14; Advisory Panel Participant List

Meeting Access Information:

WebEx link:

<https://meetings.webex.com/collabs/meetings/join?uuid=M2FWMI6181ZI5CP2LHV6MOGP13-65NX>

Call-in Number: 1-855-244-8681

Meeting Number: 197 921 546

Agenda Topic & Presenter	Approximate Times	Discussion	Action Items
Welcome and Introductions: Joe Willhoft	9:00 a.m.		
Approval of September meeting minutes (Advance Copy): Barbara Scherich	9:15 a.m.		
Update on Achievement Level Setting: Joe Willhoft	9:30 a.m.		
Achievement Level Setting/VAC report: Mike Bunch <ul style="list-style-type: none"> • Review of Plan • Online Panel results • In-Person logistics and issues • R3 results • VAC procedures and results 	9:45 a.m.		
ALS Auditor Report (Advance Copy): Joe Willhoft	12:00 p.m.		
Update on reporting to Chiefs: Joe Willhoft	12:30 p.m.		
Adjourn	1:00 p.m.		

October 28, 2014

Advisory Panel Members	
Luz Bay	The College Board
Michael Briscoe	Student Achievement Partners
Chad Buckendahl	Alpine Testing Solutions
David Conley	Smarter Balanced TAC
Brian Gong	Smarter Balanced TAC
Robert Lissitz	University of Maryland
Susan Loomis	Consultant
Susan Phillips	Consultant
Joseph Ryan	Smarter Balanced TAC
William Schmidt	Michigan State University
Stephen Sireci	University of Massachusetts Amherst

Smarter Balanced Assessment Consortium and Associated Personnel	
Joe Willhoft	Smarter Balanced Executive Director
Marty McCall	Smarter Balanced Lead Psychometrician
Joseph Martineau	Smarter Balanced Executive Committee Co-Chair
Jaci King	Smarter Balanced Director, Higher Education Collaboration
Nancy Arnold	Contract 21 Manager

Measurement Incorporated	
Michael Bunch	Senior Vice President
Barbara Scherich	Program Manager

Dear [Task Force member]:

Thank you for serving on the Smarter Balanced Achievement Level Setting Task Force. The first Task Force meeting will take place in Portland, Oregon, July 17-18, 2014, at the Sheraton Portland Airport Hotel (<http://www.sheratonportlandairport.com/>). Overnight accommodations will be at the same location.

Please contact Connie Garrett at Cardinal Travel to arrange your flights.

- E-mail: connie.g@cardinaltravelservice.com
- Telephone: 919.433.0021 (8 a.m.–5 p.m. Eastern; 5 a.m.– 2 p.m. Pacific)

Please keep the Task Force meeting schedule in mind when scheduling your flights:

- Thursday, July 17: 1 p.m.–5 p.m. (Pacific)
- Friday, July 18: 8 a.m.–12 p.m. (Pacific)

Connie Garrett will inform me when your flights have been booked. I will then reserve your overnight accommodations for the appropriate nights and send you hotel confirmation information as soon as it becomes available.

Your flights and overnight accommodations will be direct-billed to Measurement Incorporated. Costs associated with the Task Force meeting will be reimbursed by MI. Expense claim vouchers will be distributed at the meeting.

If you have any questions, please do not hesitate to contact me.

Thank you,

Mandy Hunter

Achievement Level Setting

Joe Willhoft
Mike Bunch



Achievement Level Setting Advisory Panel Meeting
October 28, 2014



Update

Joe Willhoft





Since Our Last Meeting

- Communication/Contingency Planning
- Completion of Recruiting Activities
- Completion of the Online Panel
- Completion of the In-Person Panel
- Completion of Vertical Articulation
- Preparation for Chiefs Meeting

ALS/VAC Report

Mike Bunch





Communication/Contingency Planning

Goals during contingency planning:

- Tell the truth.
- Remember that stakeholder concerns are your concerns.
- Establish Smarter Balanced as the best initial and ongoing source for information.
- Provide accurate and factual information in a timely manner.
- Reinforce key messages.
- Exercise communications discipline in a challenging environment.
- Continuously assess the intensity of the situation.

.



Communication/Contingency Planning

Smarter Balanced Contingency Response:

- Some cancellations due to health concerns
- Set goal of transparency, elevation appropriate with level of concern, and effective communication.
- Provided clear information and timely responses to panelists, state observers and other participants.
- Followed only official guidance from recognized authorities and continued to consult with hotel management, municipal authorities, and CDC to provide participants with reliable, up-to-date information
- Continuous monitoring of media, inquiries from panelists and other sources of information.



Completion of Recruiting Activities

- Online Panel
 - Over 10,000 registered
- In-Person Panel
 - 504 panelists confirmed
 - 10+ alternates confirmed
- VAC
 - 64 panelists confirmed
- Challenges/Solutions



Completion of the Online Panel

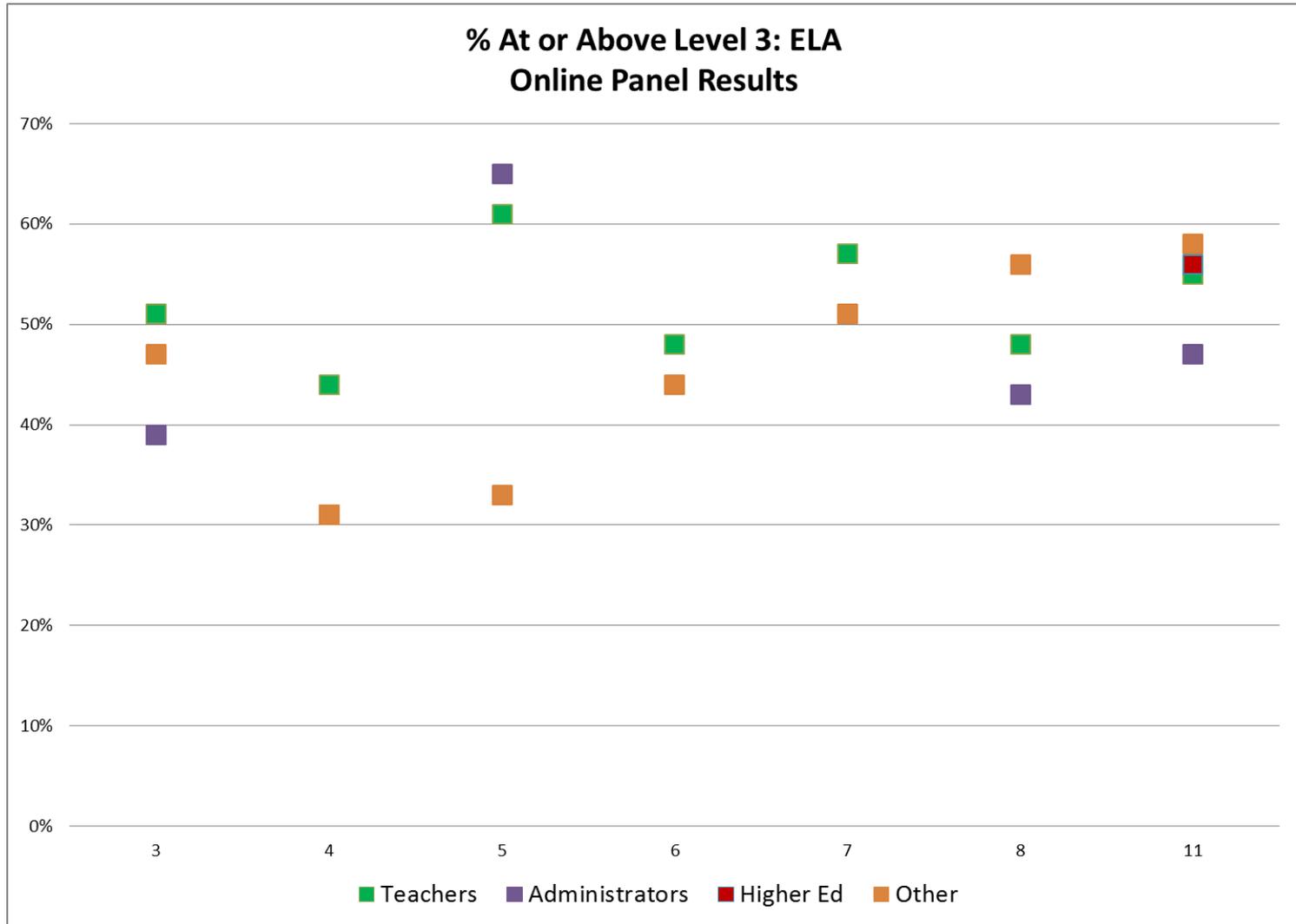
- Opening of the Window
 - October 6-17
 - Individual windows and total window extended
- Level of Participation
 - 10,099 registered
 - 5,840 logged in
 - 2,660 submitted
- Support Provided
- Results

Online Panel Participation

Grade	Content	Total
3	English Language Arts/Literacy	258
3	Mathematics	239
4	English Language Arts/Literacy	138
4	Mathematics	178
5	English Language Arts/Literacy	144
5	Mathematics	175
6	English Language Arts/Literacy	90
6	Mathematics	138
7	English Language Arts/Literacy	92
7	Mathematics	135
8	English Language Arts/Literacy	129
8	Mathematics	183
High School	English Language Arts/Literacy	321
High School	Mathematics	440
Total		2660

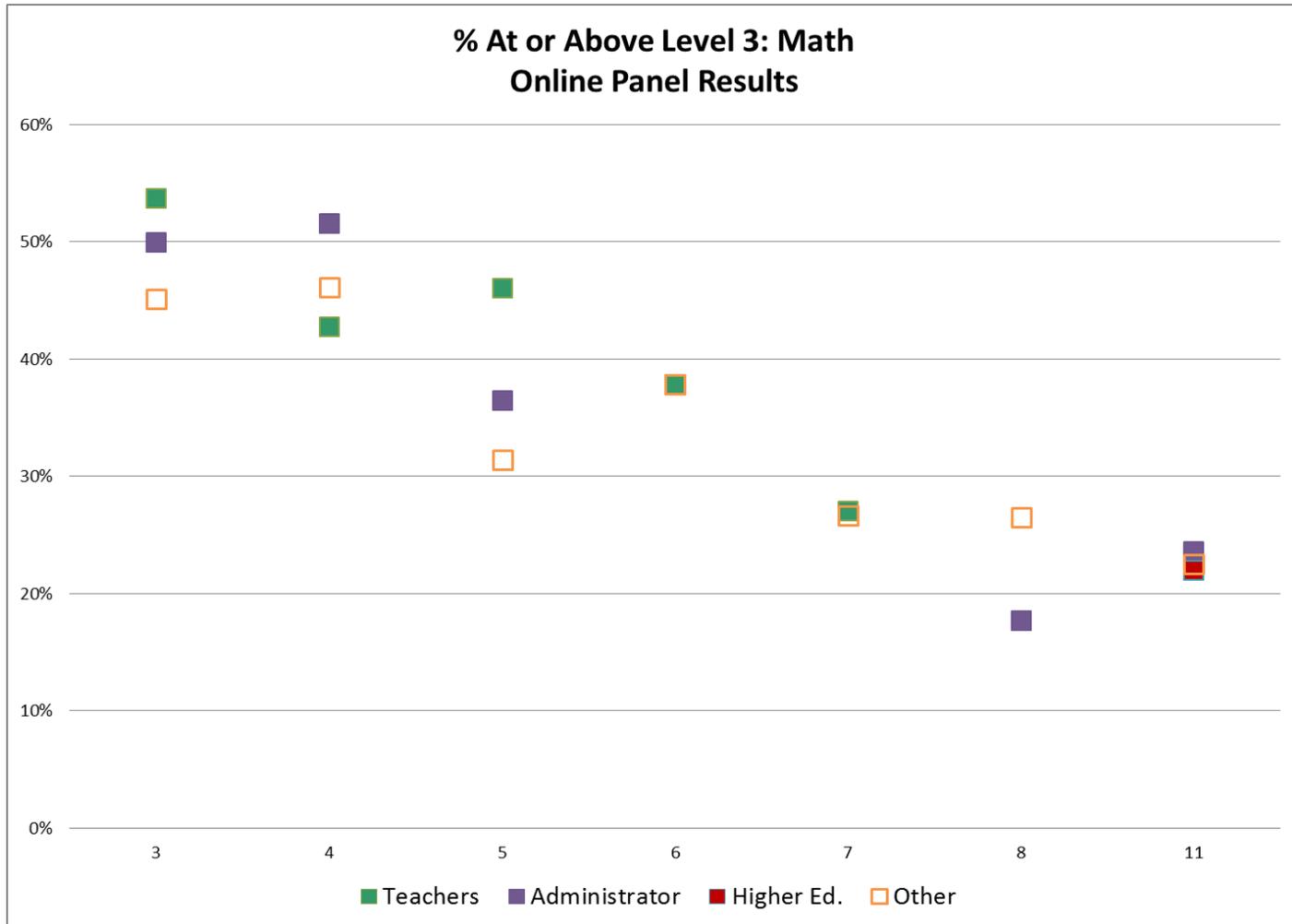


ELA





Math



Completion of the In-Person Panel

- Logistics
 - Travel
 - Computer setup/testing
 - [Traffic control](#)
 - [Breakout rooms](#)
- Training Activities
 - Software
 - Common Core and ALDs
 - Ordered Item Booklet
- Panel Activities
 - Bookmark placement
 - Discussion
- Results
 - Cut scores
 - Impact

Round 3 Results: ELA

% In Group				
Grade	Level 1	Level 2	Level 3	Level 4
3	33%	26%	21%	19%
4	36%	22%	23%	19%
5	21%	15%	47%	17%
6	33%	24%	29%	13%
7	32%	28%	34%	7%
8	24%	25%	41%	10%
11	27%	25%	36%	12%

Round 3 Results: Math

% In Group				
Grade	Level 1	Level 2	Level 3	Level 4
3	33%	28%	28%	11%
4	22%	33%	29%	16%
5	36%	32%	18%	14%
6	42%	29%	14%	16%
7	47%	34%	14%	6%
8	49%	26%	18%	7%
11	41%	35%	18%	6%

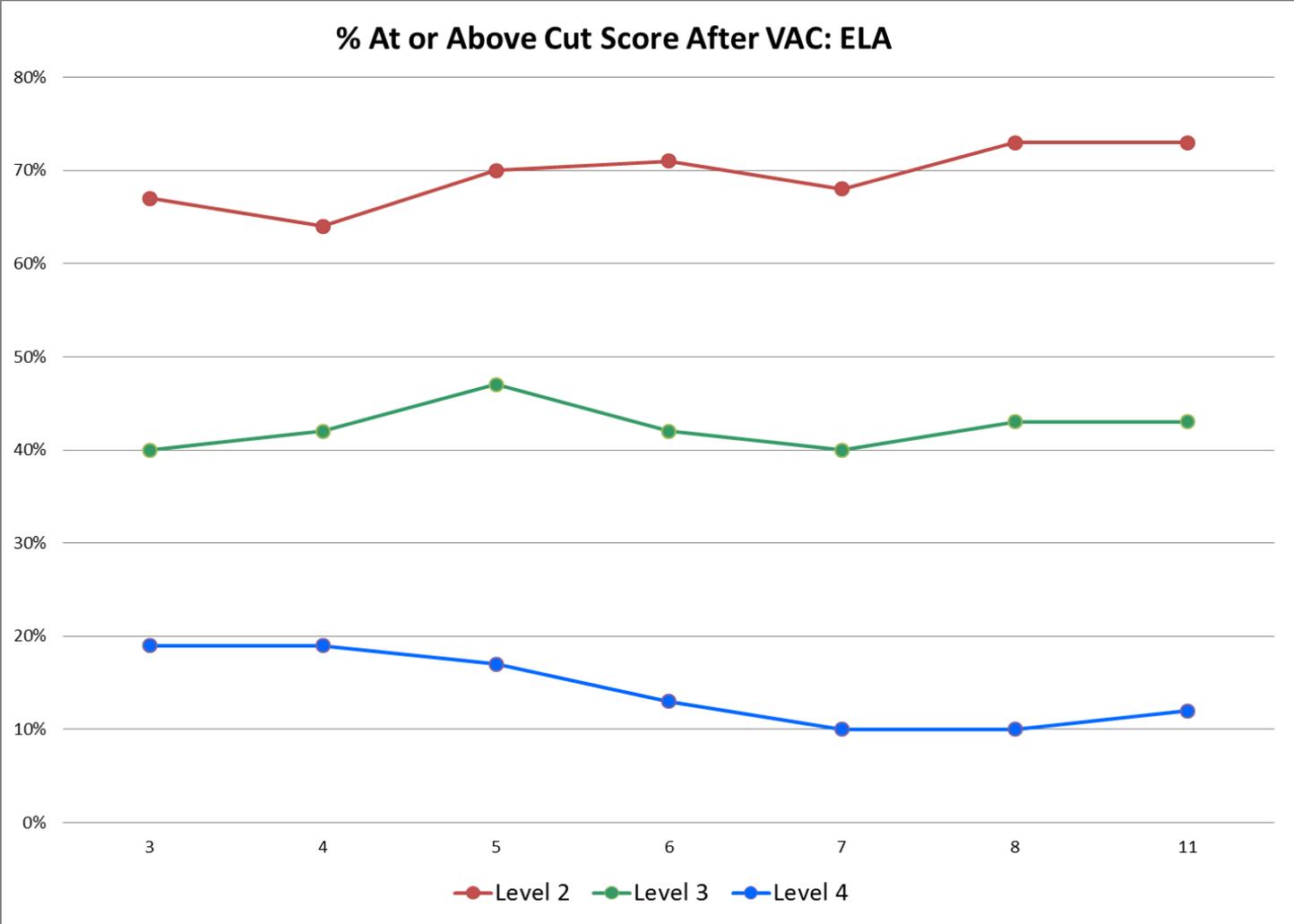
Completion of Vertical Articulation

- Training
- Procedures
- Results
 - Cut scores changed
 - Alignment across grades

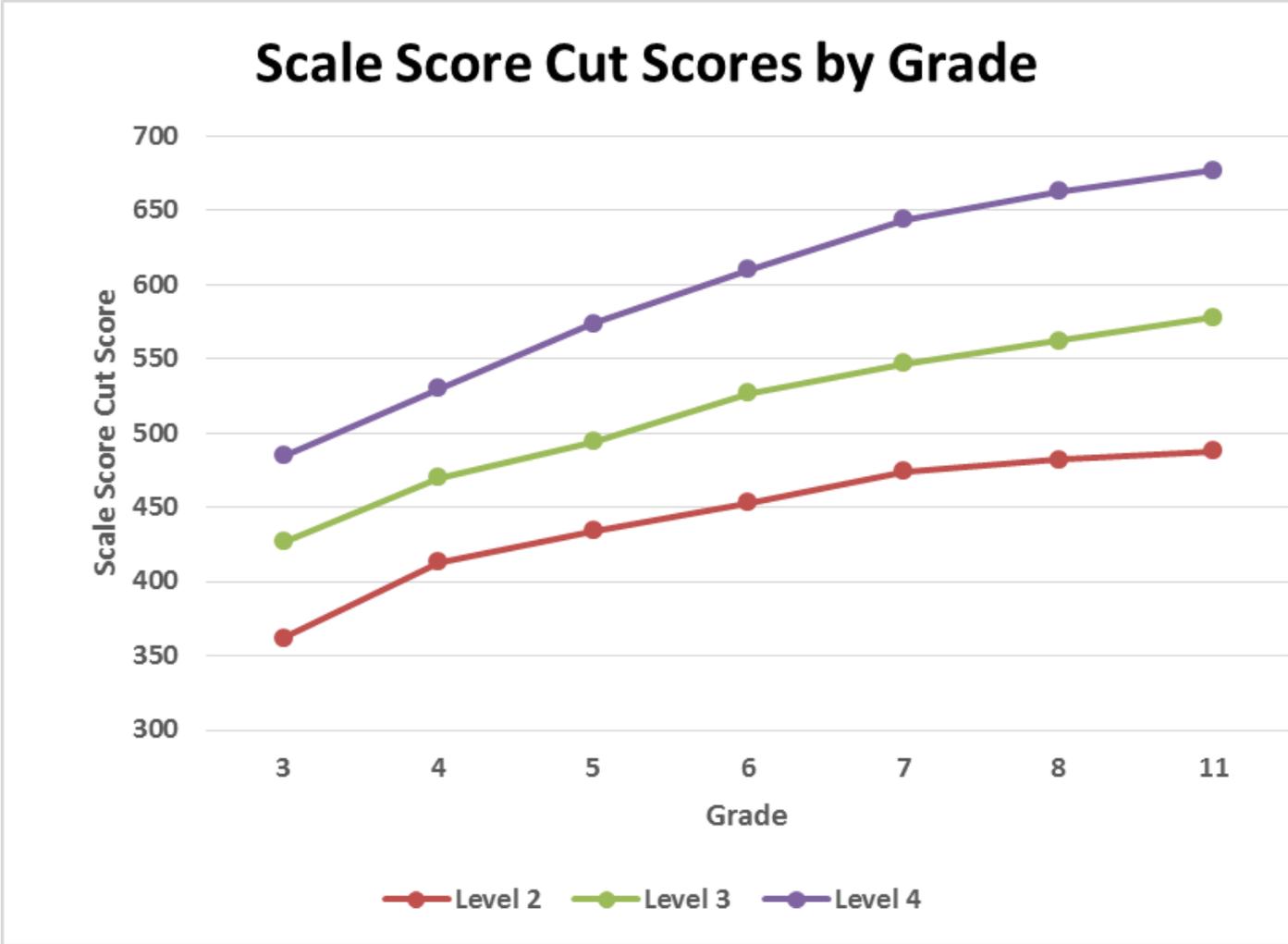
Cut Score Changes: ELA

Articulated Bookmarks			
Grade	Level 2	Level 3	Level 4
3	13	33	54
4	19	43	62
5	18	37	63
6	11	34.5	60.5
7	16	38	65
8	21	44	68
11	17	45	65

Results of Cut Score Changes: ELA



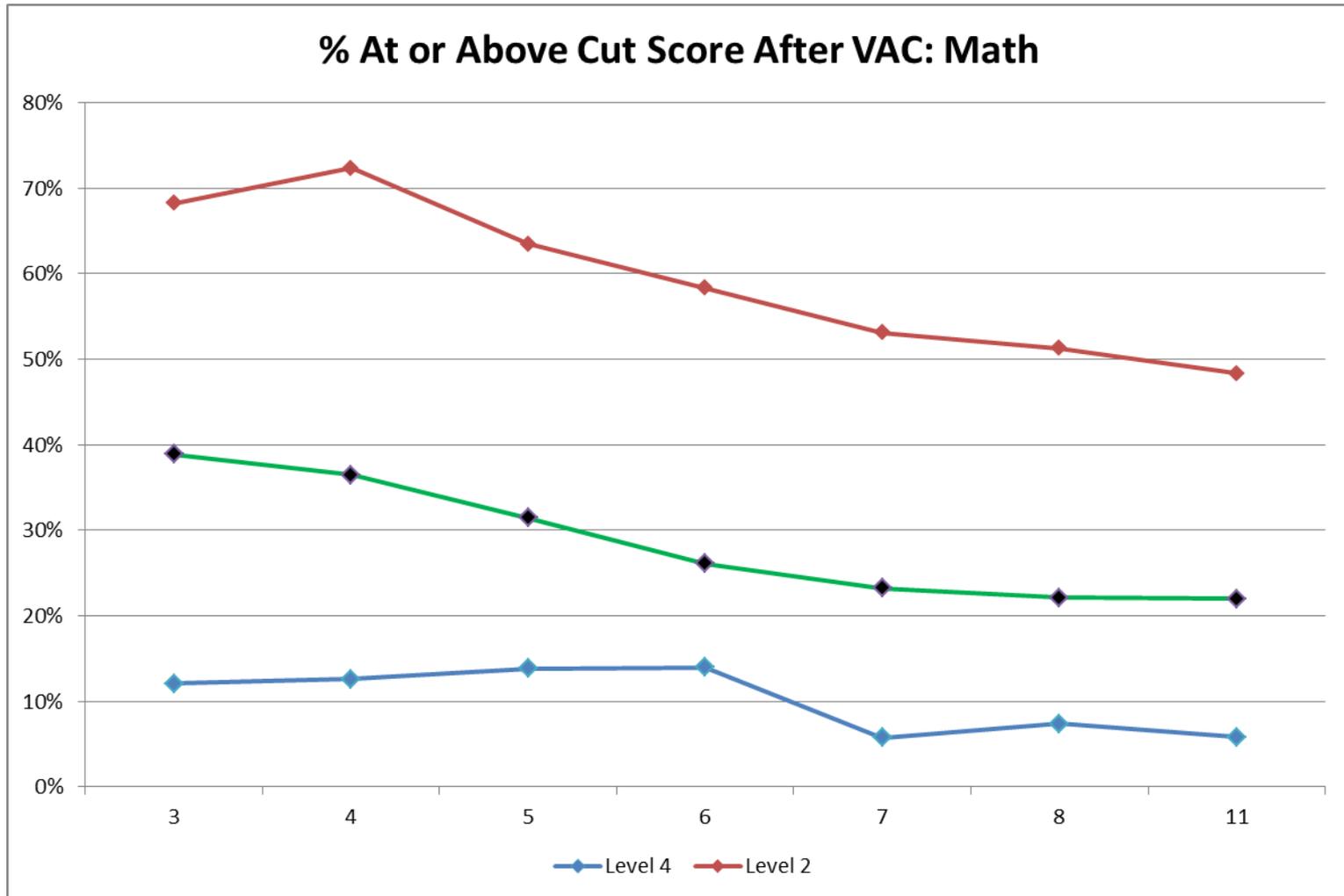
Results of Cut Score Changes: ELA



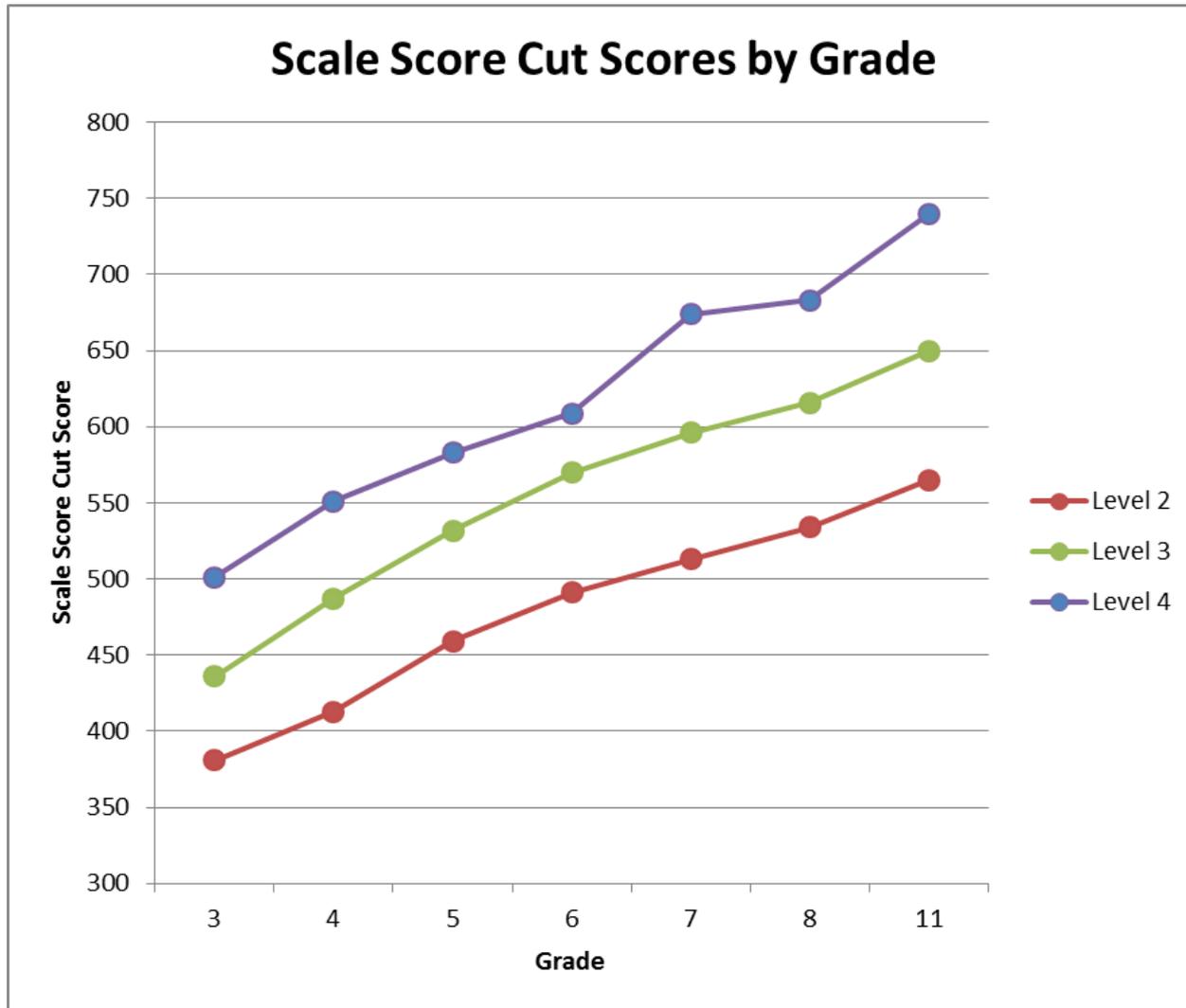
Cut Score Changes: Math

Articulated Bookmarks			
Grade	Level 2	Level 3	Level 4
3	26	47	69
4	18	49	73
5	19	50	64
6	18	48	63
7	17	40	61
8	16	41	60
11	23	50	68

Results of Cut Score Changes: Math



Results of Cut Score Changes: Math



ALS Auditor Report

Joe Willhoft



Update on Reporting to Chiefs

Joe Willhoft



To: Achievement Level Setting In-Person Panelists
From: MI_smarterbalanced21@measinc.com
Subject: Smarter Balanced Achievement Level Setting In-Person Panel

This email is being sent on behalf of the Smarter Balanced Assessment Consortium. For questions please email MI_smarterbalanced21@measinc.com.

Dear _____:

Thank you for your application to participate in the In-Person Panel for Achievement Level Setting for the Smarter Balanced Assessment Consortium. We are pleased to inform you that you have been selected as a participant on the ____ panel.

Measurement Incorporated (MI) will provide your applicable airfare, airport shuttle, hotel, and meals. Additional expenses such as parking and meals not provided during the meeting will be reimbursed. Attached you will find a sample travel reimbursement form for your reference. You will receive a hard copy of the form at the In-Person Panel.

Measurement Incorporated will also send you materials to review prior to the In-Person Panel so that you may acquire a better understanding of the tests, content standards, and achievement level descriptors prior to the Achievement Level Setting. We will also provide extensive on-site training in the tasks you will perform.

The In-Person Panel will be held in Dallas, Texas, on October 13–19, 2014. Specific details related to the logistics for travel are outlined below.

Approximately one panelist in nine will be asked to return to Dallas to serve on the Vertical Articulation Committee (VAC) on October 20, 2014. Please indicate your interest in serving on the VAC on the In-Person Panelist information form, and keep this date open on your calendar in case you are selected for the VAC. You will be notified by October (14, 16, 18) if you are selected to serve on the VAC.

Attached you will find a non-disclosure agreement that you will need to sign and return electronically before we can proceed further. We recognize that you will need to let family and friends know where you will be for three days, but we would like to avoid widespread attention being focused on this meeting. Therefore, for your own comfort and security, we ask that you not broadcast the location of this meeting. We appreciate your cooperation in this effort.

Traveling to and from Dallas:

Please follow the booking instructions outlined below to book air travel through **Cardinal Travel** to and from Dallas (see below for Cardinal Travel contact information).

- Hotel location:
 - Hilton Anatole
2201 North Stemmons Freeway
Dallas, TX 75207
(214) 761–7209
 - Cardinal Travel will inform us of your flight arrangements. Measurement Incorporated will then make your hotel reservations for you and provide you with your hotel reservation confirmation number.

- If you plan to drive to Dallas, Measurement Incorporated will reimburse your mileage up to the price of round-trip coach airfare. Please contact Mandy Hunter at MI_smarterbalanced@measinc.com if you plan to drive to the hotel.
- Airport Shuttle Information:
 - You will receive a prepaid voucher for transportation to and from the airport in Dallas.
 - Pick-up and drop-off areas are located on the voucher.
 - No other mode of ground transportation will be reimbursed unless pre-approved by Measurement Incorporated.
- Other Transportation:
 - Measurement Incorporated will provide a shuttle service serving the area within a 2-mile radius of the hotel. Details will be available on site. For excursions further afield (e.g., sightseeing, shopping), you will be on your own; Measurement Incorporated will not reimburse these transportation costs.
- Meals:
 - Please save receipts for meals you eat en route to and from the hotel on travel days. Measurement Incorporated will reimburse your meal expenses in accordance with the limits shown on the travel reimbursement form.
 - Breakfast and lunch will be provided on site daily. In addition, we will serve beverages and a light afternoon snack.
 - Measurement Incorporated will reimburse your dinner meal expenses in accordance with the limits shown on the travel reimbursement form. Receipts will be required for all reimbursed expenses.

Next Steps

- Please contact Connie Garret at Cardinal Travel to begin the air travel booking process.
 - E-mail: connie.g@cardinaltravelservice.com
 - Phone: 1.919.433.0021 during office hours (8 a.m. – 5 p.m. EDT)
 - Identify meeting as: Smarter Balanced Achievement Level Setting in Dallas
 - Be prepared to provide the following information when contacting Cardinal Travel (please make sure all information is correct before contacting Cardinal Travel):

Panel for which you have been selected	[filled in by MI, above and again here]
Departure Date and time	[filled in by panelist]
Leaving from (City and Airport)	[filled in by panelist]
Return date	[filled in by panelist]
Seating preference (window, aisle, no preference)	[filled in by panelist]

- To the extent possible, Cardinal Travel will book travel to arrive the day prior to your panel’s start date and depart the afternoon of your panel’s last day, no earlier than 3:00 p.m. If you are selected for the follow-up meeting on October 20, we will amend your flight and airport transportation plan.
- Please make all reservations by July 31, 2014.

If you are a teacher and need a substitute for your classes during your absence, we will reimburse your school or district for the actual cost of the substitute teacher. A stipend of \$150 per day for a panel session (\$450 maximum for the 3-day panel) is available for those participants who are not

employed, are self-employed, or who take an unpaid leave from work to attend. Educators will not receive a stipend for any days for which the district is reimbursed for substitute costs.

To confirm your participation on the In-Person Panel, please complete the attached In-Person Panelist information form, sign the non-disclosure agreement, and return both forms via MI_smarterbalanced@measinc.com by [1 week after we send the e-mail]. To complete and submit the information form and non-disclosure agreement, please fill in all fields electronically, save the documents for your files, and return them to Measurement Incorporated as e-mail attachments via the address above.

Attached you will find samples of the stipend/honorarium claim voucher and substitute teacher reimbursement form. Hard copies of these documents will be distributed at the In-Person Panel.

If you have any questions, please contact Mandy Hunter at MI_smarterbalanced21@measinc.com.

Thank you for your participation, and congratulations on your selection. We look forward to working with you.

Attachments:

In-Person Panelist Information Form
Non-Disclosure Agreement
Travel Reimbursement Form
Stipend/Honorarium Claim Voucher
Substitute Teacher Reimbursement Form

Smarter Balanced Resource Links

<http://www.smarterbalanced.org/smarter-balanced-assessments>

Preliminary Test Blueprints

- [ELA/Literacy Smarter Balanced Preliminary Summative Assessment Blueprint – 5/9/14 \(PDF\)](#)
- [Mathematics Smarter Balanced Preliminary Summative Assessment Blueprint – 5/9/14 \(PDF\)](#)
- [Supporting Document: Scoring Reporting and Estimated Testing Times \(PDF\)](#)

Content Specifications

- [ELA/Literacy Content Specifications Appendix B: Grade Level Tables for All Claims and Assessment Targets and Item Types \(2/4/14 draft\)](#)
- [Mathematics Content Specifications \(6/2013 draft\)](#)

Item/Task Specifications

English Language Arts/Literacy Item Specification

- [ELA CAT Item Specs Grades 3-5 \(ZIP\) \(Update 2/4/14\)](#)
- [ELA CAT Item Specs Grades 6-8 \(ZIP\) \(Update 2/4/14\)](#)
- [ELA CAT Item Specs Grades 9-11 \(ZIP\) \(Updated 2/4/14\)](#)
- [ELA PT Item Specs Opinion Grades 3-5 \(ZIP\) \(5/20/14\)](#)
- [ELA PT Item Specs Narrative Grades 3-5 \(ZIP\) \(5/20/14\)](#)
- [ELA PT Item Specs Informative Grades 3-5 \(ZIP\) \(5/20/14\)](#)
- [ELA PT Item Specs Explanatory Grades 6-8,11 \(ZIP\) \(5/20/14\)](#)
- [ELA PT Item Specs Argumentative Grades 6-8,11 \(ZIP\) \(5/20/14\)](#)
- [ELA Stimulus Specifications \(PDF\)](#)
- [Issues Related to Stimulus and Item Development \(PDF\)](#)

Mathematics

- [Mathematics Grades 3-5 \(ZIP\) \(Update 2/4/14\)](#)
- [Mathematics Grades 6-8 \(ZIP\) \(Update 2/4/14\)](#)
- [Mathematics High School \(ZIP\) \(Update 2/4/14\)](#)
- [Mathematics PT Item Specs All Grades \(PDF\) \(5/27/14\)](#)
- [Smarter Balanced Mathematics General Rubrics \(PDF\) \(DocX\)](#)

Guidelines

- [Accessibility and Accommodations Guidelines](#) (PDF)
- [ELA Audio Guidelines](#) (PDF)
- [Formulas & Conversions Guidelines](#) (PDF)
- [Calculator Availability by Grade Level](#) (PDF)
- [Mathematics Audio Guidelines](#) (PDF)
- [Scoring Guide for Selected Short-Text Mathematics Items](#) (PDF) New!
- [ELL Guidelines](#) (PDF)
- [Signing Guidelines](#) (PDF)
- [Tactile Guidelines](#) (PDF)
- [Bias and Sensitivity Guidelines](#) (PDF)
- [Style Guide](#) (PDF)

Use this link for the resources below: <http://sbac.portal.airast.org/practice-test/>

- [Calculators](#)
- [Manuals and User Guides](#) (Under Resources and Documentation)
- [Classroom Activities](#) (Under Resources and Documentation)
- [Scoring Guides](#) (Under Resources and Documentation)
- [Performance Task Writing Rubrics](#) (Under Resources and Documentation)

To: Achievement Level Setting Alternate In-Person Panelists
From: MI_smarterbalanced21@measinc.com
Subject: Smarter Balanced Achievement Level Setting In-Person Panel

This email is being sent on behalf of the Smarter Balanced Assessment Consortium. For questions please email MI_smarterbalanced21@measinc.com.

Dear _____:

Thank you for your application to participate in the In-Person Panel for Achievement Level Setting for the Smarter Balanced Assessment Consortium. We are pleased to inform you that you have been selected as an alternate for the [Grades 3–5, Grades 6–8, Grade 11] panels.

As an alternate, your role is much like that of an alternate juror. You will be assigned to a panel when you arrive. You may be needed on a panel in a different content area (English language arts/literacy or mathematics) or at a different grade level than the one for which you were nominated by your state. You will participate in all of the panel activities. Your decisions will be recorded along with those of the other panelists but will not be included in the cut score calculations unless you are called on to replace a panelist.

Measurement Incorporated (MI) will provide your applicable airfare, airport shuttle, hotel, and meals. Additional expenses such as parking and meals not provided during the meeting will be reimbursed. Attached you will find a sample travel reimbursement form for your reference. You will receive a hard copy of the form at the In-Person Panel.

Measurement Incorporated will also send you materials to review prior to the In-Person Panel so that you may acquire a better understanding of the tests, content standards, and achievement level descriptors prior to the Achievement Level Setting. We will also provide extensive on-site training in the tasks you will perform.

The In-Person Panel will be held in Dallas, Texas, on [appropriate dates], 2014. Specific details related to the logistics for travel are outlined below.

Attached you will find a non-disclosure agreement that you will need to sign and return electronically before we can proceed further. We recognize that you will need to let family and friends know where you will be for three days, but we would like to avoid widespread attention being focused on this meeting. Therefore, for your own comfort and security, we ask that you not broadcast the location of this meeting. We appreciate your cooperation in this effort.

Traveling to and from Dallas:

Please follow the booking instructions outlined below to book air travel through **Cardinal Travel** to and from Dallas (see below for Cardinal Travel contact information).

- Hotel location:
 - Hilton Anatole
2201 North Stemmons Freeway
Dallas, TX 75207
(214) 761–7209
 - Cardinal Travel will inform us of your flight arrangements. Measurement Incorporated will then make your hotel reservations for you and provide you with your hotel reservation confirmation number.
 - If you plan to drive to Dallas, Measurement Incorporated will reimburse your mileage up to the price of round-trip coach airfare. Please contact Mandy Hunter at MI_smarterbalanced21@measinc.com if you plan to drive to the hotel.

- Airport Shuttle Information:
 - You will receive a prepaid voucher for transportation to and from the airport in Dallas.
 - Pick-up and drop-off areas are located on the voucher.
 - No other mode of ground transportation will be reimbursed unless pre-approved by Measurement Incorporated.
- Other Transportation:
 - Measurement Incorporated will provide a shuttle service serving the area within a 2-mile radius of the hotel. Details will be available on site. For excursions further afield (e.g., sightseeing, shopping), you will be on your own; Measurement Incorporated will not reimburse these transportation costs.
- Meals:
 - Please save receipts for meals you eat en route to and from the hotel on travel days. Measurement Incorporated will reimburse your meal expenses in accordance with the limits shown on the travel reimbursement form.
 - Breakfast and lunch will be provided on site daily. In addition, we will serve beverages and a light afternoon snack.
 - Measurement Incorporated will reimburse your dinner meal expenses in accordance with the limits shown on the travel reimbursement form. Receipts will be required for all reimbursed expenses.

Next Steps

- Please contact Connie Garret at Cardinal Travel to begin the air travel booking process.
 - E-mail: connie.g@cardinaltravelservice.com
 - Phone: 1.919.433.0021 during office hours (8 a.m. – 5 p.m. EDT)
 - Identify meeting as: Smarter Balanced Achievement Level Setting in Dallas
 - Be prepared to provide the following information when contacting Cardinal Travel (please make sure all information is correct before contacting Cardinal Travel):

Panel for which you have been selected	[panel]
Departure Date and time	
Leaving from (City and Airport)	
Return date	
Seating preference (window, aisle, no preference)	

- To the extent possible, Cardinal Travel will book travel to arrive the day prior to your panel's start date and depart the afternoon of your panel's last day, no earlier than 3:00 p.m.
- Please make all reservations by [date], 2014.

If you are a teacher and need a substitute for your classes during your absence, we will reimburse your school or district for the actual cost of the substitute teacher. A stipend of \$150 per day for a panel session (\$450 maximum for the 3-day panel) is available for those participants who are not employed, are self-employed, or who take an unpaid leave from work to attend. Educators will not receive a stipend for any days for which the district is reimbursed for substitute costs.

To confirm your participation as an alternate for the In-Person Panel, please complete the attached Alternate In-Person Panelist information form, sign the non-disclosure agreement, and return both forms via MI_smarterbalanced21@measinc.com by [day, date]. To complete and submit the information form and non-disclosure agreement, please fill in all fields electronically, save the documents for your files, and return them to Measurement Incorporated as e-mail attachments via the address above.

Attached you will find samples of the stipend and honorarium claim voucher and substitute teacher reimbursement form. Hard copies of these documents will be distributed at the In-Person Panel.

If you have any questions, please contact Mandy Hunter at MI_smarterbalanced21@measinc.com.

Thank you for your participation, and congratulations on your selection. We look forward to working with you.

Attachments:

Alternate In-Person Panelist Information Form
Non-Disclosure Agreement
Travel Reimbursement Form
Stipend and Honorarium Claim Voucher
Substitute Teacher Reimbursement Form

To: Achievement Level Setting In-Person Panelists
From: MI_smarterbalanced21@measinc.com
Subject: Smarter Balanced Achievement Level Setting In-Person Panel

This email is being sent on behalf of the Smarter Balanced Assessment Consortium. For questions please email MI_smarterbalanced21@measinc.com.

Dear [Grades 3–5 Panelists, Grades 6–8 Panelists, Grade 11 Panelists]:

Thank you for your participation in the In-Person Panel for Achievement Level Setting for the Smarter Balanced Assessment Consortium. Outlined below you will find additional travel information as well as links to information to review prior to your arrival in Dallas.

Thank you if you volunteered to serve on the Vertical Articulation Committee (VAC) on October 20. The VAC participants have been selected. If you have not been notified that you were selected, you no longer need to reserve the October 20 date. Thanks again for your interest.

Upon your arrival in Dallas, transportation to the Hilton Anatole will be provided via Super Shuttle. Attached you will find a prepaid voucher for Super Shuttle that includes gratuity. Instructions on how to find or schedule the shuttle are included on the voucher. **Please print the voucher and take it with you to Dallas. You will need a hard copy of the voucher in order to use the shuttle service. No other form of ground transportation will be reimbursed.** The voucher is needed for your arrival only; Measurement Incorporated will arrange your return trip to the airport.

All hotel accommodations have been booked on your behalf under the Measurement Incorporated room block. No confirmation number will be necessary at check-in. Your reservation will be under your name at the hotel front desk.

If you have not contacted Connie Garrett (connie.g@cardinaltravelservice.com) at Cardinal Travel to book your flight, please do so immediately.

Breakfast and registration will begin at 7:30 a.m. on [October 13, October 15, October 17] in Atrium I. Attached you will find a hotel map. Hard copies of all forms and vouchers, along with agendas and reference material, will be included in your registration packet. Please pick up your registration packet by 8:30 a.m. Your first session will begin in the Carpenter Ballroom at 8:30 a.m. After the opening remarks you will be dismissed to your breakout rooms. The breakout rooms will contain computers, and security will be a priority. Please leave any large bags or belongings in your sleeping room, as space is limited and the breakout rooms will be locked at all times when not in use.

Also included in your registration packet will be a media release form. Smarter Balanced will be developing a video compilation to document the Achievement Level Setting process for member states, policy makers, and other interested parties. Panelists interested in providing feedback and comments on the process will have the opportunity to participate in an optional video interview for use in multimedia materials. Attached you will find a sample Interview Multimedia Permission and Release form; by completing this form you will allow Smarter Balanced to document the In-Person Panel and use your likeness in future multimedia and print materials. **Forms will be completed onsite in Dallas.**

Please access the links below prior to your arrival in Dallas to familiarize yourself with the material.

- 1) **Smarter Balanced Practice Test:** <http://www.smarterbalanced.org/practice-test/>
- 2) **Achievement Level Descriptors:** <http://www.smarterbalanced.org/achievement-levels/>

Then click:

- ELA Panelists → [ELA/literacy ALDs and College Content-Readiness Policy](#) (PDF)
- Math Panelists → [Mathematics ALDs and College Content-Readiness Policy](#) (PDF)
- All Panelists → [Achievement Level Descriptors Glossary of Terms](#) (PDF)

If you have any questions, please contact Mandy Hunter at MI_smarterbalanced21@measinc.com.

We look forward to working with you.

Attachments:

Prepaid Super Shuttle Voucher

Hotel Map

Sample Interview Multimedia Permission and Release Form

To: Achievement Level Setting In-Person Panelists
From: MI_smarterbalanced21@measinc.com
Subject: Smarter Balanced Achievement Level Setting In-Person Panel

This email is being sent on behalf of the Smarter Balanced Assessment Consortium. For questions please email MI_smarterbalanced21@measinc.com.

Dear [Grades 3–5 Panelists, Grades 6–8 Panelists, Grade 11 Panelists]:

This is a friendly reminder that the [Grade 11] In-Person Panel begins on [Monday]. Your travel arrangements should have you arriving [Sunday].

The meeting will be held at the Hilton Anatole.

- Hilton Anatole
2201 North Stemmons Freeway
Dallas, TX 75207
(214) 761–7209

Upon your arrival in Dallas, transportation to the Hilton Anatole will be provided via Super Shuttle. Attached you will find a prepaid voucher for Super Shuttle that includes gratuity. Instructions on how to find or schedule the shuttle are included on the voucher. **Please print the voucher and take it with you to Dallas. You will need a hard copy of the voucher in order to use the shuttle service. No other form of ground transportation will be reimbursed.** The voucher is needed for your arrival only; Measurement Incorporated will arrange your return trip to the airport.

All hotel accommodations have been booked on your behalf under the Measurement Incorporated room block. No confirmation number will be necessary at check-in. **Please check in at the Atrium Lobby.** Your reservation will be under your name at the hotel front desk.

We have had several inquiries about dress code. Business casual is suggested.

We understand you may be concerned about the recent Ebola case in Dallas. We are monitoring and following official guidance from the Centers for Disease Control and Prevention (CDC) and Texas public health authorities. At this time, the CDC does not recommend any change in daily routine for individuals living in or travelling to Dallas. If the situation changes, we will follow that guidance and immediately alert all Achievement Level Setting participants. For more information, please access the CDC's website at www.cdc.gov.

Breakfast and registration will begin at 7:30 a.m. on [October 13, October 15, October 17] in Atrium I. Hard copies of all forms and vouchers, along with agendas and reference material, will be included in your registration packet. Please pick up your registration packet by 8:30 a.m. Your first session will begin in the Carpenter Ballroom at 8:30 a.m. After the opening remarks you will proceed to your breakout rooms. The breakout rooms will contain computers, and security will be a priority. Please leave any large bags or belongings in your sleeping room, as space is limited and the breakout rooms will be locked at all times when not in use.

If you have any questions, experience travel delays, or need assistance, please contact Mandy Hunter at MI_smarterbalanced21@measinc.com or at (908) 902-6627.

We look forward to seeing you in Dallas!

Attachment:

Prepaid Super Shuttle Voucher

Dear Vertical Articulation Committee Panelists:

Thank you for your participation in the In-Person panel; we are excited to work with you again on the Vertical Articulation Committee to complete this important work.

For those of you that have gone home and will be returning to Dallas we have attached the Super Shuttle voucher for your ground transportation.

Panelists that have remained in Dallas are encouraged to join us for breakfast and lunch on Saturday Sunday, October 18-19.

On Monday, October 20 breakfast will be served from 7:30-8:30. We will reconvene in Batik A following breakfast.

**Evening Shuttle Service to Trinity Groves
Dining and Entertainment District
October 13, 15, and 17**



We will offer a complimentary shuttle to Trinity Groves, a 15-acre restaurant and entertainment destination just five minutes away from the hotel. Continuous shuttles will run from the Jade Lobby to Trinity Groves from 5:30 p.m. until 10:30 p.m. on October 13, 15, and 17.

Check out the Trinity Groves website here:

<http://www.trinitygroves.com/>

Email Text:

To: K-12 State Leads, Teacher Involvement Coordinators, and Higher Education Leads

CC: Executive Staff, Nancy Arnold

From: MI_smarterbalanced21@measinc.com

Subject: Messages for Achievement Level Setting In-Person Panelists Committee Members

This email is being sent on behalf of the Smarter Balanced Assessment Consortium. For questions please email MI_smarterbalanced21@measinc.com.

Dear In-Person Panelists:

Thank you for your participation in the Smarter Balanced Achievement Level Setting. Your time and efforts were critical to this essential task.

We have received many requests for several of the resources that were used in Dallas. Attached you will find the Power Point slides on the Bookmarking Procedure, Achievement Level Descriptors (ALDs) for all grades and content areas, and a list of helpful links that can be found on the Smarter Balanced website. Please note that the attached ALDs are the same ALDs that can be found on the Smarter Balanced website only formatted differently.

Also attached you will find an electronic certificate of participation for your involvement in the Achievement Level Setting activities. The certificate is a fillable PDF. Please enter your name along with the grade and content area of the panel for which you served. Once you have entered your information please save the PDF and print a copy of the certificate for your files.

If you have any questions, please contact Mandy Hunter at MI_smarterbalanced21@measinc.com.

Thank you again for your participation.

Attachments:

Bookmarking Power Point

Achievement Level Descriptors

Smarter Balanced Resources

Certificate of Participation

Email Text:

To: Smarter Balanced State Leads

CC: Michael Bunch, Barbara Scherich, Nancy Arnold

BCC:

From: MI_smarterbalanced21@measinc.com

Subject: Smarter Balanced Achievement Level Setting In-Person Panel Observers

This email is being sent on behalf of the Smarter Balanced Assessment Consortium. For questions please email MI_smarterbalanced21@measinc.com.

Dear [State Lead]:

Thank you for your interest in observing The Achievement Level Setting In-Person Panel in Dallas, Texas.

States are responsible for all observer expenses, including airfare and ground transportation, lodging, parking, tolls, substitute teacher reimbursements to districts, and dinners (breakfast and lunch will be provided).

The In-Person Panel will take place at the Hilton Anatole [<http://www3.hilton.com/en/hotels/texas/hilton-anatole-DFWANHH/index.html>] in Dallas, Texas. Measurement Incorporated has set up a room block for all incoming Observers. Please contact the Hilton Anatole at 214-748-1200 and use reservation code MIM200 to reserve your overnight accommodations.

Observers will be able to watch the proceedings but will not participate in the Achievement Level Setting process. Although space is limited in the meeting rooms, observers are welcome to move among panels in order to view activities related to English language arts/literacy and mathematics. Please refer to the attached Guidelines for Observers for additional information about the observer role.

Thank you for your participation in this important endeavor, we look forward to seeing you in Dallas.

Attachments:

Guidelines for Observers

Observer Nomination Form

Email Text:

To: Smarter Balanced State Leads

CC: Michael Bunch, Barbara Scherich, Nancy Arnold

BCC:

From: MI_smarterbalanced21@measinc.com

Subject: Smarter Balanced Achievement Level Setting In-Person Panel Observers

This email is being sent on behalf of the Smarter Balanced Assessment Consortium. For questions please email MI_smarterbalanced21@measinc.com.

Dear [Nominated Observer]:

The Achievement Level Setting In-Person Panel will convene in October in Dallas, Texas. Any state/commonwealth that had students participate in the Field Test may send one person to observe the Achievement Level Setting activities. [State Lead] has nominated you to attend, and observe the In-Person Panel.

Please refer to the attached Observer Nomination Form for the panels [State Lead] has requested you observe.

States are responsible for all observer expenses, including airfare and ground transportation, lodging, parking, tolls, substitute teacher reimbursements to districts, and dinners (breakfast and lunch will be provided).

The In-Person Panel will take place at the Hilton Anatole [\[http://www3.hilton.com/en/hotels/texas/hilton-anatole-DFWANHH/index.html\]](http://www3.hilton.com/en/hotels/texas/hilton-anatole-DFWANHH/index.html) in Dallas, Texas. Measurement Incorporated has set up a room block for all incoming Observers. Please contact the Hilton Anatole at 214-748-1200 and use reservation code MIM200 to reserve your overnight accommodations.

Observers will be able to watch the proceedings but will not participate in the Achievement Level Setting process. Although space is limited in the meeting rooms, observers are welcome to move among panels in order to view activities related to English language arts/literacy and mathematics. Please refer to the attached Guidelines for Observers for additional information about the observer role.

Thank you for your participation in this important endeavor, we look forward to seeing you in Dallas.

Attachments:

Guidelines for Observers

Observer Nomination Form

Email Text:

To: Smarter Balanced State Leads

CC: Michael Bunch, Barbara Scherich, Nancy Arnold

BCC:

From: MI_smarterbalanced21@measinc.com

Subject: Smarter Balanced Achievement Level Setting In-Person Panel Observers

This email is being sent on behalf of the Smarter Balanced Assessment Consortium. For questions please email MI_smarterbalanced21@measinc.com.

Dear Smarter Balanced State Leads:

The Achievement Level Setting In-Person Panel will convene in October in Dallas, Texas. Per Joe Willhoff's 7/23/14 email, any state/commonwealth that had students participate in the Field Test may send one person to observe the Achievement Level Setting activities.

States are responsible for all observer expenses, including airfare and ground transportation, lodging, parking, tolls, substitute teacher reimbursements to districts, and dinners (breakfast and lunch will be provided).

Observers will be able to watch the proceedings but will not participate in the Achievement Level Setting process. Although space is limited in the meeting rooms, observers are welcome to move among panels in order to view activities related to English language arts/literacy and mathematics. Please refer to the attached Guidelines for Observers for additional information about the observer role.

To nominate your state's observer, please complete the attached Observer Nomination Form and return it to Measurement Incorporated via MI_smarterbalanced21@measinc.com by August 29, 2014. To complete and submit the Observer Nomination Form, please fill in all fields electronically, save the documents for your files, and return to Measurement Incorporated as an e-mail attachment via the address above.

Measurement Incorporated will then invite your nominated observer to attend the In-Person Panel and provide logistical details.

Thank you for your assistance in this important endeavor.

Attachments:

Guidelines for Observers

Observer Nomination Form

Achievement Level Setting Guidelines for Observers

Guidelines for In-Person Panel Observers

Any state/commonwealth that had students participate in the Field Test may send one observer to the In-Person Panel for Achievement Level Setting in October. The primary role of the In-Person Panel observer is to watch and understand the Achievement Level Setting process.

Observers must not interfere with or interrupt the Achievement Level Setting activities. In order to avoid disrupting the panels and to preserve the integrity of the process, please follow these guidelines:

- Do not speak to panelists or facilitators during panel activities or breaks. You will have an opportunity to share your observations during a debriefing session. If you have an urgent question or comment, please discuss it with a non-facilitator staff person.
- You are welcome to observe several panels; please enter and exit the meeting rooms quietly. Please sit in the designated observer seats.
- Please limit use of laptops or cell phones during panel sessions. Please refrain from sending messages (via email or social media) about the workshop. Your cell phone should be placed in silent mode.
- Do not take any pictures of panelists and/or test items.

Email Text:

To: Smarter Balanced Staff

CC: Joe Willhoft, Tony Alpert, Magda Chia, Shelbi Cole, Nikki Elliott-Schuman, Jaci King, Marty McCall, Brandt Redd

BCC:

From: MI_smarterbalanced21@measinc.com

Subject: Smarter Balanced Achievement Level Setting **In-Person Panel**

Dear Smarter Balanced Staff:

You are invited to attend the Smarter Balanced Achievement Level Setting in Dallas, Texas on October 13–20, 2014.

Measurement Incorporated (MI) will provide your applicable airfare, airport shuttle, hotel, and meals. Additional expenses such as parking and meals not provided during the meeting will be reimbursed. Attached you will find a sample travel reimbursement form for your reference. You will receive a hard copy of the form at the In-Person Panel.

The In-Person Panel will be held in Dallas, Texas, October 13–19, 2014. The Vertical Articulation will be held in the same location on October 20, 2014. Specific details related to the logistics for travel are outlined below.

Traveling to and from Dallas:

Please follow the booking instructions outlined below to book air travel through **Cardinal Travel** to and from Dallas (see below for Cardinal Travel contact information).

- Hotel location:
 - Hilton Anatole
2201 North Stemmons Freeway
Dallas, TX 75207
(214) 761–7209
 - Cardinal Travel will inform us of your flight arrangements. Measurement Incorporated will then make your hotel reservations for you and provide you with your hotel reservation confirmation number.
 - If you plan to drive to Dallas, Measurement Incorporated will reimburse your mileage up to the price of round-trip coach airfare. Please contact Mandy Hunter at MI_smarterbalanced21@measinc.com if you plan to drive to the hotel.
- Airport Shuttle Information:
 - You will receive a prepaid voucher for transportation to and from the airport in Dallas.
 - Pick-up and drop-off areas are located on the voucher.
 - No other mode of ground transportation will be reimbursed unless pre-approved by Measurement Incorporated.
- Other Transportation:
 - Measurement Incorporated will provide a shuttle service serving the area within a 2-mile radius of the hotel. Details will be available on site. For excursions further afield (e.g., sightseeing, shopping), you will be on your own; Measurement Incorporated will not reimburse these transportation costs.
- Meals:
 - Please save receipts for meals you eat en route to and from the hotel on travel days. Measurement Incorporated will reimburse your meal expenses in accordance with the limits shown on the travel reimbursement form.
 - Breakfast and lunch will be provided on site daily. In addition, we will serve beverages and a light afternoon snack.

- Measurement Incorporated will reimburse your dinner meal expenses in accordance with the limits shown on the travel reimbursement form. Receipts will be required for all reimbursed expenses.

Next Steps

- Please e-mail Connie Garret at Cardinal Travel to begin the air travel booking process.
 - E-mail: connie.g@cardinaltravelservice.com office hours (8 a.m. – 5 p.m. EDT)
 - Identify meeting as: Smarter Balanced Achievement Level Setting in Dallas
 - Be prepared to provide the following information when contacting Cardinal Travel (please make sure all information is correct before contacting Cardinal Travel):

Panel(s) for which you will be present	
Departure Date and time	
Leaving from (City and Airport)	
Return date	
Seating preference (window, aisle, no preference)	
Name as it appears on ID	
Birthdate	
Cell phone number	

- Please make all reservations by September 26, 2014.

To confirm your presence at the In-Person Panel, please complete the attached information form and return via MI_smarterbalanced21@measinc.com by Friday, September 12. To complete and submit the information form please fill in all fields electronically, save the document for your files, and return to Measurement Incorporated as an e-mail attachment via the address above.

Attached you will find a sample of the travel reimbursement form. Hard copies of these documents will be distributed at the In-Person Panel.

If you have any questions, please contact Mandy Hunter at MI_smarterbalanced21@measinc.com.

Attachments:
 In-Person Panelist Information Form
 Travel Reimbursement Form

TO: Ann Wilder; Craig Deville; Winnie Reid; Amy Griswold; Dan Bowen; Jennie Bowen; Chris Dunbar; Corey Palermo; Ruth Hargis; Joe McClintock; Sheryl Grady, Arianto Wibowo; Irene Hendrawan, Eric Lewis, Ben Weber, Anthony Jordan, Ira Henderson, Lisa Johnson, Maud Eno

CC: Michael Bunch; Nelson Androes; Barbara Scherich; Chris Mayes; Kathleen Stapleton

All:

As you all know the Smarter Balanced Achievement Level Setting in Dallas is rapidly approaching. The Achievement Level Setting will take place at the Hilton Anatole in Dallas, Texas [<http://www3.hilton.com/en/hotels/texas/hilton-anatole-DFWANHH/index.html>] October 13-20. Attached you will find staff assignments. We have reserved a block of rooms, and will make arrangements on your behalf. Please contact Connie Garrett at Cardinal Travel via email [connie.g@cardinaltravelservice.com] to book your airline ticket. All travel should be charged to #43. In the coming weeks I will provide you with a prepaid voucher for shuttle service from the Dallas airport to the Hilton Anatole.

All facilitators will need to arrive the day prior to assigned panel start date for facilitator training:

Grade 11

October 13-15

Arrive by 3:00 PM on October 12

Depart After 3:00 PM on October 15

Grades 6-8

October 15-17

Arrive by 3:00 PM on October 14

Depart after 3:00 PM on October 17

Grades 3-5

October 17-19

Arrive by 3:00 PM on October 16

Depart after 3:00 PM on October 19

Breakfast and lunch will be served on site daily. We will provide an evening shuttle service to the Trinity Groves dining area [<http://www.trinitygroves.com/>] on October 13, 15 and 17. Upon arrival in Dallas you will receive a packet with information about the hotel, the area, and any materials you will need for the Achievement Level Setting.

You will have complimentary Wi-Fi access in your sleeping rooms. If you will need to checkout a MI laptop for business use please contact NetOps with your request by Friday, September 26.

Please book all airline tickets by Friday, September 19. If you have any questions please let me know.

Thank you, Mandy

NON-DISCLOSURE AGREEMENT

**Smarter Balanced Assessment Consortium
Achievement Level Setting – Dallas, TX
October 13–20, 2014**

Measurement Incorporated (MI) has been contracted by the Smarter Balanced Assessment Consortium to facilitate Achievement Level Setting activities. The security of this assessment program activity must be maintained at all times. Everyone associated with Achievement Level Setting or who has contact with secure materials must agree to maintain security by not discussing or disclosing any confidential information related to the program.

You are personally responsible for maintaining strict confidentiality of any information related to the Smarter Balanced Assessment Consortium Achievement Level Setting.

You are required to follow these security standards:

1. Do not discuss with or disclose to anyone except designated Smarter Balanced or MI staff any aspect of the assessment materials shared with you or the cut scores obtained.
2. Do not remove any secure materials from the room where you and your panel are working.
3. Do not retain, photocopy, or electronically transmit any images or notes made regarding the Smarter Balanced materials during the Achievement Level Setting. All such materials must be returned to designated Smarter Balanced or MI staff members.
4. Panelists may not work for, or consult with, any publishing company that provides or develops products for a Smarter Balanced audience that relates to any Smarter Balanced assessment program.

The Smarter Balanced Assessment Consortium and Measurement Incorporated appreciate your cooperation in this important activity. Please review this form, sign electronically below, and return it to Mandy Hunter at MI_smarterbalanced21@measinc.com.

I have read and understand these confidentiality and security standards and agree to abide by them. I acknowledge and agree that all performance tasks, items, and related materials developed for the Smarter Balanced Assessment Consortium Achievement Level Setting are highly confidential and that their contents are not to be divulged to anyone. I further understand that violation of this Non-Disclosure Agreement may be a cause for disciplinary or legal action by the Smarter Balanced Assessment Consortium and/or Measurement Incorporated.

Name (Please Print) _____

Signed _____ Date _____

STIPEND AND HONORARIUM CLAIM VOUCHER

Smarter Balanced Assessment Consortium
Achievement Level Setting – Dallas, TX
October 13–20, 2014

Type of Meeting: _____

Name of Attendee: _____

Date(s) of Attendance: _____

Signature of Attendee: _____

Social Security Number: _____
**We cannot issue payment without a SSN.*

Mailing Address: _____

City, State, and ZIP: _____

Phone: _____

Email: _____

Total Amount of Reimbursement: days × \$150/day = \$

Expenses are billed separately. Please see the Travel Reimbursement Form.

Send to:
Measurement Incorporated
423 Morris Street
Durham, NC 27701
Attn: Mandy Hunter

Please return to Measurement Incorporated no later than November 3, 2014.

Office Use Only (#43)

Office Use Only (#43)	
MI Approval	Date

SUBSTITUTE TEACHER REIMBURSEMENT FORM

Smarter Balanced Assessment Consortium
Achievement Level Setting – Dallas, TX
October 13–20, 2014

Type of Meeting: _____

Name of Attendee: _____

Date(s) of Attendance: _____

Signature of Attendee: _____

Complete by School Administration

Name of School: _____

School District: _____

School District Mailing Address: _____

City, State, and ZIP: _____

Phone: _____

Email: _____

Name of Substitute: _____

Number of Days Worked: _____

Amount Paid: _____

Total Amount of Reimbursement: _____

Administrator Name: _____

Administrator Signature: _____

Send to:
Measurement Incorporated
423 Morris Street
Durham, NC 27701
Attn: Mandy Hunter

Please return to Measurement Incorporated no later than November 3, 2014.

Office Use Only (#43)

MI Approval	Date
-------------	------

TRAVEL REIMBURSEMENT FORM

Smarter Balanced Assessment Consortium
Achievement Level Setting – Dallas, TX
October 13–20, 2014

PLEASE PRINT

Name:		
Address:		
City:	State:	Zip:
Phone:	Fax:	
E-mail:		

ITEMIZED EXPENSES

DESCRIPTION	TRAVEL DAY	OCT 13	OCT 14	OCT 15	TRAVEL DAY
Mileage reimbursement for Panelists who drive to and from the In-Person Panel site (must be pre-approved by MI)	# Miles _____	<div style="font-size: 4em; font-weight: bold;">X</div>			# Miles _____
	× 0.56				× 0.56
	Total \$ _____				Total \$ _____
Taxi, bus, train, or subway between home and airport Shuttle service will be provided in Dallas.	\$ _____				\$ _____
Miscellaneous Travel Expenses Receipts for airline baggage (charges for one checked bag will be reimbursed), lodging, airport parking or on-site parking, toll charges, and incidental expenses (limited to \$5 per day) are required and must accompany this reimbursement form.	\$ _____				\$ _____
Breakfast (\$12) and Lunch (\$18) will be reimbursed at the maximum rate of \$30 per day (with itemized receipt) for travel days only.	\$ _____				\$ _____
Dinner will be reimbursed at the maximum rate of \$36 per day (with itemized receipt). MI does not reimburse for alcoholic beverages.	\$ _____	\$ _____	\$ _____	\$ _____	\$ _____
Daily Totals	\$ _____	\$ _____	\$ _____	\$ _____	\$ _____

PLEASE NOTE: Gray boxes indicate panelists will not be reimbursed for said expenses during that time period, as these services are included in the daily program.

Total Amount Due \$ _____

Mail completed original form with original receipts attached (no later than November 3, 2014) to:

Mandy Hunter
Measurement Incorporated
423 Morris Street
Durham, NC 27701



Second Addendum to the First Addendum between Measurement Incorporated and Hilton Anatole, dated July 28, 2014

This Addendum is dated effective July 28, 2014 and relates to a certain First Addendum (“Addendum I”) signed by the Hotel on June 16, 2014 between Measurement Incorporated (“Group”) and the Hilton Anatole Hotel (“Hotel”), detailing arrangements for the Measurement Incorporated SS Meeting, Friday, October 10, 2014 through Sunday, October 20, 2014.

1. (a) ROOM ARRIVAL and DEPARTURE PATTERN PER THE CONTRACT (“Contract Block”):

Day	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Date	10/10/14	10/11/14	10/12/14	10/13/14	10/14/14	10/15/14	10/16/14	10/17/14	10/18/14	10/19/14
Standard King Bed	10	10	167	167	354	206	392	206	206	87

Note: All One Bedroom Suites count as two units in the Hotel’s overall inventory.

TOTAL SLEEPING ROOM NIGHTS ANTICIPATED: 1,805

(b) REVISED ROOM ARRIVAL and DEPARTURE PATTERN:

Day	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Date	10/10/14	10/11/14	10/12/14	10/13/14	10/14/14	10/15/14	10/16/14	10/17/14	10/18/14	10/19/14
Standard King Bed	10	10	167	167	354	206	392	206	206	87
Observer Block	0	0	25	25	25	25	25	25	25	25

Note: All One Bedroom Suites count as two units in the Hotel’s overall inventory.

TOTAL SLEEPING ROOM NIGHTS ANTICIPATED: 2,005

Room and tax charges in the Observer Block will be charged to the Individual. Room Block Performance calculations will be based upon the revised Room Block above. Except as specifically referenced herein, all other terms, conditions, policies and stipulations of the Contract will remain in full force and effect as stated in the Contract.

The undersigned are authorized to sign this Addendum:

ACCEPTED AND AGREED TO:

Measurement Incorporated

Michael Bunch, Senior Vice President

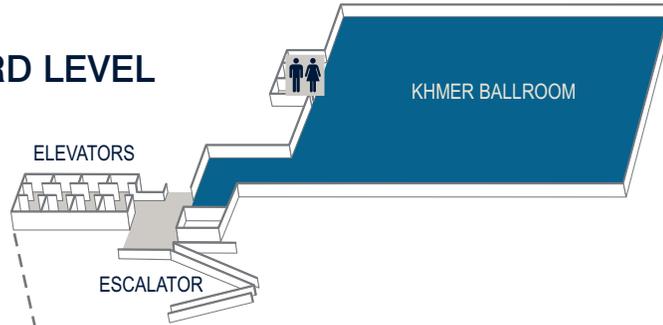
Hilton Management LLC, as manager of the Hilton Anatole Hotel

Lauren Singer, Senior Sales Manager

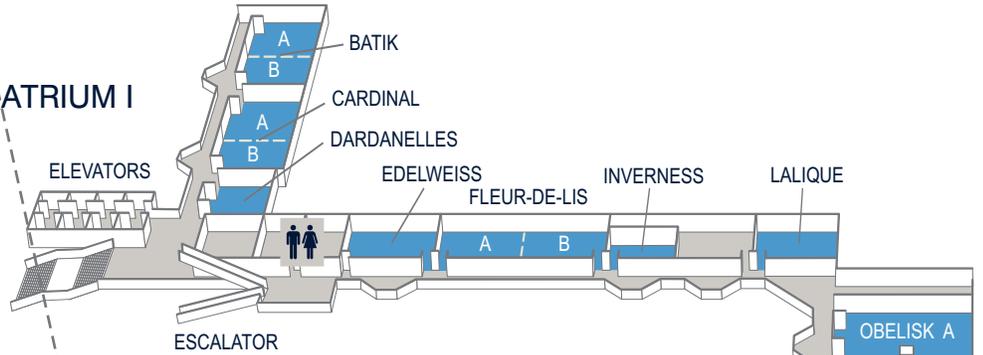
Date: August 5, 2014

Date: _____

ATRIUM THIRD LEVEL

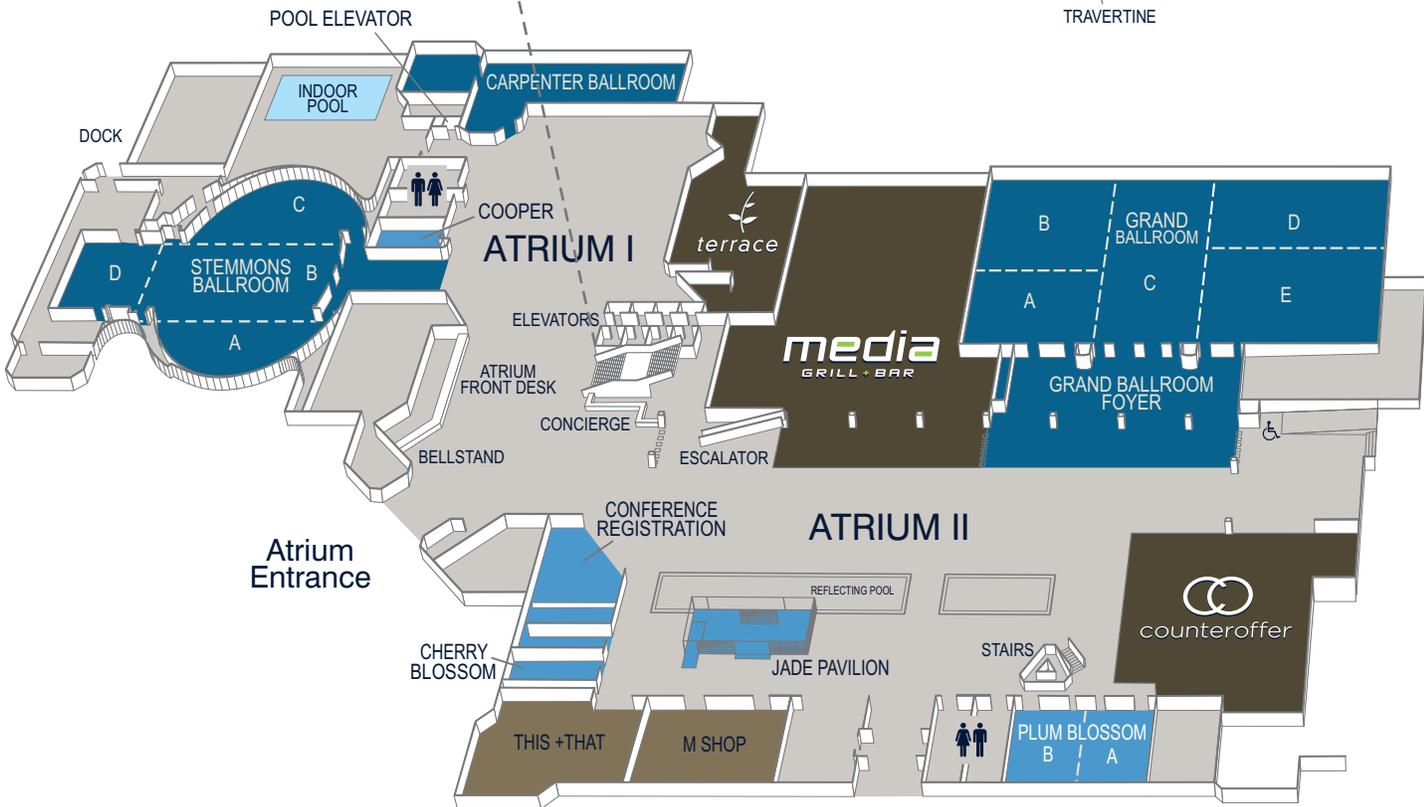
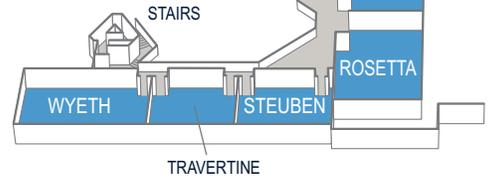


ATRIUM I



ATRIUM MEZZANINE LEVEL

ATRIUM II



ATRIUM LOBBY LEVEL

Jade Entrance

	Length	Width	Sq. Footage	Height	Banquet	Theater	Reception	Schoolroom	Conference	Hollow Square	U-Shape
ATRIUM I: LOBBY LEVEL											
Atrium I	107	92	9,844		600		800				
Carpenter Ballroom	86	31	4,557	18	240	330	364	160			
Cooper	29	17	493	9	40	53	57	29	26	30	25
Stemmons Ballroom	107	113	8,312	20	660	765	923	566	106	108	102
Stemmons Ballroom A or C	74	36	1,960	20	140	180	217	126	52	54	48
Stemmons Ballroom B	77	40	3,095	20	240	300	343	198	64	78	66
Stemmons Ballroom D (accessed via B)	31	42	1,297	16	130	175	144	112	34	54	42
Stemmons Ballroom B and D	107	40	4,392	20	380	500	488	344	100	114	102
MEZZANINE LEVEL											
Batik	57	29	1,653	9	120	172	190	90	56	60	56
Batik A or B	28	29	812	9	60	91	93	49	28	36	29
Cardinal	57	29	1,653	9	120	172	190	90	56	60	56
Cardinal A or B	28	29	812	9	60	91	93	49	28	36	29
Dardanelles	30	23	690	9	60	77	79	42	28	30	24
THIRD LEVEL											
Khmer Ballroom	139	118	16,402	22	1,500	1,860	1,885	1,080			
ATRIUM II: LOBBY LEVEL											
Conference Registration			1,390								
Cherry Blossom	12	36	432	12					24		
Grand Ballroom Foyer	100	58	5,892	9			287				
Grand Ballroom	100	174	17,400	18	1,600	1,900	2,000	1,302			
Grand Ballroom A	50	40	2,000	18	160	205	230	126			
Grand Ballroom B	60	50	3,000	18	250	351	345	187			
Grand Ballroom C	100	50	5,000	18	450	552	575	318			
Grand Ballroom D	74	50	3,700	18	320	416	425	225			
Grand Ballroom E	74	50	3,700	18	310	385	425	219			
Plum Blossom	56	36	2,016	9	150	225	200	132	48	60	50
Plum Blossom A or B	28	36	1,008	9	74	112	100	66	24	30	25
MEZZANINE LEVEL											
Edelweiss	28	28	784	9	60	80	90	50	28	36	27
Fleur-de-lis	58	27	1,566	9	120	166	180	92	56	60	56
Fleur-de-lis A or B	29	27	783	9	60	92	90	49	28	36	29
Inverness	28	19	532	9	50	63	60	36	26	30	24
Lalique	30	27	810	9	70	95	93	50	28	38	29
Obelisk	57	49	2,793	9	220	250	321	146	46	59	45
Obelisk A or B	49	28	1,372	9	110	170	158	86	46	56	45
Peridot	28	22	616	9	40	64	60	36	26	30	25
Rosetta	49	25	1,225	9	100	144	141	73	46	56	46
Steuben	29	27	783	9	60	90	90	48	24	36	24
Travertine	28	27	756	9	60	90	87	48	24	36	24
Wyeth	43	27	1,161	9	90	128	133	78	36	42	36

Materials and Equipment Needed for In-Person ALS - Dallas

Materials	Count
Table Tents	84
Name Badges (with ID on back)	571
ID labels	571
Flip charts (PostIt)	13
Notepads	571
Pens	571
ELA Common Core State Standards (hard copy)	12
Math Common Core State Standards (hard copy)	12

Equipment	Count
Computers with sound cards (1 per facilitator)	12
LCD projectors (1 per room)	12
Speakers (sets)	12
USB drives for storing data	12
Cables and connectors	375
Computers with sound cards (1 per panelist)	375
Extension cords for presenters/facilitators	14
Power strips for presenters/facilitators	14
Spare bulbs	3
Batteries (AA)	24
Earphones	504
iPad or similar tablet	2

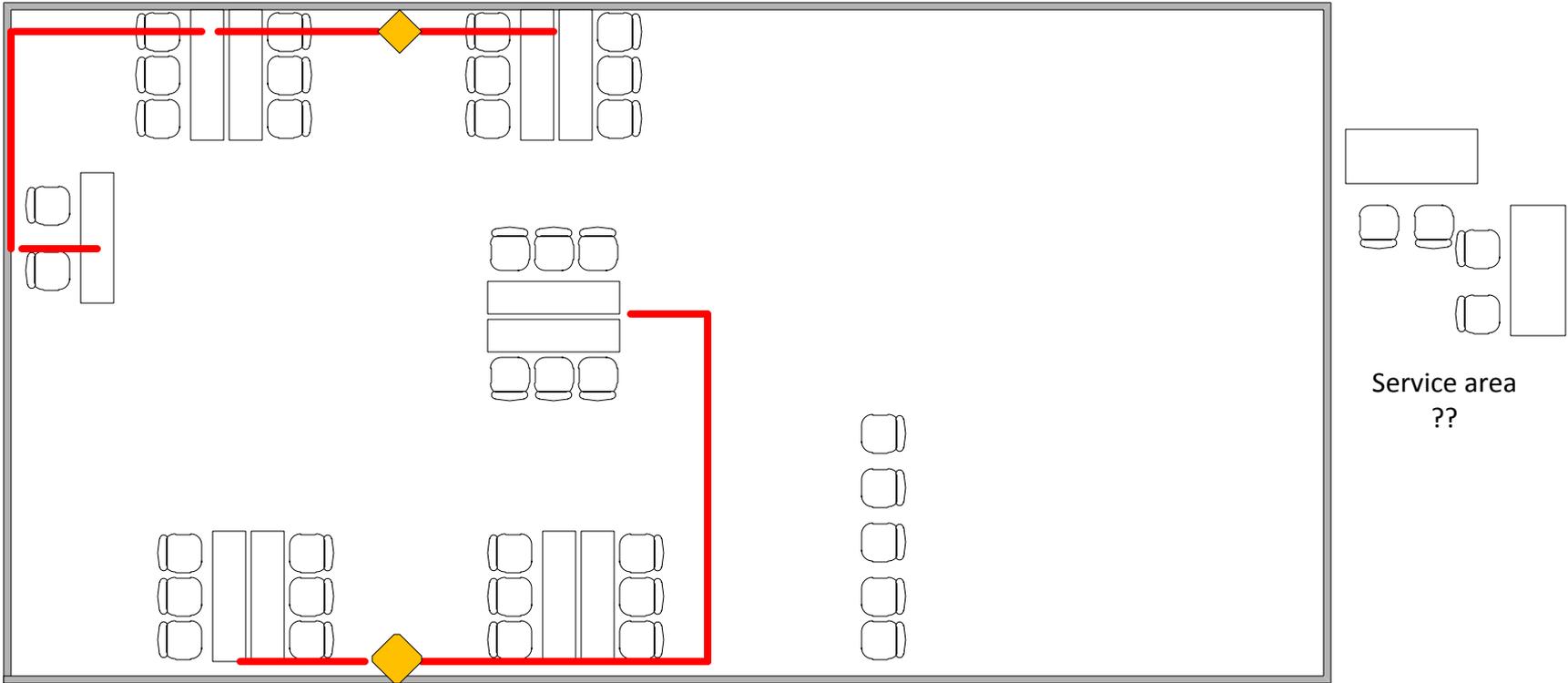
Other	Count
Internet connections for facilitators and panelists*	375
Secure storage site	1
Work room	1

To be Sent in Advance

Security Agreement
Room/Panel assignment
Log-on ID

Internet access will be required for 144 people on 10/13,
180 people on 10/15, and 180 people on 10/17, all for 1.5 hours
The lead facilitator will need internet access on 10/13,
10,15, and 10/17 for 1.5 hours.

Sample Room Set-up MI SBAC



 = Electrical Distribution Box

NetOps Breakout Room Setup/Breakdown Schedule

Sun	Mon	Tue	Wed	Thu	Fri	Sat
			1	2	3	4
5	6	7 Hardware Arrives in Dallas	8	9 NetOps Arrives in Dallas	10 8:00 AM Begin Setup: Batik A Batik B Cardinal A Cardinal B	11 Finalize Setup: Batik A Batik B Cardinal A Cardinal B
12 8:00 AM Begin Setup: Dardanelles Edelweiss Fleur De Lis A Fleur De Lis B Lalique Rosetta Steuben Wyeth	13 Continue Setup: Dardanelles Edelweiss Fleur De Lis A Fleur De Lis B Lalique Rosetta Steuben Wyeth	14 Finalize Setup: Dardanelles Edelweiss Fleur De Lis A Fleur De Lis B Lalique Rosetta Continue Setup: Steuben Wyeth	15 Continue Setup: Steuben Wyeth	16 Finalize Setup: Steuben Wyeth	17 All 12 Rooms in Session until 12:30 PM 12:30 PM Begin Breakdown: Edelweiss Lalique Rosetta Dardanelles Fleur De Lis A Fleur De Lis B	18 Continue Breakdown: Edelweiss Lalique Rosetta Dardanelles Fleur De Lis A Fleur De Lis B
19 Complete Breakdown: Edelweiss Lalique Rosetta 12:30 PM Begin Breakdown, Complete by 11:00 PM: Steuben Wyeth 12:30 PM Begin Breakdown: Cardinal A Cardinal B	20 Complete Breakdown: Fleur De Lis A Fleur De Lis B Dardanelles Cardinal A Cardinal B 5:00 PM Begin Breakdown: Batik A Batik B	21 Complete Breakdown: Batik A Batik B	22	23	24	25
26	27	28	29	30	31	Notes:

Measurement Incorporated
Measurement Incorporated ALS Meeting
Friday, October 10, 2014 – Tuesday, October 21, 2014

Booked By: Lauren Singer

Catering
Mgr.:

Julie Lecklitner

Serviced By: Julie Lecklitner

Meeting Space Recap

Date	Start Time	End Time	Function	Panel	#	Room	Setup	Number of People	Number of Computers
Fri, 10/10/14	8:00 AM	11:00 PM	SETUP	GRADE 11 ELA #1	1	Batik A	Classroom	38	37
Fri, 10/10/14	8:00 AM	11:00 PM	SETUP	GRADE 11 ELA #2	2	Batik B	Classroom	38	37
Fri, 10/10/14	8:00 AM	11:00 PM	SETUP	GRADE 11 MATH #1	3	Cardinal A	Classroom	38	37
Fri, 10/10/14	8:00 AM	11:00 PM	SETUP	GRADE 11 MATH #2	4	Cardinal B	Classroom	38	37
Sat, 10/11/14	8:00 AM	11:00 PM	SETUP	GRADE 11 ELA #1	1	Batik A	Classroom	38	37
Sat, 10/11/14	8:00 AM	11:00 PM	SETUP	GRADE 11 ELA #2	2	Batik B	Classroom	38	37
Sat, 10/11/14	8:00 AM	11:00 PM	SETUP	GRADE 11 MATH #1	3	Cardinal A	Classroom	38	37
Sat, 10/11/14	8:00 AM	11:00 PM	SETUP	GRADE 11 MATH #2	4	Cardinal B	Classroom	38	37
Sun, 10/12/14	8:00 AM	11:00 PM	SETUP	GRADE 11 ELA #1	1	Batik A	Classroom	38	37
Sun, 10/12/14	8:00 AM	11:00 PM	SETUP	GRADE 11 ELA #2	2	Batik B	Classroom	38	37
Sun, 10/12/14	8:00 AM	11:00 PM	SETUP	GRADE 11 MATH #1	3	Cardinal A	Classroom	38	37
Sun, 10/12/14	8:00 AM	11:00 PM	SETUP	GRADE 11 MATH #2	4	Cardinal B	Classroom	38	37
Sun, 10/12/14	8:00 AM	11:00 PM	SETUP	GRADE 7 ELA	5	Fleur De Lis A	Classroom	32	31
Sun, 10/12/14	8:00 AM	11:00 PM	SETUP	GRADE 7 MATH	6	Fleur De Lis B	Classroom	32	31
Sun, 10/12/14	8:00 AM	11:00 PM	SETUP	GRADE 8 MATH	7	Dardanelles	Classroom	32	31
Sun, 10/12/14	8:00 AM	11:00 PM	SETUP	GRADE 8 ELA	8	Edelweiss	Classroom	32	31
Sun, 10/12/14	8:00 AM	11:00 PM	SETUP	GRADE 6 ELA	9	Lalique	Classroom	32	31
Sun, 10/12/14	8:00 AM	11:00 PM	SETUP	GRADE 6 MATH	10	Rosetta	Classroom	32	31
Sun, 10/12/14	8:00 AM	11:00 PM	SETUP	GRADE 5 ELA	11	Steuben	Classroom	32	31
Sun, 10/12/14	8:00 AM	11:00 PM	SETUP	GRADE 5 MATH	12	Wyeth	Classroom	32	31
Mon, 10/13/14	8:00 AM	11:00 PM	IN SESSION	GRADE 11 ELA #1	1	Batik A	Classroom	38	37
Mon, 10/13/14	8:00 AM	11:00 PM	IN SESSION	GRADE 11 ELA #2	2	Batik B	Classroom	38	37
Mon, 10/13/14	8:00 AM	11:00 PM	IN SESSION	GRADE 11 MATH #1	3	Cardinal A	Classroom	38	37
Mon, 10/13/14	8:00 AM	11:00 PM	IN SESSION	GRADE 11 MATH #2	4	Cardinal B	Classroom	38	37
Mon, 10/13/14	8:00 AM	11:00 PM	SETUP	GRADE 7 ELA	5	Fleur De Lis A	Classroom	32	31

Mon, 10/13/14	8:00 AM	11:00 PM	SETUP	GRADE 7 MATH	6	Fleur De Lis B	Classroom	32	31
Mon, 10/13/14	8:00 AM	11:00 PM	SETUP	GRADE 8 MATH	7	Dardanelles	Classroom	32	31
Mon, 10/13/14	8:00 AM	11:00 PM	SETUP	GRADE 8 ELA	8	Edelweiss	Classroom	32	31
Mon, 10/13/14	8:00 AM	11:00 PM	SETUP	GRADE 6 ELA	9	Lalique	Classroom	32	31
Mon, 10/13/14	8:00 AM	11:00 PM	SETUP	GRADE 6 MATH	10	Rosetta	Classroom	32	31
Mon, 10/13/14	8:00 AM	11:00 PM	SETUP	GRADE 5 ELA	11	Steuben	Classroom	32	31
Mon, 10/13/14	8:00 AM	11:00 PM	SETUP	GRADE 5 MATH	12	Wyeth	Classroom	32	31
 									
Tues, 10/14/14	8:00 AM	11:00 PM	IN SESSION	GRADE 11 ELA #1	1	Batik A	Classroom	38	37
Tues, 10/14/14	8:00 AM	11:00 PM	IN SESSION	GRADE 11 ELA #2	2	Batik B	Classroom	38	37
Tues, 10/14/14	8:00 AM	11:00 PM	IN SESSION	GRADE 11 MATH #1	3	Cardinal A	Classroom	38	37
Tues, 10/14/14	8:00 AM	11:00 PM	IN SESSION	GRADE 11 MATH #2	4	Cardinal B	Classroom	38	37
Tues, 10/14/14	8:00 AM	11:00 PM	SETUP	GRADE 7 ELA	5	Fleur De Lis A	Classroom	32	31
Tues, 10/14/14	8:00 AM	11:00 PM	SETUP	GRADE 7 MATH	6	Fleur De Lis B	Classroom	32	31
Tues, 10/14/14	8:00 AM	11:00 PM	SETUP	GRADE 8 MATH	7	Dardanelles	Classroom	32	31
Tues, 10/14/14	8:00 AM	11:00 PM	SETUP	GRADE 8 ELA	8	Edelweiss	Classroom	32	31
Tues, 10/14/14	8:00 AM	11:00 PM	SETUP	GRADE 6 ELA	9	Lalique	Classroom	32	31
Tues, 10/14/14	8:00 AM	11:00 PM	SETUP	GRADE 6 MATH	10	Rosetta	Classroom	32	31
Tues, 10/14/14	8:00 AM	11:00 PM	SETUP	GRADE 5 ELA	11	Steuben	Classroom	32	31
Tues, 10/14/14	8:00 AM	11:00 PM	SETUP	GRADE 5 MATH	12	Wyeth	Classroom	32	31
 									
Wed, 10/15/14	8:00 AM	11:00 PM	IN SESSION	GRADE 11 ELA #1	1	Batik A	Classroom	38	37
Wed, 10/15/14	8:00 AM	11:00 PM	IN SESSION	GRADE 11 ELA #2	2	Batik B	Classroom	38	37
Wed, 10/15/14	8:00 AM	11:00 PM	IN SESSION	GRADE 11 MATH #1	3	Cardinal A	Classroom	38	37
Wed, 10/15/14	8:00 AM	11:00 PM	IN SESSION	GRADE 11 MATH #2	4	Cardinal B	Classroom	38	37
Wed, 10/15/14	8:00 AM	11:00 PM	IN SESSION	GRADE 7 ELA	5	Fleur De Lis A	Classroom	32	31
Wed, 10/15/14	8:00 AM	11:00 PM	IN SESSION	GRADE 7 MATH	6	Fleur De Lis B	Classroom	32	31
Wed, 10/15/14	8:00 AM	11:00 PM	IN SESSION	GRADE 8 MATH	7	Dardanelles	Classroom	32	31
Wed, 10/15/14	8:00 AM	11:00 PM	IN SESSION	GRADE 8 ELA	8	Edelweiss	Classroom	32	31
Wed, 10/15/14	8:00 AM	11:00 PM	IN SESSION	GRADE 6 ELA	9	Lalique	Classroom	32	31
Wed, 10/15/14	8:00 AM	11:00 PM	IN SESSION	GRADE 6 MATH	10	Rosetta	Classroom	32	31
Wed, 10/15/14	8:00 AM	11:00 PM	SETUP	GRADE 5 ELA	11	Steuben	Classroom	32	31
Wed, 10/15/14	8:00 AM	11:00 PM	SETUP	GRADE 5 MATH	12	Wyeth	Classroom	32	31
 									
Thu, 10/16/14	8:00 AM	11:00 PM	ALREADY SET	TURNOVER DAY	1	Batik A	Classroom	38	37
Thu, 10/16/14	8:00 AM	11:00 PM	ALREADY SET	TURNOVER DAY	2	Batik B	Classroom	38	37
Thu, 10/16/14	8:00 AM	11:00 PM	ALREADY SET	TURNOVER DAY	3	Cardinal A	Classroom	38	37

Thu, 10/16/14	8:00 AM	11:00 PM	ALREADY SET	TURNOVER DAY	4	Cardinal B	Classroom	38	37
Thu, 10/16/14	8:00 AM	11:00 PM	IN SESSION	GRADE 7 ELA	5	Fleur De Lis A	Classroom	32	31
Thu, 10/16/14	8:00 AM	11:00 PM	IN SESSION	GRADE 7 MATH	6	Fleur De Lis B	Classroom	32	31
Thu, 10/16/14	8:00 AM	11:00 PM	IN SESSION	GRADE 8 MATH	7	Dardanelles	Classroom	32	31
Thu, 10/16/14	8:00 AM	11:00 PM	IN SESSION	GRADE 8 ELA	8	Edelweiss	Classroom	32	31
Thu, 10/16/14	8:00 AM	11:00 PM	IN SESSION	GRADE 6 ELA	9	Lalique	Classroom	32	31
Thu, 10/16/14	8:00 AM	11:00 PM	IN SESSION	GRADE 6 MATH	10	Rosetta	Classroom	32	31
Thu, 10/16/14	8:00 AM	11:00 PM	SETUP	GRADE 5 ELA	11	Steuben	Classroom	32	31
Thu, 10/16/14	8:00 AM	11:00 PM	SETUP	GRADE 5 MATH	12	Wyeth	Classroom	32	31
Fri, 10/17/14	8:00 AM	11:00 PM	IN SESSION	GRADE 3 ELA	1	Batik A	Classroom	32	31
Fri, 10/17/14	8:00 AM	11:00 PM	IN SESSION	GRADE 3 MATH	2	Batik B	Classroom	32	31
Fri, 10/17/14	8:00 AM	11:00 PM	IN SESSION	GRADE 4 ELA	3	Cardinal A	Classroom	32	31
Fri, 10/17/14	8:00 AM	11:00 PM	IN SESSION	GRADE 4 MATH	4	Cardinal B	Classroom	32	31
Fri, 10/17/14	8:00 AM	11:00 PM	IN SESSION	GRADE 7 ELA	5	Fleur De Lis A	Classroom	32	31
Fri, 10/17/14	8:00 AM	11:00 PM	IN SESSION	GRADE 7 MATH	6	Fleur De Lis B	Classroom	32	31
Fri, 10/17/14	8:00 AM	11:00 PM	IN SESSION	GRADE 8 MATH	7	Dardanelles	Classroom	32	31
Fri, 10/17/14	8:00 AM	11:00 PM	IN SESSION	GRADE 8 ELA	8	Edelweiss	Classroom	32	31
Fri, 10/17/14	8:00 AM	11:00 PM	IN SESSION	GRADE 6 ELA	9	Lalique	Classroom	32	31
Fri, 10/17/14	8:00 AM	11:00 PM	IN SESSION	GRADE 6 MATH	10	Rosetta	Classroom	32	31
Fri, 10/17/14	8:00 AM	11:00 PM	IN SESSION	GRADE 5 ELA	11	Steuben	Classroom	32	31
Fri, 10/17/14	8:00 AM	11:00 PM	IN SESSION	GRADE 5 MATH	12	Wyeth	Classroom	32	31
Sat, 10/18/14	8:00 AM	11:00 PM	IN SESSION	GRADE 3 ELA	1	Batik A	Classroom	32	31
Sat, 10/18/14	8:00 AM	11:00 PM	IN SESSION	GRADE 3 MATH	2	Batik B	Classroom	32	31
Sat, 10/18/14	8:00 AM	11:00 PM	IN SESSION	GRADE 4 ELA	3	Cardinal A	Classroom	32	31
Sat, 10/18/14	8:00 AM	11:00 PM	IN SESSION	GRADE 4 MATH	4	Cardinal B	Classroom	32	31
Sat, 10/18/14	8:00 AM	11:00 PM	TEARDOWN		5	Fleur De Lis A	Classroom		
Sat, 10/18/14	8:00 AM	11:00 PM	TEARDOWN		6	Fleur De Lis B	Classroom		
Sat, 10/18/14	8:00 AM	11:00 PM	TEARDOWN		7	Dardanelles	Classroom		
Sat, 10/18/14	8:00 AM	11:00 PM	TEARDOWN		8	Edelweiss	Classroom		
Sat, 10/18/14	8:00 AM	11:00 PM	TEARDOWN		9	Lalique	Classroom		
Sat, 10/18/14	8:00 AM	11:00 PM	TEARDOWN		10	Rosetta	Classroom		
Sat, 10/18/14	8:00 AM	11:00 PM	IN SESSION	GRADE 5 ELA	11	Steuben	Classroom	32	31
Sat, 10/18/14	8:00 AM	11:00 PM	IN SESSION	GRADE 5 MATH	12	Wyeth	Classroom	32	31
Sun, 10/19/14	8:00 AM	11:00 PM	IN SESSION	GRADE 3 ELA	1	Batik A	Classroom	32	31

Sun, 10/19/14	8:00 AM	11:00 PM	IN SESSION	GRADE 3 MATH	2	Batik B	Classroom	32	31
Sun, 10/19/14	8:00 AM	11:00 PM	IN SESSION	GRADE 4 ELA	3	Cardinal A	Classroom	32	31
Sun, 10/19/14	8:00 AM	11:00 PM	IN SESSION	GRADE 4 MATH	4	Cardinal B	Classroom	32	31
Sun, 10/19/14	8:00 AM	11:00 PM	TEARDOWN		5	Fleur De Lis A	Classroom		
Sun, 10/19/14	8:00 AM	11:00 PM	TEARDOWN		6	Fleur De Lis B	Classroom		
Sun, 10/19/14	8:00 AM	11:00 PM	TEARDOWN		7	Dardanelles	Classroom		
Sun, 10/19/14	8:00 AM	11:00 PM	TEARDOWN		8	Edelweiss	Classroom		
Sun, 10/19/14	8:00 AM	11:00 PM	TEARDOWN		9	Lalique	Classroom		
Sun, 10/19/14	8:00 AM	11:00 PM	TEARDOWN		10	Rosetta	Classroom		
Sun, 10/19/14	8:00 AM	11:00 PM	IN SESSION	GRADE 5 ELA	11	Steuben	Classroom	32	31
Sun, 10/19/14	8:00 AM	11:00 PM	IN SESSION	GRADE 5 MATH	12	Wyeth	Classroom	32	31
Mon, 10/20/14	8:00 AM	11:00 PM	IN SESSION	ELA VAC	1	Batik A	Classroom	32	31
Mon, 10/20/14	8:00 AM	11:00 PM	IN SESSION	MATH VAC	2	Batik B	Classroom	32	31
Mon, 10/20/14	8:00 AM	11:00 PM	TEARDOWN		3	Cardinal A	Classroom		
Mon, 10/20/14	8:00 AM	11:00 PM	TEARDOWN		4	Cardinal B	Classroom		
Mon, 10/20/14	8:00 AM	11:00 PM	TEARDOWN		5	Fleur De Lis A	Classroom		
Mon, 10/20/14	8:00 AM	11:00 PM	TEARDOWN		6	Fleur De Lis B	Classroom		
Mon, 10/20/14	8:00 AM	11:00 PM	TEARDOWN		7	Dardanelles	Classroom		
Tue, 10/21/14	8:00 AM	11:00 PM	TEARDOWN		1	Batik A	Classroom		
Tue, 10/21/14	8:00 AM	11:00 PM	TEARDOWN		2	Batik B	Classroom		

Panel Grade	Panel Subject	Panelist State	Panelist Role	VAC
3	ELA	ID	Educator: General	VAC
3	ELA	MI	Educator: General	VAC
3	ELA	WV	Educator: Non-Teaching	VAC
3	ELA	CA	Educator: General	
3	ELA	CA	Educator: ELL experience	
3	ELA	WY	Educator: General	
3	ELA	CT	Educator: General	
3	ELA	WV	Educator: Non-Teaching	
3	ELA	HI	Educator: Non-Teaching	
3	ELA	WA	Educator: General	
3	ELA	DE	Educator: General	
3	ELA	DE	General Public	
3	ELA	WI	Educator: General	
3	ELA	ME	Educator: General	
3	ELA	CA	Educator: SWD experience	
3	ELA	VI	Educator: General	
3	ELA	VT	Educator: General	
3	ELA	WI	Educator: General	
3	ELA	ME	Educator: General	
3	ELA	ID	Educator: General	
3	ELA	CT	Educator: General	
3	ELA	NV	Educator: General	
3	ELA	CA	Educator: General	
3	ELA	WA	General Public	
3	ELA	IA	Educator: General	
3	ELA	WI	Educator: General	
4	ELA	WV	Educator: General	VAC
4	ELA	OR	Educator: ELL experience	VAC
4	ELA	MO	Educator: General	VAC
4	ELA	CA	Educator: ELL experience	
4	ELA	NH	Educator: General	
4	ELA	WV	General Public	
4	ELA	MI	Educator: General	
4	ELA	VI	Educator: General	
4	ELA	WA	Educator: Non-Teaching	
4	ELA	CA	Educator: General	
4	ELA	DE	Educator: General	
4	ELA	ND	Educator: General	
4	ELA	MT	Educator: General	
4	ELA	NV	Educator: General	
4	ELA	DE	Educator: ELL experience	
4	ELA	WI	Educator: SWD experience	
4	ELA	VI	Educator: Non-Teaching	
4	ELA	ME	Educator: Non-Teaching	
4	ELA	ME	Educator: General	
4	ELA	WI	Educator: SWD experience	

4	ELA	WV	General Public	
4	ELA	CA	General Public	
4	ELA	ND	Educator: General	
4	ELA	ID	Educator: ELL experience	
4	ELA	CT	Educator: General	
4	ELA	IA	Educator: General	
4	ELA	NH	Educator: General	
5	ELA	MT	Educator: SWD experience	VAC
5	ELA	CA	Educator: General	VAC
5	ELA	SD	Educator: Non-Teaching	VAC
5	ELA	CA	Educator: ELL experience	
5	ELA	DE	Educator: General	
5	ELA	WA	Educator: General	
5	ELA	ME	Educator: ELL Experience	
5	ELA	NH	Educator: Non-Teaching	
5	ELA	OR	Educator: General	
5	ELA	MT	Educator: General	
5	ELA	CA	Educator: ELL experience	
5	ELA	CA	Educator: Non-Teaching	
5	ELA	MI	Educator: General	
5	ELA	VT	Educator: Non-Teaching	
5	ELA	IA	Educator: General	
5	ELA	ND	Educator: General	
5	ELA	CT	Educator: Non-Teaching	
5	ELA	CT	Educator: Non-Teaching	
5	ELA	ME	Educator: General	
5	ELA	CT	General Public	
5	ELA	ND	Educator: General	
5	ELA	WV	Educator: General	
5	ELA	OR	Educator: General	
5	ELA	CA	Educator: General	
5	ELA	ID	Educator: General	
5	ELA	WI	Educator: Non-Teaching	
5	ELA	NC	Educator: General	
6	ELA	CT	Educator: General	VAC
6	ELA	ND	Educator: ELL experience	VAC
6	ELA	NH	General Public	
6	ELA	CA	Educator: ELL experience	
6	ELA	WI	Educator: SWD experience	
6	ELA	ME	Educator: General	
6	ELA	ID	Educator: General	
6	ELA	WI	General Public	
6	ELA	CA	Educator: Non-Teaching	
6	ELA	MT	Educator: General	
6	ELA	WA	Educator: General	
6	ELA	WY	Educator: Non-Teaching	
6	ELA	CT	Educator: Non-Teaching	

6	ELA	DE	Educator: Non-Teaching	
6	ELA	VI	Educator: Non-Teaching	
6	ELA	ME	Educator: General	
6	ELA	WV	Educator: ELL experience	
6	ELA	MI	Educator: General	
6	ELA	CT	Educator: General	
6	ELA	MT	Educator: Non-Teaching	
6	ELA	VI	Educator: Non-Teaching	
6	ELA	ME	Educator: Non-Teaching	
6	ELA	CA	Educator: SWD experience	
6	ELA	WA	Educator: Non-Teaching	
6	ELA	NH	Educator: General	
6	ELA	HI	General Public	
6	ELA	ID	Educator: General	
6	ELA	WV	Educator: Non-Teaching	
6	ELA	VI	Educator: Non-Teaching	
6	ELA	WV	Educator: General	
7	ELA	ME	Educator: General	VAC
7	ELA	NH	Educator: Non-Teaching	VAC
7	ELA	WA	Educator: General	VAC
7	ELA	OR	Educator: General	
7	ELA	CT	Educator: SWD experience	
7	ELA	CA	Educator: Non-Teaching	
7	ELA	CA	Educator: General	
7	ELA	WA	Educator: General	
7	ELA	DE	Educator: General	
7	ELA	ME	Educator: General	
7	ELA	MO	Educator: General	
7	ELA	MT	Educator: Non-Teaching	
7	ELA	CA	Educator: General	
7	ELA	WI	Educator: General	
7	ELA	WY	Educator: General	
7	ELA	WV	General Public	
7	ELA	VI	Educator: ELL experience	
7	ELA	WI	Educator: SWD experience	
7	ELA	WI	Educator: SWD experience	
7	ELA	CA	Educator: General	
7	ELA	HI	Educator: Non-Teaching	
7	ELA	ME	Educator: ELL experience	
7	ELA	CA	General Public	
7	ELA	VI	Educator: General	
7	ELA	DE	General Public	
7	ELA	CA	Educator: General	
7	ELA	IA	Educator: Non-Teaching	
8	ELA	MO	Educator: Non-Teaching	VAC
8	ELA	WY	General Public	VAC
8	ELA	WV	Educator: General	

8	ELA	CA	Educator: General	
8	ELA	DE	Educator: General	
8	ELA	VI	Educator: General	
8	ELA	CT	General Public	
8	ELA	VI	Educator: Non-Teaching	
8	ELA	NV	Educator: SWD experience	
8	ELA	MT	Educator: General	
8	ELA	VT	Educator: Non-Teaching	
8	ELA	ND	Educator: General	
8	ELA	HI	Educator: General	
8	ELA	WA	Educator: Non-Teaching	
8	ELA	ME	Educator: General	
8	ELA	ME	Educator: General	
8	ELA	VI	Educator: General	
8	ELA	DE	Educator: SWD experience	
8	ELA	WV	Educator: General	
8	ELA	CT	Educator: General	
8	ELA	CA	Educator: SWD experience	
8	ELA	ME	Educator: General	
8	ELA	MT	Educator: General	
8	ELA	OR	Educator: General	
8	ELA	WI	Educator: General	
8	ELA	ID	Educator: General	
8	ELA	CA	Educator: ELL experience	
8	ELA	CA	Educator: ELL experience	
8	ELA	ID	Educator: General	
8	ELA	MI	Educator: General	
11	ELA	HI	Educator: Non-Teaching	VAC
11	ELA	NC	Educator: General	VAC
11	ELA	SD	Educator: Non-Teaching	VAC
11	ELA	CA	Higher Education	VAC
11	ELA	WV	Higher Education	VAC
11	ELA	NC	Educator: General	VAC
11	ELA	WA	Educator: General	VAC
11	ELA	CA	Educator: Non-Teaching	VAC
11	ELA	CT	Educator: General	VAC
11	ELA	NH	Educator: General	VAC
11	ELA	CA	Educator: ELL experience	VAC
11	ELA	WI	Educator: SWD experience	VAC
11	ELA	NV	Educator: Non-Teaching	VAC
11	ELA	WV	Educator: General	VAC
11	ELA	OR	Higher Education	VAC
11	ELA	MT	Educator: Non-Teaching	VAC
11	ELA	CA	Educator: ELL experience	
11	ELA	HI	Educator: SWD experience	
11	ELA	VI	Educator: General	
11	ELA	CA	Higher Education	

11	ELA	WA	Higher Education	
11	ELA	ME	Educator: General	
11	ELA	MI	Higher Education	
11	ELA	DE	Educator: General	
11	ELA	NV	Higher Education	
11	ELA	ME	Educator: General	
11	ELA	MT	Higher Education	
11	ELA	CA	Educator: General	
11	ELA	MO	General Public	
11	ELA	DE	Educator: Non-Teaching	
11	ELA	WA	Higher Education	
11	ELA	NC	Higher Education	
11	ELA	IA	Higher Education	
11	ELA	CA	Educator: General	
11	ELA	WY	Higher Education	
11	ELA	ME	Educator: Non-Teaching	
11	ELA	NH	Educator: SWD experience	
11	ELA	ND	Educator: General	
11	ELA	IA	Higher Education	
11	ELA	OR	Educator: Non-Teaching	
11	ELA	CT	Higher Education	
11	ELA	SD	Higher Education	
11	ELA	ND	Educator: General	
11	ELA	CA	Educator: General	
11	ELA	HI	Higher Education	
11	ELA	CA	Educator: General	
11	ELA	ME	Higher Education	
11	ELA	CT	Educator: Non-Teaching	
11	ELA	MI	Higher Education	
11	ELA	WY	Educator: General	
11	ELA	ME	Educator: ELL experience	
11	ELA	ID	Higher Education	
11	ELA	DE	Educator: General	
11	ELA	OR	Higher Education	
11	ELA	OR	Higher Education	
11	ELA	WA	Higher Education	
11	ELA	OR	Educator: ELL experience	
11	ELA	IA	Educator: General	
11	ELA	WY	Educator: General	
11	ELA	MI	Higher Education	
11	ELA	ND	Higher Education	
11	ELA	WY	Educator: General	
11	ELA	WA	Educator: General	
11	ELA	CA	Educator: Non-Teaching	
11	ELA	WA	Higher Education	
11	ELA	MT	Educator: General	
11	ELA	CA	Educator: SWD experience	

11	ELA	CA	Higher Education	
3	Math	HI	Educator: General	VAC
3	Math	MI	Educator: General	VAC
3	Math	ID	Educator: Non-Teaching	
3	Math	NH	Educator: Non-Teaching	
3	Math	WI	Educator: Non-Teaching	
3	Math	MO	Educator: General	
3	Math	WA	Educator: General	
3	Math	ND	Educator: General	
3	Math	VI	Educator: Non-Teaching	
3	Math	DE	Educator: General	
3	Math	CA	Educator: ELL experience	
3	Math	CT	Educator: General	
3	Math	DE	General Public	
3	Math	WI	General Public	
3	Math	NC	Educator: Non-Teaching	
3	Math	WA	Educator: SWD experience	
3	Math	WV	Educator: General	
3	Math	ME	Educator: General	
3	Math	MT	Educator: General	
3	math	CA	Educator: ELL experience	
3	Math	NV	General Public	
3	Math	NV	Educator: Non-Teaching	
3	Math	CA	Educator: Non-Teaching	
3	Math	ME	Educator: General	
3	Math	CA	Educator: SWD experience	
3	Math	NH	Educator: General	
3	Math	IA	Educator: Non-Teaching	
3	Math	IA	Educator: General	
3	Math	VT	Educator: General	
3	Math	DE	Educator: Non-Teaching	
4	Math	MT	Educator: SWD experience	VAC
4	Math	VT	Educator: General	VAC
4	Math	WV	Educator: Non-Teaching	VAC
4	Math	MO	Educator: General	
4	Math	WI	Educator: General	
4	Math	MO	Educator: General	
4	Math	SD	Educator: Non-Teaching	
4	Math	IA	Educator: Non-Teaching	
4	Math	CT	Educator: General	
4	Math	ID	Educator: General	
4	Math	SD	Educator: General	
4	Math	WY	Educator: General	
4	Math	HI	Educator: ELL experience	
4	Math	DE	Educator: General	
4	Math	VT	Educator: General	
4	Math	WI	Educator: SWD experience	

4	Math	WA	Educator: SWD experience	
4	Math	ME	Educator: General	
4	Math	ND	Educator: General	
4	Math	VI	Educator: General	
4	Math	CT	Educator: General	
4	Math	HI	General Public	
4	Math	VT	Educator: Non-Teaching	
4	Math	OR	Educator: General	
4	Math	MT	Educator: General	
4	Math	HI	Educator: General	
4	Math	NH	Educator: General	
4	Math	NV	Educator: General	
4	Math	NH	Educator: Non-Teaching	
5	Math	CA	Educator: General	VAC
5	Math	CT	Educator: General	VAC
5	Math	ME	Educator: SWD experience	VAC
5	Math	HI	Educator: General	
5	Math	VT	Educator: General	
5	Math	WA	Educator: General	
5	Math	WI	Educator: General	
5	Math	WI	Educator: SWD experience	
5	Math	WI	Educator: General	
5	Math	HI	Educator: General	
5	Math	NV	Educator: General	
5	Math	NH	Educator: General	
5	Math	ID	Educator: Non-Teaching	
5	Math	HI	Educator: General	
5	Math	OR	Educator: General	
5	Math	ME	Educator: Non-Teaching	
5	Math	WY	Educator: General	
5	Math	NC	Educator: General	
5	Math	MO	Educator: General	
5	Math	CA	Educator: ELL experience	
5	Math	ID	Educator: General	
5	Math	VI	Educator: Non-Teaching	
5	Math	CA	Educator: ELL experience	
5	Math	ID	General Public	
5	Math	CA	Educator: ELL experience	
5	Math	WA	Educator: Non-Teaching	
5	Math	IA	Educator: Non-Teaching	
5	Math	WV	Educator: General	
5	Math	DE	Educator: ELL experience	
6	Math	ND	Educator: General	VAC
6	Math	IA	Educator: General	VAC
6	Math	NV	Educator: General	VAC
6	Math	MI	Educator: General	
6	Math	DE	Educator: General	

6	Math	MO	Educator: SWD experience	
6	Math	CA	Educator: General	
6	Math	WV	Educator: Non-Teaching	
6	Math	ID	Educator: General	
6	Math	MO	Educator: General	
6	Math	OR	Educator: General	
6	Math	ND	Educator: General	
6	Math	CA	Educator: ELL experience	
6	Math	NH	Educator: Non-Teaching	
6	Math	MT	Educator: Non-Teaching	
6	Math	HI	Educator: General	
6	Math	CA	Educator: General	
6	Math	ND	Educator: General	
6	Math	WA	Educator: General	
6	Math	IA	Educator: Non-Teaching	
6	Math	WV	Educator: Non-Teaching	
6	Math	ME	Educator: General	
6	Math	DE	Educator: General	
6	Math	CT	Educator: General	
6	Math	MT	Educator: General	
6	Math	VI	Educator: General	
6	Math	CA	Educator: SWD experience	
6	Math	OR	General Public	
6	Math	CA	Educator: ELL experience	
6	Math	WI	Educator: General	
7	Math	OR	Educator: Non-Teaching	VAC
7	Math	IA	Educator: General	VAC
7	Math	VI	General Public	VAC
7	Math	ID	Educator: General	
7	Math	MT	General Public	
7	Math	NV	Educator: Non-Teaching	
7	Math	WV	Educator: Non-Teaching	
7	Math	ME	Educator: General	
7	Math	CA	Educator: SWD experience	
7	Math	SD	Educator: General	
7	Math	WV	Educator: General	
7	Math	WA	Educator: General	
7	Math	CT	Educator: General	
7	Math	WV	Educator: General	
7	Math	HI	Educator: General	
7	Math	ND	Educator: General	
7	Math	WI	Educator: SWD experience	
7	Math	HI	Educator: General	
7	Math	CA	Educator: ELL experience	
7	Math	NH	Educator: General	
7	Math	NV	Educator: Non-Teaching	
7	Math	CA	Educator: Non-Teaching	

7	Math	WY	Educator: General	
7	Math	WA	Educator: General	
7	Math	MI	Educator: General	
7	Math	WI	Educator: Non-Teaching	
7	Math	HI	Educator: General	
7	Math	WV	General Public	
7	Math	SD	Educator: ELL experience	
7	Math	VT	Educator: General	
8	Math	DE	Educator: General	VAC
8	Math	ID	Educator: General	VAC
8	Math	HI	Educator: General	
8	Math	CA	Educator: ELL experience	
8	Math	CT	Educator: General	
8	Math	NC	Educator: General	
8	Math	ME	Educator: General	
8	Math	IA	Educator: Non-Teaching	
8	Math	NH	Educator: General	
8	Math	WV	Educator: ELL experience	
8	Math	HI	General Public	
8	Math	CA	Educator: ELL experience	
8	Math	CT	Educator: Non-Teaching	
8	Math	WV	Educator: General	
8	Math	VI	Educator: Non-Teaching	
8	Math	OR	Educator: General	
8	Math	DE	Educator: General	
8	Math	WI	Educator: SWD experience	
8	Math	OR	Educator: Non-Teaching	
8	Math	WA	Educator: General	
8	Math	ND	Educator: General	
8	Math	WY	Educator: Non-Teaching	
8	Math	MO	Educator: General	
8	Math	MT	Educator: Non-Teaching	
8	Math	ND	Educator: General	
8	Math	WI	Educator: Non-Teaching	
8	Math	NH	Educator: Non-Teaching	
8	Math	CT	Educator: Non-Teaching	
8	Math	MT	Educator: Non-Teaching	
11	Math	WV	Educator: General	VAC
11	Math	WV	Educator: General	VAC
11	Math	WA	Educator: SWD experience	VAC
11	Math	WA	Educator: General	VAC
11	Math	IA	Educator: General	VAC
11	Math	WA	Higher Education	VAC
11	Math	SD	Higher Education	VAC
11	Math	CT	Higher Education	VAC
11	Math	MO	Educator: General	VAC
11	Math	MT	Educator: ELL experience	VAC

11	Math	OR	Higher Education	VAC
11	Math	WV	General Public	VAC
11	Math	SD	Educator: Non-Teaching	VAC
11	Math	ID	Educator: General	VAC
11	Math	WI	Educator: SWD experience	VAC
11	Math	MI	Educator: General	VAC
11	Math	CA	Higher Education	
11	Math	WA	Higher Education	
11	Math	CA	Higher Education	
11	Math	ND	Higher Education	
11	Math	NC	Higher Education	
11	Math	MI	Higher Education	
11	Math	WA	Educator: General	
11	Math	ND	Educator: General	
11	Math	CA	Educator: Non-Teaching	
11	Math	CA	Educator: ELL experience	
11	Math	WA	Higher Education	
11	Math	CA	Educator: General	
11	Math	NH	Higher Education	
11	Math	WI	Educator: General	
11	Math	CA	General Public	
11	Math	NV	Educator: Non-Teaching	
11	Math	CA	Higher Education	
11	Math	MT	Educator: General	
11	Math	WY	Educator: General	
11	Math	MT	Educator: General	
11	Math	NV	Higher Education	
11	Math	ND	Educator: General	
11	Math	SD	Educator: General	
11	Math	CA	Higher Education	
11	Math	IA	Higher Education	
11	Math	NH	Educator: General	
11	Math	VT	Higher Education	
11	Math	WV	Higher Education	
11	Math	NH	Educator: General	
11	Math	CA	Higher Education	
11	Math	MI	Higher Education	
11	Math	HI	Higher Education	
11	Math	CT	Educator: General	
11	Math	DE	Higher Education	
11	Math	OR	Higher Education	
11	Math	CA	Educator: ELL experience	
11	Math	WA	Educator: SWD experience	
11	Math	NV	Educator: General	
11	Math	WI	Educator: General	
11	Math	WV	Educator: General	
11	Math	WV	Educator: General	

11	Math	DE	Educator: General	
11	Math	MT	Higher Education	
11	Math	ME	Higher Education	
11	Math	WY	Higher Education	
11	Math	MO	Higher Education	
11	Math	CA	General Public	
11	Math	HI	Educator: General	
11	Math	WA	Higher Education	
11	Math	OR	Educator: Non-Teaching	
11	Math	ID	Higher Education	
11	Math	ID	Educator: Non-Teaching	
11	Math	MO	Higher Education	
11	Math	OR	Educator: Non-Teaching	

Achievement Level Setting – Dallas, TX
 October 13–20, 2014
 Panelist Travel Schedule

	Arrival Date	Departure Date	Notification of VAC Selection
Grade 11 Panel	Sunday, October 12 th	Wednesday, October 15 th <i>*After 3:00 PM Central</i>	Tuesday, October 14 th
Grades 6,7,8 Panels	Tuesday, October 14 th	Friday, October 17 th <i>*After 3:00 PM Central</i>	Thursday, October 16 th
Grades 3,4,5 Panels	Thursday, October 16 th	Sunday, October 19 th <i>*After 3:00 PM Central</i>	Saturday, October 18 th
VAC	Sunday, October 19 th	Monday, October 20 th <i>*After 7:00 PM Central</i>	



Achievement Level Setting

Dallas, Texas
October 13–20, 2014

Registration Packet





Mandy Hunter
ELA 11A

E11A101

Front

Back

E11A101

→ Unique ID (01)

→ Table Number (1)

→ Grade (11A)

→ Subject (ELA)



**Kathleen
Stapleton**
ELA 11A

E11A102

For longer names

TABLE 1



TABLE 1



SuperShuttle
 Ground Transfer
Voucher
 Non-Transferable / Non-Refundable.

Measurement Inc.
 Hilton Anatole
 October 9-21, 2014

Name _____
 15 DFW 15 DAL
 + 18 % Gratuity

Arrival

-----tear here-----
Reservations are not necessary for arrival into D/FW Airport. Please be advised that SuperShuttle is a “shared ride” shuttle service – please allow for a 45-60 minute travel time upon boarding shuttle.

1. Claim your luggage.
2. Locate the Information Board in baggage claim and call SuperShuttle using the touch screen system; it is very important that you call from baggage claim so we know where you’re located; proceed to the “shared ride” area (upper level). Look for the orange signs that read “SuperShuttle”.
3. Look for our Guest Service Agent or Driver and let them know you’re staying at the Hilton Anatole; present arrival voucher to your driver and you will be transported to your hotel. ***Name must be printed legibly on voucher to be valid.***
4. For Love Field arrivals, reservations are required. Call 1-800-BLUE-VAN (258-3826).
5. Return Reservations: **Exclusive vans will be provided for departures.**
6. Cancellations require 4 hours notice prior to scheduled pick up for a refundable fare.

-----tear here-----

Smarter Balanced: Innovative Approaches to Achievement Level Setting



Joe Willhoft, Exec. Dir., Smarter Balanced
NCSA New Orleans June 25, 2014



Goals

- Obtain defensible cut scores
- Relate cut scores to external criteria
- Engage stakeholders in the process
- Make sense of cut scores across grades

Item Selection

- Total field test pool > 19,000 items/tasks
- “Standard Setting” sub-pools of 300 items per grade per subject (Total ~4,200)
- Std-setting pools matched grade-level blueprint

- Subset of 4.2M students in field test responded to std-setting items



Bookmark Approach

- Ordered Item Booklets (OIB) may contain up to 75 Smarter Balanced items/tasks
“Typical” operational test form will have ~40 item/tasks
- Increased size allows for an OIB with fewer gaps between items
- OIB will align to on-grade test blueprint

Sample OIB Page

Online Standard Setting

Welcome [Sample User!](#)



[Home](#)

[Instructions](#)

[Item Map](#)

OIB#: 4

Aenean eget dui mauris. Maecenas a felis quam. Maecenas ac vehicula justo. Donec dapibus arcu quis lorem ultrices convallis. Pellentesque habitant morbi tristique senectus et netus?

- A $\frac{1}{8}$
- B $\frac{2}{9}$
- C $\frac{2}{5}$
- D $\frac{7}{12}$

Key: C

[View Passage](#)

Item Statistics

RP67	-.452
p-value	.30
DOK	1
Scaled Score	235
CCSS	MA.6.EE.6:1

[Add Comment](#)

[Achievement Level Descriptor](#)

[Set Bookmark](#)

◀ Prev

Next ▶



Construct Maps and External Benchmarks

- Panelists also provided “Item Maps”
 - List of OIB items and their features
- Includes location of relevant external benchmarks
 - “Basic” and “Proficient” on 2013 NAEP Reading and Math (Gr. 4 & 8)
 - PISA Performance Levels (Gr. 11)
 - Across-state common-student college-readiness scores on ACT (Gr. 11)

Sample Construct Map

Online Standard Setting

Welcome [Sample User!](#)



Home

Instructions

Item Map

Math / Grade: 6

Submitted: 9/1/2014 10:00 AM

Page	RP67	Scaled Score	CCSS	DOK	Reviewed	Comment	Data Field 1	Data Field 2	Bookmark
1	-.477	213	MA.6.EE.7:1	1	<input checked="" type="checkbox"/>				
2	-.471	221	MA.7.NS.3:1	2	<input checked="" type="checkbox"/>				
3	-.463	229	MA.6.EE.7:2	1	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			
4	-.452	235	MA.6.EE.6:1	1	<input checked="" type="checkbox"/>				
5	-.438	241					View		
6	-.420	257	MA.6.EE.6:1	2					
7	-.416	268	MA.6.EE.6:1	1					
8	-.404	277	MA.6.EE.6:1	1					
9	-.392	283					View		

Submit



In-Person Ach. Level Setting

- Preparation
 - Endorsement of process by TAC; approved by Chiefs
 - Pilot of In-Person procedure in July
 - Recruitment of 504 participants in grade/subject panels
- In-Person Meetings (2 ½ days per grade/subject)
 - Overview and charge
 - Content Standards
 - Tests
 - ALDs
 - 3 Bookmark Rounds



Virtual “Distributed Ach. Level Setting”

- Purpose

- Provide In-Person panels with input for field at-large
- Engage more educators in process of Achievement Level Setting
- Professional development opportunity for K-12 and Higher Education partners

- Description

- Pilot test in July
- Mid-September, ~250,000 participants (~1 participant per 40 students)
- 3-hour online experience in two 90-minute sessions
 - ALD Training, Practice Test, Bookmark Training, One round/one cut
- Recommendations are aggregated by subgroup and provided as Round 1 impact data to In-Person panels



Vertical Articulation Virtual “Distributed Ach. Level Setting”

- Vertical Articulation
 - 4 members from each In-Person panel view recommendation across grade levels
 - Consider reasonableness of across-grade patterns
- Audit
 - Third party performance audit of all steps
 - Provided to TAC after Vertical Articulation for review
- November: Present recommendations to Chiefs
 - Chiefs asked to endorse recommendations
 - Promote adoption of Smarter Balanced cut scores through their own state decision-making processes

Overview of Achievement Level Setting Process

Joe Willhoft, Ph.D., Smarter Balanced
Michael B. Bunch, Ph.D., Measurement Incorporated



Smarter Balanced Chiefs' Meeting
November 6, 2014

Motions



Motion 1: Suspension of Rules

Smarter Balanced temporarily suspends its rules to allow Affiliate states to be included as voting members in matters related to Achievement Level Setting, in accordance with the voting rules described in the Smarter Balanced Governance Document as amended on August 26, 2014.

Motion 2: Endorse ALS Process

After review of the auditors' report, In-Person panelist feedback and endorsements by both the Achievement Level Setting Advisory Panel and the Smarter Balanced Technical Advisory Committee, Smarter Balanced states affirm that the Achievement Level Setting events and activities thoroughly and faithfully adhered to the plans for Achievement Level Setting that had previously been approved in April 2014.

Motion 3: Cut Scores for High School

Smarter Balanced states approve the achievement levels for Grade 11 English Language Arts/Literacy and for Grade 11 Mathematics as shown in the table below, displayed in vertically scaled logits.

Motion 4: Cut Scores for Grades 3-8

Smarter Balanced states approve the achievement levels for Grades 3-8 English Language Arts/Literacy and for Grades 3-8 Mathematics as shown in the table below, displayed in vertically scaled logits.

Overview





Completion of Recruiting Activities

- Online Panel
 - Over 10,000 registered
- In-Person Panel
 - 504 panelists confirmed
 - 10+ alternates confirmed
- VAC
 - 64 panelists confirmed



Completion of the Online Panel

- Opening of the Window
 - October 6-17
 - Individual windows and total window extended
- Level of Participation
 - 10,099 registered
 - 5,840 logged in
 - 2,660 submitted
- Support Provided
- Results Shared with In-Person Panel



Completion of the In-Person Panel

- Training Activities
 - Software
 - Common Core and ALDs
 - Ordered Item Booklet
- Panel Activities
 - Bookmark placement
 - Discussion

Agenda

Day 1

- Morning: Orientation
 - Common Core
 - Achievement Level Descriptors
 - Smarter Balanced Tests
 - Software
- Afternoon: Review of Ordered Item Booklet
 - Discuss items with others at table
 - Study additional resource materials

Agenda

Day 2

- Morning: Round 1
 - Practice in breakout rooms
 - Make bookmark recommendations individually
- Afternoon: Round 2
 - Discuss Round 1 bookmark placements at each table
 - Place Round 2 bookmarks individually

Day 3

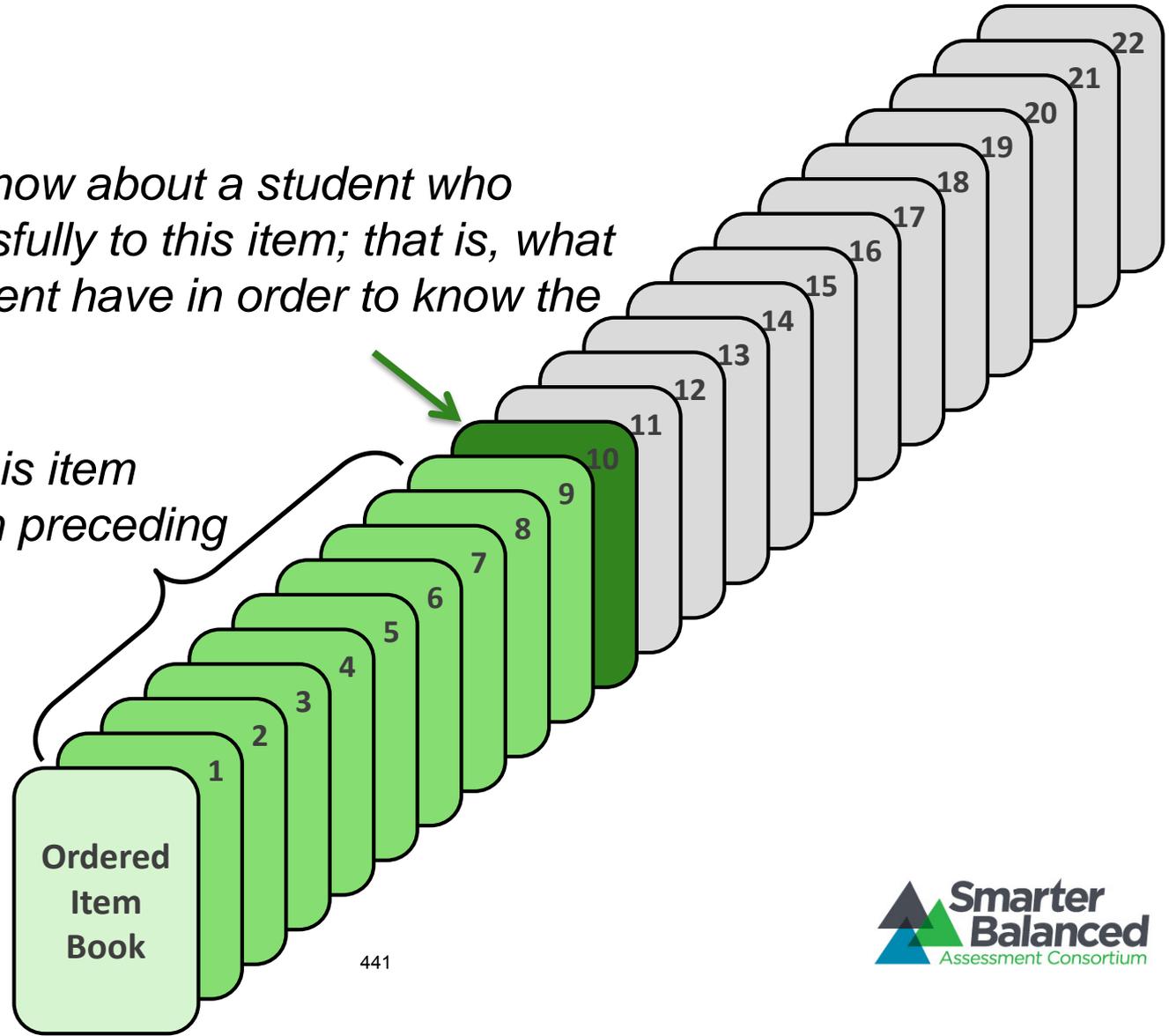
- Morning: Round 3
 - Discuss Round 2 bookmark placements for the entire room
 - View supporting data based on Round 2 bookmarks
 - Place Round 3 bookmarks individually
 - Review final recommendations
 - Evaluate the process

Training

As you study each item in the OIB, discuss two questions with your fellow panelists:

1. *What do you know about a student who responds successfully to this item; that is, what skills must a student have in order to know the correct answer?*

2. *What makes this item more difficult than preceding items?*



Ordered Item Booklet: Achievement Level Setting

English Language Arts/Literacy, 4

Page: 02

2. *What makes this item more difficult than preceding items?*

35758

Read the sentence and the question that follows.

My aunt was hungry and wanted to eat before the movie but the rest of the family
wanted to see it before dinner.

What is the **best** way to correct the error in the sentence?

- (A) Add a hyphen (-) after *movie*.
- (B) Add a comma (,) after *movie*.
- (C) Add a semicolon (;) after *movie*.
- (D) Add an exclamation mark (!) after *movie*.

Ordered Item Booklet: Achievement Level Setting

English Language Arts/Literacy, 4

Page: 60

[Back to Item Map](#)

◀ Previous

Next ▶

Item

Question Information

Passages and Other Materials **3**

Achievement Level Descriptors **1**

57202

In Source #1, the author says that "friend behavior can be important to an animal's survival." Explain what the author means and give **two** details, using **two** sources, that support your explanation. For each detail include the source title or number.

Achievement Level Descriptors

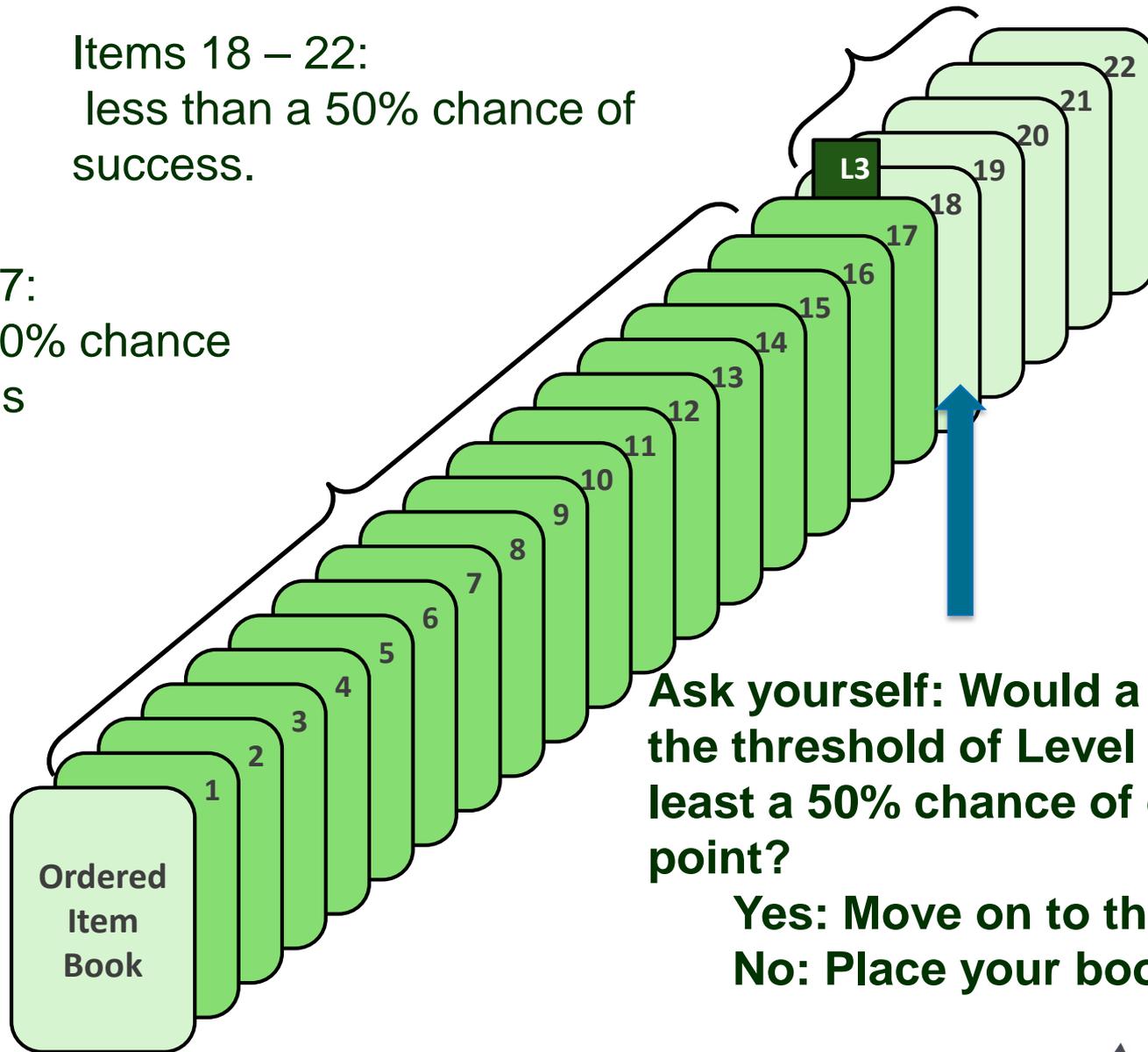
The student who just enters Level 3 should be able to:

**Targets
1–7:
Reading
Literary
Text**

- Use details and information from texts of moderate complexity to support answers and inferences.
- Identify or summarize central ideas/key events in texts of moderate complexity.
- Begin to determine the intended meanings of words, including words with multiple meanings, based on context, word relationships, word structure, and use of resources in texts of moderate complexity.
- Use supporting evidence to justify/explain own inferences in texts of moderate complexity.
- Interpret, specify, or compare how information is presented across texts of moderate complexity.
- Begin to relate knowledge of text structures, genre-specific features, or formats to obtain, interpret, explain, or connect information within texts of moderate complexity.
- Determine or interpret figurative language, literary devices, or connotative meanings of words and phrases used in context and partially explain the impact of those word choices on meaning and tone in texts of moderate complexity.

Items 18 – 22:
less than a 50% chance of
success.

Items 1-17:
At least 50% chance
of success



Ask yourself: Would a student at the threshold of Level 3 have at least a 50% chance of earning this point?

**Yes: Move on to the next item.
No: Place your bookmark here.**



Evaluations

How confident are you about the three bookmarks you just entered?

Bookmark	Very Confident	Confident	Uncertain	Very Uncertain	Total
Level 2	222 (47%)	237 (51%)	10 (2%)	0 (0%)	469
Level 3	234 (50%)	220 (47%)	15 (3%)	0 (0%)	469
Level 4	245 (52%)	217 (46%)	7 (1%)	0 (0%)	469

Overwhelming endorsement of process (92-99% positive) on 14 separate measures.



Completion of Vertical Articulation

- Training
 - Orientation
 - Groundrules
- Procedures
 - Motion
 - Second
 - Discussion
 - Vote: 2/3 majority required
- Results
 - Aligned cut scores across grades
 - Eliminated scaled score reversals

Audit





Focus of Audit

- Pilot Test
- Software
- Online Panel Process
- In-Person Workshop
- Vertical Articulation
- Adherence to Plan



Audit Report

“It is my conclusion that the standard setting activities described in this report were designed and conducted appropriately so as to yield defensible performance standards grounded in the knowledge, skills, and expectations represented by the ALDs.”

(Auditors’ Report, p. 52)

Advisory Panel Statement

After reviewing the information about the Achievement Level Setting activities and the auditors' report, the Advisory Panel confirms the design and procedures for the Achievement Level Setting and the Vertical Articulation were implemented as planned, represent a valid process that is consistent with best practices in standard setting, and support the defensibility of the content-based performance standards.

Unanimously endorsed October 28, 2014



TAC Statement

The Technical Advisory Committee concludes that the Smarter Balanced achievement level setting design and implementation reflect contemporary professional practice and represent a valid process that supports the defensibility of the content-based performance standards.

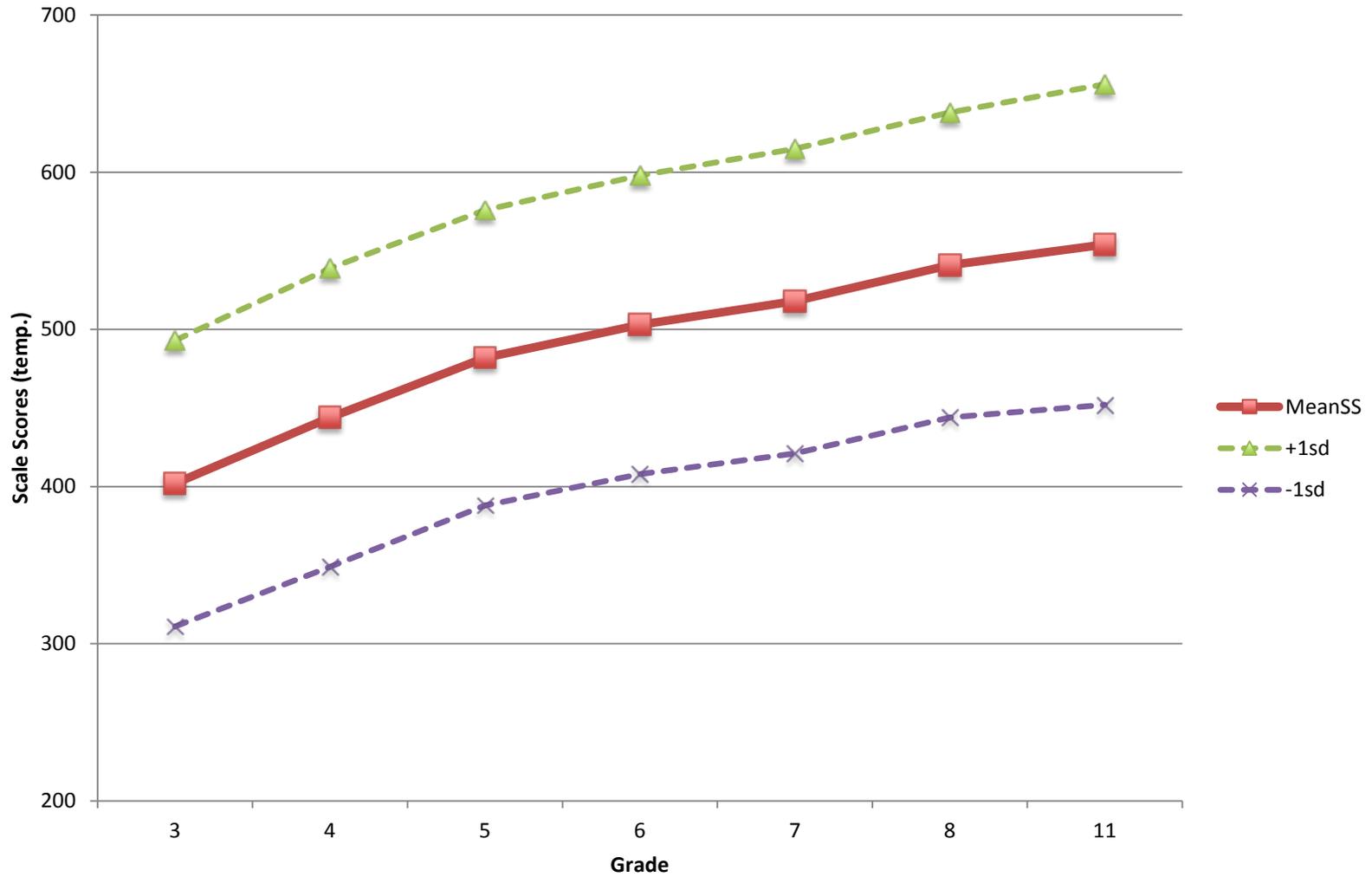
Unanimously endorsed October 30, 2014



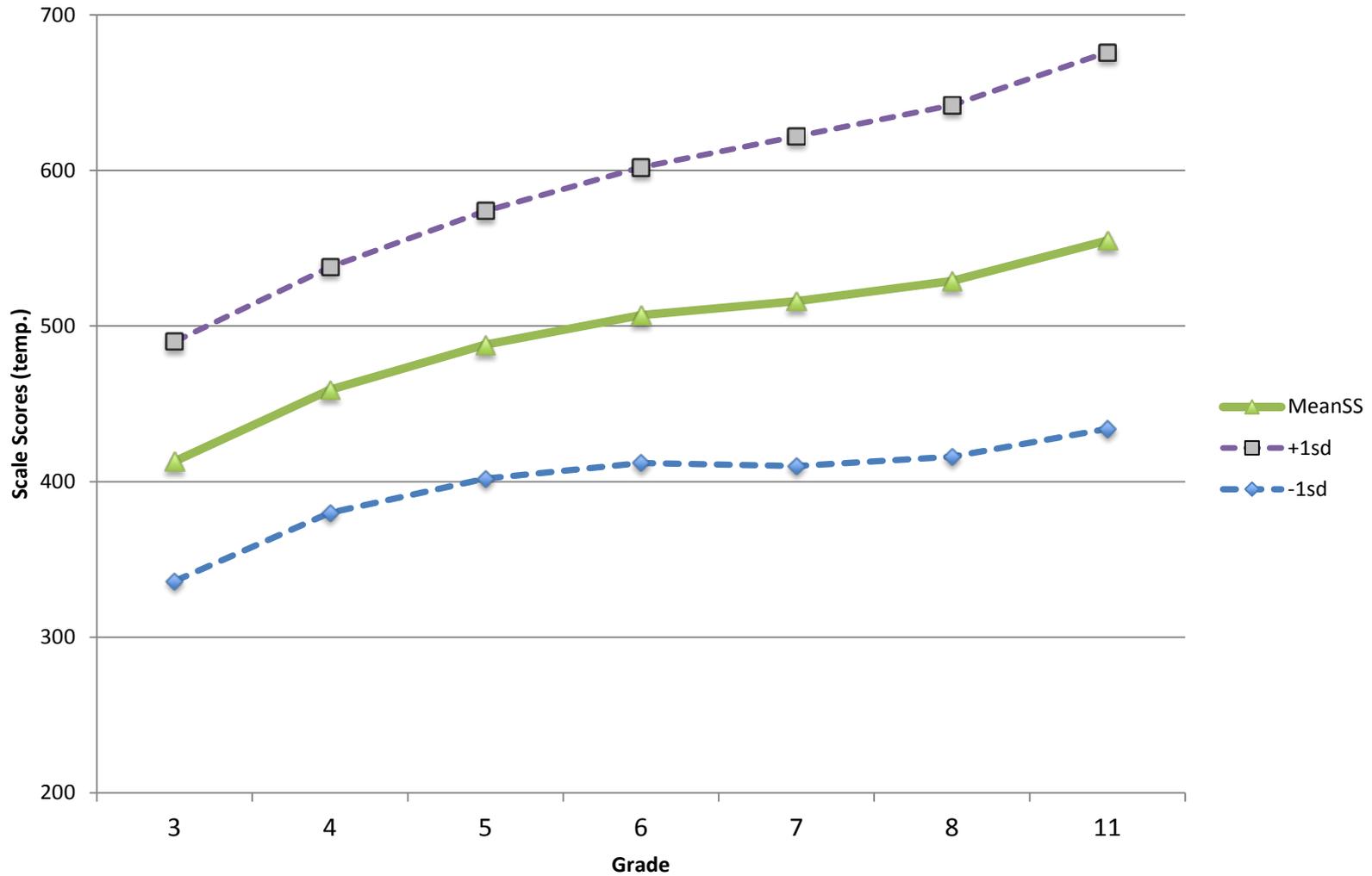
Recommendations from ALS Panels



ELA Scale Score Means Across Grades



MATH Scale Score Means Across Grades



Impact Data by Subgroup

		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
Male	11851	4346	36.7	3129	26.4	2405	20.3	1971	16.6	63.3	36.9
Female	11372	3433	30.2	2993	26.3	2483	21.8	2463	21.7	69.8	43.5
Hispanic	7411	3419	46.1	2087	28.2	1230	16.6	675	9.1	53.9	25.7
Native	404	210	52.0	96	23.7	67	16.6	31	7.7	48.0	24.3
Asian	1860	389	20.9	397	21.4	457	24.5	617	33.2	79.1	57.7
Black	1595	782	49.0	444	27.9	234	14.6	135	8.5	51.0	23.1
White	10914	2651	24.3	2828	25.9	2699	24.7	2736	25.1	75.7	49.8
Pacific Islander	703	312	44.4	188	26.7	121	17.2	82	11.7	55.6	28.9
Multi-Ethnic	1055	338	32.0	271	25.7	204	19.4	242	22.9	68.0	42.3
Individualized Education Program	2176	1355	62.3	457	21.0	220	10.1	144	6.6	37.7	16.7
English Language Learner	4356	2507	57.6	1199	27.5	474	10.9	176	4.0	42.4	14.9
Section 504	176	74	42.0	51	29.0	35	19.9	16	9.1	58.0	29.0
Economically Disadvantaged	12554	5706	45.5	3538	28.1	2143	17.1	1167	9.3	54.5	26.4

Participation Summary

	Online Panel	In-Person Panel	Cross-Grade Review Committee
Teachers	65%	64%	66%
Non-Teacher Educators & Administrators	18%	18%	19%
Higher Education	9%	12%	11%
General Public	7%	6%	5%
Total	2,660	482	64

Figures reflect the number of individuals who submitted a bookmark during the Online Panel and who attended the In-person Panel or Cross-Grade Review.

458

NAEP Comparison

Grade 4	Below Basic/1	Basic/2	Proficient/3	Advanced/4	Proficient or Higher
NAEP Reading	32%	33%	27%	8%	35%
Smarter Balanced ELA	36%	22%	23%	19%	42%
NAEP Math	17%	41%	34%	8%	42%
Smarter Balanced Math	28%	36%	24%	13%	37%

Grade 8	Below Basic/1	Basic/2	Proficient/3	Advanced/4	Proficient or Higher
NAEP Reading	22%	42%	32%	4%	36%
Smarter Balanced ELA	27%	30%	33%	10%	43%
NAEP Math	26%	38%	27%	9%	36%
Smarter Balanced Math	49%	29%	15%	7%	22%

NAEP Comparison

High School	Below Basic/1	Basic/2	Proficient/3	Advanced/4	Proficient or Higher
NAEP Reading	25%	37%	32%	5%	37%
Smarter Balanced ELA	27%	31%	31%	12%	43%
NAEP Math	35%	39%	23%	3%	26%
Smarter Balanced Math	52%	26%	16%	6%	22%

ACT/SAT Verbal

State	Number Tested	Composite Mean	English Mean	Met English Benchmark	Reading Mean	Met Reading Benchmark	Writing Mean	Met Writing Benchmark
North Carolina (ACT)	94,210	18.5	17.0	44%	18.7	31%	5.6	32%
Michigan (ACT)	105,777	19.8	18.9	57%	19.8	38%	NA	NA
Montana (ACT)	8,930	20.1	19.0	58%	20.7	50%	6.7	NA
Idaho (SAT)	16,579	1363	NA	NA	464	38%	438	26%



ACT/SAT Math

State	Number Tested	Math Mean	At or Above Benchmark
North Carolina (ACT)	94,210	19.2	30%
Michigan (ACT)	105,777	19.6	32%
Montana (ACT)	8,930	20.4	38%
Idaho (SAT)	16,579	461	36%

To: Achievement Level Setting In-Person Panelists
From: MI_smarterbalanced21@measinc.com
Subject: Smarter Balanced Achievement Level Setting In-Person Panel

This email is being sent on behalf of the Smarter Balanced Assessment Consortium. For questions please email MI_smarterbalanced21@measinc.com.

Dear Chiefs and State Leads:

Smarter Balanced will host a dinner meeting at the Council of Chief State School Officers (CCSSO) Conference on **Friday, November 14**. A buffet dinner will be served beginning at 6:30 p.m. in the Crystal Ballroom at the US Grant hotel in downtown San Diego. The discussion will run until 9:00 p.m.

For each member state, Smarter Balanced will support the travel of the chief (or designee), the K-12 Lead, and the Higher Education Lead, unless, of course, they have already arranged to be at the CCSSO conference.

As quickly as possible, but no later than noon (local) on Tuesday, each K-12 Lead needs to determine who from your state will attend the Friday dinner meeting. This information should be forwarded to Mandy Hunter (MI_smarterbalanced21@measinc.com) from Measurement Inc. so that she can coordinate travel arrangements and other logistics. Be sure to include:

State
Full Name as it appears on the identification the traveler will present to TSA
Email address
Phone (office & personal)
DOB
Gender
Role (Chief, Designee, K-12 Lead, HE Lead)

If you have already arranged to be at the CCSSO conference and plan to attend the Chiefs dinner meeting please forward that information (name, state, role) to Mandy Hunter (MI_smarterbalanced21@measinc.com) from Measurement Inc. by noon (local) on Tuesday.

We will also be providing a dial-in line for this meeting, and you will receive that information later this week.

We look forward to working with you in the days ahead as the Consortium continues to deliberate on this important decision. Please do not hesitate to contact Consortium staff with any questions or concerns.

Dear Chiefs and State Leads:

Copies to Executive Committee and Executive Staff

Attached is a summary of what staff has been working on since we met last Thursday. It's in the form of a PowerPoint, and concludes with two recommended options each for ELA and Math. In both cases, Option 1 is what you have already seen from our meeting on Thursday. Option 2 is what staff is bringing to you as our recommendations.

As I mentioned when we met, there is compelling data compiled by NAGB from multiple sources suggesting that something between 36-39% of the nation's 12th graders are prepared to enter into credit-bearing college coursework with a reasonable expectation of success. The recommendations from our VAC panels undershot that mark in Math (VAC recommends a more rigorous L3 cut score); and overshot it in ELA (VAC recommends an easier L3 cut.)

Keeping in mind that our panelists (including VAC) were focused on content, it's not much of a surprise that their recommendations would be different from the NAGB findings. Be that as it may, we probably need to take the NAGB findings into account, but also not completely dismiss the recommendations and hard work of our own content experts. So, the staff recommendations represent a "halfway" point between the VAC and NAGB results.

Take a look at the ppt deck. In hopes that we can come to closure on Friday, please do what you can to make sure K-12 and Higher Ed have a chance to dialogue.

Thanks,

Joe

Joe Willhoft, Executive Director

Smarter Balanced Assessment Consortium

(253) 381-1954

Overview of Achievement Level Setting Process

Joe Willhoft, Ph.D., Smarter Balanced
Michael B. Bunch, Ph.D., Measurement Incorporated



Smarter Balanced Chiefs' Meeting
San Diego, CA Nov. 14, 2014

Agenda

- 6:30 – 7:30 Dinner Buffet
- 7:00 – 7:15 Introductions/Overview
- 7:15 – 7:45 Achievement Level
Recommendations
- 7:45 – 9:00 Discussion
- 9:00 Next Steps / Adjourn

Motions



Motion 1: Adoption of Position Paper

Smarter Balanced member states endorse the position paper “Interpretation and Use of Scores and Achievement Levels” dated November 14, 2014, as amended, and recommend that member states attend to it when considering the use of scores from Smarter Balanced assessments.

Motion 2: Cut Scores for High School

Smarter Balanced states approve the achievement levels for Grade 11 in Mathematics and English language arts/Literacy as shown in the accompanying table, displayed in Smarter Balanced scale scores.
(staff recommendations)

Motion 3: Cut Scores for Grades 3-8

Smarter Balanced states approve the achievement levels for Grades 3-8 in Mathematics and English language arts/Literacy as shown in the accompanying table, displayed in Smarter Balanced scale scores.
(staff recommendations)

What Staff Has Done Since 11/6



1. Corrected Gr. 11 Impact Data

- A file transfer error was identified with Gr. 11 impact data. (Gr. 9 & 10 responses had been included with Gr. 11 files) (No effect on Ordered Item Booklet or on In-Person recommendations)
- Gr. 11 ELA impact data had been 42.6% “At or above level 3”; is now 42.8%
- Gr. 11 Math impact data had been 22.0% “At or above level 3”; is now 26.4%

2. Created Smarter Balanced Scale Scores

- Scale scores on Smarter Balanced assessments
 - Use a **four digit number**
 - Range roughly from **2000-to-3000**
- Avoids confusion with other, commonly-used scales such as: percentile ranks, percent correct, ACT, SAT, Lexiles, etc.
- Provides adequate “space” on the scale to see growth in whole-number units

3. Position Statement About Use of Ach. Levels & Scale Scores

- Collaboration among staff and advisors for purpose of articulating need for and pitfalls of using achievement levels
- Will be proposed for adoption by Smarter Balanced states on Nov 12

4. Developed Three Principles to Guide Staff Recommendations

1. Honor the work of the panelists who contributed their content expertise to the process
1. Give consideration to external evidence of student readiness for credit-bearing college work: Notably NAGB Grade 12 college preparedness data
1. Maintain a “system perspective” by using information from all grades (and both content areas) to make recommendations that support a coherent system

5. Developed Options for Consideration

- 2 options in Math and ELA respectively.
 - Original recommendation from VAC
 - Option 2 is the staff recommendation.
- The staff recommendation:
 - Adheres to the “Guiding Principles”
 - Closes the gap between Panel recommendations and NAGB estimates of “College Preparedness” – giving equal weight to Panel recommendations and NAGB findings
 - Leaves other Panel recommendations (L2 and L4) as is, with minor exceptions to separate L4 from L3

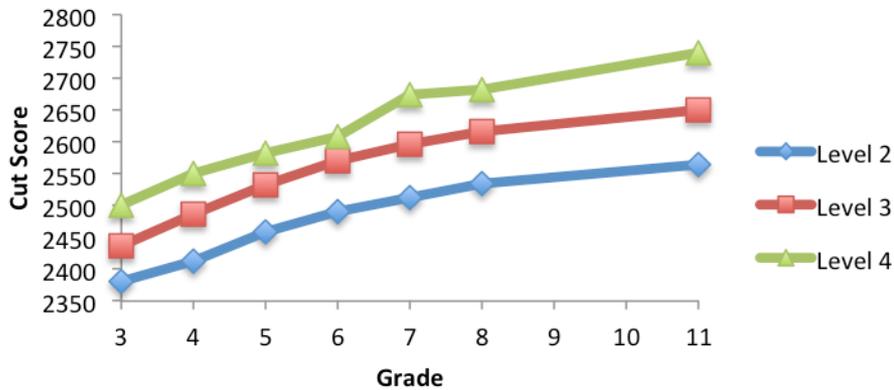
Recommendations -- Math



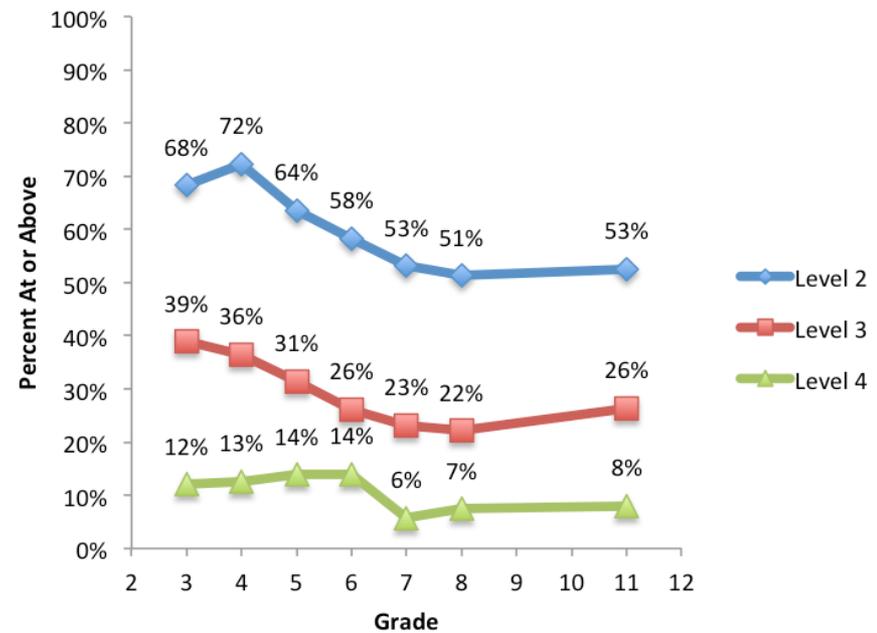
Math Option 1

VAC Recommendations

**Cut Scores - Math
(VAC Recommended)**



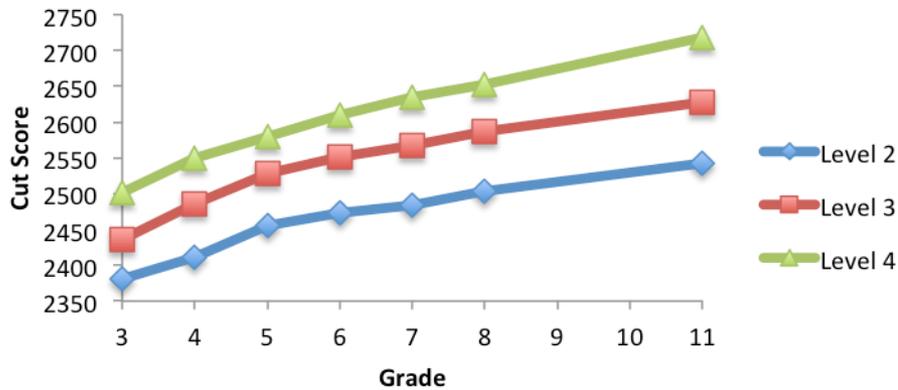
**Pct. At or Above - Math
(VAC Recommended)**



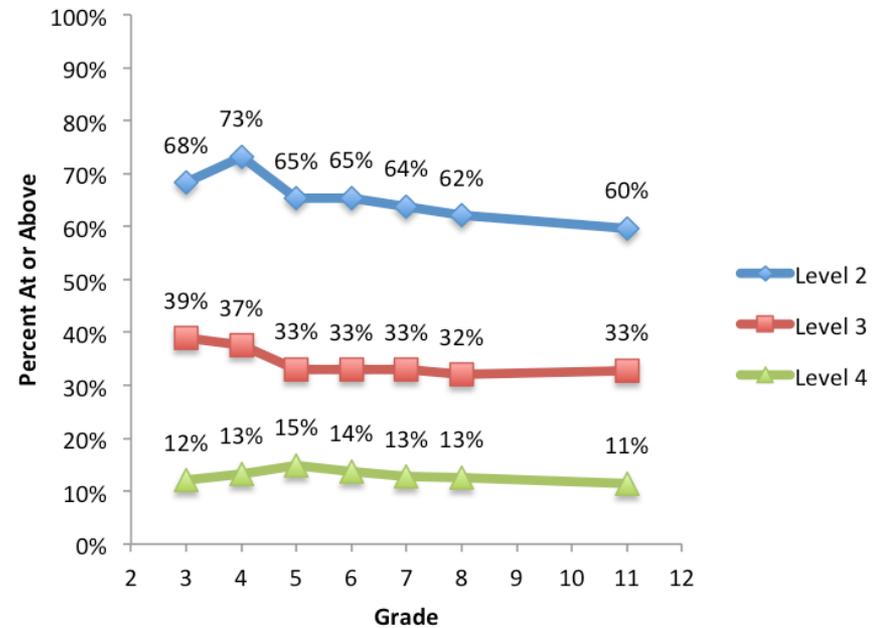
Math Option 2

Staff Recommendations

**Cut Scores - Math
(Staff Recommended)**



**Pct. At or Above - Math
(Staff Recommended)**



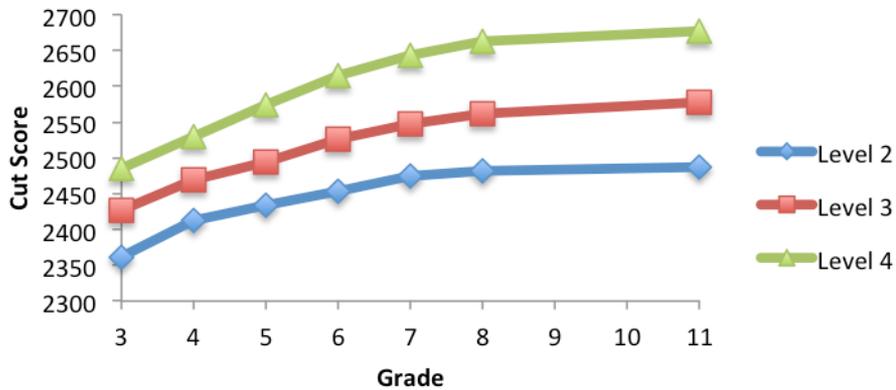
Recommendations -- ELA



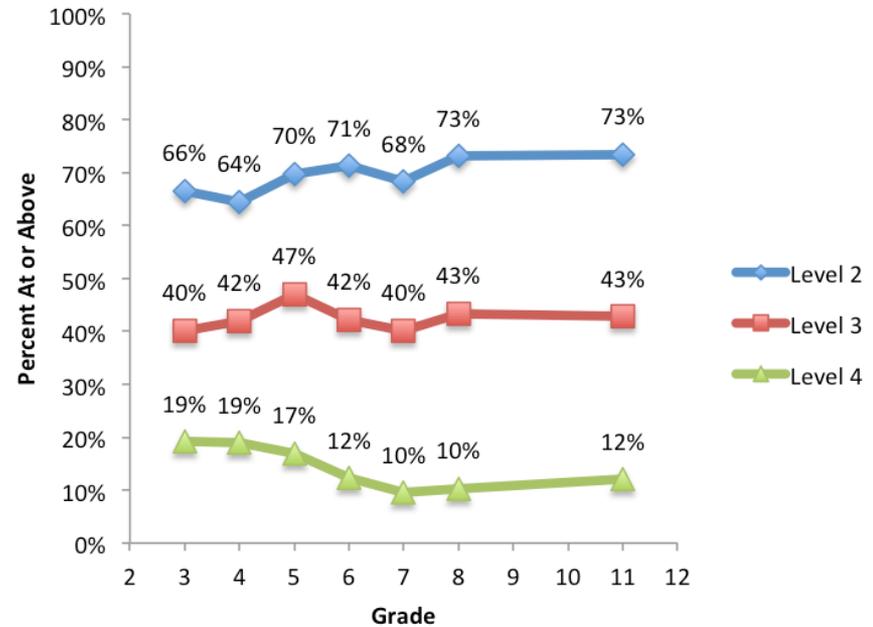
ELA Option 1

VAC Recommendations

**Cut Scores - ELA
(VAC Recommended)**



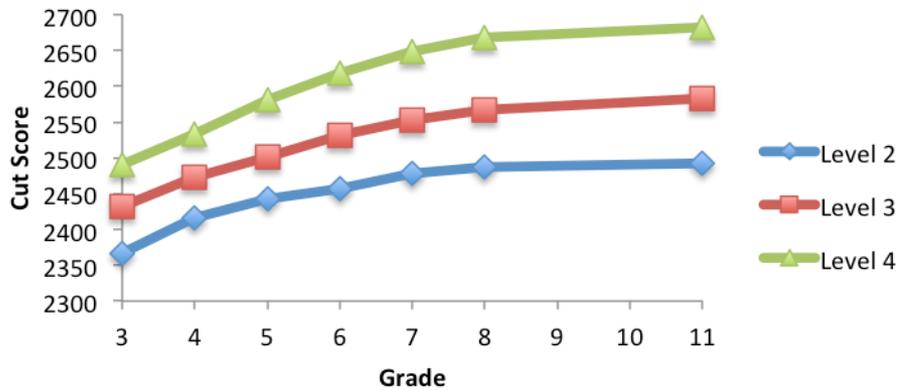
**Pct. At or Above - ELA
(VAC Recommended)**



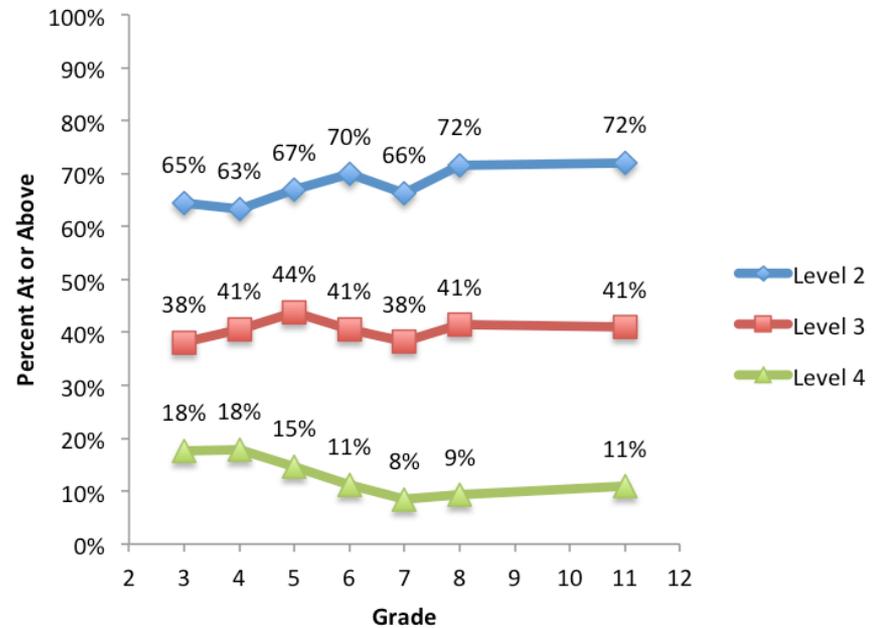
ELA Option 2

Staff Recommendations

**Cut Scores - ELA
(Staff Recommended)**



**Pct. At or Above - ELA
(Staff Recommended)**



Discussion



Gr. 11 Recommended Cut Scores (VAC and Staff)

Grade 11 (VAC)

Math Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
11	2565	2650	2740

ELA Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
11	2488	2578	2677

Grade 11 (Staff)

Math Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
11	2543	2628	2718

ELA Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
11	2493	2583	2682

Gr. 3-8 Recommended Cut Scores (VAC and Staff)

Grades 3-8

(VAC)

Math Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
3	2381	2436	2501
4	2413	2487	2551
5	2459	2532	2583
6	2491	2570	2609
7	2513	2596	2674
8	2534	2616	2683

ELA Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
3	2362	2427	2485
4	2413	2470	2530
5	2434	2494	2574
6	2453	2527	2614
7	2474	2547	2644
8	2482	2562	2663

Grades 3-8

(Staff)

Math Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
3	2381	2436	2501
4	2411	2485	2549
5	2455	2528	2579
6	2473	2552	2610
7	2484	2567	2635
8	2504	2586	2653

ELA Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
3	2367	2432	2490
4	2416	2473	2533
5	2442	2502	2582
6	2457	2531	2618
7	2479	2552	2649
8	2487	2567	2668

**Motions and Actions of Smarter Balanced Chiefs and Delegates
November 14, 2014**

Motion 1: Adoption of Position Paper

Smarter Balanced member states endorse the position paper “Interpretation and Use of Scores and Achievement Levels” dated November 14, 2014, as amended, and recommend that member states attend to it when considering the use of scores from Smarter Balanced assessments.

Motion passed 20-0 with no abstentions.

Motion 2: Cut Scores for High School

Smarter Balanced states approve the achievement levels for Grade 11 in Mathematics and English language arts/Literacy as shown in the accompanying table, displayed in Smarter Balanced scale scores. (staff recommendations)

Grade 11 (Staff)

Math Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
11	2543	2628	2718

ELA Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
11	2493	2583	2682

Motion passed 18-0 with 2 abstentions.

Motion 3: Cut Scores for Grades 3-8

Smarter Balanced states approve the achievement levels for Grades 3-8 in Mathematics and English language arts/Literacy as shown in the accompanying table, displayed in Smarter Balanced scale scores. (staff recommendations)

Grades 3-8

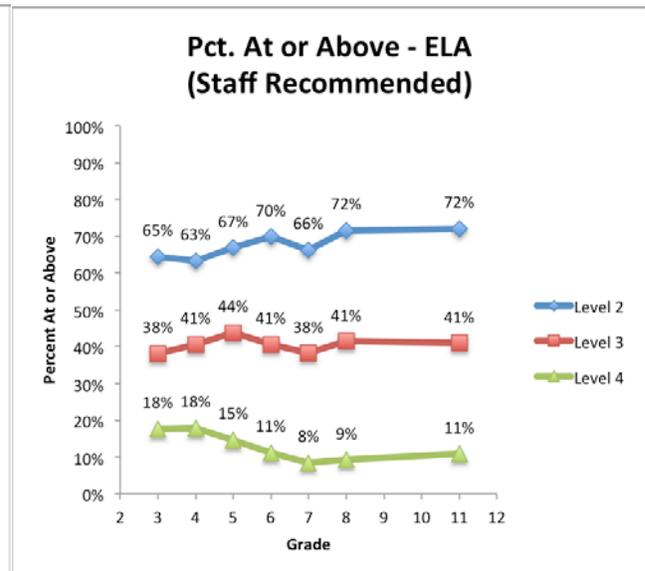
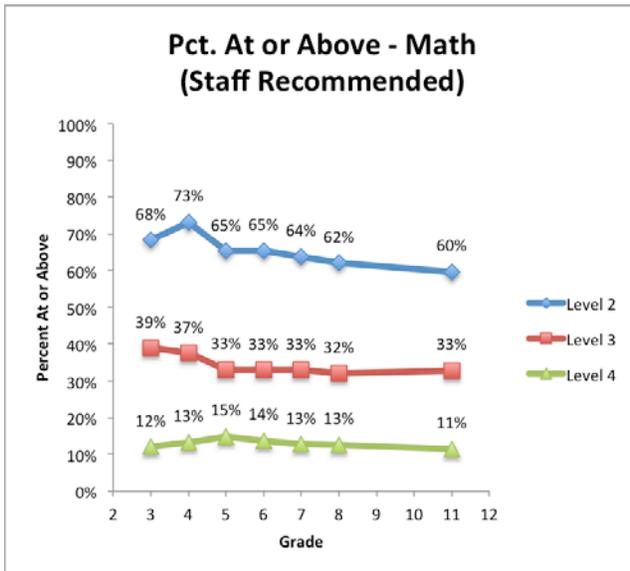
(Staff)

Math Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
3	2381	2436	2501
4	2411	2485	2549
5	2455	2528	2579
6	2473	2552	2610
7	2484	2567	2635
8	2504	2586	2653

ELA Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
3	2367	2432	2490
4	2416	2473	2533
5	2442	2502	2582
6	2457	2531	2618
7	2479	2552	2649
8	2487	2567	2668

Motion passed 18-0 with 2 abstentions.

Impact of Motions 2 and 3:



Appendix C – Audit Report and Related Documentation

- Auditors' Report 489

Observation and Report on
Smarter Balanced Standard Setting

October 12–20, 2014

Prepared by:

Gregory J. Cizek, PhD
and
Heather Koons, PhD

October 23, 2014

I. Overview

The Smarter Balanced Assessment Consortium (hereafter *Smarter Balanced*) desired to establish achievement levels on newly developed assessments in English Language Arts (ELA) and Mathematics. Large and representative groups of qualified participants were empaneled to perform the standard setting procedures for high school assessments (11th grade) on October 13-15, 2014, for middle grades (6-8) assessments on October 15-17, 2014, and for elementary (grades 3-5) assessments on October 17-19, 2014. The three two and one-half day sessions for each test were implemented using an adaptation of the Bookmark procedure (Lewis, Mitzel, Mercado, & Schulz, 2012). The in-person sessions were preceded by an on-line input gathering process and followed by a vertical articulation process comprising a stakeholder review of the performance standards recommended by the participants. The standard setting activities were conducted by contractors Measurement, Incorporated (MI) and CTB/McGraw-Hill (CTB).

The authors of this report responded to a request by Smarter Balanced that an independent, external auditor be present for the sessions to verify that the standard setting process and procedures followed the plan that had previously been reviewed by the Smarter Balanced Technical Advisory Committee and approved by Smarter Balanced governing states, and to submit a report of observations and findings. The auditors selected, Dr. Gregory J. Cizek and Dr. Heather H. Koons, have expertise and extensive experience in the area of setting performance standards (see, e.g., Cizek, 2001, 2012; Cizek & Bunch, 2007; Cizek, Bunch, & Koons, 2004). Dr. Cizek was the lead auditor and was present to perform observations of the in-person sessions on October 13-16 and 18-19 and the vertical articulation session on October 20th. Because of a prior commitment, Dr. Cizek was unavailable to observe on October 17th; Dr. Koons observed on that day. Drs. Cizek and Koons coordinated their observations to include pre-workshop

communications, an on-site briefing when Dr. Koons arrived at the sessions; a short period of overlapping observations on October 16th; a second debriefing on the morning of October 18th; and an additional time of overlapping observations on October 18th. Both auditors contributed to the preparation of this report summarizing their observations, and both reviewed interim report drafts. The conclusions expressed in the Summary and Recommendations section of this report are those of Dr. Cizek, in consultation with Dr. Koons.

This report provides a description of the standard setting activities, some recommendations, and a summary evaluation. The report is organized into five sections: 1) Overview; 2) On-line Procedures; 3) In-person Observations; 4) Vertical Articulation; 5) Summary and Recommendations; and 6) References.

II. On-line Procedures

Demonstrations of software used for the on-line procedures (and to be used in the operational standard setting), and software used for automated assembly of the Ordered-Item Booklets (OIBs) were provided at a meeting on October 7, 2014 at the MI headquarters in Durham, North Carolina. The meeting provided information on the materials made available to the on-line participants, which included a 12-page document on how to use the on-line tool; links to various Smarter Balanced assessment resources, and a practice activity. A field test of the on-line materials and procedures was conducted by MI in Durham, North Carolina in August 2014.

The on-line component of the standard setting was opened for participation on Monday, October 6, 2014 and the last window closed on Saturday, October 18, 2014 at 11:59 p.m. Eastern time. Participants using the on-line tool were able to review an OIB and provide their judgment regarding placement of a single bookmark for Level 3. A practice activity and various resources were provided for participants. Among the reference materials available to participants were the achievement level descriptors (ALDs), the content specifications for the test, bookmark placement instructions, and relevant practice tests. Participants were then provided with a six-item orientation round OIB which was provided to help familiarize them with the bookmarking task. For each item in the OIB, certain information was provided to participants, including target and claim identification, Depth of Knowledge (DOK) addressed, stimulus materials, and key/rubric/scoring information.

During the practice activity, set one bookmark for Level 3. It was possible, however, for participants to have skipped completing the practice OIB, although the on-line tool recorded if the practice activity was performed; it was not recorded if participants reviewed the ALDs, if they reviewed the on-line tool information provided, or if they took practice test before actually engaging in the OIB review and providing their judgments. Participants then completed the

operational OIB, comprising approximately 70 items, and again placed one (Level 3) bookmark. Participants did not need to review every item; they could skip to a portion of OIB where they believed the Level 3 bookmark would most likely be appropriate. Some NAEP and PISA items were embedded in the OIBs and identified as such for participants; of the NAEP and PISA items viewed, at least on NAEP item (Grade 11 ELA) appeared to be missing a stimulus, and a PISA item included spellings that might not be familiar to students accustomed to North American English (e.g., “labour”).

During their OIB review, participants were permitted to take a break at any time. At the end of the activity, participants who submitted their ratings were offered the opportunity to receive a certificate of completion; the certificate was automatically provided to any participant who completed their bookmark placements using three hours or more; however, there was some confusion about the certificates with some participants who took less than three hours being asked about their desire for a certificate, but who were not provided with one.

Overall, the on-line tool appeared to be well-designed and appropriate for the task of collecting on-line judgments. For the future, or as research activities for the present standard setting, some suggestions would include:

- 1) It may be desirable to produce an FD record of how many items were reviewed by on-line participants.
- 2) It may be desirable to improve or simplify on-screen navigation. Performance tasks with stimuli seemed somewhat difficult to navigate between several screens, especially considering the need to view the item, the stimuli/passages, scoring guides, ALDs, exemplar responses for constructed-response items, and so on.
- 3) It would also seem desirable to provide participants with scoring rules for all multi-part

items (e.g., Must all be correct to get credit? Is each element scored separately?).

- 4) It may be desirable to keep the text with explicit directions for the bookmarking task on screen at all times when actual item judgments are being made.

In addition to review of the on-line tool, an interview and demonstration was conducted with the creators of OIBs. An automated assembly program was used to create the 14 OIBs needed for the operational standard setting. Overall, it appeared that the procedures were very successful in creating OIBs that appropriately covered the test specifications.

III. In-person Observations

Procedures for the in-person standard setting workshops and vertical articulation were piloted during the week of August 18, 2014 in Durham, NC. Dr. Koons was present to observe the piloting from from 1:30 p.m. through the end of the day on August 18th and from 1:30-3:30 p.m. on August 19th. Dr. Koons provided feedback to Dr. Bunch of Measurement Inc., based on her observations.

Prior to the operational sessions, the auditors were provided with several documents for review. These materials included:

- * Smarter Balanced Achievement Level Setting Plan
- * Smarter Balanced Achievement Level Setting Plan Supplement
- * Selecting Items for the ALS
- * Achievement Level Setting Agenda
- * In-person Standard Setting Facilitator Scripts
- * In-person Workshop PowerPoint Slides

The operational sessions took place from October 12-19, 2014 in Dallas, Texas. Three two and one-half days sessions were conducted for ELA and mathematics: one for high school (grade 11); one for middle grades (6-8); and one for elementary grades (grades 3-5).

The first sessions (grade 11) were held on October 13-15, 2014. The evening prior to this session, a brief orientation was provided for selected participants who were identified to serve in the role of table leaders. The following morning, the operational standard setting began. A target sample of 72 panelists for each subject area was desired; however some participants cancelled plans to attend the workshop in the days immediately preceding the event. The grade 11 participants (as well as participants on other grade level panels) were chosen to be representative

of the teaching population of Smarter Balanced member states; final demographic characteristics of all participants will be documented in a forthcoming technical report produced by the standard setting contractor. For grade 11, a total of 68 participants in ELA and 70 participants for mathematics were empaneled in four rooms as follows:

- * 34 ELA panelists in Room A (ELA-A)
- * 34 ELA panelists in Room B (ELA-B)
- * 35 Mathematics panelists in Room A (Math-A)
- * 35 Mathematics panelists in Room B (Math-B)

For grades 6, 7, and 8, a total of 87 participants in ELA and 89 participants for mathematics were empaneled in six rooms as follows:

- * 30 Grade 6 ELA panelists
- * 27 Grade 7 ELA panelists
- * 30 Grade 8 ELA panelists
- * 30 Grade 6 Mathematics panelists
- * 30 Grade 7 Mathematics panelists
- * 29 Grade 8 Mathematics panelists

For grades 3, 4, and 5, a total of 80 participants in ELA and 88 participants for mathematics were empaneled in six rooms as follows:

- * 26 Grade 3 ELA panelists
- * 27 Grade 4 ELA panelists
- * 27 Grade 5 ELA panelists
- * 30 Grade 3 Mathematics panelists

* 29 Grade 4 Mathematics panelists

* 29 Grade 5 Mathematics panelists

The following sections provide the auditors' observations for each of the grade band sessions. One caveat should be noted. At all times, one auditor was present to view the activities. However, at any given time, there were four or six breakout rooms working on the standard setting tasks. It was not possible to view all rooms at the same time, nor did it seem advisable to observe only one of the rooms for an entire session. Instead, auditors attempted to spend smaller blocks of time in each room. The advantage of this strategy was that the auditors gained a fairly good sense the approaches and particularities of each group. A disadvantage of this strategy was that the activities of a single group were not typically observed from start to finish. Thus, details of how sessions progressed were sometimes not fully apprehended until the observations were pieced together across several days and sessions. As a consequence of this, the details of some aspects of the sessions are dispersed throughout the report. That is, some aspects of the sessions were able to be fully reported on Day 1 of a panel's activities, some on Day 2, and some on Day 3; some aspects were fully observed in one grade level band, other aspects were fully observed in another grade level band. Readers are encouraged to consider the entire report to gain a complete picture of the entire process.

Grade 11 Sessions, Day 1: Morning Activities (Monday, October 13, 2014)

Grade 11 session activities began with participant registration and breakfast the morning of the first day. All registration and other logistics appeared to proceed without any issues. After breakfast, participants met in a full group (i.e., ELA and mathematics, combined) for an orientation session in a separate ballroom. The whole group orientation session occurred from 8:30-9:30 a.m. on the first day. It began with a welcome and overview presentation by Dr. Joe

Willhoft, Executive Director of Smarter Balanced. The presentation included an overview of the general purpose of the sessions, provision of some background and context, questions, solicitation of participants' concerns (e.g., role of Smarter Balanced at college entry; role of Smarter Balanced in career training readiness; test taker need for keyboarding skills), the introduction of key Smarter Balanced content leadership in ELA and mathematics and announcement of the presence of state and other observers. Finally, a "parking lot" was announced, whereby participants could indicate concerns or questions to their individual room facilitators, with those concerns and questions to be addressed subsequently in a whole group session.

The orientation continued with a presentation by Dr. Mike Bunch, Senior Vice President of MI, who described the agenda and goals for the workshop, along with brief attention to background on the common core state standards (CCSS), the ALDs, the practice test, the bookmark procedure, the concept of an OIB, performance tasks, evaluations, and "key shifts" embodied in the CCSS such as less breadth but deeper, more conceptual understanding, emphasis on cross grade coherence and application in mathematics, and more complex texts and increased presence of nonfiction texts in ELA. Dr. Bunch reminded participants of the focus on claims across subjects, with the four ELA claims focusing on 1) Reading, 2) Writing, 3) Speaking/Listening, and 4) Research/Inquiry; and mathematics claims focusing on 1) Concepts and Procedures, 2) Problem Solving, 3) Modeling and Data Analysis, and 4) Communication and Reasoning. The presentation introduced participants to the four types of ALDs—Policy, Range, Threshold, and Reporting—and indicated that the focus in the workshop would be on the threshold ALDs describing the knowledge and skills of students just entering Levels 2, 3, and 4. The presentation also included a brief overview of the question and test development history and processes, the various item formats comprising Smarter Balanced assessments, including selected-response (SR), constructed-response (CR), and technology enhanced (TE) formats, as well as information

on the mode of administration (computer-based), and available supports and accommodations for SWDs. Finally, Dr. Bunch noted a few “ground rules” for the workshop, including the fact that they were assembled to recommend, not set, performance levels, the requirement that participants sign a security/confidentiality agreement, and that participants engage in a group process where each panelist can freely contribute to discussions.

Dr. Bunch then introduced Deborah Sigman, co-chair of the Smarter Balanced executive committee, who thanked participants for their involvement and their efforts with Smarter Balanced to improve teaching and learning for students. To conclude the session, Dr. Bunch provided information on the room assignments for the mathematics and ELA groups and introduced the room facilitators:

HS ELA-A: Craig Deville (MI) and Anne Wilder (MI)

HS ELA-B: Gretchen Shultz (CTB) and Ric Mercado (CTB)

HS Math-A: Winne Reid (MI) and Maude Eno (MI)

HS Math-B: Judy Hickman (CTB) and Jennifer Lord-Besson (CTB)

Participants were then dismissed to their content area breakout rooms. At the conclusion of the session, Dr. Willhoft asked observers to remain in the ballroom, where he conducted a brief session to instruct the state-representative observers in the ground rules for their observations. (Smarter Balanced member states had been afforded the opportunity to send observers to the in-person standard setting activities.). Dr. Willhoft welcomed the representatives and admonished them to engage in non-intrusive movement between rooms, to avoid engaging in conversations with panelists while in the session rooms (although OK to do so at breaks or afterward), prohibiting them from speaking after the standard setting sessions about the results of the sessions, about specific content on the assessment, or about individual panelists, but encouraging them to

feel free to speak freely after the sessions with colleagues and others about the standard setting process.

Participants began working in their breakout groups beginning at 9:45 a.m.. Rooms were arranged in rectangular tables of six panelists, with each panelist having his/her own computer monitor, keyboard, and (for ELA) headsets. Participants were instructed to stay at the same table for the session, but they could change seats within a table group if they desired. In general, the meeting spaces seemed somewhat cramped, with little work surface for participants, tables that were fairly close to each other, and height of monitors that sometimes made it difficult for participants on opposite sides of a table to engage in discussion. (Some participants could more easily communicate with persons next to them or at an adjacent table than with other participants at their own tables.) Overall, however, the arrangements did not appear to present an overwhelming obstacle to the conduct of the sessions and seemed conducive for participants to perform their tasks.

The breakout sessions began with a welcome from the facilitators, individual panelist introductions, housekeeping information (such as non-disclosure forms, multimedia permission release forms (for audio, video, and photographic recording of participants), and a panelist information questionnaire. An introduction was then provided to the software to be used in the sessions. The software appeared to work well; isolated minor issues were addressed quickly and effectively by MI personnel. Facilitators then reinforced and extended information on the claims, targets and standards covered by the Smarter Balanced assessments, and provided additional information on the ALDs, with a focus on the threshold ALDs. Participants were reminded that they would be setting a single bookmark for Level 3 during an orientation session, and they would be setting three bookmarks during the operational bookmarking, beginning with the Level 3 bookmark, followed by Level 4, then Level 2. The rooms each adjourned for a break at

approximately 10:40 a.m., and resumed at approximately 10:55 a.m.

When they reconvened, the groups engaged in several tasks to prepare them for operational standard setting, including taking the Smarter Balanced "training test" that allowed them time to explore SR, CR/performance tasks, and TE item formats, experience with the keyboarding/mousing activities that students would need to apply, and experience with the sound and headphones used (for ELA). The review of on-line and technology features appeared to go smoothly in each room; the monitors appeared to be large and clear enough for ease of use, with stimuli and items both appearing on the same screen and limited scrolling required. Participants then reviewed the Smarter Balanced "practice test." It was somewhat unclear how the purposes of the "training tests" and "practice tests" differed, although the practice test appeared to give participants an opportunity to review some of the available accommodations for SWD (e.g., text to speech).

At approximately noon on Day 1, a room-level activity was conducted, with participants solicited for their reactions to the item reviews. Participants were generally favorable in their perceptions, noting that the items and tasks appeared to reflect rigorous, real-life, and transferable knowledge and skills. Some less favorable observations touched on the limitations of the test interface that did not allow cut-and-pasting from source documents, the cognitive demand of "topic jumping" from one challenging task to another within the assessment, and the general impression that the rigor of the assessments was not well aligned to the current abilities of many students and the lack of instructional resources available in schools to support instruction covering the CCSS. The groups then broke for lunch at approximately 12:25 p.m., with instruction to reconvene in the ballroom for a whole group activity at 1:20 p.m.

Grade 11 Sessions, Day 1: Afternoon Activities (Monday, October 13, 2014)

At the end of the lunch break on Day 1, in a whole group session, Dr. Willhoft and Mr.

Tony Alpert, Chief Operating Officer for Smarter Balanced, read and answered "parking lot questions" that had arisen in each of the rooms. The lunch break/parking lot activity ended at 1:20 p.m., and participants transitioned to the ballroom for a whole-group activity.

At approximately 1:35 p.m., Dr. Dan Lewis (CTB) initiated a whole-group presentation in the ballroom. The presentation began with an overview of remainder of the day and the processes in which participants would engage, including an introduction to the interface that panelists would be using, the resources, keys/scoring guides/rubrics that would be available, the OIB, general information on the bookmarking process, and specific directions on the cognitive analysis of each item they would be performing. Dr. Lewis provided specific information on the two questions participants must consider when engaging in the cognitive analysis: 1) What knowledge and skills must a student have in order to know the correct answer to an item? and 2) What characteristics of the item make it more challenging than the preceding item(s). Although the slides used by Dr. Lewis were produced using a small font that was somewhat difficult for the participants to see, the explanations were clear and participants appeared to grasp the key aspects of the presentation.

Participants were then dismissed to return to their content area rooms. From approximately 2:00 p.m. through the end of Day 1, participants worked in the breakout rooms on the task of reviewing their item map/OIBs and gaining an integrated conceptualization of the content of the Smarter Balanced assessment. Using the on-line interface, participants worked through the OIB, answering the two main questions, and adding comments on items to note the knowledge and cognitive skills tapped by each items. This activity continued throughout the afternoon, although many participants did not appear to have sufficient time to complete the activity by the end of the day.

At 2:15 p.m. on Day 1, CTB personnel provided one of the auditors (Cizek) with a demonstration of Bookmark Pro software—the software that would be used to derive performance

standards for all of the Smarter Balanced assessments under consideration. The software actually performs multiple functions, including the collection of panelist judgments, data analysis, and production of graphic and tabular results at the end of each round. Overall, the software appeared to be a very streamlined and useful utility for standard setting.

All participants were dismissed for the day between approximately 5:00 and 5:30 pm on Day 1. At approximately 5:30 p.m., a debriefing session was conducted in the ballroom, attended by contractor staff, Smarter Balanced representatives, and the auditor. Two major issues were raised: 1) some software concerns that arose early on Day 1, and 2) the insufficient time allocated for participants to complete the cognitive analyses of the OIBs. With respect to the first issue, facilitators noted that they had made slight adjustments in their presentations, such that the issue did not appear to be consequential. A considerable amount of time was spent discussing the second issue. It was decided to alter the Day 2 schedule to allow time for participants to complete the Day 1 activity in the morning by inserting some "catch up" time and slightly shortening the time allocation for Round 1 bookmark placements.

Grade 11 Sessions, Day 2: Morning Activities (October 14, 2014)

The second day of grade 11 standard setting consisted of a whole-group opening session beginning at 8:30 a.m. in the ballroom. The session commenced with Dr. Bunch announcing the adjustments in the agenda to allow participants to complete their cognitive analyses of the OIBs. The next portion of the session consisted of a presentation by Dr. Lewis who provided a more detailed description of the bookmark procedure and specific information and instructions on threshold students, bookmark placement for Level 3, Level 2, and Level 4. Dr. Lewis then answered questions from the total group and offered clarifications on issues such as: Why are some items repeated in the OIB? What about performance of threshold students on the items after the

bookmark? What is the relationship between threshold performance and specific targets? What about items that appear later in the OIB, but that appear to assess very fundamental skills? Is Level 3 passing? The whole group session ended at approximately 9:00 a.m. and participants were dismissed to go to their breakout rooms.

From 9:00-9:45 a.m., participants engaged in the “catch up” activity to complete the cognitive analysis of the OIBs. By approximately 9:45 a.m., all groups had begun a practice activity in which they placed a Level 3 (only) bookmark using a six-item orientation-round OIB and other resources. During this activity, one ELA room took early break (10:05 a.m.) because of some connection issues experienced by participants who were not able to continue the practice activity. (Apparently, the connectivity issue was not limited to one room or to the practice activity. However, the on-site IT personnel worked to resolve issues each time they arose and issues appeared to decline as the Day 2 activities continued.) The group reconvened at approximately 10:20 a.m. and all participants finished the practice ratings. Ratings were then collected and facilitators polled the room to see where bookmarks had been set. Software for presenting results to the group projected on a screen did not appear to work, so medians, etc. could not be presented. (The software worked correctly in another room observed.) The facilitator addressed the issue by asking for a show of hands for bookmark placements on each page. Discussion ensued as to participants' rationales for bookmark placements on selected pages.

Operational Round 1 bookmark placements then began at approximately 10:45 a.m. In one room observed, panelists were directed to place their bookmarks independently, first for Level 3, then Level 4, then Level 2; in another room observed, panelists were told that they could work together to place their Round 1 bookmarks. In all rooms observed, panelists were also directed to complete the first questionnaire upon closing out of the practice round before beginning their Round 1 operational bookmark placements.

In at least two of the rooms observed, participants appeared to have continuing, sporadic difficulty with connectivity of their computers. Some participants completed their task beginning at approximately 11:30 a.m. and were instructed to return for the next activity following lunch at 12:50 p.m. Subsequently, participants were instructed to return at 2:00 p.m.

Grade 11 Sessions, Day 2: Afternoon Activities (October 14, 2014)

The Day 2 afternoon session began at 2:00 p.m., with participants meeting in breakout groups for an introduction to the process that would be used to discuss the Round 1 results and to place the Round 2 bookmarks. Some technological difficulties were again experienced; however, facilitators were still able to present relevant information on the Round 1 bookmark placements for the group to consider, and technical specialists were sought out to help remedy the problems. A block of time was then allocated for participants to discuss the rationales for their Round 1 bookmark placements in table groups, beginning with their placement of the Level 3 bookmark. (The same process was used for Grade 11 and subsequent grade level sessions; namely, discussions following the setting of Round 1 bookmarks were at the table level; discussions following Round 2 were at the room level.) At the conclusion of the discussion, participants were dismissed for a short afternoon break at approximately 3:10 p.m.

The groups reconvened at approximately 3:25 p.m. They were presented with information on the on-line panel's results and an end-of-round questionnaire. At one point during the afternoon activities, Dr. Willhoft and Mr. Alpert visited the panel rooms to answer specific content area questions that were more relevant to the breakout groups than to the whole group and were not addressed during the lunch-hour "parking lot" presentation. The groups then were instructed to finish completing the end-of-round 1 questionnaire, submit the questionnaire, and the Round 2 activities began.

To complete the Round 2 bookmark placements, the groups were instructed to go back to their item maps. The item maps used for Round 2 contained new information on item scale locations. The scale location information allowed participants to discern relative differences in OIB item difficulties so that, if electing to move their bookmarks for Round 2, the relative magnitude of impact of a move on the resulting cut scores could be seen. New information on NAEP and PISA item difficulty was also presented on the item map used for Round 2. Facilitators attempted to describe how the NAEP and PISA performance should be taken into account, although some participants seemed unclear as to precisely how to use the information. To complete their Round 2 ratings, participants were instructed to again begin with a review of their Level 3 bookmark placement, then proceed to review their Level 4 bookmark, then their Level 2 bookmark. In two of the groups, some participants appeared to complete their Round 2 bookmark placements by approximately 4:15 p.m. with nearly all participants completing their work by 4:30 p.m. In the other two rooms, most participants did not finish until approximately 5:15 p.m. or later. At the end of Day 2, all participants were instructed to complete the Round 2 questionnaire and were dismissed individually for the day, noting that the activities for Day 3 would be begin at 8:30 a.m. in grade/content area breakout groups, not as a whole group as had been the case for the first two days.

At the conclusion of Day 2, another debriefing session took place at 5:30 p.m. in the ballroom, involving contractor staff, Smarter Balanced representatives, and the auditor. No major issues from the day were identified by the group.

Grade 11 Sessions, Day 3: Morning Activities (October 15, 2014)

The final day of grade 11 standard setting began at 8:30 a.m. with subject area groups meeting in their breakout rooms to begin Round 3. The sessions began with the facilitator

providing information related to logistics (reimbursement forms, airport shuttles, etc.), followed by some questions and answers about the process to date and what would be the next steps for the results.

The facilitator then presented a graphic display of participants' page judgments for the three achievement levels. A tabular breakdown of results (median bookmark placements) by table and room overall was then presented. Another slide was then shown to participants, providing them with impact data, based on their median Round 2 recommendations, and based on Smarter Balanced assessment field test data. A discussion ensued regarding the group's perspectives on the impact, what the distribution of student performance should look like, and other issues.

In one of the rooms observed (an ELA panel), the distributions of page judgments across the levels appeared to be fairly unimodal and with modest variability. The group appeared in general to find the impact to be realistic, with the possible exception of the Level 2 results. Some panelists noted that they put forth greatest effort in their work on the Level 3 and Level 4 cut scores, experiencing some fatigue by the time they reached the Level 2 cut, combined with some sense that the Level 2 threshold description was somewhat less clear for them in terms of operational utility. In another group (Math), there seemed to be polarization in the results, which were uniformly bimodal across the three levels. The discussion in this group appeared to reveal a lack of clarity about the appropriate referent for the standard setting task, with one viewpoint (from those in the higher modal area of the distributions) being that the content of the items accurately represented what the threshold description indicated, and the other viewpoint (from those in the lower modal area of the distributions) being that their judgments reflected more appropriate expectations for what "real students" can do.

Observation of one other room suggested what might be a moderate concern: Although in earlier rounds, there appeared to be an understanding among panelists about grounding judgments

exclusively in the threshold descriptions, as more impact and other data were provided to participants, the more some appeared to “drift” from grounding their judgments exclusively in the threshold and range ALDs. More than one panelist articulated what were clearly wrong strategies for how they placed their bookmarks: one was wrong in the sense that they were not explicitly attempting to translate the threshold descriptions into page judgments based on the content of the items; another was wrong in the sense that they placed their bookmark at the first item where they thought the threshold Level N student fell below the RP50 criterion, as opposed to going beyond that item to see if other items beyond the identified item were still within the RP for the Level N student. It is possible that the drift might, to some extent, be attributable to a facilitator effect. (Over the course of the sessions, it seemed clear that facilitators in different rooms had differing degrees of experience and comfort leading standard setting workshops, an observation which will be elaborated upon later in this report.) One strategy for dealing with this drift might be to keep the bookmarking task posted at all times when panelists are making judgments. For example, during the cognitive analysis of the OIB activity, a slide with the two questions to be answered was kept on the screen throughout the entire activity so that panelists could constantly refer to it. Similarly, posting the precise bookmarking task directions might also help; for example:

- 1) “Referring to the threshold Level N student, place your bookmark on the first page where the threshold Level N student would have less than a 50% chance of answering the item correctly (or obtaining that score point or greater).”
- 2) “Don’t stop at the first item where the chance of answering correctly drops below 50%, but examine items following the one you identify to be sure you have identified an appropriate bookmark location.”

In all groups, an additional slide was then projected, showing participants the impact data

that would result, based only on the on-line participants' judgments. A final slide showed results separately for ethnic groups, gender, LEP status, and other subgroups.

Participants then completed a questionnaire and began their Round 3 review and judgments. Before beginning their Round 3 judgments, the facilitators highlighted a new piece of information on the item maps: information on the range of items on the OIB item maps related to an ACT benchmark score. Discussion of the relevance of the ACT information ensued, and many participants questioned the relevance, representativeness, and usefulness of the ACT data. Participants' work on Round 3 began at approximately 9:30 a.m. Panelists took a break of approximately 40 minutes while the data from Round 3 were submitted for analysis and prepared for presentation. Three slides were shown to panelists: one showing the final median bookmark placements for the room, one showing the percentages (impact) that would fall into each category implied by the final group median placements, and one showing the percentages as a histogram. Panelists then completed a final questionnaire. The facilitators thanked the panelists for their work, provided additional information about lunch, airport transportation, reimbursement forms, and other logistics. As they were adjourned, panelists' materials were collected.

Grades 6-8 Sessions, Day 1: Morning Activities (October 15, 2014)

At the same time that the Grade 11 groups were completing their Round 3 ratings, the standard setting session activities for Grades 6, 7, and 8 ELA and mathematics began. The agenda for grades 6-8 mirrored that of grade 11, with table leader orientation occurring the evening of October 14, 2014 and general participant registration and breakfast the morning of October 15, 2014. After breakfast, grade 6-8 participants met in a full group (i.e., ELA and Mathematics, combined) for an orientation session in a separate ballroom. The whole group orientation session began with a welcome and overview presentation by Dr. Willhoft including an overview of the

general purpose of the sessions, provision of some background and context, questions, solicitation of participants' concerns. The orientation continued with a presentation by Dr. Bunch, who described the agenda and goals for the workshop, along with brief attention to background on the common core state standards (CCSS) including claims and targets, the item and test development process, assessment item formats, the ALDs, the practice test, the bookmark procedure, the concept of an OIB, performance tasks, evaluations, "key shifts" embodied in the CCSS, and the four types of ALDs with an emphasis on the threshold descriptions. Finally, Dr. Bunch noted the "ground rules" for the workshop, including the fact that participants were assembled to recommend, not set, performance levels, the requirement that participants sign a security/confidentiality agreement, and that participants engage in a group process where each panelist can freely contribute to discussions. To conclude the session, Dr. Bunch provided information on the room assignments for the mathematics and ELA groups, and introduced the room facilitators:

Grade 6 ELA - Chris Dunbar and Corey Palermo (MI)

Grade 6 Math - Kelly Bolton and Rick Mercado (CTB)

Grade 7 ELA - Molly Buck and Sarah Hagge (CTB)

Grade 7 Math - Lisa Johnson and Jennie Bowen (MI)

Grade 8 ELA - Amy Griswold and Dan Bowen (MI)

Grade 8 Math - John Upchurch and Jennifer Lord-Bessen (CTB)

Participants were then dismissed to their content area rooms where they began working in their breakout groups. The rectangular arrangement of tables that had been used for the grade 11 sessions was maintained for the the grades 6-8 sessions. One less table was needed per room, however, as only 30 panelists had been invited to participate for each grade/subject combination;

this fact permitted somewhat greater working space within the rooms and somewhat greater distance between tables. The set-up of the monitors and keyboards remained the same—an arrangement that posed a minor difficulty for some participants in terms of their ability to interact with other panelists seated across from them at their tables.

The breakout sessions again began with a welcome from the facilitators, individual panelist introductions, housekeeping information (i.e., non-disclosure forms, multimedia permission release forms (for audio, video, and photographic recording of participants), and a panelist information questionnaire. An introduction was then provided to the software to be used in the sessions. The software appeared to work well, although some intermittent connectivity issues were still noted. All participants again completed a training test and practice test. The groups then broke for lunch at approximately 12:30 p.m., with instruction to reconvene in the ballroom for a whole group activity at 1:20 p.m. The “parking lot” activity held previously at the end of the lunch hour was moved to more grade/content specific answers to questions presented by Dr. Wilhoft and Mr. Alpert who went to individual grade/subject rooms and addressed questions relevant to those grades and subjects.

Grades 6-8 Sessions, Day 1: Afternoon Activities (October 15, 2014)

The whole group activity to begin the afternoon consisted of the same presentation on the afternoon of the first day for the grade 11 panelists: Dr. Lewis presented an overview of the item map and the tasks that participants would be performing in the afternoon, with an emphasis on answering the two main questions about each item (i.e., what the item measures/sources of challenge, and what makes an item in the OIB more challenging than the preceding items. A change implemented was the inclusion of an additional aid and some explication of the specific bookmarking task in the presentation by Dr. Lewis. The slides projected in each room were also

modified to include “Bookmark Placement Instructions” that clearly stated the bookmarking task:

“Ask yourself: Would a student at the threshold have at least a 50% chance of earning this point? If YES, move on to the next item. If NO, place your bookmark here.”

Panelists were then dismissed to their breakout rooms and engaged in the cognitive analysis of the OIB for the remainder of the afternoon. Some grade/content groups finished the task by the end of the day; others were still completing the task when their groups were dismissed for the day at approximately 5:15 p.m.

As was done at the end of the first two days of the grade 11 sessions, a debriefing session was held with contractor staff, Smarter Balanced representatives, and the auditor in attendance. No major concerns were identified in the session. Some facilitators again observed that some members of their grade/content groups needed for additional time to complete their analysis of the OIB. To address that issue, the schedule for grades 6-8 Day 2 was again adjusted as it was for grade 11 sessions to add 45 minutes at the beginning of the next day for participants to complete that task. In addition, some facilitators noted intermittent connectivity issues, although the issue did not appear to impede any of the panelists’ work. Finally, Dr. Lewis indicated his desire to include an additional slide in his opening Day 2 presentation to clarify how the term “consistently” in ALDs should be operationalized when panelist are making judgments about items using RP50. Discussion of the idea was generally favorable; some suggestions were made as to how the clarification should be presented, and the addition of the new slide was approved by the group.

Grades 6-8 Sessions, Day 2: Morning Activities (October 16, 2014)

The second day of grade 6-8 standard setting consisted of a whole-group opening session beginning at 8:30 a.m. in the ballroom. The session again commenced with Dr. Bunch announcing

the adjustments in the agenda to allow participants to complete their cognitive analyses of the OIBs. The next portion of the session consisted of a presentation by Dr. Lewis who provided a more detailed description of the bookmark procedure and specific information and instructions on threshold students, bookmark placement for Level 3, Level 2, and Level 4. New slides were added; among them, one with the specific “Bookmark Placement Instructions” and the additional slide and explication of “consistently” that had been discussed in the debriefing session the previous afternoon with attention was given to issue of "consistently" vs. 50% chance of answering correctly. It was not clear that the issue was totally cleared up for all participants (although in observations of subsequent grade/subject discussions and activities, any lack of clarity did not appear to affect participants' ratings).

After dismissal from the whole group meeting, participants returned to their grade/content rooms to review/complete the cognitive analysis of the OIBs. The additional 45 minutes allocated to the task appeared to be helpful to the groups. After a break, the groups then began the practice activity in which they placed a Level 3 (only) bookmark using a six-item orientation-round OIB and other resources. Results from the practice were shown to the groups, and a review of how to place the bookmark, discussion of the content of the item that contributed to the placements, and clarifications were offered.

In the grade/subject rooms, the facilitators then provided directions on how to set the operational bookmarks for all levels, beginning with Level 3, then Level 4, then Level 2. Panelists were instructed to place their bookmarks independently at their tables and to complete the first questionnaire upon closing out of the practice round before beginning their Round 1 operational bookmark placements. The Round 1 bookmark placements continued through a break for lunch at approximately noon.

Grades 6-8 Sessions, Day 2: Afternoon Activities (October 16, 2014)

The afternoon session of Day 2 began with most grade/content groups finishing up Round 1 ratings or, for groups that had finished prior to lunch, beginning a review of the Round 1 results and engaging in table discussions about where individual panelists had placed their bookmarks. In each of the groups, the item maps showed a column called “Location” which showed the scale locations of each item in the OIB. The way this information was treated across the groups appeared to vary, with some facilitators providing minimal attention to that column, whereas at least some of the facilitators provided information to panelists on how to use the information. In at least one group (Grade 8 ELA), the facilitator slowly and explicitly helped the panelists in that room to see how moving a bookmark several pages in an area where the items were close in their scale locations would minimally affect the resulting cut score, whereas moving a bookmark even one or two items in a location where the items differed more substantially in their locations would have a more pronounced effect on the cut score for that level.

As part of the Round 1 review, table groups were instructed in how to determine their median benchmark score and were asked to compare their table median to the whole group median and discuss rationale for placing bookmarks. Discussion focused on Level 3 bookmark first, and then Levels 2 and 4. After a thorough discussion of bookmark placements, results from the online panel were presented and discussed. The online panel results were grouped according to participant self-designated category: Teachers, Administrators, Other (parents, etc.). Medians and interquartile ranges were shown for each major group and the overall group. Participants discussed the online participant results, but were directed to use them as another point of reference only, not something that should distract their focus from the ALDs for setting their bookmarks. Participants then completed a readiness questionnaire before beginning their Round 2 bookmark placements. Round 1 discussion times varied among groups, but all groups took a short break between 3:00 and

3:35 p.m. All groups finished their Round 2 bookmark placements before 5:00 p.m.

At the end of the day, a debriefing session took place attended by contractor staff, Smarter Balanced staff, and the auditor. No major issues were raised. Dr. Lewis indicated that the addition of the “Bookmark Placement Instructions” had been beneficial and had helped participants clarify and focus their efforts. He also provided guidance for Day 3 by asking facilitators to present all of the information on the slides in the morning before going into deep discussion on any one slide. There was also a brief discussion about strategies for presenting the impact data and focusing participants on determining appropriate cut scores based on the ALDs and not letting them get too focused on the impact data and “chasing numbers.”

Grades 6-8 Sessions, Day 3: Morning Activities (October 17, 2014)

The morning session of Day 3 for grades 6-8 began at 8:30 a.m. with facilitators convening their grade level groups. After going over logistics such as shuttle times to the airport and checkout time, the whole group began review of Round 2 data and impact data. All major technology issues seem to have been resolved as none were observed in any of the sessions. Normative and impact information provided were the same as the data provided to the grade 11 groups, with minor exceptions. The complete information provided to panelists consisted of the same six slides (described below) for their grade levels. For each group in the grade 6-8 band, a seventh slide which showed the impact data for Round 3 of the grade above. (The exception was grade 8, which saw the impact data from grade 11, Round 3.)

Slide 1: Histogram of individual bookmark placements. In many groups, there appeared to be good separation between levels and no “reversals” (e.g., instances where some group members placed their Level 3 bookmarks at locations that were lower than other panelists placed their Level 2 bookmarks). Discussions focused on the range of bookmark

placements within levels. In some groups there was overlap between levels, with the high end of one level overlapping with the low end of the next level. In these instances, the facilitator focused discussion on the overlapping points and having the panelists who placed their bookmarks in this area explain their rationale. In all rooms, the focus was brought back to the ALDs and what the ALDs said about what the threshold student should be able to do.

Slide 2: Table and overall group medians and ranges. In a couple of groups, there was a table that generally set their bookmarks lower than the rest of the group and one that set their bookmarks higher than the rest of the group. It seemed that in these cases, the groups had identified an outlier item that matched an ALD description, but was not a part of a group of similar level items.

Slide 3: Impact data for Round 2. Data used were student performance on Spring 2014 field test and showed the percentage of students who would achieve. Level 1, 2, 3 or 4 based on panelists' bookmark placements. In general, panelists did not seem surprised by the impact data. They raised questions about how this impact data reflect typical first-time standard setting versus later performance. They expected that performance would improve as students and teachers become more familiar with the standards and tests,

Slide 4: Histogram of impact data

Slide 5: Online panel (Teachers, Higher Ed, Administrator, Other, All) impact data levels for students achieving Levels 1&2 and Levels 3&4

Slide 6: Histogram of students at Level 3 disaggregated by various demographics (ALL, males, females, Hispanic, American Indian, Alaskan Native, Asian, African American, White/Caucasian, Hawaiian/Pacific Islander, Multi-ethnic, IEP, LEP/ELL, 504 Plan, Economically Disadvantaged)

Slide 7: Impact data for the grade above (e.g., Round 3 grade 11 for the grade 8 group, Round 2 grade 8 for the grade 7 group, and so on)--Levels 1&2 percentage and Levels 3&4 percentage

In most small groups, the information in all of the slides was presented and explained first. The full group discussion/analysis of the information took place after the presentation of the last slide. In some groups, the impact data raised concerns about potential uses of the test scores. For example, if too few students achieve Level 3 or 4 in some states, teachers may lose their jobs. Facilitators emphasized that the participants are making recommendations only and their recommendations are just one step in the standard setting process. Whenever such concerns were raised, facilitators brought the focus back to their task of placing bookmarks at appropriate points based on the ALDs. In one room, a participant made the distinction between thinking about where they want college and career ready students to be (what the ALDs indicate) and where their students are now (impact data). Other participants discussed motivational factors possibly suppressing student performance on the field test and recognition that in many field test states, the standards had not been fully implemented, so students may not have seen some of the material on the test. These discussions all helped participants contextualize the impact data showing fewer students at Level 3 and 4 than they would have liked to see.

After thorough discussion of Round 2 data and rationales for bookmark placement, participants completed a readiness survey and then commenced Round 3. (Participants were allowed to check out from the hotel during the interval between completing Round 3 and presentation of Round 3 results.) Facilitators stressed the confidential nature of the material presented and discussed during the meeting and reminded participants to leave their packets in the room. In one room, during the interval between submitting Round 3 bookmarks and receiving the final results, the facilitator led a discussion of “take-aways” from the session, what participants

might be asked by people in their home states, and how they might respond to difficult questions without divulging confidential information. Finally, each group was shown their group's Round 3 bookmark (median of the group) and impact data for the round then asked to complete a final questionnaire. By approximately 12:00, all sessions had adjourned.

Grades 3-5 Sessions, Day 1: Morning Activities (October 17, 2014)

Concurrently with the Grade 6-8 panels completing their Round 3 ratings, the standard setting session activities for Grades 3, 4, and 5 ELA and mathematics began. The activities for grades 3-5 followed the same agenda for table leader training on the prior evening (October 16, 2014) and for the daily activities as had been followed for grades 6-8 and 11. After an opening breakfast, grade 3-5 participants met as a full group for the orientation session in the ballroom. The whole group orientation session began with a welcome and overview presentation by Dr. Willhoft including an overview of the general purpose of the sessions, provision of some background and context, questions, solicitation of participants' concerns; the session continued with a presentation by Dr. Bunch, who described the agenda and goals for the workshop, along with brief attention to background on the common core state standards (CCSS) including claims and targets, the item and test development process, assessment item formats, the ALDs, the practice test, the bookmark procedure, the concept of an OIB, performance tasks, evaluations, "key shifts" embodied in the CCSS, and the four types of ALDs with an emphasis on the threshold descriptions. Dr. Bunch ended his presentation by noting the "ground rules" for the workshop and by providing information on the room assignments and facilitators for the mathematics and ELA groups; these included:

Grade 3 ELA - Craig Deville and Sheryl Grady (MI)

Grade 3 Math - Heather Farina and Rick Mercado (CTB)

Grade 4 ELA - Kelly Connelly (MI) and Sarah Hagge (CTB)

Grade 4 Math - Winnie Reid and Lisa Johnson (MI)

Grade 5 ELA - Joe McClintock and Ruth Hargis (MI)

Grade 5 Math - Judy Hickman and Jennifer Lord-Bessen (CTB)

In the breakout sessions, facilitators provided background information on the Smarter Balanced assessment framework, CAT testing and three types of ALDS: policy, range, and threshold. Next, participants were shown how to access reference materials on the computer and were asked to bring up the CCSS and DOK pages. Facilitators distributed copies of the threshold ALDs and discussed their purpose. Table groups spent 10-20 minutes discussing the ALDs and, after the small group discussions, table leads shared key points that arose in their small groups. Table groups approached the ALDs discussions differently: Some focused on understanding a target/standard as it is articulated across the levels (e.g. fractions at level 2, then 3 and 4); others honed in on one ALD level and worked to understand all aspects of that level. In G4, math participants commented that some standards (e.g. time) were addressed in ALD 2 and 4, but not ALD 3. Some participants commented that they would like to see examples of what is meant by some of the terminology (e.g. low-to-moderate text complexity); other participants asked about how the reading level of the texts was determined.

In general, between 10:00 and 10:30 a.m., all groups took a short break and then reconvened to engage in preparation for the short practice test. The practice test is publicly available on the Smarter Balanced web site as a way to familiarize potential test takers with the variety of items types that will appear on the operational assessments. The practice test again provided participants with an opportunity to view the variety of online supports available to students. After taking the practice test, participants discussed the items in small groups and then

participated in a large group discussion of their observations. During this subsequent discussion, participants commented that students would need to be trained on test taking strategies and become familiar with appropriate technology. Examples such as knowing how to scroll, ability to type, knowledge of the tools available to them (e.g. glossary) came up as things students would need to learn to do to perform well on the tests. After the group discussion and before adjourning for lunch at 12:30 p.m., participants took a brief readiness survey.

Grades 3-5 Sessions, Day 1: Afternoon Activities (October 17, 2014)

After lunch, participants reconvened in the large ballroom. The whole group activity in the afternoon began with Dr. Willhoft giving brief comments and informational updates. The session then included the same presentation by Dr. Lewis as was provided on the afternoon of the first day for the grades 6-8 and 11 panelists. That presentation, “Understanding What the Test Measures,” provided an overview of the item map and the tasks that participants would be performing in the afternoon, and emphasized the importance of answering the two main questions about each item (i.e., what the item measures/sources of challenge, and what makes an item in the OIB more challenging than the preceding items.) Dr. Lewis indicated that, over the next couple of days, participants would use this understanding to place Level 2, 3, and 4 bookmarks according to the ALD information; he also emphasized that participants should focus on the content of each item in the OIB with the goal of understanding what they know about what a student can do if the student gets the item correct. As part of his presentation, Dr. Lewis walked through two performance items to illustrate the various pieces of information provided about each item (e.g. DOK, target standard, rubric, student exemplars) and to emphasize that each performance item stimulus would appear for each score point. For example, an item worth two points would appear in the OIB twice: once for score point 1 and once for score point 2.

At approximately 2:00 p.m., panelists were dismissed to work in their grade/subject area groups where they began the process of examining the OIB. At the beginning of the small group session, participants asked a number of questions. Many were answered immediately, such as general housekeeping items and basic questions about Smarter Balanced, composition of ALD-development groups, and computer adaptive testing. More detailed questions were referred to Smarter Balanced staff (e.g. weighting of components of test). In some cases, participants brought up questions about the various uses of the tests in their states (e.g. teacher evaluations, promotion decisions). Facilitators focused participants back on their task of understanding the OIB well and placing appropriate bookmarks for the thresholds based on the ALDs.

During the afternoon session, a Smarter Balanced staff person (Nikki Elliott-Schuman, Smarter Balanced ELA content lead) circulated to all of the ELA small group sessions to address questions collected from participants in the morning session. (The Smarter Balance mathematics content lead, Dr. Shelbi Cole was also present and supported the mathematics panels.) Getting questions answered seemed to reassure participants who have come to this process from a wide variety of background experiences with the CCSS and Smarter Balanced assessments. In response to participant questions, Ms. Elliott-Schuman also shared information about enhancements that are being explored/developed such as a training test to help students understand and use the various accommodations available on the tests.

Facilitators then walked participants through the process of reviewing the OIB, had them log in, bring up the booklet and together review one or two questions. Facilitators picked example items to illustrate various aspects of items that participants would encounter and showed them how to enter comments to refer back to over the next few days as they work on placing bookmarks. It was strongly emphasized that the comments were for participants' eyes only; facilitators and other staff could not read them. In the ELA sessions, facilitators reviewed the presentation of a

performance assessment item, which appears six times in the OIB, twice for the conventions dimension (points 1, 2) and four times for the organization/use-of-evidence dimension (points 1, 2, 3, 4). For each score point, participants could access a rubric and student exemplars with annotations explaining why the response earned the specific score. Facilitators addressed all questions about the afternoon activity and made sure that participants were navigating smoothly within the OIB. Participants used the rest of the afternoon—until approximately 5:00 p.m.—to work through all of the items and make comments on them. Although the session officially ended at 5:00, participants in some rooms were allowed to review items until 5:30. In several other rooms, participants were not able to complete their initial review of the OIB. Some groups ended at 5:00 and none went beyond 5:30 p.m.

At 5:30 p.m., the end-of-day debriefing session took place, attended by contractors, Smarter Balanced staff, and the auditor. The time required to complete the initial review of the OIB was again raised as an issue; it was reported that only the grade 4 and grade 5 math groups had come close to finishing their initial review. In the previous sessions, it had worked well to allow time in the morning of Day 2 to continue review of the OIB; the same procedure was suggested for this group. It was decided that, after the morning large group session, each grade-level room would continue OIB review for 30-60 minutes before moving on to their next task.

Grades 3-5 Sessions, Day 2: Morning Activities (October 18, 2014)

The second day of the Grade 3-5 sessions began with panelists meeting in their grade/content groups to finish the cognitive analysis of their OIBs. The schedule adjustment adding this time appeared to be appropriate, as many panelists used the time to complete the work they began from the previous day. After completing the analysis, all groups experienced the training test and practice test activities, including discussions of why they placed their practice

Level 3 bookmarks as they did and other content focused discussions. Placement of Round 1 bookmarks began before lunch on Day 2, with some groups finishing the Round 1 bookmark placements before lunch and others taking a break for lunch and finishing up the Round 1 placements immediately after lunch. All groups who completed Round 1 bookmarking also completed a questionnaire and viewed information on Level 3 bookmark placements generated by the on-line panel for their grade and subject. In general, it appeared that connectivity issues were cleared up and few tech disruptions occurred.

Grades 3-5 Sessions, Day 2: Afternoon Activities (October 18, 2014)

The afternoon session of Day 2 began with most grade/content groups finishing up Round 1 ratings or, for groups that had finished prior to lunch, beginning a review of the Round 1 results and engaging in table discussions about where individual panelists had placed their bookmarks. In one room, there was somewhat of a lengthy delay in getting the on-line data to review. The facilitator sought help from the IT support staff and encouraged panelists to continue discussions of their Round 1 bookmark placements during the interval. In addition to the on-line information, the group median and range of bookmark placements were also provided in tabular and graphic formats. The rest of afternoon was spent setting Round 2 bookmarks, with panelists again instructed to set their bookmarks for all three levels in the appropriate sequence, submit their judgments, then fill out the end-or-round questionnaire when they were finished. Many panelists appeared to complete the Round 2 bookmark tasks early, with all finishing the tasks by 4:00 p.m. They were instructed that the activities for Day 3 would begin at 8:30 a.m., and asked to report to the individual grade/subject breakout rooms for the start of Day 3.

As had been done previously, Day 2 for grades 3-5 ended with a debriefing session involving contractor staff, Smarter Balanced representatives, and the auditor. Because all

participants had finished their Round 2 ratings somewhat early, the debriefing for grade 3-5 was convened a bit earlier in the afternoon on Day 2 at approximately 4:15 p.m. The group was informed that one participant had a family emergency and needed to leave early to return home; some attention was also given to logistics, such as moving the lunch break and shuttle service for Day 3 a bit earlier. Dr. Bunch queried the group for any common concerns; no major issues requiring attention or modification were identified. Because it was anticipated that the groups would have ample working time on Day 3, Dr. Bunch directed the facilitators to take time reviewing bookmark placements, to encourage thoughtful group review and discussion of any overlaps or “reversals” in bookmark placements (e.g., situations in which a group member's Level 3 bookmark was placed at an earlier location than another member's Level 2 bookmark).

Grades 3-5 Sessions, Day 3: Morning Activities (October 19, 2014)

The activities of grades 3-5 on Day 3 consisted of a review of Round 2 results, a questionnaire, placement of Round 3 bookmarks, presentation of impact data and a final questionnaire. Two groups were observed in depth on this morning, and the observations provided the basis for our conclusion that facilitators in different rooms had differing degrees of experience and comfort leading standard setting workshops. In one of the groups observed, the facilitation was masterful. The facilitators explained all information clearly; they responded accurately and appropriately to all participants' questions; they anticipated concerns; and made highly appropriate and useful time out of intervals when they were awaiting data analysis results. When participants had already completed their Round 3 judgments and were waiting for their data to be analyzed and returned, the analysis took a fairly long time—perhaps as much as 30 minutes—however, the facilitators made very effective use of the wait time, thanking participants for their sacrifice and effort, for giving time over the weekend, and reminded participants to share

back home about the ALS process, but no to provide information about people, items, or results. The facilitators also provided information on logistics (shuttles, lunch, etc.). When the data analysis was complete, results for final review and evaluation were projected, including the room's median Round 3 Level 2, Level 3, and Level 4 bookmarks, and impact data for all four levels (table and bar chart). There was some discussion of the results, but participants generally seemed to feel that the results were appropriate. After completing the final questionnaire, participants returned their paper materials, and were dismissed for lunch and return trips home.

In another of the groups observed on this morning, the facilitation was less skilled and seemed more labored. The session began with presentation of a histogram showing Round 2 bookmark placements for the group, along with the group's interquartile range and median. A second slide was shown that provided the Round 2 median cut scores for each achievement level for the five table groups and for the total grade/content area group. The lead facilitator in the room seemed to have difficulty understanding and explaining the median values shown on the slide, asking out loud if higher medians or lower medians reflected higher or lower expectations. The third slide shown provided impact data (percent of students that would be categorized at each achievement level) for the total room based on field test performance. The lead facilitator read the interpretation of the results shown on the slide; the second facilitator elaborated on the interpretation and aided the group in understanding how they should use the impact data. A panelist asked how the impact percentages would change if the cut score for a level were increased by one raw score point correct. The facilitator indicated that it was not possible to tell based on this data. Discussion was then prompted regarding reactions to the impact data, with participants in general reacting to a comparatively larger percentage of students being classified as Level 1 and a comparatively smaller percentage of students being classified as Level 2. The facilitator then reposted the histogram (slide 1) and related the spread of bookmark locations to those

percentages--a comparison that was essentially irrelevant and potentially misleading, conflating two results (i.e., range of panelists' judgments and categorical impacts). Again, the second facilitator in the room attempted to provide some additional information that may have helped panelists understand the results. One panelist contributed an interpretation that was accurate and provided a clear way for panelists to think about how they would move their bookmarks to affect the percentage of students in a category. However, the panelist also suggested that the task of the group was to obtain a "standard normal distribution" of results and that the current impact did not reflect that—a suggestion that the facilitators did not correct.

The next slide shown provided the medians and interquartile range from the on-line panel for the Level 3 cut, showing the impact (percentage of students that would be classified as Level 3 or above). Another slide containing impact for subgroups (percent that would be classified as Level 3 or above) was also shown. Finally, the group was shown the impact (percent in Levels 1 and 2, and percent in Levels 3 and 4) based on the upper grade (Grade 6) round 3 final recommendations. The facilitator asked the group to "gasp in unison" at the discrepancy between the results for the two grades (i.e., grade 5 and 6). Impact data based on the lower grade was not provided. The facilitator then explained that the next step in the process was that a vertical articulation panel would review and potentially smooth out differences in impact across grades and the panelists were encouraged to focus on content. A panelist reinforced that the task of the group was to "focus on the ALDs and how those are translated into their expectations instead of trying to match another grade's results."

Panelists were then instructed to open their OIBs and review their bookmark placements as shown on the first slide (histogram) and to share their rationales for why they placed their bookmarks at those locations, particularly for the Level 2 cut. As the discussion started to drift, a panelist reminded the group that the focus should not be on what they think the percentages should

be, so much as what the ALDs indicate. The panelists then discussed the other levels in like manner. For the level 4 discussion, a panelist defended her placement of a bookmark distinguishing level 3 and level 4 by saying that she thought the upper level 3 students would be able to get the item correct, but was asked by the facilitator, "would at least 50 percent of them get it correct?" Although some of the strategies and understandings revealed in the discussion seemed incorrect, overall the group appeared to have a good grasp of how to place their bookmarks. The group began a break at 10:35 a.m. then returned to complete a questionnaire, setting of Round 3 bookmarks, reviewing Round 3 results; they were then thanked for their participation and dismissed.

IV. Vertical Articulation

A vertical articulation session (called “Cross Grade Review”) was held on Monday, October 20, 2014. The session began at 8:30 a.m. Dr. Bunch welcomed the group and introduced Deb Sigman and Joseph Martineau (Smarter Balanced Executive Committee Co-Chairs) who both also welcomed the group, thanked them for their work, and gave brief remarks on the purpose of the session—ensuring a coherent system across the grades. Dr. Bunch also introduced the group facilitators for the ELA and mathematics panels (himself and Dr. Lewis) and he introduced Dr. Willhoft who also gave welcoming remarks and provided an overview of the next steps in the ALS process. The articulation panelists were distributed as follows:

ELA Cross Grade Panel (32 total panelists) Mathematics Cross Grade Panel (32 total panelists)

Grade 11 - 16 panelists

Grade 11 - 16 panelists

Grade 8 - 2 panelists

Grade 8 - 2 panelists

Grade 7 - 3 panelists

Grade 7 - 3 panelists

Grade 6 - 2 panelists

Grade 6 - 3 panelists

Grade 5 - 3 panelists

Grade 5 - 3 panelists

Grade 4 - 3 panelists

Grade 4 - 3 panelists

Grade 3 - 3 panelists

Grade 3 - 2 panelists

Dr. Bunch then presented a series of slides and information that provided participants with an historical overview of each phase of the process performed to date and those to come, beginning with the on-line panel, the in-person workshop, a TAC review, and culminating with review and approval by the chief state school offices on November 6, 2014, and production of a technical

report. The next slide showed the progression of review from individual, to table, to grade/content, to cross-grade. The third slide provided answers to the question “Why are you here?” and he described the efforts made to ensure broad and qualified representation. The fourth slide outlined the participants’ tasks: 1) examining cross-grade bookmarks; 2) establishing cross-grade coherence and reasonableness; and 3) reviewing impact data and scaled scores. The next slides provided elaboration on the definition, requirements, and typical patterns of cross-grade articulation, along with an example of an “unexpected” pattern of cross-grade standards.

The next slide introduced the rationale for articulating impact data and vertically-scaled scores across grades, as opposed to bookmarks. Graphs showing the mean scaled scores on the vertical scale across grades that resulted from the field test data were then shown for ELA and mathematics. The presentation then turned to “The Tools We Will Use” which explained the presence of variability within grade/content panels and introduced box and whisker plots to represent that variability. Tables showing hypothetical data of grade level scaled scores, impact data, and the relationship to bookmark placements were presented to familiarize participants with the data they would be using for their cross-grade articulation task. Dr. Bunch reviewed materials that panelists would have for reference, including all of the grade/content OIBs and their grade/content ALDs.

The specific steps that participants should follow were then presented. The steps included review of current bookmark placements, consideration of moving bookmarks at grade levels, and voting procedures for group acceptance of a proposed bookmark change. Procedures included the requirement for making and recording of formal motions to suggest a change, a second for the motion, and a 2/3 majority required for approval of a suggested change with a recorded tally of “yeas”, “nays” and abstentions. It was not clear if a vote would be required if a panel chose to ratify the original bookmark location. (This auditor mentioned this issue to Mr. Alpert, who

indicated that it would be addressed in the breakout groups.) Dr. Willhoft reviewed the requirements for who would be eligible to vote. A slide was then shown that illustrated the voting log that groups would be asked to complete. There was some lack of clarity regarding who was eligible to make a motion; Dr. Bunch indicated participants he would “prefer” for motions to be made by someone from the grade level under consideration, although he also indicated that a motion could be made by anyone eligible to do so from the total group. It was also not clear if a motion was required to keep a bookmark location as originally recommended by a grade/content panel.

A short time of questions and answers followed at the end of Dr. Bunch’s presentation. Dr. Willhoft answered largely policy and procedural questions from participants related to how the chief state school officers would make their decisions on final cut score adoptions and the information that they would be provided to aid in that decision. A brief logistics update was provided, covering time frames for the agenda, breaks, and assignments for the breakout rooms. The whole group was then dismissed at approximately 9:45 a.m. to begin work in their breakout rooms, one for ELA and one for mathematics.

In the breakout groups, panelists introduced themselves and gave brief descriptions of their roles. The ELA group was facilitated by Dr. Bunch and Dr. Craig Deville from MI; the mathematics group was facilitated by Dr. Lewis and Dr. Rick Mercado of CTB/McGraw-Hill. In general, it appeared that the optimal way in which facilitators led their groups was for one of the facilitators to lead the process, while the other facilitator processed input into the computer and ensured that the relevant materials was displayed.

Participants in each group were then shown the cross grade distributions of impact data; the graphs were shown with points along the grades indicated by a point showing the panel median with accompanying measure of variability (box plots with interquartile ranges and whiskers

extending from P10 to P90). The data and graphics were presented in a dynamic fashion, so the effect of any suggested changes could immediately be seen in terms of impact and scaled scores.

Discussion of the results followed. In both content areas, the groups began their work by examining the articulation between the upper-most grades (i.e., grade 11 and grade 8), then continued their review working downward across the grades. The groups appeared to understand the nature of the data presented to them and the task they were empaneled to perform. However it also appeared that panelists believed that they needed to smooth out all bumps and dips in the impact data. Groups began the work of considering changes and broke for lunch at approximately noon, returning to complete their work at approximately 1:00 p.m. The groups began by working on articulation of Level 3 across the grade levels. After completing that work, they proceeded to consider articulation of Level 4, followed by Level 2.

Overall, the facilitation in the breakout rooms seemed to be highly effective. Facilitators answered panelists' questions clearly and kept the process on track. In one of the rooms, at the beginning of the breakout session, at least some panelists commented that they felt somewhat rushed in their work. To a small degree, in both rooms there was some disagreement among panelists about the appropriate referent for their decisions—the ALDs, the state of implementation of the common core in their states, and where student performance is currently and the impact on students of the decisions. However, for the most part, panelists considered adjusting impact, grounding their judgments in the content of the OIBs and the ALDs. The discussions were largely content focused; all participants appeared to be comfortable contributing to them; and no personal agendas were evident. For the most part, the content-based adjustments made by the panels fell within the P10 to P90 range from the original panels. In rare instances, the panels exceeded those values. Whenever a change was voted upon, the groups generally worked to endorse a change that was broadly acceptable; it was typical for votes to be unanimous within a group, or nearly so. All groups

finished their work by 5:00 p.m. An evaluation questionnaire was not administered to the vertical articulation panelists. They were thanked for their contributions and dismissed.

V. Summary and Recommendations

Based on observations of the procedures and processes used to obtain recommended performance standards, it is my opinion that the standard setting activities implemented for the Smarter Balanced summative assessment standard setting were, overall, conducted in a manner consistent with sound psychometric practices.

Few issues arose during the in-person standard setting sessions or the vertical articulation session; issues that arose were comparatively minor and, in my opinion, would not have substantially affected the validity or reproducibility of the results. Although there are several aspects of the process that should be considered for revision and incorporation for future standard setting procedures, none of them could be considered to be a fatal flaw in the process. In the following subsections, some specific strengths and recommendations are presented.

Strengths

There were a number of strengths observed during the standard setting activities for the Smarter Balanced assessments. A thoughtful standard setting plan was developed and reviewed by the Smarter Balanced Technical Advisory Committee and formally approved by Smarter Balanced governing states. Overall, there was a well-organized and faithful implementation of that plan by the contractors responsible for conducting the standard setting workshops. The contractors provided adequate resources and personnel to ensure that the standard setting was conducted professionally and paced appropriately.

For the in-person sessions, it appeared that all panelists had strong qualifications for participation; they appeared to be highly qualified and they seriously, conscientiously, and

enthusiastically engaged in their tasks; they generally understood how to complete their tasks and they engaged in relevant, content-focused discussions. All participants appeared to work diligently during the sessions; no issues regarding domination of discussion in groups/tables were apparent, and no participants appeared to exert personal agendas.

With the exception of some of the feedback regarding external data sources (e.g., PISA, NAEP, ACT), participants appeared to understand the nature of the feedback provided to them (i.e., normative and impact information). In discussions they appeared to reference and consider the information appropriately. Overall, participants identified as table leaders appeared to function well in both the initial standard setting sessions and the subsequent vertical articulation session.

For the vertical articulation process, the panelists were similarly representative, engaged, deliberate, and thoughtful in their work. The orientation session appeared to be carefully planned and well delivered; the breakout sessions were also facilitated exceptionally well. At the beginning of the breakout sessions, some panelists commented that they felt somewhat rushed in their work. For the most part, all the discussions were content focused; all participants appeared to be comfortable contributing to them; and no personal agendas were evident. There was some disagreement among panelists about the appropriate referent for their decisions—the ALDs, the state of implementation of the common core in their states, and where student performance is currently and the impact on students of the decisions.

In addition, the following specific strengths were observed:

- 1) All procedures were pilot tested and adjusted in response to experience from the pilot;
- 2) Improvements in the operational standard setting based on lessons learned from the pilot testing were incorporated (e.g. hard copies of ALDs provided, doing the practice test, more

focus on making participants aware of item aspects, and so on);

3) The process was marked by substantial transparency (e.g., parking lot answers to any questions raised by the groups, whether directly related to standard setting or not);

4) The technology generally worked well, including computers, headphones, projectors, and software used for the sessions. Comparatively minor technology issues were quickly and effectively addressed;

5) The groups were divided into tables to facilitate more ready and deeper discussions;

6) A cognitive analysis of the full OIBs was included; it allowed panelists to become deeply familiar with the test content on which they would be making their judgments;

7) A practice activity was included for panelists to gain greater familiarity and experience with the bookmarking procedure;

8) Adequate time was allocated to allow the procedures to be completed at an appropriate pace. The schedule adjustment to allow additional time for all groups on Days 2 and 3 was a helpful adjustment, allowing participants to complete their cognitive analysis of items in the OIB. Had this adjustment not been made, many participants may not have become familiar with the items at the end of the OIB--which would represent the most challenging content and, likely, content tapping some targets/standards not represented previously in the OIB;

9) Nightly debriefing sessions occurred to gather feedback from facilitators on the day's activities and to plan for appropriate adjustments;

10) Appropriate evaluations ("Reflection Questionnaires") were administered at relevant

junctures in each in-person workshop;

11) The TAC-reviewed and state-approved plan was essentially followed faithfully; only minor deviations from the plan as written were observed. Some of the deviations were unintended (e.g., one group did the practice test out of order) and some were intended (e.g., inclusion of new slides on bookmark task); none of the deviations appeared to be consequential in terms of affecting rigor/integrity of process, results;

12) The orientation and training covered some basic test development information and the role of the ALDs was emphasized throughout the process;

13) All procedures went well on Day 1 of the first session; all procedures improved over the course of the three workshops as minor improvements made things run more smoothly.

14) The logistics were handled exceptionally well. The hotel accommodations, staff, food service, and other aspects of the venue were without any issues. The contractor staff responsible for logistics also helped the meeting proceed without a hitch: panelists always knew where to go, when, had relevant materials, and were assisted with all arrangements, such as transportation and special needs (e.g., a larger monitor was quickly obtained for panelist with visual impairment);

15) There appeared to be capable and responsive technical and practical support that provided accurate and timely data analysis and feedback for panelists;

16) The materials, forms, and documents appeared to be well-designed and easy for participants to use; and

17) There was appropriate concern for and attention to confidentiality and security of

materials and results, including the collection of all paper materials at the conclusion of the panels' deliberations.

Limitations/Considerations for the Future

A few areas may warrant attention either as information for policy makers, or as information that Smarter Balanced or contractors may wish to consider when planning or implementing future standard setting activities. Some observations and recommendations include:

- 1) The computer interface seemed somewhat challenging for participants to use, often requiring clicking back and forth among screens or documents or several click sequences to get to desired materials. In the current standard setting, participants appeared to adapt to the interface, but greater intuitive look and ease of use would seem desirable.
- 2) The facilities for the present standard setting were marked by small rooms, rectangular tables, and equipment on the table (monitors) that made it difficult for groups to work independently and difficult for participants at a table to converse among themselves. Again, participants adapted in various ways (e.g., by relocating their monitors, moving their chairs into circles away from their tables, etc.), but larger rooms, lower footprint monitors (laptops?), or other feasible configuration changes would enhance independence among groups and discussions within groups.
- 3) There appeared to be recurring, intermittent connectivity issues. Although the participants' work was not substantially impeded and technical support staff always quickly and efficiently addressed the issues, it would seem that more stable/reliable connectivity should be investigated for future studies.

4) The auditors were not able to see all activities for all rooms. It is recommended that at least two auditors be assigned the task of performing observations when future standard setting workshops involve multiple groups working concurrently.

5) There appeared to be fairly wide variability in how ancillary information (e.g., NAEP, PISA, ACT, on-line data) was presented and interpreted in the groups. Information was presented and described at different times, in different ways by different facilitators. Some facilitators indicated that participants should "feel free to use, incorporate, review, ignore [the information] as you see fit;" others facilitated thoughtful discussions on the data. For the future, it would be desirable for workshop planners to carefully consider precisely *how* participants should consider and incorporate any such data sources and more uniform scripting and implementation of this aspect of the process would be desirable.

6) There appeared to be substantial variability in facilitation of the breakout groups: some facilitators appeared to be highly experienced, thoroughly understood the process and concepts, took charge, elaborated on script to help participants fully understand/perform their tasks, led very effective discussions and so on; others seemed to be much less experienced, lacked a level of comfort in conducting the sessions, were less familiar with the methodology, materials, agenda, software, output, stayed very close to scripts or read scripts verbatim, and lacked skill in facilitating deep discussions when results were presented. This concern was exacerbated when rooms were staffed by two less experienced/comfortable facilitators. For the future, it is recommended that attention be paid to ensuring at least one facilitator in each room should be highly skilled and comfortable with the activities; in no case should two less-seasoned facilitators be assigned to the same group.

7) There was some variability in how table leaders functioned. Some table leaders facilitated more effective discussions than others; some groups functioned more or less democratically, and so on. For the future, it may be desirable to provide greater training to table leaders and for facilitators to have a debriefing session each day with table leaders to reinforce table leader roles, identify successful practices and to recommend strategies for addressing concerns raised by the table leaders.

8) In the cross-grade articulation process, it may have been desirable to also explain to panelists that mild to moderate fluctuations across the grades could be acceptable. In addition, some panelists commented that the data presented on some of the slides was too small to read. In the future, it would seem desirable to increase the font size for the presentations, or, if feasible, another strategy might be to “push” non-dynamic versions of the screens to participants’ monitors such as is done in a webinar format.

9) In the cross-grade articulation panel, it would seem desirable to administer a final evaluation questionnaire. Such a data collection would have the potential to serve as additional validity evidence for the process.

Conclusions

Overall, I can provide a positive evaluation of the standard setting activities for the Smarter Balanced summative assessments. On the one hand, there were a few aspects of the process that represented concerns that weakened the confidence that could be gained from the process. On the other hand, there were a number of strengths apparent in the standard setting procedures. Importantly, the plan for setting the performance standards was developed in conjunction with the advice of the Smarter Balanced Technical Advisory Committee. The procedures were

implemented with good fidelity to that plan. Further, it is my opinion that the procedures and processes used to derive recommended achievement levels followed sound, best practices of the psychometric profession. All participants--educators, policy makers, public representatives, and contractor staff--appeared to take a serious and conscientious approach to the tasks. The inclusion of a vertical articulation activity provided an important “check” on the coherence and reasonableness of the panelists’ recommendations.

It is my conclusion that the standard setting activities described in this report were designed and conducted appropriately so as to yield defensible performance standards grounded in the knowledge, skills, and expectations represented by the ALDs. Only minor issues arose during the standard setting and vertical articulation processes; in my judgment, none of them present a major barrier to the integrity of the results. Because the procedures and processes used to derive recommended performance standards appeared to follow sound psychometric practices. Unless analyses of the panelists’ evaluations were to indicate otherwise (I did not review these prior to submitting this report), I conclude that the panelists' cut score recommendations should be considered to be valid and reliable estimates of appropriate, content-referenced cut scores to define the four performance categories on the Smarter Balanced summative assessment. I believe that policy makers can have confidence that the recommendations from the standard setting panelists are based on sound procedures, producing trustworthy and defensible results.

VI. References

- Cizek, G. J. (Ed.) (2001). *Setting performance standards: Concepts, methods, and perspectives*. Mahwah, NJ: Lawrence Erlbaum.
- Cizek, G. J. (Ed.) (2012). *Setting performance standards: Foundations, methods, and innovations*. New York: Taylor and Francis.
- Cizek, G. J., & Bunch, M. (2007). *Standard setting: A practitioner's guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: SAGE.
- Cizek, G. J., & Bunch, M. B., & Koons, H. H. (2004). Setting performance standards: Contemporary methods. *Educational Measurement: Issues and Practice*, 23(4), 31-50.
- Lewis, D. M., Mitzel, H. C., Mercado, R., Schulz, E. M. (2012) The Bookmark standard setting procedure. In G. J. Cizek (Ed.) *Setting performance standards: Foundations, methods, and innovations* (pp. 225-254). New York: Taylor and Francis.

Appendix D – Recruitment Documentation

Online Panel

- D01_OP K-12 Leads Online Panel Recruitment 543
- D02_OP K-12 Leads Recruitment Information 557
- D03_OP TIC Information 559
- D04_OP HE Lead Information 561
- D05_OP Educator FAQ 563
- D06_OP He Leads E-mail 567
- D07_OP K-12 Educator E-mail 568
- D08_OP HE recruitment Reminder 569
- D09_OP K-12 Recruitment Reminder 570
- D10_OP Outreach E-mail Newsletter Higher Education 571
- D11_OP Outreach E-mail Newsletter K-12 573
- D12_OP Parent Outreach E-mail 575

In-Person Workshop/Cross-Grade Review Panel

- D13_IPP HE Leads E-mail 577
- D14_IPP K-12 TIC E-mail 579
- D15_IPP HE Leads Information 581
- D16_IPP K-12 Information Packet 586
- D17_IPP HE FAQ 594
- D18_IPP K12 Educator FAQ 597
- D19_IPP HE Educator E-mail 601
- D20_IPP K-12 Educator E-mail 603
- D21_IPP HE Nomination List Template 605
- D22_IPP Parent-Community E-mail 606
- D23_IPP K-12 Participant Application 608
- D24_IPP State Nomination List Template 609

Achievement Level Setting

Recruitment for Online Panel
Joe Willhoft, Executive Director



K-12 Leads and TICs
April 22, 2014

Achievement Level Setting Dates

- Online Panel: October 6-17
- In-Person Panel: October 13-19
 - October 13-15 (Grade 11)
 - October 15-17 (Grades 6-8)
 - October 17-19 (Grades 3-5)
- Vertical Articulation Committee: October 20-21

Online Panel

'Crowd-Sourced' Approach to Achievement Level
Score Recommendations



Success Factors for the Online Panel

- Collect input from a large group of educators
- Collect input from a diverse group of other participants to recommend achievement level scores
- Raise awareness about process for developing achievement levels
- Engage communities and stakeholder groups from member states
- Clearly explain process to those stakeholder groups and outside entities (i.e., media, researchers, policy makers)
- Maintain scientific validity by providing data to inform the Vertical Articulation Committee

Online Panel Overview

- October 6-17 window
- Open to the public – parents, teachers, college faculty, administrators, business and community leaders
- No qualifications required – any interested person is welcome to participate
- Two-stage recruitment process:
 - Through states during school year
 - Through professional organizations, media over the summer

Online Panel Registration

- Participants will
 - receive an informational email from a State Lead or TIC
 - log onto smarterbalanced.measinc.com
 - choose a content area and grade level
 - provide email and demographic information
 - select a two-day window in which to complete the activity
 - load the window selected into their calendar

Notification Process

Process	Responsible Party	Date(s) / Timing
Registration	Panelist	4/22/2014 – 9/19/2014
Panelist Verification Email Sent	Smarter Balanced	Immediately after signing up with valid email address
Panelist Verifies Email and Selects Window	Panelist	After receiving verification email
Panelist Confirmation Email Sent	Smarter Balanced	Immediately after selecting a window
Panelist Receives Email Reminder	Smarter Balanced	Shortly before selected window opens

Online Panel Activity

- In October, participants will
 - receive email reminders of the event
 - log onto smarterbalanced.measinc.com during the window they selected
 - take the test for the grade and content area they selected
 - review the Common Core State Standards and Achievement Level Descriptors
 - view a test booklet and place a bookmark to recommend a score for Achievement Level 3

Online Panel FAQ

- **Cost:** Open to the public. Participation is free.
- **Privacy:** Personally identifiable information of individual participants will not be released to the general public.
- **Accessibility:** Accessible through virtually any computing device that connects to the Internet, including tablets (Smart Phones not recommended due to screen size).
- **Reminder:** Panelists will receive reminders before their two-day window opens.
- **Time Commitment:** Up to 3 hours over a two-day window.
- **Log in/Log Out:** Ability to save work and log out, and then log back in to complete session within the two-day window.

Recruitment Timeline

Activity	Responsible Party	Date(s)
Contact Participants	TICs, K-12 Leads, Higher Ed Leads	April 22–September 19, 2014
Register for Online Panel	Participants	April 22–September 19, 2014
Complete Online Panel Achievement Level Setting	Participants	October 6–17, 2014 (two-day window selected during registration)

Online Panel Recruitment

- Two-stage recruitment process, overlapping with recruitment for the In-Person Panel
- April-May:
 - Communication to recruit educators
 - Outreach through schools and initial outreach to partners, direct and ‘owned’ media (TICs, K-12 and higher ed leads, the Smarter Balanced newsletter)
 - Message testing/audience research with parent focus group
 - Feedback to states on number of registrants by role
 - Additional communication tools to states if needed
- May-September:
 - Public outreach through partner organizations, associations, and earned and social media

Communication Package for Online Panel

Document Name	Description
Online Panel TIC Recruitment Information	Describes the Achievement Level setting activity and registration process
Online Panel FAQ	Provides frequently asked questions and responses about the panel to inform participants and TICs
Online Panel Outreach Emails	Provide sample recruitment emails for teachers and other educators
Optional Recruitment Messages	Engage parents, community leaders, and statewide associations



Higher Ed Leads will receive similar materials.

State Updates

- Regular Online Panel registration updates
 - Numbers of registrants
 - Breakdowns by role
- Additional recruitment or communication tools provided as needed

Questions?

State Updates

- Regular Online Panel registration updates
 - Numbers of registrants
 - Breakdowns by role
- Additional recruitment or communication tools provided as needed

Questions?

Recruitment Communication Overview for K-12 Leads Meeting

Standard-setting is the process for setting common expectations for student readiness/proficiency on the Smarter Balanced assessment. **Teachers in the K-12 community are key to the success of this process.** This is a critical opportunity to collect input from a large and diverse population of participants and establish measures of progress. The standard-setting process will rely on the professional judgment and classroom experience of educators in order to develop performance standards that are **rigorous, fair and accurate.**

A note on nomenclature: While the term “standard setting” is well understood within the assessment community, for most people it creates confusion since standards were “set” when the Common Core was adopted. We are in the process of testing out alternate terms and will have those ready to roll out when recruitment of participants begins in April.

Distributed standard setting and in-person standard-setting

- Our goal is to create a process that is as inclusive as possible, representative of views across the Consortium, transparent and scientifically rigorous.
- The ***“distributed standard-setting”*** panel will be open to the public and allow for mass participation by parents, teachers, college faculty and administrators, and business/community leaders. The distributed standard setting will enable hundreds of thousands of interested educators, parents and other stakeholders across the country to participate in the online process, which will take approximately three hours to complete.
- The ***“in-person standard-setting”*** workshop will allow for a group of more than 500 participants, nominated by states, to engage in a comprehensive review of questions on the assessment. The workshop, which will occur over several days this fall at a location to be determined, will give participants the opportunity to review the difficulty of questions and define what students performing at each of four achievement levels can reasonably be expected to answer correctly at each grade level. For grades 3 to 8, the panels will be comprised primarily of content specialists and educators. The high school panels will include higher education faculty in addition to K–12 educators.
- Finally, a ***Vertical Articulation Committee*** composed of 60 panelists from the in-person standard setting will review data and results from both the in-person and distributed processes and **make final recommendations to state leadership.** This step will help ensure appropriate vertical articulation across the grade span of the assessment system.

Role of member states/teacher involvement coordinators (TIC’s)

Working with Teacher Involvement Coordinators (TICs), member states will have the opportunity to nominate qualified candidates for the in-person standard setting. In April and May, we will work with the TICs, who will conduct outreach through the schools to solicit nominations for in-person standard

Recruitment Communication Overview for K-12 Leads Meeting

setting and enrollment in distributed standard setting. Higher education participants will be solicited through the state higher education leads.

Recruitment process

- Recruitment will begin for the in-person standard-setting workshop on April 15 and on April 22 for the distributed standard setting. A separate packet of material for that recruitment effort also will be distributed to K–12 leads and TIC’s. The contents of that packet are still under development. See the table below for dates and deadlines. **Nominations are due for the in-person workshop on May 2.**
- For the distributed standard-setting, we will launch an awareness campaign through intensive outreach coordinated through state education agencies before the end of the school year. In addition to K-12 participation, we also will be coordinating outreach to the higher education community through the higher education leads. Registration for the distributed standard-setting event will begin in May and continue through August to allow for maximum participation.
- To fill any spots for the distributed standard-setting event not reserved during the school year, we will conduct outreach to teachers, parents, higher education faculty and the business community through professional organizations and select media channels (e.g. trade publications) over the summer.

Recruitment Timeline

Activity	In-Person Standard Setting	Distributed Standard Setting
Information Package(s) to TIC, State Leads, and Higher Ed Leads	April 15 Materials provided to TIC’s will include: <ul style="list-style-type: none"> • Recruitment e-mail template • Standard-setting educator recruitment letter, • Standard-setting FAQ • Standard-setting TIC information page 	April 22 Recruitment will focus on outreach by TICs/higher education leads until the end of the academic year, then continue through professional organizations and select media during the summer.
Nominations due/ Participant Registration deadline	May 2	Late August (Date TBD)

Attachment 1: TIC Information Sheet for the Online Panel

This fall, K–12 educators, postsecondary educators, parents, and business/community leaders from Smarter Balanced governing states will collaborate to develop common achievement levels that are rigorous, fair, and accurate for each of the Smarter Balanced assessments. Smarter Balanced is asking for the assistance of Teacher Involvement Coordinators (TICs) with the recruitment of K–12 educators, parents, and business/community leaders to participate in this process. Smarter Balanced is asking Teacher Involvement Coordinators from each of the Smarter Balanced member states to recruit educators for the Online Panel.

Developing Achievement Levels will occur in a four-stage process:

The *“Online Panel”* will be open to the public and allow for broad participation by parents, teachers, school and district administrators, college faculty and administrators, and business/community leaders. The Online Panel will enable hundreds of thousands of interested constituents and members of the general public across the country to participate in the process. An FAQ and other information will be available online for those who register to participate in the Online Panel.

The *“In-Person Panel”* will allow for a group of more than 500 participants, nominated by Smarter Balanced governing states, to engage in a comprehensive review of questions on the assessment and make recommendations for operationalizing common expectations. For grades 3–8, the panel will be composed primarily of content specialists and educators. The high school panels will include higher education faculty in addition to K–12 educators.

A *“Vertical Articulation Committee”* composed of 60 panelists selected from those participating in the in-person workshop will review data and results from both panels and make Achievement Level recommendations. This step will help ensure a balanced view of achievement and proficiency levels for test-takers from Grade 3 through high school. No additional recruitment is required for this activity.

Finally, the recommendations from the Online Panel, the In-Person Panel, and the Vertical Articulation Committee will be presented to the chief school officers in Smarter Balanced governing states for their consideration and endorsement, in order to establish a common set of performance standards for mathematics and English language arts/Literacy across grades 3–8 and high school.

Purpose of the Online Panel

This is a critical opportunity to collect input from a diverse group of participants and establish consistent measures of progress for the Smarter Balanced interim and summative assessments. By participating in the online session, panelists will recommend an achievement level score that demonstrates how much students should know or be able to do in order to be proficient at the grade-level standards and to be on track for eventual college and career readiness. It is important to have as many participants as possible, from as many backgrounds as possible, provide these recommendations.

Participants in this panel will not set the final scores. The results of the panel will help inform the overall recommendations to the Smarter Balanced Assessment Consortium. Leaders from Smarter Balanced member states will carefully consider these recommendations in arriving at their decisions. The achievement levels endorsed by member states also will be subject to existing approval processes within individual states.

Logistics for the Online Panel

The panels will occur online from October 6 to 17, 2014. Panelists will select a 2-day window for participation during the registration process. It will take panelists up to 3 hours to complete the orientation process, review test questions, and recommend a score for Achievement Level 3.

Registration for the Online Panel

For more information and to register, interested participants will visit [this link](#) which will direct them to an application page. Panelists will be asked to choose the grade and content area (English language arts/literacy or mathematics) and grade in which they want to participate. Panelists will provide an email address, role, and demographic information. Panelists will then be asked to verify their email and select a 2-day window between October 6 and 17, 2014, for participation.

Registration must be submitted by September 19, 2014.

The timeline for the Online Panel is summarized below.

Activity	Responsible Party	Date(s)
Contact Participants	TICs, K-12 Leads, Higher Ed Leads	April 22-September 19, 2014
Register for Online Panel	Participants	April 22-September 19, 2014
Complete Online Achievement Level Setting	Participants	October 6-17, 2014 (two-day window selected during registration)

The information on the remaining attachments is designed to support recruitment efforts by Teacher Involvement Coordinators, and may be distributed as part of state recruitment efforts.

Attachment 1: Higher Education Leads Information Sheet for the Online Panel

This fall, K–12 educators, postsecondary educators, parents, and business/community leaders from Smarter Balanced governing states will collaborate to develop common achievement levels that are rigorous, fair, and accurate for each of the Smarter Balanced assessments. Smarter Balanced is asking for the assistance of Higher Education Leads with the recruitment of faculty and administrators to participate in this process.

Developing Achievement Levels will occur in a four-stage process:

The *“Online Panel”* will be open to the public and allow for broad participation by parents, teachers, school and district administrators, college faculty and administrators, and business/community leaders. The online panel will enable hundreds of thousands of interested constituents and members of the general public across the country to participate in the process. An FAQ and other information will be available online for those who register to participate in the online panel.

The *“In-Person Panel”* will allow for a group of more than 500 participants, nominated by Smarter Balanced governing states, to engage in a comprehensive review of questions on the assessment and make recommendations for operationalizing common expectations. For grades 3–8, the panel will be composed primarily of content specialists and educators. The high school panels will include higher education faculty in addition to K–12 educators.

A *“Vertical Articulation Committee”* composed of 60 panelists selected from those participating in the in-person workshop will review data and results from both panels and make Achievement Level recommendations. This step will help ensure a balanced view of achievement and proficiency levels for test-takers from Grade 3 through high school. No additional recruitment is required for this activity.

Finally, the recommendations from the Online Panel, the In-Person Panel, and the Vertical Articulation Committee will be presented to the chief school officers in Smarter Balanced governing states for their consideration and endorsement, in order to establish a common set of performance standards for mathematics and English language arts/Literacy across grades 3–8 and high school.

Purpose of the Online Panel

This is a critical opportunity to collect input from a diverse group of participants and establish consistent measures of progress for the Smarter Balanced interim and summative assessments. By participating in the online session, panelists will recommend an achievement level score that demonstrates how much students should know or be able to do in order to be proficient at the grade-level standards and on track for eventual college and career readiness. It is important to have as many participants as possible, from as many backgrounds as possible, provide these recommendations.

Participants in this panel will not set the final scores. The results of the panel will help inform the overall recommendations to the Smarter Balanced Assessment Consortium. Leaders from Smarter Balanced member states will carefully consider these recommendations in arriving at their decisions. The achievement levels endorsed by member states also will be subject to existing approval processes within individual states.

Logistics for the Online Panel

The panels will occur online from October 6 to 17, 2014. Panelists will select a 2-day window for participation during the registration process. It will take panelists up to 3 hours to complete the orientation process, review test questions, and recommend a score for Achievement Level 3.

Registration for the Online Panel

For more information and to register, interested participants will visit [this link](#) which will direct them to an application page. Panelists will be asked to choose the grade and content area (English language arts/literacy or mathematics) and grade in which they want to participate. Panelists will provide an email address, role, and demographic information. Panelists will then be asked to verify their email and select a 2-day window between October 6 and 17, 2014, for participation.

Registration must be submitted by September 19, 2014.

The timeline for the Online Panel is summarized below.

Activity	Responsible Party	Date(s)
Contact Participants	TICs, K-12 Leads, Higher Ed Leads	April 22-September 19, 2014
Register for Online Panel	Participants	April 22-September 19, 2014
Complete Online Achievement Level Setting	Participants	October 6-17, 2014 (two-day window selected during registration)

The information on the remaining attachments is designed to support recruitment efforts by Teacher Involvement Coordinators, and may be distributed as part of state recruitment efforts.

Educator Frequently Asked Questions (FAQ)

Smarter Balanced is recruiting educators from grades 3-12 to participate in the workshop to set Student Readiness Expectations. The purpose of the workshop is to set common expectations for student readiness/proficiency on the Smarter Balanced assessments. This process is also known as “standard-setting.” This is a critical opportunity for educators to ensure that the score results from these Common Core-aligned assessments are based on rigorous, yet fair expectations for students. This FAQ will answer many of the questions that educators involved in the process may have about their participation. For any additional questions or more information on the workshop and Smarter Balanced, contact CTB_smarterbalanced21@ctb.com.

The workshop to set Student Readiness Expectations will take three days and involves orientation to the tests, thorough training on a well-established review procedure, and three rounds of review and discussion among participants.

Frequently Asked Questions

Q: What is the workshop to set Student Readiness Expectations?

A: This is a critical opportunity to collect input from a diverse group of participants and establish consistent measures of progress. The process will rely on the professional judgment and classroom experience of educators in order to develop performance standards that are rigorous, fair and accurate. The workshop will allow for a group of more than 500 participants, nominated by states and selected by Smarter Balanced, to engage in a comprehensive review of questions on the assessments. The workshop, which will occur over several days this fall at a location to be determined, will give participants the opportunity to review the difficulty of questions and define what students performing at each of four achievement levels can reasonably be expected to answer correctly at each grade level.

Panelists participating in the workshop will use their professional judgment and experience in the classroom to make a recommendation about the minimum test scores expected for a student to be considered performing at each of the defined achievement levels on the Smarter Balanced assessments. The recommendations will be based on the Smarter Balanced definition of achievement, educators’ expertise in their content area, the claims and targets defined for each content area and grade level, educators’ experience with students in their classrooms and/or educators’ experience with the objectives of these assessments.

Q: Will the in-person workshop set the final scoring on these assessments?

Participants of this workshop will not set the final scores. This panel will make recommendations to the Smarter Balanced Assessment Consortium. Other technical and policy groups will also make recommendations. Smarter Balanced leadership and the chief state school officers of Smarter Balanced states will carefully consider this panel's recommendation in arriving at final decisions

Q: What are the dates of the workshop?

A: The workshop will be held over the following dates. Selected participants will be assigned to a specific panel depending on their content area expertise and grade levels taught.

- High School Language Arts and Mathematics: September 15–17, 2014
- Grades 6–8 Language Arts and Mathematics: September 17–19, 2014
- Grades 3–5 Language Arts and Mathematics: September 19–21, 2014

Q: How are educators selected to participate in the workshop to set Student Readiness Expectation?

A: Smarter Balanced will select educators from among those nominated by the states to serve on the workshop committees as participants or alternates. The distribution of educators will represent a balance of geographic regions, grade levels, years of experience, and other relevant factors.

Q: When will I know if I have been selected to participate in the workshop to set Student Readiness Expectations workshop?

A: Recruitment for the workshop will occur across all Smarter Balanced Governing States. Selected participants and alternates will receive notification regarding their status of participation by the end of June.

Once participants are selected, Measurement Incorporated will provide expected attendees with information related to travel and logistics. Measurement Incorporated will pay for travel, lodging, meals, and other allowable expenses for educators selected to participate in the workshop for setting Student Readiness Expectations.

Q: Where does the workshop occur?

A: The workshop for setting Student Readiness Expectations will be held in XX.

Q: Will my expenses be paid?

A: Measurement Incorporated will pay for travel, lodging, meals, and other allowable expenses for educators selected to participate in the workshop for setting Student Readiness Expectations.

Q: How much time can I expect to spend on this project in total?

A: Each workshop will occur over three consecutive days. The workshop day is approximately eight hours, with morning, lunch, and afternoon breaks.

Q: What qualifications do I need to apply to participate in the Student Readiness Expectation Setting process?

A: Participants must meet the following qualifications:

- Currently certified or licensed to teach ELA/literacy and/or mathematics in a K-12 public school.
- Currently teaching in a public school or are currently employed by a public school or district or state education entity, including higher education, located within a Smarter Balanced Governing State.
- Taught ELA/literacy and/or mathematics in grades 3–8 and/or high school within the past three years or have worked in a classroom content support role such as a literacy or mathematics coach, district or state content specialist, etc.
- Previously reviewed part or all of the Common Core State Standards for the content area in which they are interested in reviewing items and/or performance tasks.
- Previous experience with standard-setting workshops and activities (preferred but not required).

Q: Are recently retired teachers eligible to participate in the in-person workshop?

A: Participation in the workshop is limited to school employees who currently hold a position within an educational institution. Participation in the Student Readiness Expectation Setting

process is also meant to serve as a professional development activity for current school and district employees. The Smarter Balanced Participant Qualification Guidelines define the following criteria:

- Currently certified or licensed to teach ELA/literacy and/or mathematics in a K-12 public school.
- Currently teaching in a public school or currently employed by a public school or district or state education entity, including higher education, located within a Smarter Balanced Governing State.

Q: Will participants be able to share or discuss the results of the workshop with professional peers?

A: Individual items or questions on the Assessment are confidential since they may be used for the test in the future. Those selected will be asked to sign a confidentiality/non-disclosure agreement.

Email Text:

To: Smarter Balanced Higher Education Leads and Delegates

CC: Executive Staff, Nancy Arnold

From: CTB/McGraw-Hill (CTB_smarterbalanced21@ctb.com)

Subject: Action: Recruitment for Online Panel to Recommend Achievement Levels for Smarter Balanced Assessments

This email is being sent on behalf of CTB/McGraw-Hill. For questions please contact Jennifer Stegman at CTB_smarterbalanced21@ctb.com.

Dear Smarter Balanced Higher Education Leads and Delegates:

We request your assistance with the recruitment of higher education faculty and administrators to participate in the process to recommend scores for Achievement Levels on Smarter Balanced assessments. The Online Panel, which is open to the public and allows for broad participation, creates a critical opportunity to collect input from a diverse group of participants and establish consistent measures of progress for the Smarter Balanced interim and summative assessments. Participants will help ensure that the score results of the new Common Core-aligned assessments are based on rigorous and fair expectations for students.

By participating in the online session, panelists will recommend an achievement level score that demonstrates how much students should know or be able to do in order to be proficient at the grade-level standards and to be on track for eventual college and career readiness. It is important to have as many participants as possible, from as many backgrounds as possible, provide these recommendations.

This email and the included attachments provide information about the registration process for faculty and administrators for participation on the Online Panel. The information is designed to support recruitment efforts by higher education leads and delegates and may be distributed to educators as part of your recruitment efforts.

The attached documents consist of the following information:

- Information for higher education leads and delegates about the Online Panel to recommend Achievement Levels;
- A sample recruitment email;
- A list of Frequently Asked Questions that provides information for potential participants about the panel logistics and activities.

Please request all participants to register [here](#) by September 19, 2014.

For questions please contact Jennifer Stegman at CTB_smarterbalanced21@ctb.com.

Thank you for your assistance in this important endeavor.

Sample Educator Email

Dear Educator:

The Smarter Balanced Assessment Consortium is developing a common assessment system for mathematics and English language arts/literacy aligned to the Common Core State Standards to ensure all students are prepared for college and/or career. CTB/McGraw-Hill, on behalf of Smarter Balanced, is recruiting K–12 educators and administrators to participate in an Online Panel for recommending Achievement Levels for the assessments. This process is often referred to as “Standard Setting.”

The Online Panel, which is open to the public and allows for broad participation, creates a critical opportunity to collect input from a diverse group of participants and establish consistent measures of progress for the Smarter Balanced interim and summative assessments.

By participating in the online session, panelists will recommend an achievement level score that demonstrates how much students should know or be able to do in order to be proficient at the grade-level standards and to be on track for eventual college and career readiness. It is important to have as many participants as possible, from as many backgrounds as possible, provide these recommendations.

The process of setting Achievement Levels will rely on the professional judgment and experience of educators and administrators in order to develop performance standards that are rigorous, fair, and accurate. The recommendations will be based on Achievement Level Descriptors written and approved by Smarter Balanced Governing States last year, the claims and targets defined for each content area and grade level, as well as educators’ expertise in their content area and experience with students. Acceptance of recommended scores is subject to existing approval processes within individual states.

The Online Panel to recommend Achievement Levels will occur during a 2-day window within the October 6–17, 2014, time period. It will take panelists up to 3 hours to complete the orientation process, review test questions, and recommend a score for Achievement Level 3.

Interested participants can register [here](#).

On the registration site, you will be asked to choose the grade and content area (English language arts/literacy or mathematics) and grade in which you want to participate. You will provide an email address, role, and demographic information. You will then be asked to verify your email and select a 2-day window for participation.

Registrations must be submitted online by September 19, 2014.

Email Text:

To: Smarter Balanced Higher Education Leads and Delegates

CC: Executive Staff, Nancy Arnold

From: CTB/McGraw-Hill (CTB_smarterbalanced21@ctb.com)

Subject: Reminder 6-13: Recruitment for Online Panel for Achievement Level Setting for Smarter Balanced Assessments

This email is being sent on behalf of CTB/McGraw-Hill. For questions please contact Jennifer Stegman at CTB_smarterbalanced21@ctb.com.

Dear Smarter Balanced Higher Education Leads and Delegates:

We have attached the weekly Online Panel registration report showing the number of participants who have registered by state. You will receive these reports each Friday.

Smarter Balanced is requesting Higher Education Leads and Delegates to encourage faculty and administrators to register early for this opportunity in order to receive first choice of participation windows.

As a reminder, participants should visit [this link](#) and complete the easy four-step registration process. Panelists will choose a 2-day window between October 6 and 17, 2014, for participation.

Access to the Online Panel registration site is also available through the Smarter Balanced home page at www.smarterbalanced.org.

All registrations must be submitted by September 19, 2014.

Questions about the registration process can be directed to Jennifer Stegman at CTB_smarterbalanced21@ctb.com.

Thank you for your assistance in this important endeavor.

Attachment: Weekly Online Panelist Registrations as of June 13, 2014

Email Text:

To: Smarter Balanced State Leads and Teacher Involvement Coordinators

CC: Executive Staff, Nancy Arnold

From: CTB/McGraw-Hill (CTB_smarterbalanced21@ctb.com)

Subject: Reminder 6-13: Recruitment for Online Panel for Achievement Level Setting for Smarter Balanced Assessments

This email is being sent on behalf of CTB/McGraw-Hill. For questions please contact Jennifer Stegman at CTB_smarterbalanced21@ctb.com.

Dear Smarter Balanced State Leads and Teacher Involvement Coordinators:

We have attached the weekly Online Panel registration report showing the number of participants who have registered by state. You will receive these reports each Friday.

Smarter Balanced is requesting states to encourage educators to register early for this opportunity in order to receive first choice of participation windows.

As a reminder, participants should visit [this link](#) and complete the easy four-step registration process. Panelists will choose a 2-day window between October 6 and 17, 2014, for participation.

Access to the Online Panel registration site is also available through the Smarter Balanced home page at www.smarterbalanced.org.

All registrations must be submitted by September 19, 2014.

Questions about the registration process can be directed to Jennifer Stegman at CTB_smarterbalanced21@ctb.com.

Thank you for your assistance in this important endeavor.

Attachment: Weekly Online Panelist Registrations as of June 13, 2014

Higher Education Faculty and Administrator Email Template and Stock Newsletter Item

- 1. Email Template:** This customizable email template for higher education faculty and administrators is intended for distribution to colleges and universities as well as state and local professional associations. This tool can be used for direct email outreach to inform higher education faculty and administrators about the opportunity to participate in the Online Panel. The template can be customized to the needs of individual states.

For colleges and universities, it is vital that high school graduates arrive ready for credit-bearing course work. For this reason, higher education faculty and administrators have a sizable stake in the future of K-12 standards and assessments.

The **[Common Core State Standards/name of state college- and career-ready standards]** help ensure that students experience real academic rigor at the K-12 level. The Smarter Balanced Assessment Consortium is developing new assessments aligned to these higher standards. The tests, administered at Grades 3 through 8 and 11, will provide a meaningful and fair measurement of student achievement and progress toward college and career readiness.

To ensure that the tests are rigorous, fair, and accurate for all students, Smarter Balanced is inviting **[higher education faculty and administrators in your state/members of your association]** to participate in an Online Panel to Recommend Achievement Levels on the assessments.

The perspective of higher education faculty and administrators is a critical part of this process. Participants will recommend an achievement level score that demonstrates how much students should know or be able to do in order to meet the grade-level standards and to be on track for eventual college and career readiness. Panelists can participate from the comfort of their own home or office, at a time that works best for them.

The Online Panel to Recommend Achievement Levels will occur October 6–17, 2014. It will take participants up to 3 hours to complete the orientation process, review test questions, and recommend a score. Each participant will have two days to complete the entire process.

Visit the Smarter Balanced website to [register and learn more](#). Registration closes on September 19, 2014.

We appreciate your assistance with this important endeavor.

- 2. Institution/Association Newsletter Item:** This customizable newsletter item for higher education faculty and administrators is intended for distribution to institutions and state and local professional associations for publication as an announcement/article in print and email newsletters. The template can be customized to the needs of individual organizations.

For colleges and universities, it is vital that high school students arrive ready for credit-bearing course work. For this reason, higher education faculty and administrators have a sizable stake in the future of K-12 standards and assessments.

The [Common Core State Standards/name of state college- and career-ready standards] help ensure that students experience real academic rigor at the K-12 level. The Smarter Balanced Assessment Consortium is developing new assessments aligned to these higher standards. The tests, administered at Grades 3 through 8 and 11, will provide a meaningful and fair measurement of student achievement and progress toward college and career readiness.

[Higher education faculty and administrators in your state/members of your association] are invited to participate in the Online Panel to Recommend Achievement Levels for the Smarter Balanced assessments. This process is often referred to as “standard setting.” By participating in the online session, panelists will recommend an achievement level score that demonstrates how much students should know or be able to do in order to meet the grade-level standards and to be on track for eventual college and career readiness.

The Online Panel to Recommend Achievement Levels will occur October 6–17, 2014. It will take participants up to 3 hours to complete the orientation process, review test questions, and recommend a score. Each participant will have two days to complete the entire process.

Visit the Smarter Balanced website to [register and learn more](#). Registration closes on September 19, 2014.

Educators/School Administrators Email Template and Stock Newsletter Item

- 1. Email Template:** This customizable email template for educators/school administrators is intended for distribution to districts and state and local professional associations. This tool can be used for direct email outreach to inform educators/school administrators about the opportunity to participate in the Online Panel. The template can be customized as needed.

Nothing is more important than giving all students an education that prepares them for success. That is the goal of the [Common Core State Standards/name of State College- and career-ready standards]. The standards help ensure that all students leave school prepared for success after graduation and also provide clear, consistent guidelines for educators and school administrators.

The Smarter Balanced Assessment Consortium is developing new assessments aligned to these higher standards. The tests provide teachers and school leaders with information to improve instruction and help students meet high expectations through innovative, computer-adaptive assessments that are challenging and engaging.

To ensure that the tests are rigorous, fair, and accurate for all students, Smarter Balanced is inviting [Educators/administrators in your state/members of your association] to participate in an Online Panel to Recommend Achievement Levels. Participants will recommend an achievement level score that demonstrates how much students should know or be able to do in order to meet the grade-level standards and to be on track for eventual college and career readiness. Panelists can participate from the comfort of their own home or office, at a time that works best for them.

The Online Panel to Recommend Achievement Levels will occur October 6–17, 2014. It will take participants up to 3 hours to complete the orientation process, review test questions, and recommend a score. Each participant will have two days to complete the entire process.

Visit the Smarter Balanced website to [register and learn more](#). Registration closes on September 19, 2014.

We appreciate your assistance with this important endeavor.

- 2. Association Newsletter Item:** This customizable newsletter item for educators/school administrators is intended for distribution to districts and state or local professional associations for publication through print and email newsletters. The template can be customized as needed.

[Educators in your state/members of your association] are invited to participate in the Online Panel to Recommend Achievement Levels for the Smarter Balanced assessments. This process is often referred to as “standard setting.” By participating in the Online Panel, educators will recommend an achievement level score that demonstrates how much students should know or be able to do in order to meet the grade-level standards and to be on track for eventual college and career readiness.

Educators will have the opportunity to review and analyze the test’s items and structure and recommend an achievement level score. This opportunity will allow a large group of educators and administrators to



Outreach Templates for Online Panel

help establish consistent measures of progress for the Smarter Balanced interim and summative assessments.

The Online Panel to Recommend Achievement Levels will occur October 6–17, 2014. It will take participants up to 3 hours to complete the orientation process, review test questions, and recommend a score. Each participant will have two days to complete the entire process.

Visit the Smarter Balanced website to [register and learn more](#). Registration closes on September 19, 2014.

Parent Outreach Email Templates

Version 1

Are you a parent of a school-aged child? Would you like to weigh in on new assessments being introduced across the U.S.?

The Smarter Balanced Assessment Consortium is developing a new testing system for English and math aligned to the Common Core State Standards.

Developed with input from teachers and parents, these computer-based assessments will be accessible to students of all abilities and backgrounds and will measure student progress effectively and accurately. Our goal is to accurately identify core knowledge and skills—not just test whether students recognize patterns and memorize facts.

We know parents want their children to get an education that prepares them for success. That's why we are seeking your help in setting achievement levels that will measure whether students are ready for the next grade and for college and a career after high school.

In order to make sure that our process is inclusive and reflects parent input, we are organizing an online panel. You can participate from the comfort of your own computer, giving us feedback on realistic expectations for students.

Help ensure that these tests are rigorous and fair and give us a clear picture of whether students are ready for success after high school. The online panel will take about three hours of your time, on a date of your choice between October 6 and October 17, 2014.

For more information and to register, please click here. (Link to Smarter Balanced page)

Version 2

You want your children to have an education that will prepare them for college and career success. By raising standards for all, we can challenge schools, teachers and students to do better.

Teachers agree that assessment provides important information about student achievement, which will help improve results for students across the board.

That's why we are inviting parents to participate in an online panel to set achievement levels for new Common Core aligned tests. We are seeking your help in setting achievement levels to measure whether students are ready for the next level, whether it's 3rd grade or on to college or a career. In order to get the widest possible level of input, we are organizing an online panel, and you can participate from the comfort of your own computer.



Parent Outreach for Online Panel Email Templates

We want you to ensure that new tests aligned to the Common Core will provide a realistic roadmap for your child's success. That's why the Smarter Balanced Assessment Consortium is inviting parents, teachers and community leaders to take part in the largest achievement level setting event to-date.

This online event will require about three hours of your time, on a date of your choice between October 6 and October 17, 2014.

Join parents from across the country in this critical opportunity to participate and give feedback on these new assessments.

For more information and to register, please click here. (Link to Smarter Balanced page)

Version 3

Parents,

All children deserve the best education possible. The new college and career ready academic standards will help ensure that all students are prepared for success after graduation.

The Smarter Balanced Assessment Consortium is developing new assessments that will provide a meaningful and fair 'academic checkup' on English and math for students. The tests will measure real-world skills like critical thinking and problem solving using digital software accessible to students of all backgrounds.

We know you are dedicated to your child's success in school and want him or her to have the brightest possible future with options for college or a career. That's why we are inviting parents to participate in an online panel to preview the Smarter Balanced assessments and help us set achievement levels.

Smarter Balanced is seeking your help in setting achievement levels that will reflect whether students are ready for the next grade and for college and a career after high school. In order to get the broadest input possible, we are organizing an online panel, and you can participate from the comfort of your own computer.

Smarter Balanced is inviting parents to give their input on the achievement level setting for the assessments in Grades 3 through 11. This online event will require about three hours of your time on the date of your choice between October 6 and October 17, 2014.

For more information and to register, please click here. (Link to Smarter Balanced page)

Email Text:

To: Smarter Balanced Higher Education Leads and Delegates

CC: Executive Staff, Nancy Arnold

From: CTB/McGraw-Hill (CTB_smarterbalanced21@ctb.com)

Subject: Action: Recruitment for In-Person Panel to Recommend Achievement Levels for Smarter Balanced Assessments

This email is being sent on behalf of CTB/McGraw-Hill. For questions please contact Jennifer Stegman at CTB_smarterbalanced21@ctb.com.

Dear Smarter Balanced Higher Education Leads and Delegates:

We request your assistance with the nomination of qualified higher education faculty and administrators to participate in the process to recommend cut scores for Achievement Levels on Smarter Balanced Assessments. This is a critical opportunity to collect input from a diverse group of participants and establish consistent measures of progress for the Smarter Balanced interim and summative assessments. Participants will help ensure that the score results of the new Common Core-aligned assessments are based on rigorous and fair expectations for students. The process for recommending Achievement Levels will rely on the professional judgment and experience of educators in order to develop performance standards that are rigorous, fair and accurate. The recommendations from the In-Person Panel will be submitted for endorsement to the chief school officers in governing states. Our goal is to follow a process that is as inclusive as possible, representative of views across member states, transparent and scientifically rigorous.

The In-Person Panel will allow for a group of more than 500 participants, nominated by states, to engage in a comprehensive review of questions on the assessments. The High School panel, which will convene October 13–15, 2014, at a location to be determined, will give participants the opportunity to review the difficulty of questions and define what students performing at each of four Achievement Levels can reasonably be expected to answer correctly at the Grade 11 level.

Higher education representatives will primarily participate in the segment of the panel devoted to setting Achievement Levels on the Grade 11 assessments. This message and the accompanying materials are focused on nominating individuals for Grade 11. Higher education leads and delegates also may want to work with their K–12 counterparts to identify faculty who might be nominated by K–12 for inclusion in the work on earlier grades.

This email and the included attachments provide information about the nomination of faculty and administrators for the Grade 11 segment of the In-Person Panel, part of the process for recommending Achievement Levels on the Smarter Balanced assessments to governing state chiefs. The information is designed to support recruitment efforts by higher education leads and delegates and may be distributed to educators as part of your recruitment efforts.

The attached documents consist of the following information:

- Information for higher education leads and delegates about the In-Person Panel to recommend Achievement Levels, including state recruitment counts;



In-Person Panel for Setting Achievement Levels

- A sample recruitment email;
- A list of Frequently Asked Questions that provides information for potential nominees about the panel logistics and activities; and
- A nomination list template.

Please request all nominees to complete the online Nominee Registration Form by **Friday, May 23, 2014**, and submit your list of nominees by the same date to Jennifer Stegman at CTB_smarterbalanced21@ctb.com.

Questions about the nomination process also can be directed to Jennifer Stegman at CTB_smarterbalanced21@ctb.com.

Thank you for your assistance in this important endeavor.

Email Text:

To: Smarter Balanced State Leads and Teacher Involvement Coordinators

CC: Executive Staff, Nancy Arnold

From: CTB/McGraw-Hill (CTB_smarterbalanced21@ctb.com)

Subject: Action: Recruitment for In-Person Panel to Recommend Achievement Levels for Smarter Balanced Assessments

This email is being sent on behalf of CTB/McGraw-Hill. For questions please contact Jennifer Stegman at CTB_smarterbalanced21@ctb.com.

Dear Smarter Balanced Teacher Involvement Coordinators and State Leads:

We request your assistance with the nomination of qualified educators, administrators, parents and business/community leaders to participate in the process to recommend cut scores for Achievement Levels on Smarter Balanced assessments. This is a critical opportunity to collect input from a diverse group of participants and establish consistent measures of progress for the Smarter Balanced interim and summative assessments. Participants will help ensure that the score results of the new Common Core-aligned assessments are based on rigorous and fair expectations for students. The process for recommending student Achievement Levels will rely on the professional judgment and experience of educators, administrators, parents and community leaders in order to develop performance standards that are rigorous, fair and accurate. The recommendations from the In-Person Panel will be submitted for endorsement to the chief school officers in governing states. Our goal is to follow a process that is as inclusive as possible, representative of views across member states, transparent and scientifically rigorous.

The In-Person Panel will allow for a group of more than 500 participants, nominated by states, to engage in a comprehensive review of questions on the assessments. The panel, which will convene October 13–19, 2014, at a location to be determined, will give participants the opportunity to review the difficulty of questions and define what students performing at each of four Achievement Levels can reasonably be expected to answer correctly at each grade level. Selected panelists will each participate for three days of the panel.

This email and the included attachments provide information about the nomination of educators, administrators, parents and community members for the In-Person Panel, part of the process for recommending Achievement Levels on the Smarter Balanced assessments to governing state chiefs.

The attached documents consist of the following information:

- Information for State Leads and TICs about the In-Person Panel to recommend Achievement Levels that includes state recruitment counts and required educator qualifications;
- Sample recruitment emails for educators and community members to support the TICs' state recruitment;
- A list of Frequently Asked Questions that provides information about the workshop logistics and activities;
- A sample form that TICs can use to collect information about potential nominees; and
- A nomination list template.



In- Person Panel for Setting Achievement Levels

Please request all nominees to complete the online Nominee Registration Form by **Friday, May 23, 2014**, and **submit your list of nominees by the same date to Jennifer Stegman at CTB_smarterbalanced21@ctb.com**.

Thank you for your assistance in this important endeavor.

Attachment 1: Higher Education Lead Information Sheet

This fall, K-12 educators, post-secondary educators, parents, and community leaders from Smarter Balanced governing states will collaborate to develop common Achievement Levels and performance standards that are rigorous, fair, and accurate for each of the Smarter Balanced assessments. Smarter Balanced is asking for the assistance of higher education leads and delegates with the recruitment and nomination of qualified faculty and administrators to participate in this process.

Developing Achievement Levels will occur in a four-stage process:

The “*In-Person Panel*” will allow for a group of more than 500 participants, nominated by states, to engage in a comprehensive review of questions on the assessment and to make recommendations for operationalizing common expectations. For grades 3–8, the panels will be composed primarily of K–12 content specialists and teachers. The high school panels will include higher education faculty and administrators in addition to K–12 educators.

The “*Online Panel*” will be open to the public and allow for broad participation by parents, teachers, school and district administrators, college faculty and administrators, and business/community leaders. The Online Panel will enable thousands of interested educators and community leaders across the country to participate in the process. Higher education leads and delegates will receive information on April 22 about recruiting participants for the Online Panel.

A “*Vertical Articulation Committee*” composed of 60 panelists selected from those participating in the In-Person Panel will review data and results from both the In-Person Panel and the Online Panel and make Achievement Level recommendations. This step will help ensure a balanced view of achievement and proficiency levels for test-takers from Grade 3 through high school. No additional recruitment is required for this activity.

Finally, the recommendations from the In-Person Panel, the Online Panel, and the Vertical Articulation Committee will be presented to the chief school officers in Smarter Balanced governing states for their consideration and endorsement, in order to establish a common set of performance standards for mathematics and English language arts/Literacy across grades 3–8 and high school.

Purpose of the In-Person Panel

The purpose of the In-Person Panel is to develop common expectations for student performance on the Smarter Balanced assessments. Panelists will make recommendations to the Smarter Balanced Assessment Consortium about the minimum test scores aligned to four defined Achievement Levels.

The recommendations will be based on Achievement Level Descriptors written and approved by Smarter Balanced Governing States last year, the claims and targets defined for each content area and grade level, as well as participants’ expertise in their content area and experience with students in their classrooms.

Logistics for the In-Person Panel

Panelists will participate in the High School segment of the panel October 13–15, 2014 (see Table 1). Participants will review tests and recommend score points matching the Achievement Levels, using a process consistent with industry best practices. During the orientation session on the first day, participants will take a practice test in their subject and grade level, review content standards and Achievement Level Descriptors for their subject and grade level, and receive detailed instruction in a test review procedure that leads to recommended scores for identifying students at each of the Achievement Levels. Once participants have undergone training, they will engage in three rounds of test item review. Between rounds of item review, participants will discuss results of the previous round with their colleagues. The panel facilitators will provide the results of each round of review and initiate the discussions. Throughout the process, each participant will have an opportunity to express his or her opinions about appropriate expectations for students at each Achievement Level and hear the opinions of others in the group.

Table 1. Key Dates for In-Person Panel

Activity	Start Date	End Date
High School English Language Arts/Literacy and Mathematics	October 13	October 15
Grades 6–8 English Language Arts/Literacy and Mathematics	October 15	October 17
Grades 3–5 English Language Arts/Literacy and Mathematics	October 17	October 19

Nomination and Selection Process for the In-Person Panel

The nomination process for higher education participants in the high school segment of the panel will follow these steps:

1. Higher education leads or delegates identify potential nominees. The number of participants needed from your state is listed at the end of this document. In order to ensure that a representative panel can be selected, we request that Higher Education Leads nominate at

least the number of nominations indicated. Higher education participants should include faculty and administrators with responsibility for placement and/or first-year curricula (e.g. vice presidents for enrollment management, directors of admissions, registrars, deans of undergraduate education).

Higher education nominees must possess the following qualifications:

- Currently employed at a college or university in a Smarter Balanced state;
 - Familiar with the Common Core State Standards for high school; and
 - Has strong familiarity with first-year course requirements and curricula.
2. The higher education lead or delegate contacts potential nominees (a sample recruitment email and FAQs are included in this information packet).
 3. Once individuals confirm their interest in participating in the panel, the higher education lead or delegate should direct them to the online Higher Education Nominee Registration Form at <https://www.surveymonkey.com/s/IPSSRE>.
 4. The higher education lead or delegate should send a list of confirmed nominees to CTB_smarterbalanced21@ctb.com (using the template provided as Attachment 4). Smarter Balanced will check the registration forms completed against the list of nominees to ensure that it has received completed forms for each nominee.

Both the registration forms and the list of nominees must be submitted by May 23, 2014.

Smarter Balanced is requesting more nominees than there will be spaces in the panel to ensure that the panel is diverse in terms of types of institutions represented as well as participants' personal characteristics and professional background/expertise. Smarter Balanced will select participants from each state and a small number of "at-large" alternates based on these characteristics.

Alternates will be asked to attend the panel and will fill in for participants from any state who are unexpectedly unable to attend (time permitting, Smarter Balanced will work with states to identify a replacement for any earlier cancellations). Even if all participants are able to attend, the alternates will be invited to participate so their time is not wasted. Higher education leads and delegates will receive advance notification of selected participants and alternates.

Once participants are selected, they will receive an invitation with all necessary logistical and other background information. Smarter Balanced will pay for travel, lodging, meals, and other allowable expenses for all those selected to participate in the panel. An honorarium may be provided as applicable. Table 2 summarizes the recruitment timeline.

Table 2. Recruitment Timeline

Process	Responsible Party	Date(s)
Recruitment	Higher Education Leads and Delegates	4/15/2014 - 5/23/2014
Deadline to Submit Recommended Panelists	Higher Education Leads and Delegates	5/23/2014
Panelist Selection and Higher Education Lead Notification	Smarter Balanced	5/27/2014 - 6/9/2014
Panelist Notification	Smarter Balanced	6/10/2014 - 6/30/2014

The information on the remaining attachments is designed to support recruitment efforts by higher education leads and delegates, and may be distributed as part of state recruitment efforts.

Table 3. Number of Higher Education Participants for Grade 11 by State and Content Area

State	English language arts/literacy	Mathematics
California -CCC	1	2
California -CSU	1	2
California - UC	2	0
Connecticut	1	1
Delaware	1	1
Hawaii	1	1
Idaho	1	1
Iowa	1	1
Maine	1	1
Michigan	3	3
Missouri	1	1
Montana	1	1
Nevada	1	1
New Hampshire	1	1
North Carolina	1	1
North Dakota	1	1
Oregon	2	2
South Carolina	2	2
South Dakota	1	1
Vermont	1	1
Washington	3	3
West Virginia	1	1

In-Person Panel for Setting Achievement Levels

Wyoming	1	1
Total	30	30

Please submit at least the number of nominations indicated.

Note: There may be faculty who have interest and appropriate expertise to participate in the segments of the panel for earlier grades. Higher education leads may work with their K-12 counterparts to identify faculty who K-12 may want to nominate for inclusion as “non-teacher educators” for earlier grades.

Attachment 1: TIC Information Sheet

This fall, K–12 educators, post-secondary educators, parents, and business/community leaders from Smarter Balanced governing states will collaborate to develop common Achievement Levels that are rigorous, fair, and accurate for each of the Smarter Balanced assessments. Smarter Balanced is asking for the assistance of Teacher Involvement Coordinators (TICs) with the recruitment and nomination of qualified K–12 educators, administrators, parents and business/community leaders to participate in this process. Smarter Balanced will request the higher education leads from the Smarter Balanced states to recruit and nominate post-secondary educators.

Developing Achievement Levels will occur in a four-stage process:

The “*In-Person Panel*” will allow for a group of more than 500 participants, nominated by member states, to engage in a comprehensive review of questions on the assessment and to make recommendations for operationalizing common expectations. For grades 3–8, the panel will be composed primarily of content specialists and educators. The high school panel will include K–12 educators and higher education faculty. A small number of parents, administrators, and community leaders will also be included.

The “*Online Panel*” will be open to the public and allow for broad participation by parents, teachers, school and district administrators, college faculty and administrators, and business/community leaders. The Online Panel will enable thousands of interested educators and other experts across the country to participate in the process. TICs will receive information on April 22 about recruiting participants for the Online Panel.

A “*Vertical Articulation Committee*” composed of 60 panelists selected from those participating in the In-Person Panel will review data and results from both the In-Person Panel and the Online Panel and make Achievement Level recommendations. This step will help ensure a balanced view of achievement and proficiency levels for test-takers from Grade 3 through high school. No additional recruitment is required for this activity.

Finally, the recommendations from the In-Person Panel, the Online Panel, and the Vertical Articulation Committee will be presented to the chief school officers in Smarter Balanced governing states for their consideration and endorsement, in order to establish a common set of performance standards for mathematics and English language arts/Literacy across grades 3–8 and high school.

Purpose of the In-Person Panel

The purpose of the In-Person Panel is to develop common expectations for student performance on the Smarter Balanced assessments. Panelists will make recommendations to the Smarter Balanced Assessment Consortium about the minimum test scores aligned to four defined Achievement Levels.

The recommendations will be based on Achievement Level Descriptors written and approved by Smarter Balanced Governing States last year; the claims and targets defined for each content area and grade level; educators’ expertise in their content area and experience with students in their classrooms; and parent and community knowledge of students.

Logistics for the In-Person Panel

The panel will take place October 13–19, 2014. Educators will participate in one of three segments of the panel based on their grade level and subject area expertise (see Table 1). Participants will review tests and recommend score points matching the Achievement Levels, using a process consistent with industry best practices. Each segment will include an orientation session on the first day, during which participants will take a practice test in their subject and grade level, review content standards and Achievement Level Descriptors for their subject and grade level, and receive detailed instruction in a test review procedure that leads to recommended scores for identifying students at each of the Achievement Levels. Once participants have undergone training, they will engage in three rounds of test item review. Between rounds of item review, participants will discuss results of the previous round with their colleagues. The panel facilitators will provide the results of each round of review and initiate the discussions. Throughout the process, each participant will have an opportunity to express his or her opinions about appropriate expectations for students at each Achievement Level and hear the opinions of others in the group.

Table 1. Key Dates for In-Person Panel

Activity	Start Date	End Date
High School English Language Arts/Literacy and Mathematics	October 13	October 15
Grades 6–8 English Language Arts/Literacy and Mathematics	October 15	October 17
Grades 3–5 English Language Arts/Literacy and Mathematics	October 17	October 19

Nomination and Selection Process for the In-Person Panel

The nomination process will follow these steps:

1. The TIC and/or State Lead identifies potential nominees. The number of participants needed from your state for each segment of the panel is listed as Appendix A of this document. In order to ensure that a representative panel can be selected, we request that TICs submit at least the number of nominations indicated in the tables in Appendix A.

Educator Nominees should meet **all** of the following criteria:

- Currently certificated or licensed to teach ELA/literacy and/or mathematics in a K–12 public school;
- Currently teaching in a public school or (for the non-teacher educator category) are currently employed by a public school or district (including higher education), located within the nominating Smarter Balanced state;
- Has taught ELA/literacy and/or mathematics in grades 3–8 and/or high school within the past three years or has worked in a classroom content support role such as a literacy or mathematics coach, district or state content specialist, administrator, etc.; and
- Previously reviewed part or all of the Common Core State Standards for the content area in which they are interested in reviewing items and/or performance tasks.

We recommend that Parent and Community Nominees should meet **at least one** of the following criteria:

- Business leaders with experience hiring and managing personnel;
 - Parents, especially those involved with education groups such as PTA; or
 - Community leaders, including elected officials, school board members, legislators, and recognized civic leaders.
2. The TIC contacts potential nominees (sample recruitment emails for educators and parents/community members and FAQs are included in this information packet). A sample form that TICs may use to collect information about potential nominees is included in this document as Appendix B and is also included as Attachment 5.
 3. Once the member state has determined the state’s final list of nominees, those individuals should be referred to the online Nominee Registration Form at <https://www.surveymonkey.com/s/IPSSRE>.
 4. The TIC should then send a list of nominees to CTB_SmarterBalanced21@ctb.com using the template included as Attachment 6. Smarter Balanced will check the registration forms completed against the list of nominees to ensure that it has received completed forms for each nominee.

Both the online registration forms and the list of nominees must be submitted by May 23, 2014.

Smarter Balanced will use the information on the online Nominee Registration Form to select participants for the In-Person Panel. Participants will be selected to achieve coverage of the necessary content expertise and state representation. Smarter Balanced also will select a small number of “at-large” alternates. Alternates will be asked to attend the panel and will fill in for

participants from any state who are unexpectedly unable to attend (time permitting, Smarter Balanced will work with states to identify a replacement for any earlier cancellations). Even if all participants are able to attend, the alternates will be invited to participate so their time is not wasted. TICs will receive advance notification of selected participants and alternates.

Once participants are selected, they will receive an invitation confirming the time and place of the panel. Smarter Balanced will pay for travel, lodging, meals, and other allowable expenses for educators, parents, and community leaders selected to participate in the panel. Substitute reimbursement or a stipend may be provided as applicable. Table 2 summarizes the recruitment timeline.

Table 2. Recruitment Timeline

Process	Responsible Party	Date(s)
Recruitment	State Teacher Involvement Coordinator	4/15/2014 - 5/23/2014
Deadline to Submit Recommended Participants https://www.surveymonkey.com/s/IPSSRE	State Teacher Involvement Coordinator	5/23/2014
Participant Selection and TIC Notification	Smarter Balanced	5/27/2014 - 6/9/2014
Participant Notification	Smarter Balanced	6/10/2014 - 6/30/2014

The information on the remaining attachments is designed to support recruitment efforts by Teacher Involvement Coordinators, and may be distributed as part of state recruitment efforts.

Appendix A. Minimum Number of Requested Nominees by Content Area, Role, Grade Level and State

- Table A1 outlines the minimum number of nominations requested from your state for English Language Arts/Literacy by role and grade level; and
- Table A2 outlines the minimum number of nominations requested from your state for Mathematics by role and grade level.

Table A1: California Minimum Number of Requested Nominations for English Language Arts/Literacy by Role and Grade Level

Grade Level	On-Grade Teacher	*Above-Grade Teacher	**Teacher: ELL Experience	***Teacher: SWD Experience	Educator: Non-Teaching	Parents and Business/Community Leaders	Total Number of Nominations Requested by Grade Level
Grade 3	1	1	0	0	0	0	2
Grade 4	0	0	0	0	0	1	1
Grade 5	1	0	0	0	0	0	1
Grade 6	0	0	0	1	0	0	1
Grade 7	1	0	0	0	0	0	1
Grade 8	0	1	0	0	0	0	1
High School	1	0	1	0	0	0	2
Total Number of Nominations Requested by Role	4	2	1	1	0	1	9

*Smarter Balanced wants to make sure that teachers from the next grade up are involved in recommending achievement levels at each grade (e.g., a nomination for an Above-Grade Teacher identified in the Grade 8 row should be a Grade 9 teacher).

**Educators certified to teach English language learners (ELL) and who have taught ELL students in the last 3 years.

***Educators certified to teach Students with Disabilities (SWD) and who have taught SWD students in the last 3 years.

Note: Please submit at least the number of nominations indicated in the table above. Additional nominations of highly qualified educators of any grade level, parents, or business/community leaders are encouraged as well. The minimum numbers are requested in order to ensure adequate coverage of individuals representing various roles.

Table A2: California Minimum Number of Requested Nominations for Mathematics by Role and Grade Level

Grade Level	On-Grade Teacher	*Above-Grade Teacher	**Teacher: ELL Experience	***Teacher: SWD Experience	Educator: Non-Teaching	Parents and Business/Community Leaders	Total Number of Nominations Requested by Grade Level
Grade 3	0	0	0	0	1	0	1
Grade 4	1	0	0	0	0	0	1
Grade 5	0	1	0	0	0	0	1
Grade 6	1	0	0	0	0	0	1
Grade 7	0	0	0	0	1	0	1
Grade 8	1	0	0	0	0	0	1
High School	1	0	0	0	1	0	2
Total Number of Nominations Requested by Role	4	1	0	0	3	0	8

*Smarter Balanced wants to make sure that teachers from the next grade up are involved in recommending achievement levels at each grade (e.g., a nomination for an Above-Grade Teacher identified in the Grade 5 row should be a Grade 6 teacher).

**Educators certified to teach English language learners (ELL) and who have taught ELL students in the last 3 years.

***Educators certified to teach Students with Disabilities (SWD) and who have taught SWD students in the last 3 years.



In-Person Panel for Setting Achievement Levels

Note: Please submit at least the number of nominations indicated in the table above. Additional nominations of highly qualified educators of any grade level, parents, or business/community leaders are encouraged as well. The minimum numbers are requested in order to ensure adequate coverage of individuals representing various roles.

Appendix B. Sample Participant Application

Application for In-Person Panel to Set Achievement Levels	
First Name:	
Last Name:	
Primary Email Address [Work]:	
Secondary Email Address [Personal – For summer contact]:	
Contact Phone Number:	
Alternate Phone Number where you can be contacted during the summer:	
Gender:	<input type="checkbox"/> Male <input type="checkbox"/> Female
Ethnicity (X):	<input type="checkbox"/> Hispanic or Latino <input type="checkbox"/> Not Hispanic or Latino
Race (X): (Mark all that apply)	<input type="checkbox"/> Native American or Alaska Native <input type="checkbox"/> Asian <input type="checkbox"/> Black or African American <input type="checkbox"/> Native Hawaiian or Other Pacific Islander <input type="checkbox"/> White
State:	
Please select the Role that best fits your experience	<input type="checkbox"/> Parent <input type="checkbox"/> Business or Community Leader <input type="checkbox"/> Educator: General <input type="checkbox"/> Educator: English language learner experience <input type="checkbox"/> Educator: Students with disabilities experience <input type="checkbox"/> Educator: Administration/Non-Teaching
This Section Is To Be Filled Out By Educators	
Name of District or Education Entity for 2013-2014:	
Name of School for 2013-2014:	
School Type (X): <input type="checkbox"/> Urban <input type="checkbox"/> Suburban <input type="checkbox"/> Rural	
Content Area(s) Taught or Supported within the last three years:	
Grade Level(s) Taught or Supported within the last three years:	
Are you certified to teach English language learners? (X)	<input type="checkbox"/> Yes <input type="checkbox"/> No
If yes, have you taught English language learners in the last three years? (X)	<input type="checkbox"/> Yes <input type="checkbox"/> No
Are you certified to teach students with disabilities? (X)	<input type="checkbox"/> Yes <input type="checkbox"/> No
If yes, have you taught students with disabilities in the last three years? (X)	<input type="checkbox"/> Yes <input type="checkbox"/> No

Attachment 2: Higher Education Frequently Asked Questions (FAQ)

Smarter Balanced is recruiting higher education faculty and administrators to participate in a panel for setting Achievement Levels for its K–12 assessments in English language arts/Literacy and mathematics. This process is often referred to as “standard-setting.” This is a critical opportunity for educators to ensure that the score results from these Common Core-aligned assessments are based on rigorous, yet fair expectations for students. This FAQ will answer many of the questions that invited nominees may have about their participation. For any additional questions or more information on the panel and Smarter Balanced, contact CTB_smarterbalanced21@ctb.com.

The In-Person Panel for setting Achievement Levels will take three days for each participant and involves orientation to the tests, thorough training on a well-established review procedure, and three rounds of review and discussion among participants.

Q: What is the In-Person Panel?

A: This is a critical opportunity to collect input from a diverse group of participants and establish consistent measures of student progress toward college and career readiness. The process will rely on the professional judgment and expertise of participants in order to develop performance standards that are rigorous, fair and accurate. The panel will be composed of more than 500 participants, nominated by states and selected by Smarter Balanced, to engage in a comprehensive review of questions on the assessments. The panel, which will convene between October 13–19, 2014, at a location to be determined, will give participants the opportunity to review the difficulty of questions and define what students performing at each of four Achievement Levels can reasonably be expected to answer correctly at each grade level. Higher education participation will occur principally during the segment of the panel devoted to setting Achievement Levels for the Grade 11 assessments (October 13–15, 2014).

Participants in the In-Person Panel will use their professional judgment and experience to make a recommendation about the minimum test scores expected for a student to be considered performing at each of the defined Achievement Levels on the Smarter Balanced assessments. The recommendations will be based on Achievement Level Descriptors written and approved by Smarter Balanced Governing States last year, the claims and targets defined for each content area and grade level, as well as educators’ expertise in their content area and experience with students in their classrooms.

Q: Will the In-Person Panel set the final operational scores on these assessments?

A: Participants in this panel will not set the final operational scores. A “Vertical Articulation Committee” will make recommendations to the chief school officers in Smarter Balanced

Governing States using a variety of information sources. These sources will include the recommendations of the In-Person Panel, recommendations from an Online Panel which will engage in a similar – but less comprehensive – process as the In-Person Panel, and external assessment information. Acceptance of recommended scores is subject to existing approval processes within individual states.

Q: What are the dates of the In-Person Panel?

A: The segment of the panel devoted to setting Achievement Levels for the Grade 11 assessments will occur October 13–15, 2014.

Q: Where will the In-Person Panel be held?

A: The location has not yet been determined. Selected participants will be notified of the panel location once it is finalized.

Q: How are panelists selected to participate in the In-Person Panel?

A: Smarter Balanced will select participants from among those nominated by the states. A small number of alternates also will be selected. The distribution of higher education participants will represent a balance of content area, institution type, role (faculty or administrators in relevant roles such as registrar or dean of undergraduate education), and demographic diversity.

Q: When will I know if I have been selected to participate in the In-Person Panel?

A: Selected participants and alternates will receive notification regarding their status of participation by the end of June.

Q: Will my expenses be paid?

A: Smarter Balanced will pay for travel, lodging, meals, and other allowable expenses for panelists selected to participate in the In-Person Panel for setting Achievement Levels. An honorarium may be provided as applicable.

Q: How much time can I expect to spend on this project in total?

A: Each segment of the panel will occur over three consecutive days. The panel day is approximately eight hours, with lunch as well as morning and afternoon breaks. An additional day may be required for travel. We encourage participants to review information about the Smarter Balanced Achievement Level Descriptors and Common Core State Standards available on the Internet prior to participating in the panel, but we will also provide a thorough review at the panel.

Q: Where can I find information about the Smarter Balanced Assessment and the Common Core State Standards?

- A:** Smarter Balanced has developed many tools and resources to inform educators and the community of the next generation assessment system. These resources can be found on the Smarter Balanced website at <http://www.smarterbalanced.org/>.
- Q:** **What qualifications do I need to apply to participate in the In-Person Panel?**
- A:** Higher education nominees must possess all of the following qualifications:
- Currently employed at a college or university in a Smarter Balanced state;
 - Familiar with the Common Core State Standards for high school; and
 - Has strong familiarity with first-year course requirements and curricula.
- Q:** **Are recently retired faculty or administrators eligible to participate in the In-Person Panel?**
- A:** Participation in the panel is limited to individuals who are currently employed by a college or university; however, retired faculty or administrators can participate in the Online Panel. Information about Online Panel recruitment will be distributed on April 22.
- Q:** **Will participants be able to share or discuss the results of the In-Person Panel with professional peers?**
- A:** Though we will encourage participants to share with colleagues the *process* of how the In-Person Panel was conducted, we will require participants *not to divulge any of the content* or results of their work. Individual questions on the assessments are confidential since they may be used on assessments in the future. Those selected to participate will be asked to sign a confidentiality/non-disclosure agreement.

Attachment 2: Frequently Asked Questions (FAQ)

Smarter Balanced is recruiting educators from grades 3–12, administrators, parents and community/business leaders to participate in a panel for setting Achievement Levels for its assessments. This process is often referred to as “standard-setting.” This is a critical opportunity for participants to ensure that the score results from these Common Core-aligned assessments are based on rigorous, yet fair expectations for students. This FAQ will answer many of the questions that invited nominees may have about the process. For any additional questions or more information on the panel and Smarter Balanced, contact CTB_smarterbalanced21@ctb.com.

The In-Person Panel for setting Achievement Levels will take three days for each participant and involves orientation to the tests, thorough training on a well-established review procedure, and three rounds of review and discussion among participants.

Q: What is the In-Person Panel?

A: This is a critical opportunity to collect input from a diverse group of participants and establish consistent measures of student progress toward college and career readiness. The process will rely on the professional judgment and experience of educators in order to set achievement levels that are rigorous, fair and accurate.

The panel will be composed of more than 500 participants, nominated by states and selected by Smarter Balanced, to engage in a comprehensive review of questions on the assessments. The panel, which will convene between October 13–19, 2014, at a location to be determined, will give participants the opportunity to review the difficulty of questions and define what students performing at each of four Achievement Levels can reasonably be expected to answer correctly at each grade level.

Participants in the In-Person Panel will use their professional judgment and experience to make a recommendation about the minimum test scores expected for a student to be considered performing at each of the defined Achievement Levels on the Smarter Balanced assessments. The recommendations will be based on Achievement Level Descriptors written and approved by Smarter Balanced Governing States last year, the claims and targets defined for each content area and grade level, educators’ expertise in their content area and experience with students in their classrooms, and parent and community knowledge of students.

Q: Will the In-Person Panel set the final operational scores on these assessments?

A: Participants in this panel will not set the final operational scores. A “Vertical Articulation Committee” will make recommendations to the chief school officers in Smarter Balanced Governing States using a variety of information sources. These sources will include the

recommendations of the In-Person Panel, recommendations from an Online Panel which will engage in a similar – but less comprehensive – process as the In-Person Panel, and external assessment information. Acceptance of recommended scores is subject to existing approval processes within individual states.

Q: What are the dates of the In-Person Panel?

A: The panel will be held over the following dates. Selected participants will be assigned to a specific segment depending on their content area expertise, experience, and/or grade levels taught.

- High School English Language Arts/Literacy and Mathematics: October 13–15, 2014
- Grades 6–8 English Language Arts/Literacy and Mathematics: October 15–17, 2014
- Grades 3–5 English Language Arts/Literacy and Mathematics: October 17–19, 2014

Q: Where will the In-Person Panel be held?

A: The location has not yet been determined. Selected participants will be notified of the panel location once it is finalized.

Q: How are panelists selected to participate in the In-Person Panel?

A: Smarter Balanced will select educators, administrators, parents and community members from among those nominated by the states to serve as participants or alternates. The distribution of educators will represent a balance of content areas, states, grade levels, expertise, and other relevant factors.

Q: When will I know if I have been selected to participate in the In-Person Panel?

A: Selected participants and alternates nominated by the states will receive notification regarding their status of participation by the end of June.

Q: Will my expenses be paid?

A: Smarter Balanced will pay for travel, lodging, meals, and other allowable expenses for panelists selected to participate in the In-Person Panel. Substitute reimbursement or a stipend may be provided as applicable.

Q: How much time can I expect to spend on this project in total?

A: Each segment of the panel will occur over three consecutive days. The panel day is approximately eight hours, with morning, lunch, and afternoon breaks. An additional day may be required for travel. We encourage participants to review information about the Smarter Balanced Achievement Level Descriptors and Common Core State Standards available on the Internet prior to participating in the panel, but we will also provide a thorough review at the panel.

Q: Where can I find information about the Smarter Balanced Assessment and the Common Core State Standards?

A: Smarter Balanced has developed many tools and resources to inform schools and the community of the next generation assessment system. These resources can be found on the Smarter Balanced website at <http://www.smarterbalanced.org/>.

Q: What qualifications do I need to apply to participate in the In-Person Panel?

A: Educator Nominees should meet **all** of the following criteria:

- Currently certificated or licensed to teach ELA/literacy and/or mathematics in a K–12 public school;
- Currently teaching in a public school or (for the non-teacher educator category) are currently employed by a public school or district (including higher education), located within the nominating Smarter Balanced state;
- Has taught ELA/literacy and/or mathematics in grades 3–8 and/or high school within the past three years or has worked in a classroom content support role such as a literacy or mathematics coach, district or state content specialist, administrator, etc.; and
- Previously reviewed part or all of the Common Core State Standards for the content area in which they are interested in reviewing items and/or performance tasks.

We recommend that Parent and Community Nominees should meet **at least one** of the following criteria:

- Business leaders with experience hiring and managing personnel;
- Parents, especially those involved with education groups such as PTA; or
- Community leaders, including elected officials, school board members, legislators, and recognized civic leaders.

Q: Are recently retired teachers eligible to participate in the In-Person Panel?

A: Educator participation in the panel is limited to school employees who currently hold a position within an educational institution. Participation in the panel is also meant to serve as a professional development activity for current school and district employees. However, retired educators can participate in the Online Panel. Information about Online Panel recruitment will be distributed on April 22.

Q: Will participants be able to share or discuss the results of the In-Person Panel with professional peers?

- A:** Though we will encourage participants to share with colleagues the *process* of how the In-Person Panel was conducted, we will require participants *not to divulge any of the content* or results of their work. Individual questions on the assessments are confidential since they may be used on assessments in the future. Those selected to participate will be asked to sign a confidentiality/non-disclosure agreement.

Attachment 3: Sample Participant Email

Dear Colleague:

The Smarter Balanced Assessment Consortium is developing a common assessment system for K–12 Mathematics and English language arts/Literacy aligned to the Common Core State Standards to ensure all students are prepared to succeed in college and/or career. I am writing to solicit your interest in participating in a panel to recommend Achievement Levels for the Grade 11 assessments.

The purpose of the In-Person Panel is to determine common expectations for student readiness/proficiency on the Smarter Balanced assessments through a transparent, consensus-driven process with extensive input from educators and business/community leaders. Participants in the Panel will make recommendations about the minimum test scores they believe to be necessary for a student to be considered performing at each of the defined Achievement Levels on the Smarter Balanced assessments.

The process of setting Achievement Levels will rely on the expertise and professional judgment of participants in order to develop performance standards that are rigorous, fair and accurate. The recommendations will be based on Achievement Level Descriptors written and approved by Smarter Balanced Governing States last year (with extensive input from higher education faculty), the claims and targets defined for each content area and grade level, as well as educators' expertise in their content area and experience with students in their classrooms.

Participants in this panel will not set the final operational scores. A "Vertical Articulation Committee" will make recommendations to the chief school officers in Smarter Balanced Governing States using a variety of information sources. These sources will include the recommendations of the In-Person Panel, recommendations from an Online Panel which will engage in a similar – but less comprehensive – process as the In-Person Panel, and external assessment information. Acceptance of recommended scores is subject to existing approval processes within individual states.

The In-Person Panel will take place October 13–19, 2014, at a location to be determined. The first three days of this panel (October 13–15) will be devoted to establishing Achievement Level recommendations for Grade 11. The Grade 11 recommendations will then be considered as teams develop Achievement Level recommendations for earlier grades.



In-Person Panel for Setting Achievement Levels

Please respond by Friday, May 16, 2014, to confirm your interest in being nominated for this opportunity. I will then send you a link to an online registration form to complete by Friday, May 23. Smarter Balanced will select higher education representatives to serve as panelists from among the nominations submitted. The distribution of participants will represent a balance of states, content areas, institution types, and other relevant factors. You will be notified by Smarter Balanced about panel selections in June.

Smarter Balanced will pay for travel, lodging, meals, and other allowable expenses for all participants. An honorarium may be provided as applicable. Please see the attached FAQ for additional information.

Attachment 3: Sample Educator Email

Dear Educator:

The Smarter Balanced Assessment Consortium is developing a common assessment system for mathematics and English language arts/Literacy aligned to the Common Core State Standards to ensure all students are prepared for college and/or career. CTB/McGraw-Hill, on behalf of Smarter Balanced, is recruiting qualified educators and administrators to participate in an In-Person Panel for setting the Achievement Levels for the assessments. This process is often referred to as “Standard Setting”.

The purpose of the In-Person Panel is to determine common expectations for student readiness/proficiency on the Smarter Balanced assessments through a transparent, consensus-driven process with extensive input from educators. Participants in the Panel will make recommendations about the minimum test scores they believe to be necessary for a student to be considered performing at each of the defined Achievement Levels on the Smarter Balanced assessments. Input from educators with rich teaching experience will be a key factor in the setting of the Achievement Levels.

The process of setting Achievement Levels will rely on the professional judgment and experience of educators and administrators in order to develop performance standards that are rigorous, fair and accurate. The recommendations will be based on Achievement Level Descriptors written and approved by Smarter Balanced Governing States last year, the claims and targets defined for each content area and grade level, as well as educators’ expertise in their content area and experience with students.

Participants in this Panel will not set the final operational scores. A Vertical Articulation Committee will make recommendations to the chief school officers in Smarter Balanced Governing States using a variety of information sources. These sources will include the recommendations of the In-Person Panel, recommendations from an Online Panel which will engage in a similar – but less comprehensive – process as the In-Person Panel, and external assessment information. Acceptance of recommended scores is subject to existing approval processes within individual states.

The In-Person Panel to determine Achievement Levels will take place October 13–19, 2014. Educators will participate in one of three segments of the panel based on their grade level and subject area expertise. The date of each segment can be found below:

Table 1. Key Dates for In-Person Panel

Activity	Start Date	End Date
High School English Language Arts/Literacy and Mathematics	October 13	October 15
Grades 6–8 English Language Arts/Literacy and Mathematics	October 15	October 17
Grades 3–5 English Language Arts/Literacy and Mathematics	October 17	October 19

Interested candidates can apply by <insert State information>.

Smarter Balanced will select educators to serve as panelists from among the nominations submitted by Smarter Balanced states. The distribution of educators will represent a balance of states, grade levels, years of experience, and other relevant factors.

Smarter Balanced will pay for travel, lodging, meals, and other allowable expenses. Substitute reimbursement or a stipend may be provided as applicable. Please see the attached FAQ for additional information.

Attachment 4: Sample Parent/Community Email

Dear Parent/Community Member:

The [Smarter Balanced Assessment Consortium](#) is developing a common assessment system for mathematics and English language arts/Literacy aligned to the Common Core State Standards to ensure all students are prepared to succeed in college and/or career. This next generation assessment system incorporates a variety of testing methods, including written and verbal responses, and has been developed by an inclusive team of educators, researchers, parents and community members.

CTB/McGraw-Hill, on behalf of Smarter Balanced, is recruiting parents and community leaders as well as qualified educators to participate in an In-Person Panel for determining the Achievement Levels for the assessments. The purpose of the Panel is to set common expectations for student readiness/proficiency on the Smarter Balanced assessments through a transparent, consensus-driven process with extensive input from parents, business and community leaders, educators, researchers, and policymakers. Participants in the workshop will make recommendations to Smarter Balanced about the minimum test scores aligned to four defined Achievement Levels. Input from an inclusive panel (i.e. educators, researchers, parents and the community) will be a key factor in the development of these Achievement Levels.

The process of determining Achievement Levels will rely on parent and community knowledge of students in addition to the professional judgment and experience of educators in order to develop performance standards that are rigorous, fair and accurate.

Participants in the In-Person Panel will make recommendations to Smarter Balanced. Other technical and policy groups will also make recommendations. Smarter Balanced leadership and the chief state school officers of Smarter Balanced states will carefully consider this panel's recommendations in arriving at final decisions.

Committees will meet on-site for the In-Person Panel. Each panelist will participate in only one segment of the panel. The dates of each segment can be found below:

Table 1. Key Dates for In-Person Panel

Activity	Start Date	End Date
High School English Language Arts/Literacy and Mathematics	October 13	October 15
Grades 6–8 English Language Arts/Literacy and Mathematics	October 15	October 17
Grades 3–5 English Language Arts/Literacy and Mathematics	October 17	October 19

Interested candidates can apply by <Insert State Information>.

Smarter Balanced will select participants to serve as panelists from among the nominations submitted by Smarter Balanced states. The distribution of participants will represent a balance of geographic regions, grade levels, and other relevant factors.

Smarter Balanced will pay for travel, lodging, meals, and other allowable expenses for participants. A stipend may be provided as applicable. Please see the attached FAQ for additional information.

Attachment 5. Participant Application

Application for In-Person Panel to Recommend Achievement Levels	
First Name:	
Last Name:	
Primary Email Address [Work]:	
Secondary Email Address [Personal – For summer contact]:	
Contact Phone Number:	
Alternate Phone Number where you can be contacted during the summer:	
Gender: <input type="checkbox"/> Male <input type="checkbox"/> Female	
Ethnicity (X): <input type="checkbox"/> Not Hispanic or Latino	
Race (X): (Mark all that apply)	
<input type="checkbox"/> Native American or Alaska Native	<input type="checkbox"/> Asian
<input type="checkbox"/> Black or African American	<input type="checkbox"/> Native Hawaiian or Other Pacific Islander
<input type="checkbox"/> White	
State:	
Please select the Role that best fits your experience	
<input type="checkbox"/> Educator: General	<input type="checkbox"/> Parent <input type="checkbox"/> Business or Community Leader
<input type="checkbox"/> Educator: Students with disabilities experience	<input type="checkbox"/> Educator: English language learner experience
<input type="checkbox"/> Educator: Administration/Non-Teaching	
This Section Is To Be Filled Out By Educators	
Name of District or Education Entity for 2013-2014:	
Name of School for 2013-2014:	
School Type (X): <input type="checkbox"/> Urban <input type="checkbox"/> Suburban <input type="checkbox"/> Rural	
Content Area(s) Taught or Supported within the last three years:	
Grade Level(s) Taught or Supported within the last three years:	
Are you certified to teach English language learners? (X) <input type="checkbox"/> Yes <input type="checkbox"/> No	
If yes, have you taught English language learners in the last three years? (X) <input type="checkbox"/> Yes <input type="checkbox"/> No	
Are you certified to teach students with disabilities? (X) <input type="checkbox"/> Yes <input type="checkbox"/> No	
If yes, have you taught students with disabilities in the last three years? (X) <input type="checkbox"/> Yes <input type="checkbox"/> No	

Appendix E – Online Panel Documentation

- E01_Achievement Level Setting Plan 611
- E02_Achievement Level Setting Plan Supplement 629
- E03_Technical Recommendations 639
- E04_Selecting Items for the ALS 647
- E05_Bookmark Pro Description..... 650
- E06_Readiness Form 652
- E07_ Online Panel Results..... 653

Project 21 Achievement Level Setting Plan

Executive Summary

In October 2014, Measurement Incorporated (MI) and CTB/McGraw-Hill (CTB), under contract to the Smarter Balanced Assessment Consortium, will conduct an In-Person Workshop involving over 500 panelists from around the country. Thousands more will participate in an Online Panel. Afterward, 60 panelists from the In-Person Workshop will convene via webinar to conduct a Vertical Articulation. The outcome of the In-Person, Online, and Vertical Articulation activities is to recommend cut scores to Smarter Balanced states on 14 Smarter Balanced assessments. During the summer, a field test of all aspects of the three activities will be implemented. The plan described here provides a step-by-step description of the achievement level setting process that includes training, item review, data analysis, feedback, and discussion.

The Smarter Balanced approach to achievement level setting (often referred to as: “standard setting”) will be threefold: Online Panel, In-Person Workshop, and Vertical Articulation. Data from the spring 2014 field test of Smarter Balanced assessments in English language arts/literacy and mathematics in grades 3–8 and high school will be used. The method will use a Bookmark procedure, perhaps the most widely used procedure for setting achievement level cut scores on standardized tests of student achievement.

The Online Panel (October 6–17) will allow thousands of teachers, educational administrators, higher education faculty and administrators, and members of the general public to evaluate a test and enter a cut score for Level 3 (considered the college/career readiness level for high school and the on-target level for grades 3–8). Registered educators and members of the public in Smarter Balanced states will log in to take a practice test for a selected grade and content area, review content standards and Achievement Level Descriptors, and examine items for that grade/content test. They will place a bookmark to identify the Level 3 cut score.

The In-Person Workshop will occur October 13–19 and will involve 504 teachers, administrators, higher education faculty and administrators, and members of the general public. This workshop will be conducted in waves, (October 13–15 for grade 11, October 15–17 for grades 6–8, and October 17–19 for grades 3–5). Each wave will undergo rigorous training, studying the content standards, Achievement Level Descriptors, and the Bookmark procedure. They will then work through three rounds of bookmarking in small groups and discuss the results between rounds in groups of 30–36 individuals.

The Vertical Articulation Committee will meet online on October 22 and 23 for four hours each day. This group will examine all results from the previous week plus the results of the Online Panel to make final recommendations that take into account transitions from grade to grade.

Following the Vertical Articulation Committee meeting, staff of MI and CTB will meet with the Technical Advisory Committee (TAC) as well as with Smarter Balanced Staff, in preparation for a final presentation to the Smarter Balanced Governing States. At an in-person meeting, Governing State Chiefs will be asked to review and endorse the recommended cut scores that will define Smarter Balanced achievement levels for each grade and content area.

Throughout the enterprise, the contractors will receive advice and recommendations from the TAC, a Task Force, and a work group. In addition, an external auditor will evaluate all training materials and procedures and the conduct of the In-Person Workshop and Vertical Articulation Committee meeting.

Overview

A three-pronged approach to achievement level setting will include an Online Panel, an In-Person Workshop, and a Vertical Articulation Committee. Thousands of teachers, educational administrators, higher education faculty and administrators, and members of the general public will be invited to evaluate a test and enter a cut score for Level 3. Individuals participating in the Online Panel will typically take up to three hours to complete the activity. The In-Person Workshop will be much more involved—panelists will work for 2.5 days in groups of 30–36 people. They will study the Common Core State Standards and Achievement Level Descriptors for their particular grade and subject, learn how to apply the Bookmark procedure, and then work through three rounds of item rating and discussion with their fellow panelists. The Vertical Articulation Committee will meet via webinar to review all results of the Online Panel and In-Person Workshop. Whereas panelists for the Online Panel and In-Person Workshop will examine one test at one grade level, the panelists of the Vertical Articulation Committee will review all cut scores for all grades for one subject, making sure that cut scores do not result in wild swings in percentages of students scoring at each level from one grade to the next, and will consider the reasonableness of the system of cut scores across grades.

Item Mapping

A Bookmark procedure (Lewis, Mitzel, Mercado, & Schulz, 2012) will be used, including item maps, sometimes referred to as construct maps. Item maps are graphic organizers that show at a glance the arrangement of items in an ordered item booklet (OIB) with the statistical and other information about each item next to it, such as shown in Figure 1.

Figure 1. Sample Item Map

Online Standard Setting Welcome [Sample User!](#) 

[Home](#) [Instructions](#) [Item Map](#)

Round: 2 Math / Grade: 6
 Round 1 Round 2
 Submitted: 9/1/2014 10:00 AM

Page	RP67	Scaled Score	CCSS	DOK	Reviewed	Comment	Data Field 1	Data Field 2	Cut
1	-.477	213	MA.6.EE.7:1	1	<input checked="" type="checkbox"/>				
2	-.471	221	MA.7.NS.3:1	2	<input checked="" type="checkbox"/>				
3	-.463	229	MA.6.EE.7:2	1	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			
4	-.452	235	MA.6.EE.6:1	1	<input checked="" type="checkbox"/>				
5	-.438	241					View		Cut 1
6	-.420	257	MA.6.EE.6:1	2					
7	-.416	268	MA.6.EE.6:1	1					
8	-.404	277	MA.6.EE.6:1	1					
9	-.392	283					View		

This particular example shows an item map with page numbers 1–9 showing (there are many more), along with statistical and other information about the item on each page. The RP67 value (second column) refers to the ability level needed in order to have about a 2/3 chance of answering the item

correctly¹. Scaled Score refers to a score associated with that particular ability level. CCSS refers to the specific Common Core State Standards content standard measured by the item, and DOK refers to the cognitive demand (or “depth of knowledge”) that the item requires of the student. The next four columns are for user comments and other data, some of which may not be directly associated with an item in the test. For example, on page 5, there is no entry in the CCSS column, indicating that there is no item on the page or that the location on the Scale Score refers to an external data source, such as one of the performance levels on NAEP or PISA. The final column allows the user to enter a cut score. In this example, the user has entered Cut 1 on page 5, indicating that an ability level of $-.438$ or a scaled score of 241 is the minimum amount of ability required for a student to make the first cut (in the case of a Smarter Balanced test, the cut between Level 1 and Level 2).

In August MI and CTB will receive field test data from Educational Testing Service (ETS), one of the Smarter Balanced contractors. MI and CTB will use the field test data to calculate the difficulty of each item in each field test and then arrange those items from least to most difficult. They will construct new test booklets using test blueprints similar to those used to construct operational tests, but will place the items in order of increasing difficulty. For each such OIB, a map such as the one shown above will be constructed.

The Bookmark Procedure

The bookmark procedure is so named because panelists identify cut scores by entering markers in a specially designed test booklet. The test booklet consists of a set of items placed in difficulty order, easiest items first and hardest items last. In between, selected-response (SR) and constructed-response (CR) items are intermingled in order of their difficulty. Each CR item appears several times in the booklet, once for each of its score points. For a given CR entry, the item prompt and the rubric for a particular score point appear, along with sample responses illustrating that score point. The method has become quite popular because of its ability to present SR and CR items at the same time and because of its use of item response theory (IRT) analyses.

The difficulty-ordered booklet can be composed of any collection of items spanning the range of content, item types, and difficulty represented in a typical test, and need not consist only of items that have appeared in an intact test. This booklet can have more items or fewer items than a regular test booklet. As noted above, MI and CTB will use items from the spring 2014 field test. There will be hundreds of items for each subject/grade combination from which to choose, so they will be able to select those that offer the best combination of content coverage and statistical properties.

Each page of the OIB has one item and its information, as shown in Figure 2. On this particular page, the item appears on the left side of the page and the statistical and other information on the right. This information is identical to that found in the item map. As they review the OIBs, panelists have two opportunities to review the data, by itself on the item page, and in the context of other items on the item map.

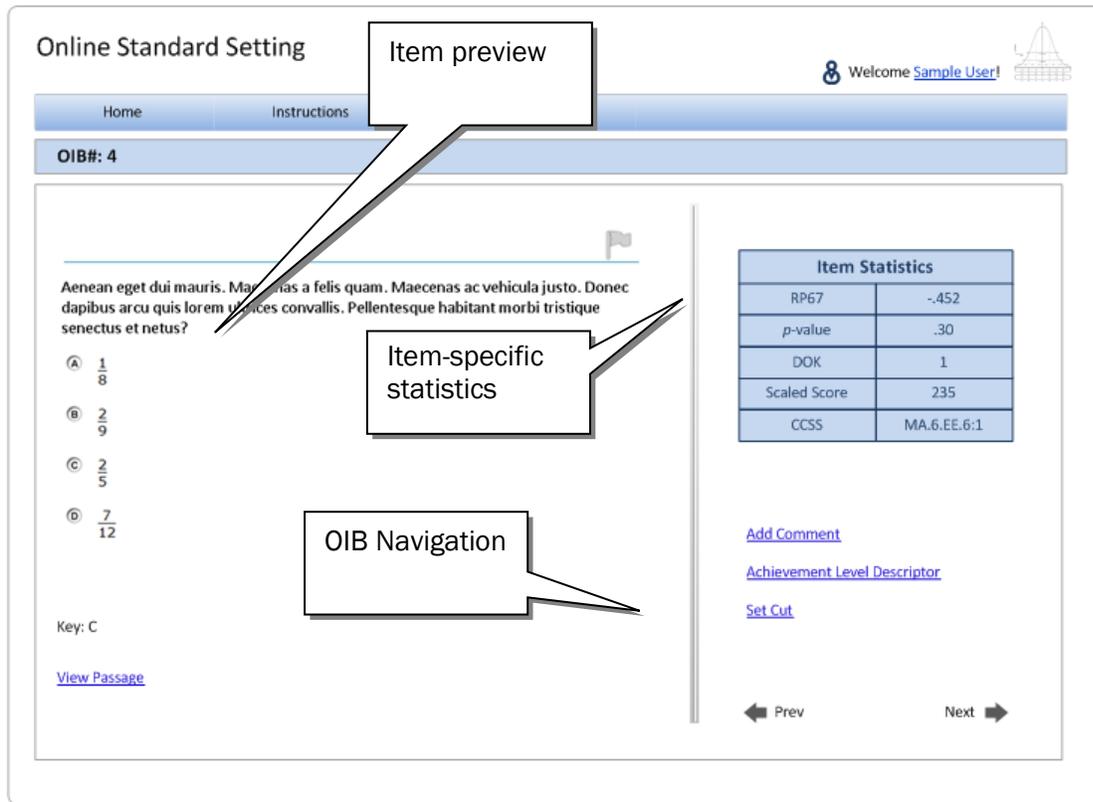
Central to the notion of the Bookmark procedure, as with all achievement level setting procedures that involve the rating of items, is the notion of “just barely.” A cut score marks the beginning of a level, not the average performance of all students at that level. Thus, panelists who are looking for the cut that demarks Level 1 from Level 2 (referred to as the “cut score for Level 2”) will consider students who just barely meet the qualifications listed in the Achievement Level Descriptor for Level 2. As panelists work through the OIB, they ask themselves a set of questions about the items on each page:

¹ We use RP67 (response probability = 67%) throughout this plan; however the exact value of RP is a subject for discussion with the Smarter Balanced Technical Advisory Committee and will be finalized well in advance of creation of ordered item booklets.

- What knowledge or skill is required to earn this point?
- What makes later items more difficult than earlier items?
- Think of a large group of students at the cut score for this level (i.e., just barely at this level). Would at least 2/3 of them earn this point?

Starting on page 1, and thinking about the requirements of Level 2, it is usually very easy to answer Yes to this final question. As a panelist moves on through the OIB and items become more challenging, it becomes more difficult to answer Yes. Eventually, the panelist must answer No. The ability level listed on the last page on which the panelist could answer Yes becomes the de facto cut score because that is the ability level associated with a 2/3 chance of answering the item correctly.

Figure 2. Sample OIB Page



The screenshot shows the Online Standard Setting (OIB) interface. At the top, there are navigation links for 'Home' and 'Instructions', and a 'Welcome Sample User!' message. The OIB number is 4. The main content area displays an item preview with a passage of Latin text and four multiple-choice options (A, B, C, D). The key is C. To the right, there is an 'Item Statistics' table with the following data:

Item Statistics	
RP67	-.452
p-value	.30
DOK	1
Scaled Score	235
CCSS	MA.6.EE.6:1

Below the table are links for 'Add Comment', 'Achievement Level Descriptor', and 'Set Cut'. At the bottom, there are 'Prev' and 'Next' navigation buttons. Callouts in the image identify the 'Item preview' (the question and options), 'Item-specific statistics' (the table), and 'OIB Navigation' (the navigation buttons).

The Bookmark procedure is based on the fact that for tests constructed in accordance with item response theory, the difficulty indices of test items and the ability indices of students can be laid out on a common scale. The probability of a given student answering a given item correctly can be expressed as a relatively straightforward function of the student's ability and the item's difficulty. Once field tests have been scored, MI and CTB will know the difficulty of each item and the ability estimate for each student whose field test responses were used in the analysis. MI and CTB will use that information to calculate RP values for each item and then order those items in terms of their RP values.

They will be using the same OIBs for both the Online Panel and the In-Person Workshop. This approach is being taken to minimize the number of items in the item pool that will be exposed during achievement level setting. Given the public nature of the Online Panel in particular, Smarter Balanced assumes that all items used in achievement level setting are effectively released to the public.

Bookmark software. Figures 1 and 2 illustrate a computer-based approach to the Bookmark procedure. MI is creating the software to translate paper-based Bookmark materials (item maps and OIBs) they have used over the past several years into computer-based products, and will provide computers for In-Person panelists to use. One advantage of this approach is that it allows additional materials to be linked together, such as reading passages, ALDs, content standards, and external data directly to the item map and/or OIB. Users of the system will have all necessary support documents and data at their fingertips throughout the achievement level setting activity. The software, which will support both the Online Panel and In-Person Workshop, will be ready for testing in late July to early August.

The Online Panel

MI and CTB, along with Hager Sharp, a communications subcontractor, have begun recruiting participants for the Online Panel through K-12 Leads, Higher Education Leads, state Teacher Involvement Coordinators (TICs), and others. They are using a variety of communication tools to reach an extremely wide audience and hope eventually to have thousands of participants.

As they register for the event, each Online Panelist receives a two-day window sometime between October 6 and October 17 to log on, review a test, and enter a bookmark. They will also be encouraged to review the content standards and Achievement Level Descriptors for that test and participate in a self-paced introduction to the Bookmark procedure. Each Online Panelist will enter a bookmark for Level 3 only. At the end of the activity, all Online Panelists will be asked to complete a short evaluation of the experience and their confidence in the appropriateness of the cut score they entered.

Pilot test. MI will conduct a pilot test of the Online Panel in August and September, the objectives of which will be to test the loadbearing capacity of the system (thousands of simultaneous or nearly simultaneous users could be expected), the clarity of the directions, and the ease of navigation of the item map and OIB. They will test loadbearing capacity through in-house simulation programs. To test clarity and ease of navigation, they will use groups of readers at one or more of their scoring centers located throughout the United States. Specifically, they will invite readers matching a demographic profile (i.e., some with teaching experience, some without) to log on, complete the exercise, enter a bookmark, and report on their experience to scoring center managers. They will modify training materials and procedures as necessary, in consultation with the external auditor and Smarter Balanced staff.

The In-Person Workshop

The In-Person Workshop will take place in a single centrally located facility. There will be three waves of panels from October 13 through October 19. The first wave, grade 11, will begin on Monday morning, October 13 and work through noon October 15. The second wave, grades 6–8, will begin on Wednesday morning, October 15 and work through noon October 17. The final wave, grades 3–5, will begin on Friday morning, October 17 and work through noon October 19. Table 1 summarizes the numbers of panelists by subject and grade. The Appendix contains a detailed agenda for each day of the workshop.

Table 1. In-Person Workshop Panelists by Subject and Grade

Grade	English Language Arts/Literacy	Mathematics
3	1 panel of 30	1 panel of 30
4	1 panel of 30	1 panel of 30
5	1 panel of 30	1 panel of 30
6	1 panel of 30	1 panel of 30
7	1 panel of 30	1 panel of 30
8	1 panel of 30	1 panel of 30
High School	2 panels of 36	2 panels of 36
Total	252	252

Recruitment and selection of panelists. Recruitment of panelists for the In-Person Workshop began April 15. State K-12 Leads, Higher Education Leads, and Teacher Involvement Coordinators received communication tools developed by the contractor and approved by Smarter Balanced to enable them to recruit teachers (general as well as teachers of English language learners and students with disabilities), school administrators, higher education faculty, business and community leaders, and the general public. Each Smarter Balanced state will have 20–25 positions to fill, and each state will have at least one representative for each of the 14 tests.

Preparation of materials. The success of any achievement level setting activity is dependent on good training materials. Staff of MI and CTB will prepare the following training materials:

- Introductory PowerPoint presentation to orient panelists to the goals and tasks of the workshop
- Common Core State Standards – up-to-date versions of the subject/grade-specific content standards as well as guidelines to their use in the achievement level setting activity
- Achievement Level Descriptors – up-to-date versions of the ALDs for the specific subject and grade for each panel
- Practice Test – using the latest version of the Smarter Balanced practice tests for each grade and subject
- Orientation to the Bookmark procedure – PowerPoint presentation designed to show panelists how Bookmark works and specifically how panelists are to implement the procedure in a computer-based environment
- Bookmark Practice Round – an exercise involving 6–8 items that panelists review prior to entering a bookmark and discussing their placements in a large-group setting.
- Readiness Form – a multipart form that asks panelists at several key points during the process how well they understand the process they are implementing and how ready they are to proceed to the next step
- Evaluation Form – a series of statements about the training, environment, and conduct of the workshop that the panelist responds to on a graded scale (such as Strongly Agree to Strongly Disagree)

MI and CTB staff will draft all training materials and submit them to Smarter Balanced staff and the external auditor for review in advance of the workshop. Final versions of all training materials will reflect the comments and recommendations of these reviews and be approved by Smarter Balanced leadership prior to use.

Orientation and training. Using the training materials approved by Smarter Balanced, MI and CTB staff will provide large-group and small-group training. For the opening session, they have enlisted Dr. Joe Willhoft to give the welcome and charge. Specific training on the content standards, ALDs, and practice tests will follow in order, provided by various MI and CTB staff. Dr. Michael Bunch, lead facilitator, will provide the orientation to the Bookmark procedure. At the end of each training session, panelists will complete a portion of the Readiness Form.

In-Person Workshop panelists will be encouraged to review the appropriate ALDs and CCSS standards prior to coming to the workshop. However, it will not be assumed that all have done so and panelists will be provided an opportunity not only to review the materials on site but to discuss them in a large-group setting. They will have an opportunity to indicate on the Readiness Form just how familiar they are with those materials. No panelist will be permitted to advance to item review without indicating familiarity with the ALDs and content standards and indicating readiness to proceed.

Following the orientation to the Bookmark procedure, panelists will break into their small groups to gain first-hand experience in setting a bookmark through a practice exercise. This exercise will consist of a 6-item OIB with items of varying difficulty. Each panel will have access to one or two small-group facilitators who will orient panelists to the computers and software and show them how to navigate the practice OIB. Panelists will then have several minutes to review the six items and enter a bookmark. The facilitator will then lead a discussion focusing on how many panelists chose each item to place their bookmarks. Following this discussion, panelists will once more complete a section of their Readiness Forms, indicating their readiness to begin Round 1.

Round-by-round item review and discussion. General orientation to the Bookmark procedure will take place in the large-group meeting area. After that orientation, panelists will break into their respective rooms to begin Round 1. As panelists begin Round 1, they will be seated near five other panelists—their table. Panelists will be invited to work through their on-screen OIBs and discuss the items with others at their table. The purpose of the first “live” engagement will be to allow panelists to place a single bookmark for Level 2. Panelists can discuss their opinions with one another at their table as much as they like, but when they enter a bookmark, it will be their bookmark, not that of the table.

After placing one bookmark, panelists will be dismissed for the day. They will resume Round 1 the next morning. The morning activity will begin with a review and discussion of the previous afternoon’s activity and placement of the Level 2 bookmark. Facilitators will conclude the discussion and direct panelists to complete Round 1, placing bookmarks for Levels 3 and 4. There is no need to place a bookmark for Level 1, since the full range of Level 1 will be defined as soon as the bookmark for Level 2 is placed.

Panelists will work through the morning entering bookmarks for Levels 3 and 4, discussing their rationales and ideas with others at their table as they wish. Once again, panelists will enter their own bookmarks, submit them, and log out of the session. Round 1 ends when the last panelist’s three bookmarks have been submitted.

Round 2 begins the afternoon of the second day. Panelists return from lunch, log back in, and have an opportunity to see the results of Round 1. The facilitator will present the results and initiate a discussion. This discussion will focus on the range of page numbers for each bookmark and the rationales for their placement. All panelists will be encouraged to contribute to this room-wide

discussion. The facilitator will then draw panelists’ attention to impact data—percentages of students who would be at each level, based on the Round 1 cut scores—and discuss them.

After this discussion, panelists will complete the next section of the Readiness Form and begin Round 2. They will proceed as in Round 1, conferring with others at their table but entering their own bookmarks. When they enter three bookmarks and submit them, they are free to log out for the day.

Panelists return the morning of the third day to see the results of Round 2. Once again, the facilitator leads a discussion of the range of bookmark placements, corresponding cut scores, and percentages of students classified at each level, based on the Round 2 cut scores. Following this room-wide discussion, panelists complete the final section of their Readiness Forms and begin Round 3.

Data analysis and reporting. As panelists enter bookmarks and submit them, the data flow directly from their computers to servers MI will set up prior to the start of the workshop. Staff from CTB, using BookmarkPro software, will receive the data, analyze them, and produce reports that facilitators will share at the beginning of the next round. Figures 3–5 illustrate some of the reports generated by BookmarkPro.

Figure 3. Sample Bookmark Placement Report

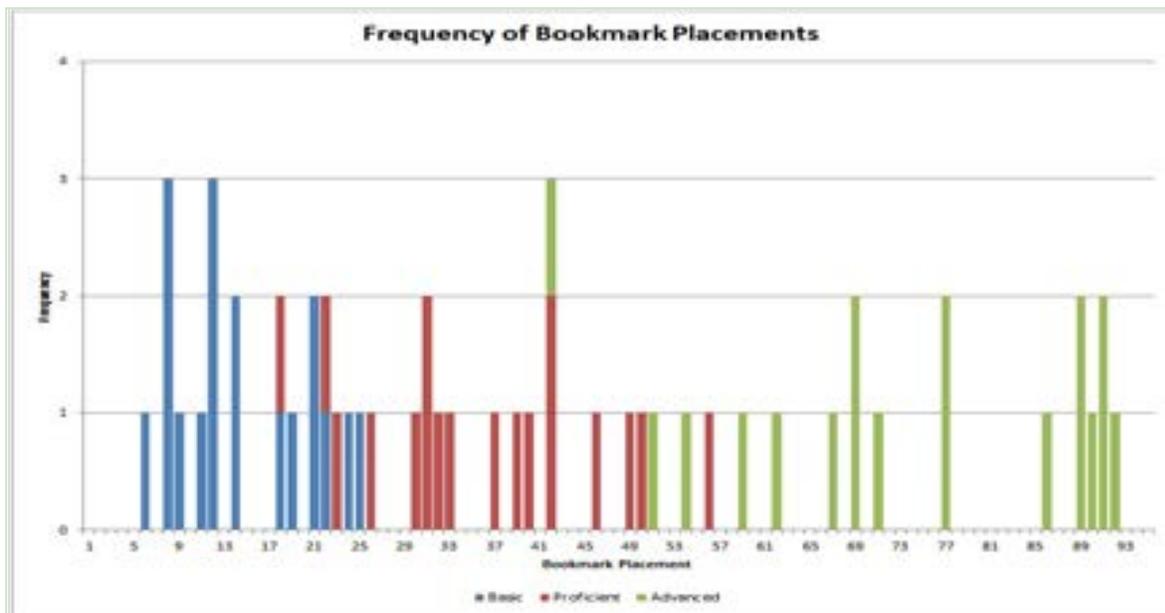


Figure 4. Sample Cut Score Dispersion Report

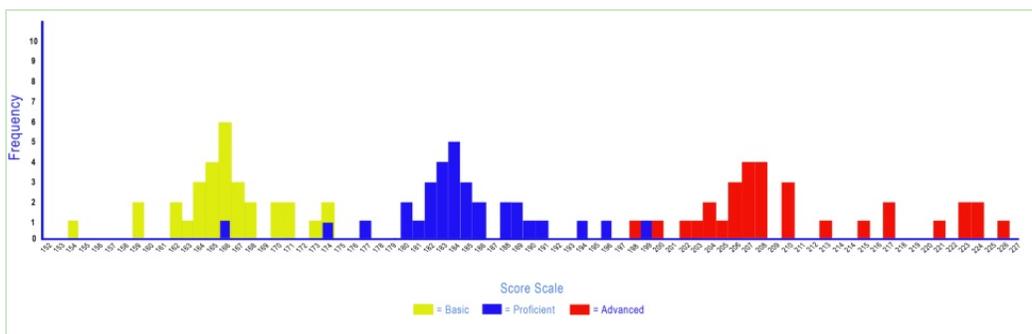
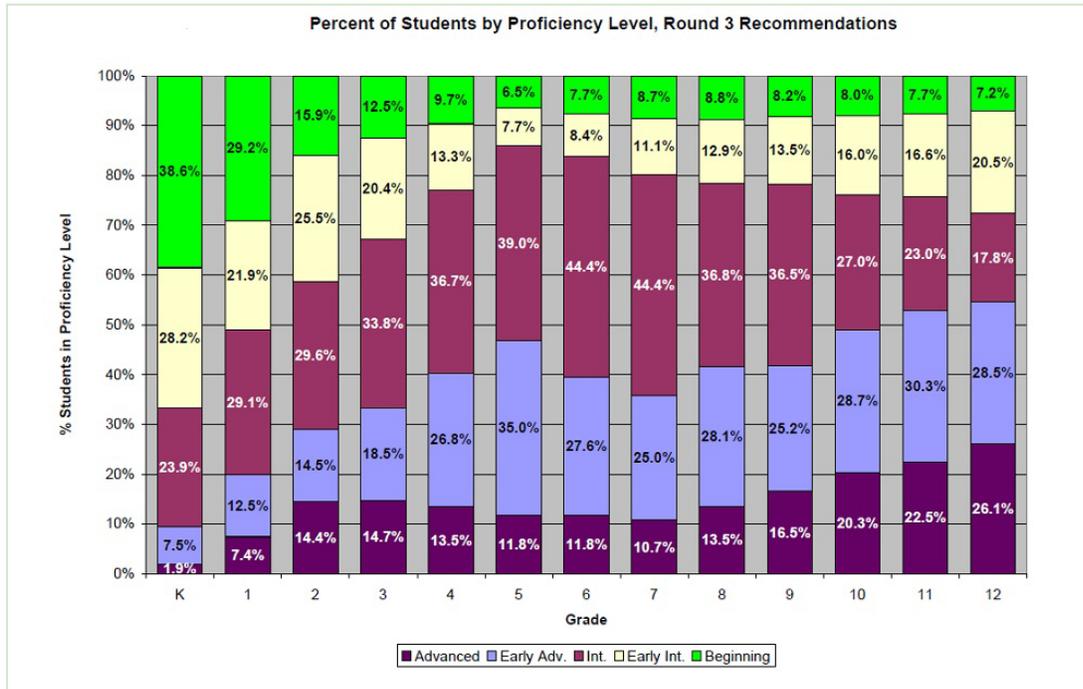


Figure 5. Sample Across-Grade Impact Data



Pilot test. MI will pilot test the software and training materials in August, employing readers at one of its scoring facilities. The MI plan is to engage 40 readers, some with teaching experience and some without, in four panels (grades 4, 6, 8 and high school) of 10 for one subject. The purpose of the pilot test will be to test the adequacy of training materials as well as the ease of navigation of the software in a more demanding setting than that required for the Online Panel (i.e., multiple cut scores, revisions, and presentation of complex reports after each round).

The Vertical Articulation Committee

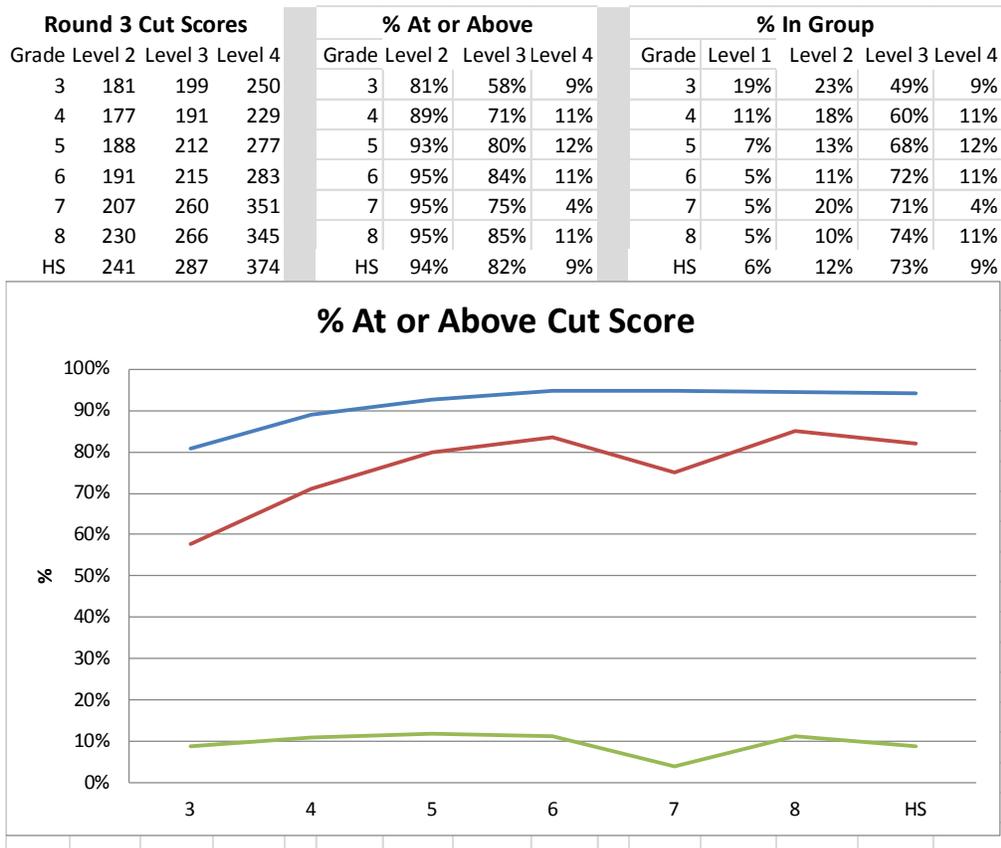
MI and CTB will conduct a vertical-articulation activity for the English language arts/literacy and mathematics tests across Grades 3–8 and 11. They will invite a total of 60 panelists from across the grades from the In-Person Workshop to participate, 30 from English language arts/literacy and 30 from mathematics for each of two vertical articulation committees. There will be four members from each grade-level’s In-Person panel for grades 3–8 and six members from the grade 11 panel. Members of the English language arts/literacy and mathematics committees will be identified during the In-Person Workshop, and will have a composition similar to the grade-level In-Person panels.

The purpose of vertical articulation is to produce a set of cut scores that yield plausible impacts across grades. For example, if recommended cut scores would make 75 percent of Grade 4 students at Level 3 or better but only 50 percent of Grade 5 students as Level 3 or better, and if there were no plausible explanation for why this condition should be, it would be necessary to reexamine both sets of cut scores in the context of realistic expectations at the two grades. Having vertical articulation is often found to be particularly useful because the recommendations from the In-Person panels are each developed in isolation; the articulation committees are able to look across all grades to consider the reasonableness of the all recommendations when considered together.

Figure 6 shows a typical articulation display for a VAC meeting. This single-view spreadsheet/graphic is used to help panelists arrive at moderated cut scores efficiently and fairly, employing simple majority voting for each recommended cut-score change.

For presentations such as the one shown in Figure 4, the data tables and graph would be linked together so that a change in any cut score automatically updates both the impact table and the graph. Thus, a panelist suggesting a change in any cut score will see the effect immediately. Committee members will be reminded that the objective of the exercise is not necessarily to make the lines straighter or “prettier” but to arrive at cut scores the panelists can explain to policy makers and that policy makers can explain to the public.

Figure 6. Sample Vertical Articulation Data Presentation



In this example, panelists may want to examine the Level 3 cut score for grade 7. While the general trend for proficiency is flat over time, grade 7 stands out. Performances of grade 6 and 7 students align with those of other grades, but the “dip” at grade 7 is difficult to explain. In a situation like this, the anomalies would be pointed out, and explanations would be asked for. Occasionally, an explanation makes sense to all the panelists, and there’s no need to have additional discussion. If the anomaly cannot be explained, panelists for the affected and adjacent grades might be asked to discuss their expectations and interpretations of the ALDs. After some discussion, the committee is asked for a recommendation for a cut-score change, following Robert’s Rules of Order. If there is a motion for a change, the recommended change is implemented into the tables and graph for review. After discussion a vote is taken and recorded.

The articulation committees will be conducted as online meetings. In this virtual setting, each panelist will have access to the data, as well as notes from three rounds of in-person achievement level setting and commentary provided by facilitators of the in-person achievement level setting. Chat function with other VAC panelists will also be available. Voting in a virtual environment will be by mouse click. The plan, therefore, is to conduct a live, real-time virtual articulation event, the details of which are spelled out below.

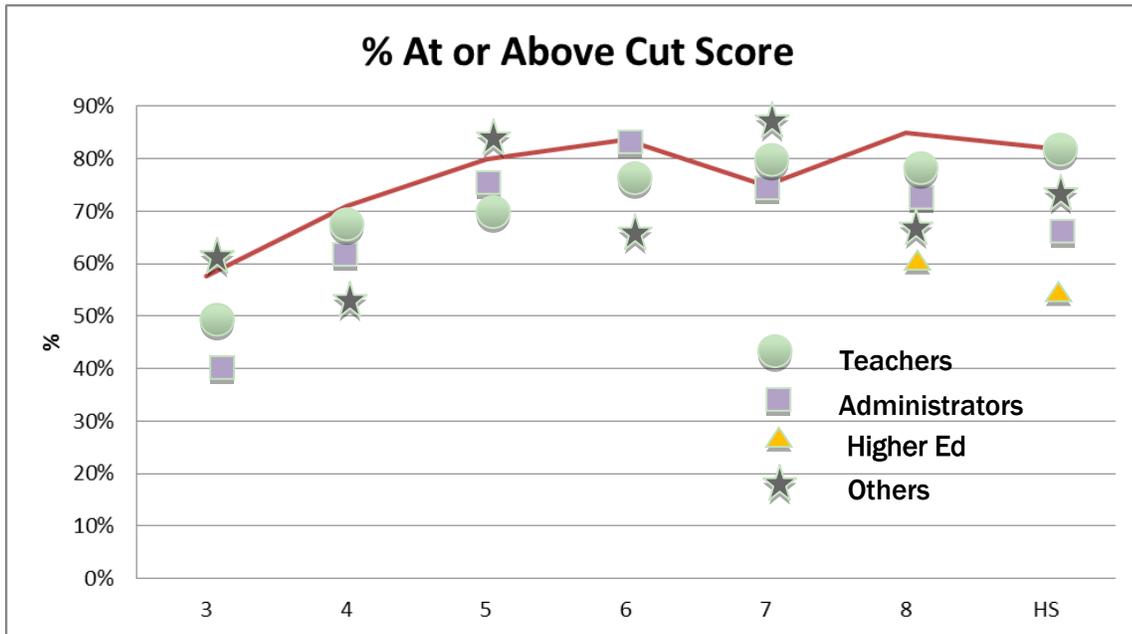
All VAC panelists will be selected from the various In-Person achievement level setting panelists. By virtue of that, they will have received training and should be thoroughly knowledgeable about the tests, the ALDs, the CCSS, and the achievement level setting methodology. All that will remain for successful completion of this task will be an understanding of the purpose and nature of vertical articulation and orientation to the ground rules by which it will be conducted. MI staff will prepare a PowerPoint presentation to orient panelists to the process and present it at the outset of the meeting. Panelists will have an opportunity to ask questions and receive answers from experts. Drs. Bunch and Lewis will conduct the sessions. Both have extensive experience in leading such groups.

The VAC will take place over two days, four hours per session. The schedule has been designed to accommodate panelists in multiple time zones. On the first day, panelists will focus on the Level 3 cut scores across grades, and move on to Level 2 if time permits. On the second day, panelists will review results of the first day prior to reviewing additional cut scores. Both panels will have access not only to their own results of the first day but to those of the other panel as well. A complete schedule is included in the Appendix.

VAC panelists will complete evaluation forms in much the same way that Online Panel and In-Person achievement level setting panelists complete them, following training and between rounds. The evaluation forms will be constructed in the same way as previously constructed (i.e., by focusing on the critical elements of the session).

Use of Online Panel data. The VACs will be conducted on October 22 and 23. The Online Panel will conclude on October 17. MI will collect data from the Online Panel and calculate the mean and interquartile ranges of the cut score for Level 3, not just for all the thousands of individuals who have logged in but also by subgroup: teachers, administrators, higher education staff, and others. For each grade, these means/ranges on the Round 3 cut scores will be superimposed onto results from the In-Person Workshop so that Figure 6 becomes Figure 7.

Figure 7. VAC Presentation with Online Panel Data



Pilot testing. The pilot test of the online VAC will begin immediately upon completion of the pilot test of the In-Person Workshop. The focus of the pilot test will be the clarity of the directions, the timing of the presentation of cut scores and alternatives, and the overall flow of the process. Twelve of the 40 In-Person Workshop pilot panelists will be selected to participate in a two-day webinar following the agenda shown in the Appendix. They will view cut scores for grades 4, 6, 8, and high school and recommend changes, following the procedures described above. Dr. Bunch will conduct the webinar, collect the responses, and then discuss the process with the panelists. He will share these responses with Smarter Balanced staff and the external auditor and modify the process as necessary in advance of the live VAC in October.

Reference

Lewis, D.M., Miztel, H.C., Mercado, R.L., & Schulz, E.M. (2012). The Bookmark Standard Setting Procedure. In G.J. Cizek (Ed.), *Setting Performance Standards: Concepts, Methods, and Perspectives* (2nd Ed.). New York: Routledge.

Appendix – Agendas

Agenda for In-Person Workshop	
Day 1	
Time	Event
7:30 a.m.	Registration opens (Ballroom Lobby) <ul style="list-style-type: none"> • Badges, place cards, and room and table assignments • Continental breakfast
8:30 a.m.	Opening session (Carpenter Ballroom) <ul style="list-style-type: none"> • Keynote address and charge • Overview of schedule of events for the In-Person Panel • Housekeeping matters
9:15 a.m.	Review of CCSS and ALDs (Carpenter Ballroom) <ul style="list-style-type: none"> • Review of materials sent in advance • Discussion of CCSS and ALDs • Completion of Readiness Form Part 1
10:15 a.m.	Break (adjourn to breakout rooms)
10:30 a.m.	Introduction to the test <ul style="list-style-type: none"> • Overview of test development and scoring • Practice Test administration • Table-wide discussion of tests • Closing comments on test construction • Completion of Readiness Form Part 2
12:30 p.m.	Lunch (Ballroom Lobby)
1:30 p.m.	Introduction to the bookmark procedure (Carpenter Ballroom) <ul style="list-style-type: none"> • PowerPoint presentation • Practice round • Questions and answers
3:15 p.m.	Completion of Readiness Form Part 3
3:30 p.m.	Adjourn to breakout rooms
3:45 p.m.	Begin Round 1 <ul style="list-style-type: none"> • Panelists work in 6-person groups to review OIB and item map, discuss possible bookmarks, and enter bookmarks individually
4:45 p.m.	Wrap-up for Day 1

Agenda for In-Person Workshop	
	<ul style="list-style-type: none"> Panelists log out; computers secured
5:00 p.m.	Adjourn
5:15 p.m.	Smarter Balanced, MI, CTB, and auditor debriefing
6:00 p.m.	Secure meeting rooms and depart
Day 2	
7:30 a.m.	Registration opens (Ballroom Lobby) <ul style="list-style-type: none"> Badges, place cards, and room and table assignments Continental breakfast
8:30 a.m.	Convene in breakout rooms <ul style="list-style-type: none"> Review Round 1 process Questions and answers on work thus far for Round 1
9:15 a.m.	Resume Round 1 as on previous afternoon
11:30 a.m.	Lunch (Ballroom Lobby) <ul style="list-style-type: none"> MI/CTB staff analyze data and prepare reports while panelists eat lunch. Confer with Smarter Balanced staff to approve Round 1 analyses and reports Identify Vertical Articulation Committee panelists
1:00 p.m.	Review of Round 1 <ul style="list-style-type: none"> Grade group facilitators review results of Round 1; lead discussion of bookmark placements, impact data, and rationales for bookmark placements
2:30 p.m.	Completion of Readiness Form Part 4
2:45 p.m.	Begin Round 2 <ul style="list-style-type: none"> Panelists work in 6-person groups as in Round 1
4:45 p.m.	Wrap-up for Day 2 <ul style="list-style-type: none"> Panelists log out; computers secured
5:00 p.m.	Adjourn <ul style="list-style-type: none"> MI/CTB staff analyze data and prepare reports. Meet with Vertical Articulation Committee panelists.
5:30 p.m.	Smarter Balanced, MI, CTB, and auditor debriefing
6:00 p.m.	Secure meeting rooms and depart
Day 3	
7:30 a.m.	Registration opens (Ballroom Lobby) <ul style="list-style-type: none"> Badges, place cards, and room and table assignments

Agenda for In-Person Workshop	
	<ul style="list-style-type: none"> Continental breakfast
8:30 a.m.	Convene in breakout rooms <ul style="list-style-type: none"> Review Round 2 results; grade group facilitators review results of Round 1; lead discussion of bookmark placements, impact data, and rationales for bookmark placements
9:45 a.m.	Completion of Readiness Form Part 5
10:00 a.m.	Begin Round 3 <ul style="list-style-type: none"> Panelists work in 6-person groups as in Rounds 1 and 2.
11:30 a.m.	Wrap-up for Day 3 <ul style="list-style-type: none"> MI/CTB staff process and share Round 3 results for this group plus others that have completed previously (not applicable to Grade 11; grade 6–8 groups will see only their results and those for Grade 11; grade 3–5 groups may see all results).
11:45 a.m.	Completion of Evaluation Forms; log out; all computers secured
Noon	Adjourn for lunch (Ballroom Lobby) <ul style="list-style-type: none"> Panelists may stay for lunch or leave if they have successfully logged out. MI/CTB staff process Round 3 data confer with Smarter Balanced staff and auditor, update achievement level setting report.

Agenda for Vertical Articulation Committee	
October 22	[All times are eastern.]
1:00 p.m.	Welcome and introductions
1:15 p.m.	Orientation to VAC (PowerPoint presentation)
2:00 p.m.	Questions and Answers
2:45 p.m.	Completion of VAC Readiness Form Round 1
3:00 p.m.	Presentation of Round 3 cut scores and impact data
3:15 p.m.	Discussion of Level 3 (CCR) cut across grades with Online Panel data (using voting feature of WebEx)
3:45 p.m.	<ul style="list-style-type: none"> • Formal voting on cut score changes for Level 3 • Panelist recommends change; change receives a second (message to whole group) • Facilitator opens floor for comments • Facilitator calls for the vote • Panelists enter Yes/No vote on specific motion on the floor • Facilitator announces results of vote • Facilitator opens the floor for motions for more changes as above • Facilitator calls for motion to approve all changes for Level 3 • Panelist moves for approval; another seconds • Facilitator calls for discussion, then vote on approval of changes
5:00 p.m.	Adjourn Day 1
October 23	
1:00 p.m.	Review of Day 1; completion of VAC Readiness Form Round 2
1:15 p.m.	Presentation of Round 3 cut scores and impact data updated from Day 1
2:00 p.m.	Discussion of Level 2 cut scores across grades

Agenda for Vertical Articulation Committee	
2:30 p.m.	Formal voting on cut score changes for Level 2 as above
3:30 p.m.	Completion of VAC Readiness Form Round 3
3:45 p.m.	Discussion of Level 4 cut scores across grades
4:00 p.m.	Formal voting on cut score changes for Level 4 as above
5:00 p.m.	Adjourn

Project 21 Achievement Level Setting Plan Supplement

The Achievement Level Setting Plan presented to the Chiefs on April 29, 2014, was approved unanimously on April 30, 2014. However, during the presentation of the Plan, members of the Technical Advisory Committee (TAC) and the Chiefs themselves requested written clarification or amplification of certain elements of the plan. The purpose of this Supplement is to document responses to questions raised on April 29 and 30 and to provide further amplification of some of those responses. Key elements addressed in this Supplement include the following:

- Development of Ordered Item Booklets
- Development and Presentation of Item Maps
- Orientation to and Use of External Data
- Use of Practice Test in Panelist Training
- Demonstration of Universal Tools, Designated Supports, and Accommodations Available During Training
- Presentation of Online Panel Recommendations
- Establishment of a Score Scale

Development of Ordered Item Booklets

Basic assumptions. Development of ordered item booklets (OIBs) begins and ends with the operational blueprints (approved April 30) for each subject and grade. MI and CTB staff will use calibrated, field-tested items to fill cells of the blueprints, such that the end product will have a rectangular distribution of item difficulty and content characteristics (target, claim, and item type distribution) of the blueprint. Specific assumptions underlying the assembly of these items are provided below.

- Up to 75 score elements will populate each ordered item book per grade/content area.
- Item difficulty will be uniformly and broadly distributed (i.e., there will be a rectangular distribution of IRT difficulty indices that includes low, medium, and high values). However, the distribution might be established to reflect the pool of items formed by a typical CAT administration for a fixed number of students of varying abilities (e.g., low, moderate, and high).
- Whenever two or more items are available for a given content standard, the item with an IRT difficulty needed to fill a gap or avoid a duplication of difficulty indices will be preferred.
- OIB developers will select items to support the broadest *practical* range of item difficulty.
 - Easier items that allow bookmark placement supporting a cut score at, for example, the 10th percentile (to support the lowest cut score)
 - Difficult items that allow a bookmark placement supporting a cut score at, for example, the 95th percentile (to support the highest cut score)
- The response probability (RP) used to locate items on the scale will be determined after an analysis is conducted. For instance, RP67 may not provide sufficient floor for difficult tests. An analysis will be conducted after item parameters and student ability frequency distributions are available, and recommendations will be jointly provided by MI and CTB, with input from the TAC and approval by Smarter Balanced.

Items tied to passages. For English language arts/literacy tests, many items are passage based. MI/CTB will pay special attention to these passages to produce OIBs that include a reasonable number of items per passage. Specific steps are outlined below.

- Select passages and include at least the number of items that a student would be expected to respond to on an operational CAT. More items than might be experienced in a typical CAT administration can be selected for the OIB if there is a strategic reason to do so, such as to fill in scale score gaps. This assumes that there are more items written to a passage than would be typically provided to a given student.
- Select passages with items that span a broad range of difficulty.
- Select passages that represent the full range of passage types for that grade.

Operational approach. Steps for assembling each OIB are outlined below.

1. Place all available items in the item map template.
2. Integrate impact data on item maps to see whether there are enough on-grade item pools at low, medium, and high degrees of difficulty. If not, make recommendation to Smarter Balanced leadership regarding the use of off-grade items for OIB construction. If approved for use by Smarter Balanced, proceed with the next step, and select off-grade items to support needs after exhausting on-grade item pools.
3. Remove redundant items.
 - a. Remove items that measure the same standards as retained items with similar locations (i.e., items that do not add to the OIB by filling missing parts of the blueprint or reducing gaps in scale coverage).
 - b. In English language arts/literacy, consider the joint effect of removing all items associated with a given passage.
 - c. Continue to remove items to leave an item set that meets all constraints, including the target number of score points, match to the blueprint, optimal coverage of the scale, and impact sufficiency requirements.
4. If off-grade items are necessary and approved, select the minimum number of items from off-grade pools to fulfill the various targets.
5. Pass the completed OIB to content specialists, who will review it in unordered (i.e., normal or operational) layout for overall alignment to the blueprint.
6. Pass the completed OIB to Smarter Balanced for review and approval.
7. Repeat Steps 2-6 as necessary.
8. Approve final OIB.

Development and Presentation of Item Maps

The digital item maps provide a broad range of information. The item map is represented as a series of rows, with each row providing information and supporting interaction with a single item (or polytomous item score point). The item map provides information to serve the following functions:

1. A guide to the ordered item booklet, including the following information:

- Order of difficulty (Items are ordered from easiest to hardest based on the item IRT RP location [RP value to be determined]).
 - Item ID
 - Point of origination. ID information will indicate the field test form, and item number within that form.
 - Item details. For example, if an item has multiple parts, ID information would indicate which part was currently being considered (e.g., Item 1 Part A or Item 1 Part B).
 - Polytomous item score point identifier. For polytomous items, each non-zero score point is represented on a separate row of the item map. For example, for an item scored from 0-2, there would be 2 rows in the item map, representing points 1 and 2, such that the score point of 1 for the item was located at the point on the test scale where a student of that ability would have a RP% likelihood of scoring a 1 or higher (where RP is the response probability). For a response probability of RP67, an examinee at that scale location would have a 67% likelihood of scoring *at least* a 1 on the item.
 - Item attributes
 - Item format. The item type for each item/score point will be indicated. Item types include:
 - Selected response
 - Constructed response
 - Performance Task
 - Technology Enhanced
 - Other
 - Scoring information. Correct score key (A, B, C, or D) for selected-response items.
 - Content standard alignments. Abbreviated content standard indicators will be provided via hyperlink.
2. A hyperlink to the OIB. Clicking the order of difficulty (page number) hyperlink will direct the user to the selected item's OIB page where the item and link to supplemental item information may be viewed (e.g., CR item scoring rubrics and exemplars).
 3. A notepad:
 - The item map has a column to organize comments about the items as panelists review them individually and with fellow panelists. Clicking on a comments cell in a row will allow the panelist to enter comments about the item. Hovering over a comment will display the full text of the comments the panelist entered for a specific item. Panelists are asked to discuss two questions to promote deep thinking as they review the ordered item booklet. The questions are intended to stimulate discussion and help panelists understand what an item measures and the contributors to item difficulty. The two questions are:
 - i. What does this item measure? That is, what do you know about a student who can respond successfully to this item or achieve at least the given score point?
 - ii. What makes this item more difficult than the preceding items?
 - The item map has a column to organize comments about possible cut scores panelists are considering. Clicking on the cut score comments cell in a row allows panelists to enter comments regarding their rationale for possible cut scores being considered. Hovering over a comment will display the full text of the comments.
 4. A reference to external benchmarks that provide context that may support the validity of panelists' judgments. A column can be provided for each external benchmark, or multiple benchmarks may be combined into a single or several columns. Benchmarks will be represented

by shading multiple rows distributed evenly about the point estimate for each benchmark. For example, a link to the NAEP Proficient Achievement Level for grade 8 might be represented by shading five contiguous rows in the appropriate column, with the center row representing the best point estimate of the link. External benchmarks may be established to link the Smarter Balanced scales to NAEP, PISA, ACT, SAT, or other appropriate benchmarks. Specific uses of these data sources are outlined below.

- NAEP – NAEP items were embedded in the field tests for grades 4 and 8. Those specific items will not be included in the OIBs for these grades, but their scale values will be translated onto the Smarter Balanced score scale, linking that scale to the NAEP scale. The approximate value of the corresponding NAEP scale will be represented by a range of shaded rows.
- PISA – PISA items were embedded in the high school field tests. Those specific items will not be included in the OIBs for these grades, but their scale values will be translated onto the Smarter Balanced score scale, linking that scale to the PISA scale. The approximate value of the corresponding PISA scale will be represented by a range of shaded rows.
- ACT – MI is working with Michigan and will work with other states to obtain ACT scores for high school students who also participated in the Smarter Balanced field tests. Using those data, they will establish ACT scale equivalents for Smarter Balanced scale scores. ACT has established scale score cutoffs representing college readiness. As with NAEP and PISA scale score equivalents, these cut scores will be translated onto the item maps for the high school tests as regions (i.e., ranges of rows associated with ACT scale values).
- SAT – MI staff will work with Maine and other states with high percentages of SAT takers who also participated in the Smarter Balanced field test to establish scale equivalents as with ACT, following the same procedures.
- Other data sources – The National Assessment Governing Board (NAGB) recently released results of a study indicating that 38 percent of twelfth graders are prepared for college-level English courses and that 39 percent of twelfth graders are prepared for college-level mathematics (Fields, 2014). The study reported NAEP scale scores associated with college readiness as 302 in Reading (the current grade 12 Proficient cut score) and 163 in Mathematics (between the cut scores for Basic and Proficient for grade 12). While it will not be possible to produce grade 12 NAEP scale score equivalents, it will be possible to mark off these percentages using field test impact data. Other reports, such as those by David Conley and his associates (cf. Conley, 2007), also provide estimates of the percentages of high school graduates who are ready for college and/or career. Results of those studies will also be used to the extent possible (considering date of publication, population studied, and definition of readiness).
- The use of NAEP and PISA data will entail not only the calibration of items from those tests embedded in Smarter Balanced field tests but the availability of item parameters and original (i.e., source) scale data for those items as well. Specifically, in order for NAEP and PISA items to be useful, it will be necessary to obtain from original sources the b values of each item on its original (NAEP/PISA) scale and either the full theta to scale score table for each test or the algorithms used to convert theta to scale score. The objective is to calculate the RP value for each NAEP or PISA item on the Smarter Balanced scale and relate that scale score to a NAEP or PISA scale score using the same RP algorithm but with NAEP/PISA values, thereby identifying markers on those scales that correspond to points (or regions) on the Smarter Balanced scales.

5. A rating form to make and submit panelist cut score judgments. Bookmark training will be provided to the panelists prior to making their first round of judgments. Training materials will be available to panelists for review via hyperlinks in the cut point rating column. Each achievement level setting round will have a range of valid theta values specific to each grade/content area (i.e., the total range of obtainable theta values for that grade/subject). The panelist will see an error message if he or she makes a rating outside the valid range or makes the same cut score rating for multiple achievement levels.
6. A link to relevant information:
 - detailed descriptions of content standards
 - Threshold Achievement Level Descriptors (Threshold ALDs)
7. An OIB review progress monitor. This allows panelists to indicate and review their progress as they navigate through the ordered item booklet pages. Panelists can click the “reviewed” cell in a given row to indicate that they have reviewed the given item/score point so that they can monitor their progress. Online panelists may review the OIB in multiple test sessions, so providing a mechanism to monitor their progress will be especially important to them.
8. Record of the judgments made by round. A drop-down menu will allow panelists to select and view their current results as well as the results from previous rounds.

Orientation to External Data

The depth and focus of the orientation to external data will vary by group. Online panelists will receive a relatively brief orientation, while In-Person panelists will receive extensive training.

- **Online panelists** – MI/CTB will prepare a brief, self-paced, narrated PowerPoint overview of the source and purpose of each external data source for the affected grades (currently, grades 4, 8, and high school). These presentations will outline the relevance of the data source and give an approximation of the percentages of students found to be at various proficiency levels (e.g., percent of students at or above NAEP Proficient for grade 4 and how that level relates to the Smarter Balanced scale). Each presentation will include a caveat regarding overinterpretation.
- **In-Person panelists** – Starting with the PowerPoint presentations prepared for the Online panelists, MI/CTB will prepare resource materials to discuss during training. In addition, lead facilitators will be prepared to answer questions during training and will provide follow-up training for panel facilitators who will continue to answer questions during and between rounds.

Use of Practice Test in Panelist Training

Online panelists will be encouraged to log onto the appropriate practice test prior to completing the Bookmark assignment. Their readiness form will include a question confirming that they took the practice test.

The Achievement Level Setting Plan calls for In-Person panelists to take the Practice Tests as well. After the general orientation and orientation to the ALDs and CCCS, panelists will take the appropriate level of the Practice Test and discuss the experience with others in their small groups. The Practice Tests will not only give the panelists an understanding of the types of items students encounter; the experience will also expose them to the various technology enhancements, supports, and accommodations.

Demonstration of Universal Tools, Designated Supports, and Accommodations Available During Training

Smarter Balanced items were developed using principles of universal design. Beyond these universal applications, there are specific supports and accommodations, applicable to students with disabilities and English language learners (ELLs). MI/CTB will demonstrate some of the universal tools (e.g., calculator, notepad, highlighter, strikethrough) as well as accommodated versions of a sample of items during the discussion of the tests after test administration. Specific embedded supports to be demonstrated will include translated test directions and glossary for ELL students and American Sign Language and Braille for students with disabilities.

Presentation of Online Panel Recommendations

Online panelist recommendations for the proficiency cut score will be shared with In-Person Workshop panelists as well as with Vertical Articulation Committee panelists. During the In-Person Workshop, MI/CTB will use whatever Online Panel data are available to share with panelists after Round 1. During the Vertical Articulation Committee meeting, all Online Panel data will be available to share with VAC panelists.

Online panelists will enter a single cut score – for Level 3 (proficiency). Online panelists will complete a demographic survey and indicate their role. MI/CTB will aggregate these cut scores by panelist category (K-12 teacher, other K-12 educator, higher education, and others) and calculate the impact of each cut score. The “Other” category will include parents, business leaders, and others. Given the range of cut scores recommended by members of each group and the differing sizes of these groups, MI/CTB will represent the cut scores and impact as ranges, rather than as point estimates, using graphic representations such as box-and-whisker plots or other graphics that illustrate range. The plan calls for up to 250,000 online panelists, which could mean 4,000 to 5,000 participants for any category for any test. Group means based on 100 or more responses will be reported.

Establishment of a Score Scale

Smarter Balanced assessments will be built using an IRT model. The field tests were constructed to allow items from adjacent grades to be embedded in each grade’s tests. Thus, it will be possible to construct a vertical scale, linking tests from grade 3 through high school.

Field test data will be provided by ETS in late August. Those data will include item difficulty estimates (b values) and step functions for items and ability estimates (theta) for students. Panelists will work with theta (θ) values. However, the theta values will be modified from their original form such that the lowest theta value will have a value of +1. Thus, for example, if the vertical scale ranges from -6.00 to +6.00, all theta values will be increased by 7.00 logits, causing the lowest value to be +1.00 and the highest value to be +13.00. Panelists typically adapt well to score values that are not whole numbers, but they sometimes have difficulty adapting to negative values for student ability estimates. Therefore, adjusted theta values will be retained throughout Online Panel, In-Person Panel, and Vertical Articulation Committee review of items and tests. Conversion to scale scores will occur after approval of cuts by the Chiefs. Keeping the focus on theta values, rather than scale scores, throughout this process has the added advantage of securing the final scale score cuts until Chiefs have approved them.

Following the Chiefs' approval, theta-cut scores will be converted to scale scores. The goal is to set two key cut scores that are near the extremes of the overall scale and which will have memorable scale values, such as 0 and 100 on the Celsius scale, corresponding to freezing and boiling points of water. Therefore, the cut score for high school Level 3 will be set at 1250, and the grade 3 Level 3 cut will be set at 250. The following example illustrates how this conversion can be accomplished.

For this example, the original theta scale ranges from -6.00 to +6.00. Thus, the adjusted theta scale ranges from +1.00 to +13.00. For the sake of simplicity, assume that the grade 3 Level 3 cut has an adjusted theta value of +2.00, and the high school Level 3 has an adjusted theta value of +12.00. Slope and intercept for the conversion can now be calculated using the following table:

Cut Scores	Adj. Theta	Scale	Slope	Intercept
Grade 3 Level 3	2	250	100	250
High School Level 3	12	1250		

With slope and intercept set, all theta values can be converted to scale scores (SS_x) using the following algorithm:

$$SS_x = 250 + b(\theta_x - \theta_{3,3})$$

$$\text{where } b = (1250 - 250) / (\theta_{HS3} - \theta_{3,3}) = 1000 / 10 = 100,$$

θ_{HS3} = Adjusted Theta cut for High School Level 3, and

$\theta_{3,3}$ = Adjusted Theta cut for Grade 3 Level 3

The intercept (250) is defined as the grade 3 Level 3 cut score. The following table illustrates how all remaining theta-cut scores would be converted to scale scores, given these two anchor points.

Hypothetical Final Cut Scores

	Level 2		Level 3		Level 4	
	Adj. Theta	Scale	Adj. Theta	Scale	Adj. Theta	Scale
Grade 3	1.09	159	2.00	250	3.00	350
Grade 4	2.75	325	3.39	389	4.86	536
Grade 5	4.52	502	5.67	617	7.01	751
Grade 6	6.70	720	7.72	822	8.61	911
Grade 7	8.44	894	9.33	983	10.24	1074
Grade 8	9.78	1028	10.94	1144	11.47	1197
High School	10.96	1146	12.00	1250	12.80	1330

LOSS **150**

HOSS **1350**

Lowest Obtainable Scale Score (SS for lowest θ)

Highest Obtainable Scale Score (SS for highest θ)

This score scale has the advantage of differing from existing score scales, summarized below:

Test	Population	Scale Score Minimum	Scale Score Maximum
ACT	High school students applying to college	1	36
NAEP	Grades 4, 8, and 12 general achievement	0	500
PISA	International assessment of 15-year olds	200	800
SAT	High school students applying to college	200	800

References

Conley, D. T. (2007). *Toward a More Comprehensive Conception of College Readiness*. Eugene, OR: Educational Policy Improvement Center.

Fields, R. (2014). *Towards The National Assessment of Educational Progress (NAEP) as an Indicator of Academic Preparedness for College and Job Training*. Washington, DC: National Assessment Governing Board.

Smarter Balanced Assessment Consortium: Technical Recommendations for the Bookmark Standard Setting Procedure

Measurement Incorporated and CTB/McGraw-Hill

March 28, 2014



Technical Recommendations for the Bookmark Standard Setting Procedure

The purpose of this white paper is to provide recommendations with regard to the item response theory (IRT) model to be used in standard setting, as well as the response probability (RP) value to use in directing standard-setting panelists to place bookmarks.

Model Considerations

The Bookmark standard setting procedure (Lewis, Mitzel, Mercado, & Schulz, 2012) is an item response theory-based item mapping procedure developed in 1996 in response to the need for a robust standard setting procedure for high-stakes assessments of mixed format. Since 1996, it has become the most widely used procedure for setting cut scores on statewide assessments and other high stakes educational assessments. Its psychometric foundation is well documented (e.g., Cizek & Bunch, 2007), and its usefulness has been well established through adoption of cut scores produced by Bookmark-based standard-setting activities.

In a typical Bookmark procedure, each item in an ordered item booklet (OIB) has been mapped to an underlying construct in terms of the amount of that construct the examinee must possess in order to have a reasonable chance of answering the item correctly (in the case of a selected-response item) or obtaining a given score point or higher (in the case of a constructed-response item or performance task).

In creating OIBs, psychometricians have a choice of IRT models: 1, 2, and 3-parameter logistic models. Each of these models includes a difficulty parameter, but the 2-parameter model includes an item discrimination parameter, and the 3-parameter model includes a pseudo-guessing parameter. The choice of model for creating OIBs typically depends on the model used in the construction, administration, and scoring of the operational test.

In the three-parameter logistic (3PL) model, the Bookmark procedure relies on the basic relationship between person ability (θ) and item difficulty (b), discrimination (a), and pseudo-guessing (c), where the probability of answering an item correctly (P) can be expressed as shown in equation (1):

$$P_j(X=1|\theta) = c_j + (1 - c_j)/\{1+\exp[-1.7a_j(\theta - b_j)]\}, \quad (1)$$

where P_j is the probability of answering correctly, c_j is the pseudo-guessing index, θ is the ability required, a_j is the item discrimination index, \exp is the exponential function, and b_j is the item difficulty index. The way that guessing is accounted for is critical to the mapping. For most bookmark procedures, the c (pseudo-guessing) parameter is set to zero, so that the response probability specified is associated with the likelihood of a student knowing the correct response without guessing, as shown in equation (2):

$$P_j(X=1|\theta) = 1/\{1 + \exp[-1.7a_j(\theta - b_j)]\}. \quad (2)$$

The most common one-parameter IRT model is the Rasch model, in which the relationship between examinee ability and item response is expressed as in equation (3):

$$P_j(X=1 | \theta) = 1/[1+\exp(b_j - \theta)], \quad (3)$$

where all terms are as defined in equations (1) and (2).

Given that the typical application of the 3-parameter model reduces the pseudo-guessing parameter to 0, the choice of models for creation of OIBs is actually between the 1-parameter and the 2-parameter models. Measurement Incorporated and CTB/McGraw-Hill psychometricians are well-acquainted with all three models and have software capable of producing the necessary outputs whether the model has 1, 2, or 3 parameters.

RP Value Considerations

In preparation for Bookmark standard setting, it is customary to create an OIB whose items are in difficulty order. Specifically, items are placed in order of the ability level (θ value) required to have a particular likelihood or probability of answering correctly. This value is referred to as the response probability or RP value.

In the Bookmark procedure, panelists are typically asked to find an item that a certain percentage of examinees at a critical threshold will be able to answer correctly. The cut score is identified at the ability level associated with that point in an ordered item booklet beyond which panelists can no longer say that the target group would have the specified likelihood of answering correctly. The choice of that percentage is critical not only to defining the group of examinees but to defining the threshold between adjacent ability groups. This percentage is commonly called the RP value. In practice, users of the Bookmark procedure have employed 50 percent, 60 percent, 67 percent, and other values.

The use of RP67 is, in the experience of MI and CTB and as recognized in the literature (Karantonis & Sireci, 2006), the most commonly used RP value in state standard settings. However, there are circumstances when the measurement properties of specific assessments call for the consideration of other RP values. We routinely examine each assessment's measurement properties and student achievement distributions (i.e., the "targeting of items to the student achievement distribution) to support the efficacy of the default RP67 response probability criterion prior to standard setting. Circumstances have arisen in which we have recommended and used other RP values based on such review.

Given the expected difficulty of the Smarter Balanced assessments, RP values lower than RP67 will be considered. If a test is difficult, an RP67 response probability may result in a minimum cut score (e.g., a cut score associated with placing a Bookmark after the first page of the resulting ordered item booklet—the lowest expectation possible for any level of achievement) associated with an unacceptable percentage of students falling below the cut. For instance, for a very difficult test it is possible that the use of RP67 could result in 50% of students being in the lowest level of achievement even if the Bookmark is placed after the first item in the OIB (only having to demonstrate mastery of the easiest item to meet the cut score). In such cases lower RP values should be considered.

Given the expected difficulty of the test, we recommend the consideration of three RP values, RP67, RP60, and RP50, as described below. It is important to note that if the RP value changes, we must also change the instructions to standard setting participants so that they make their judgments with the appropriate RP value taken into consideration. The RP value used also will affect the ALDs. We

discuss the three RP values below, and how the choice of RP value affects the way the judgment tasks are communicated to participants.

RP67. RP67 indicates that the item is located on the scale at the scale location θ such that a student with ability θ would be expected to have a .67 likelihood of successfully responding to a selected-response item or have a .67 likelihood of achieving a given score point or higher for a polytomous item. Participant instructions to place the Bookmark associated with, for example, the Level 2 cut score, can be summarized as

“Place a Bookmark in the ordered item booklet after the first item such that a student who demonstrates mastery of the items prior to the Bookmark has minimally demonstrated the knowledge, skills, and abilities associated with the Level 2 ALD.”

Extensive training is done on the meaning of the Bookmark, regardless of the RP value, but in this case, a succinct summary of the RP training is “Students show mastery of an item when they have at least a 67% chance of answering an item correctly.”

RP60. RP60 indicates that the item is located at the scale location θ such that a student with ability θ would be expected to have a .60 likelihood of successfully responding to a selected-response item or have a .60 likelihood of achieving a given score point or higher for a polytomous item.

The instructions for participants are similar to those described above for RP67 and are not repeated here. The judgment task may still be considered under the context of mastery using RP60. That is, a student with a .60 probability of answering an item correctly can be assumed to have mastery of the item (though this may be debated). Nonetheless, as long as training on the value of .60 with respect to the mastery concept has occurred, the context of the instructions to participants need not be substantially modified from those under RP67.

RP50. The use of RP50 has precedent. Lewis and Mitzel (1995) describe the use of RP50 in implementations of the Bookmark Procedure. However, the notion of mastery is not appropriate when RP50 is used. That is, the use of RP50 requires participants to consider the point of indifference with respect to the minimally qualified candidate. In this case, the student has a .50 likelihood of responding successfully to the item, and it seems inappropriate to assume that students have mastered a concept that they demonstrate with a 50% success rate.

Choosing an RP value. Solving equation (2) for θ produces equation (4):

$$\theta = b_j + \ln((1-P_j)/P_j)/(-1.7a_j) \quad (4)$$

where \ln is the natural logarithm, and other values are as defined for equation (2). For any value other than 50%, the value for $+\ln((1-P_j)/P_j)$ is nonzero. However, when $P_j = .50$, the value of $+\ln((1-P_j)/P_j)$ reduces to $\ln(1)$, which is 0. At this value, item discrimination plays no part in the determination of the threshold ability level, and the value of θ reduces to the item difficulty b_j . The three values and their effects are shown below. In effect, when RP is set equal to 50, the expected θ value is equal to the item’s b value, the same as in the Rasch model when $RP=50$. At this point, the model actually becomes a Rasch model, at least as far as creation of the OIB is concerned.

RP	2PL Value of θ	Rasch Value of θ
50	b_j	b_j
60	$b_j + .405/1.7a_j$	$b_j + .405$
67	$b_j + .708/1.7a_j$	$b_j + .708$

Use of the three RP values above will result in different item maps. The ordering of items will likely be somewhat different for each of the maps, though the ordering is expected to be highly correlated. The results of Beretvas' (2004) study on IRT models and RP values indicated that the rank ordering of items varied with the RP value for common IRT models (except, of course, the use of the 1PL or Rasch models with dichotomous items).

The main effect, and rationale for the use of one RP value over the other for difficult assessments, is that lower RP values will allow participants to place their Bookmarks further in the ordered item booklets than higher RP values but result in the same impact data. Most importantly, the use of a lower RP value (e.g., RP50 or RP60) will somewhat ameliorate the problem of using RP67 with a difficult test resulting in an unacceptably high percentage of students falling below the cut score even when the Bookmark is placed in the beginning of an ordered item booklet.

Adjustment for Guessing. Each of the three RP values can be considered in the context of selected response items) with and without an adjustment for guessing. Typically, the use of RP67 is adjusted for guessing (RPGA67), which means student mastery is defined in terms of “knowing the correct response to a selected response item without guessing.” In practice the adjustment for guessing is achieved by setting the pseudo-guessing parameters (or c parameter) to zero when using item parameters to locate items on the scale. However, this adjustment is not necessary, and in the case of difficult tests, will provide similar results as lowering the RP value. Thus, we will initially consider six options in the evaluation: RPGA67 (the default value used in most operational standard settings), RP67, RPGA60, RP60, RPGA50, RP50. The first of these—RPGA67—will produce the fewest students meeting or exceeding the cut score associated with a given page in the ordered item booklet. The last—RP50—will produce the most students meeting or exceeding the cut score associated with a given page in the ordered item booklet.

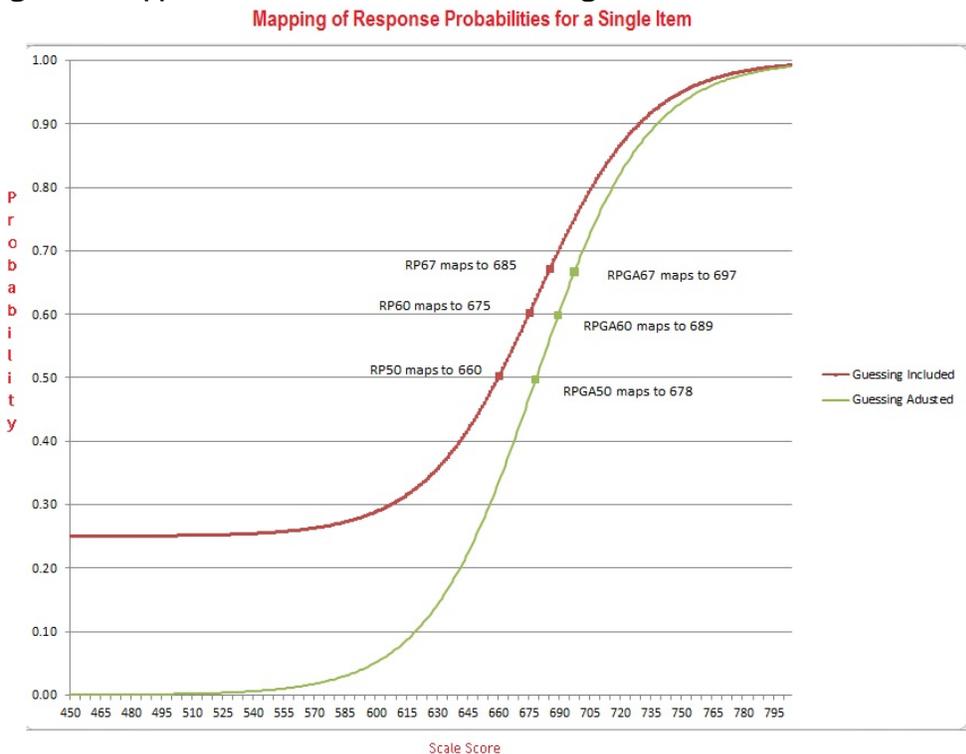
Item mapping allows individual items to be located along the scale score continuum so that interpretations about what students know and can do at individual scale score points may be facilitated. Item mapping is a component in the process of setting performance standards in the Bookmark Standard Setting Procedure (BSSP; Lewis, Mitzel, & Green, 1996). Though item mapping is based in item response theory, Loricé (2001) aptly noted that item mapping requires the human judgmental process because the response probability (RP; the likelihood that a person answers the item correctly) must be determined in order to align an item with a specific score point. Numerous psychometricians (Kolstad, 1996; Ferrara, Perie, Johnson, 2001; Mitzel, Lewis, Patz, & Green, 2001; Zwick, Senturk, Wang, & Loomis, 2001; Beretvas, 2004) have stated that as RP levels are varied, the item location may shift up or down the scale score continuum. When using the 3-parameter IRT model, items may shift in rank order as well.

However, RP50 can be appropriately used by modifying the judgment task instructions to standard setting participants. The use of RP50 has been conceptually described to Bookmark participants through the following analogy. Participants were asked to consider a scale from 1 to 100 formed by a group of 100 individuals (say chess players) of regularly increasing ability with the 1st individual being a complete chess novice and the 100th being an expert chess player. Participants were asked how they would locate themselves as chess players on this continuum—the answer being to find the chess player (from 1 to 100) that they could beat exactly 50% of the time. If that was chess player 40, for example (locating them at 40 on the scale from 1 to 100), then that means they would beat chess player 39 more than half of the time (since that player had less skill than player 40), and they would beat chess player 41 less than half of the time. This approach is understandable to participants, with training, and relates to games they play in everyday life (chess, checkers, basketball, golf, etc., can be used to help participants relate).

This analogy is used to convey the judgment task for the standard setting participants. By analogy, their judgment task required to identify, for example, the Level 2 cut score, is to identify the item in the ordered item booklet that a minimally qualified (just at the cut score and no higher) Level 2 candidate would have a .50 likelihood of successfully responding to a selected response item or of achieving at least the given score point for a polytomous item.

This can be explained in training a variety of ways, such as if the student were presented with 100 similar items measuring the same skill with the same difficulty, then they would be expected to respond successfully to about 50 of the 100 items. This situation is depicted in Figure 1.

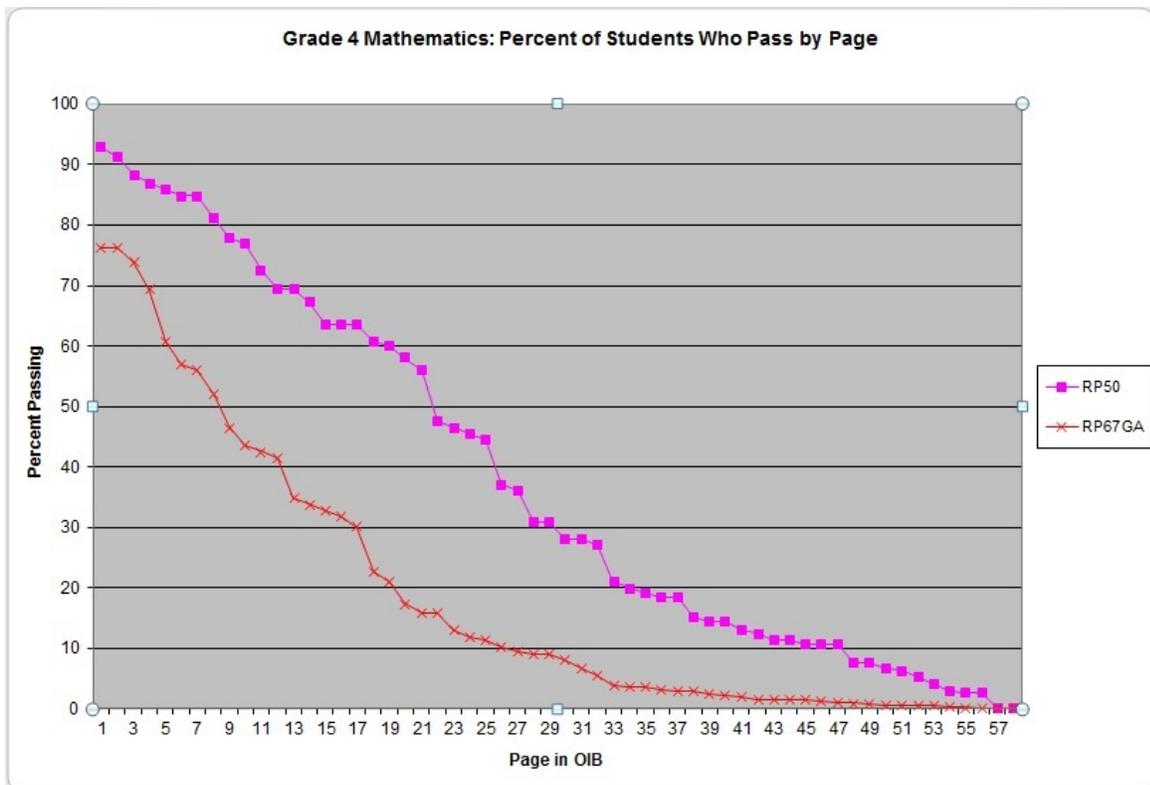
Figure 1. Mapped Scale Score Location for a Single Item with and without Guessing



As described above, changing the RP value has an effect on the ordering of the items in the OIB and on the impact data associated with a given location in the ordered item booklet. It is highly desirable to have the full range of “acceptable” impact data associated with each level of achievement defined by the cut scores. Thus, the impact data associated with all possible Bookmark placements in each ordered item booklet will be represented in graphics for each of the RP values discussed above (and any others recommended by Smarter Balanced or their TAC members). This will identify problems associated with floor or ceiling effects and gaps in the ordered item booklets.

For instance, if RP67 guessing adjusted resulted in only 70% of students passing associated with a Bookmark on OIB page 4 but RP50 resulted in 87% of students passing associated with a Bookmark on OIB page 1, this might argue for the use of RP50 over RPGA67. As may be seen in Figure 2 for RPGA67 beginning with OIB page 26 the remaining items in the OIB only describe the highest performing 10% of student. By OIB page 47 increases in cut scores barely change the impact data. The items are generally difficult for the students. Lowering the RP value to RP50 allows OIB pages 26 to the end to describe the skills of the highest performing 37% of students with increases in cut scores steadily changing the impact data. For Smarter Balanced, this information will be examined for each possible Bookmark location in each ordered item booklet (cut scores and impact data associated with Bookmarks after page 1, 2, 3, 4, ..., through the last page of the ordered item booklet).

Figure 2. Impact data associated with a specified set of items and two different RP values



Conclusions and Recommendations

1. With regard to IRT model to be used for creation of OIBs, we believe it should be up to the Psychometric Contractor to pick a model that will guide construction, delivery, and scoring of the Smarter Balanced tests. We are prepared to apply any of the three standard IRT models.
2. With regard to choice of RP value, we had indicated in our proposal that we recommend a value of .67. We have also considered values of .60 and .50. Our concern about the value of .50 is that it actually reduces the model from a 3PL model (with the c parameter set to 0, or, in effect, a 2PL model) to a 1PL model. We therefore recommend .67 as the RP value unless our analysis of the field test data supports the use of .60.

References

- Beretvas, S. N. (2004). Comparison of Bookmark difficulty locations under different item response models. *Applied Psychological Measurement, 28*(1), 25-47.
- Cizek, G. J., & Bunch, M. B. (2007). *Standard Setting: A Guide to Establishing and Evaluating Performance Standards on Tests*. Thousand Oaks, CA: Sage.
- Ferrara, S., Perie, M., & Johnson, E. (2002). *Matching the judgmental task with standard setting panelist expertise: The item-descriptor (ID) matching procedure*. Washington, DC: American Institutes for Research.
- Karantonis, A., & Sireci, S. G. (2006). The Bookmark standard setting method: A literature review. *Educational Measurement Issues and Practice, 25*(1), 4-12.
- Kolstad, A. J. (1996). *1992 National Adult Literacy Survey: Literacy levels and the 80 percent response probability convention*. National Centre for Education Statistics: Washington, DC.
- Lewis, D. M., & Mitzel, H. C. (1995). An item response theory based standard setting procedure. In D. C. Green (Chair), *Some uses of item response theory in standard setting*. Symposium conducted at the annual meeting of the California Educational Research Association, Lake Tahoe, NV.
- Lewis D.M., Mitzel, H. C., Green, D. R. (1996). *Standard Setting: A Bookmark Approach*. In D. R. Green (Chair), *IRT-Based Standard-Setting Procedures Utilizing Behavioral Anchoring*. Symposium presented at the 1996 Council of Chief State School Officers 1996 National Conference on Large Scale Assessment, Phoenix, AZ.
- Lewis, D. M., Mitzel, H. C., Mercado, R. L., & Schulz, M. W. (2012). The Bookmark standard setting procedure. In G. J. Cizek (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*. New York: Routledge.
- Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The Bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting Performance Standards: Concepts, Methods, and Perspectives*. Mahwah, NJ: Lawrence Erlbaum & Associates.
- Zwick, R., Senturk, D., Wang, J., & Loomis, S. C. (2001). An investigation of alternative methods for item mapping in the National Assessment of Educational Progress. *Educational Measurement: Issues and Practice, 20*(2), 15-25.

Selecting Items for the Ordered Item Booklets (OIBs) for the Smarter Balanced Achievement Level Setting (ALS)

Panelists at the Smarter Balanced Achievement Level Setting (ALS) will examine purposefully selected sets of items as part of the Bookmark Standard Setting Procedure. The items for each grade and content area combination will be presented in order of difficulty as part of ordered item booklets (OIBs).

It is expected that 100–200 items per grade and content area combination will survive psychometric analyses and be available for inclusion in the OIB. Measurement Incorporated (MI) will select the items for the OIBs, and these selections will be reviewed by McGraw-Hill Education CTB. MI will select a subset of items for inclusion in each OIB based on the following two factors.

- 1) **Coverage of the Common Core State Standards:** selection should result in an appropriate balance of item alignments, scoring models, and depth of knowledge (DOK) levels
- 2) **Coverage of the test scale:** included items should appropriately cover the test scale

This document describes the process that will be used to select items for the OIBs to be used at the Smarter Balanced ALS.

Beginning the Selection Process

Each OIB will comprise approximately 70 items. Each polytomous item will be represented in the OIB once for each non-zero score point associated with the item. For example, an item with four non-zero score levels will appear four times: score point 1 will appear once in the OIB, score point 2 will appear later, and so forth. Accordingly, the total number of pages in each OIB will likely be approximately 80–95 pages.

To begin the selection process, all of the eligible items for each OIB will be mapped onto the test scale using a response probability (RP) criterion approved by Smarter Balanced. The RP criterion is discussed in the document “RP Criterion Considerations for Smarter Balanced Achievement Level Setting.” Item maps comprising the items, the RP-adjusted scale locations, and the item metadata will be prepared for use in the item selection process.

For convenience, the RP-adjusted scale locations will be transformed onto a temporary scale. Specifically, the RP-adjusted theta values will be transformed linearly onto a temporary scale with a mean of 100 times the grade level (e.g., 600 for grade 6) and with a standard deviation of 50. If a vertical scale is developed in time for the online or in-person ALS, a single transformation, as opposed to the grade-by-grade transformations will be used to develop an interim scale across all grades.

Criterion 1: Coverage of the Common Core State Standards

MI will examine the coverage of the Common Core State Standards among the items in the initial selection. Ideally, the distribution of Standards alignments for the items selected for each OIB will match those specified in the test blueprints, as summarized in “Item Pool Specifications.” This document shows the desired distribution of test items for the Smarter Balanced item pool in terms of claims, targets, item types, and DOK levels.

Given the complexity of the Standards and the limited number of items to be included in the OIB, it cannot be expected that the Standards alignment of the OIB will match the Item Pool Specifications exactly. However, MI will seek to select items such that the number of items aligned to each Claim approximates the proportions stated in the Specifications.

MI will select items that are scored both dichotomously and polytomously. To promote adequate inclusion of both types of items, MI will first select dichotomous items for each OIB, such that approximately 60 dichotomous items are selected for each OIB. MI will then select one performance task (PT) and approximately nine polytomously scored items to each OIB, as indicated by the Item Pool Specifications.

Criterion 2: Coverage of the Test Scale

The items for the OIB will be selected so that (a) they are well-distributed across the test scale, (b) there are not large gaps in scale location between consecutive items, and (c) they have locations that cover a broad range of student achievement. With items selected in this way, ALS panelists will be able to recommend cut scores with an appropriate level of precision across the test scale.

Unlike fixed-form test development, in which item selection may seek to select items that are more highly concentrated in the center of the student ability distribution (which increases overall precision), for standard setting purposes, it is more common to develop OIBs such that a rectangular distribution of items results across the student ability distribution. To achieve this result, the test scale is divided into three partitions based on the student achievement distribution. The first partition begins with the lowest obtainable test score and ends with the point on the test scale associated with the 33rd percentile of students' scores. The second partition ranges from the 33rd to the 66th percentile. The third partition ranges from the 66th percentile to the highest obtainable test score.

Approximately equal numbers of items (and score points) will be selected for each of the three partitions such that an acceptable overall number of items and score points results. Selected this way, items will be selected across the observed range of scores, especially in ranges where many students have scored.

MI will make an initial selection of items for the OIB with approximately equal number of items per scale partition. In making this initial selection, MI will purposefully select items with different scale locations within the partition to moderate scale location gaps between consecutive items. Selecting adjacent items with different scale locations assures a change in cut score recommendation when there is a change in bookmark placement. Selecting adjacent items with modest gaps in scale location supports a modest change in cut score associated with a modest change, say a single item, in bookmark placement.

Final Checks and Transmittal to Smarter Balanced for Approval

MI will complete its item selections with reviews by both researchers and content experts to inspect the proposed item selections on psychometric and content-based considerations.

The proposed item selections for the OIBs will be transmitted to Smarter Balanced for review and approval. To assist Smarter Balanced in its review, the following materials will be provided:

- Item metadata for each of the items selected for each OIB
- Item-level statistics for selected items
- Approximately 10 additional items per OIB that could be selected for each OIB, should Smarter Balance staff choose to replace items selected by MI

Acceptance Criteria

MI and CTB suggest that Smarter Balanced evaluate the selection of items for each OIB using these criteria.

- Do the items cover the claims of the Common Core State Standards, as described in the blueprint?

Selecting Items for the ALS

- Do the items represent the various item scoring models appropriately?
- Are the DOK levels represented appropriately?
- Are the RP-adjusted scale locations distributed along the test scale?
- Are there few “gaps” in locations of consecutive items?
- Are there few or no consecutive items with the same locations?

MI and CTB will look to Smarter Balanced to approve the final item selections for all OIBs according to the project schedule, provided separately.

Description of Bookmark Pro

Bookmark Pro is McGraw-Hill Education CTB's proprietary software that produces materials and reports for standard settings based on the Bookmark Standard Setting Procedure.

Bookmark Pro is prepared for a workshop by loading associated input files including metadata (information about the test items), item parameter files, and frequency distributions for the various grade and content area combinations. The names of the achievement levels and the response probability used to locate items on the scale must be specified.

At the In-Person Smarter Balanced Achievement Level Setting, panelists electronically recorded their three bookmarks (as page numbers of the ordered item books) for Level 2, Level 3, and Level 4 in Measurement Incorporated's standard setting software. Each panelist's three bookmark judgments were exported from the MI software into a text file that was imported into *Bookmark Pro*. The individual bookmarks were transformed into the associated scale scores, and the median scale scores were used to determine the impact data and to create the tabular and graphical output for that panel.

The procedure for the Online Panel was similar to that of the In-Person Panel. However, only one bookmark was placed by participants in the Online Panel (the Level 3 bookmark). In addition, each grade and content area's bookmarks were processed for only one round for the Online Panel in contrast to the three rounds for the In-Person Panel.

The following list describes the work done prior to a workshop and at the workshop.

Prior to Workshop

- Specification of grade content areas (GCA)
- Specification of response probability (RP) value (requires the files below to do RP analysis)
- Naming conventions (name of workshop, etc., for output headers)
- Input files
 - Metadata file (item attributes) in specified format
 - Parameter files (item response theory metric specified) in specified format
 - Frequency distributions (same metric as parameters) in specified format
 - Standard error of measurement (SEM) file in specified format
 - Names of achievement levels (from lowest to highest)
- Create GCAs and load input files
- Output files
 - Graphic of impact data associated with items in order of difficulty for one or more RP values; multiple RP values used in RP study
 - Item map including order of difficulty, item locations (based on specified RP value), specified metadata

At Workshop

- Input files (by GCA and round)
 - Text file of panelists' bookmark pages in specified format:
e.g., (grade.content.table.panelistID.Level2R1.Level3R1.Level4.R1)
- Output files
 - Median bookmark and scale score summaries (by table and overall)
 - Median bookmarks and scale scores
 - Impact data
 - Impact data bar chart (overall)
 - Histogram of panel bookmark placements
 - Achievement levels color coded
 - 25th, 50th, 75th percentiles (median and interquartile range)

Smarter Balanced Achievement Level Setting Readiness Form

Part 1: Common Core State Standards and Achievement Level Descriptors

[To be completed by online panelists after reviewing CCSS and ALDs]

I have reviewed the Common Core State Standards for the grade and subject assigned to me. Yes No

I understand the relationship between the content of the standards and what students are expected to know and be able to do in this subject at this grade level. Yes No

I understand how the Common Core State Standards relate to the tasks I will be asked to perform. Yes No

I have reviewed the Achievement Level Descriptors for the grade and subject assigned to me. Yes No

I understand the progression of student achievement from Achievement Level 1 to Achievement Level 4. Yes No

I understand how the Achievement Level Descriptors relate to the tasks I have been asked to perform for the Bookmark Achievement Level Setting Process . Yes No

For High School (Grade 11) only: I understand that Level 3 for grade 11 indicates conditional content readiness for college and career, while Level 4 for grade 11 indicates content readiness. Yes No

[Advance to Practice Test; allow panelists to opt out if they have already completed a Practice Test for this grade and subject in the past 3 months.]

Part 2: Practice Test

I have taken the Practice Test. Yes No

[Return to Home Page: Setting Achievement Levels Using the Bookmark Procedure.]

Part 3: Bookmark Round 1

I have reviewed the document "Setting Achievement Levels Using the Bookmark Procedure." Yes No

I have completed the Orientation Round Booklet and set one bookmark. Yes No

Which bookmark did you set? Level 2 Level 3 Level 4

I understand how my bookmark will be combined with the bookmarks of other panelists to derive a cut score. Yes No

I understand the procedure I am to follow and am ready to review the ordered item booklet for the subject and grade assigned to me and set one bookmark. Yes No

Online Panel Bookmark Results

Table D.3.1 Bookmark Medians and IQR, Grade 3

Panelist Category	ELA	Math
	Level 3 bookmark median	Level 3 bookmark median
Teachers	23.0	35.0
Administrators	35.0	38.0
Other	26.5	43.0
All online panelists	26.5	35.0
IQR	16.0–53.0	22.0–55.0

Table D.3.2 Bookmark Medians and IQR, Grade 4

Panelist Category	ELA	Math
	Level 3 bookmark median	Level 3 bookmark median
Teachers	40.0	40.5
Administrators	48.0	32.0
Other	48.0	36.5
All online panelists	42.0	40.0
IQR	13.0–60.0	17.75–62.0

Table D.3.3 Bookmark Medians and IQR, Grade 5

Panelist Category	ELA	Math
	Level 3 bookmark median	Level 3 bookmark median
Teachers	28.0	31.0
Administrators	24.0	43.0
Other	49.0	50.0
All online panelists	27.5	42.0
IQR	9.5–51.5	25.0–63.0

Table D.3.4 Bookmark Medians and IQR, Grade 6

Panelist Category	ELA	Math
	Level 3 bookmark median	Level 3 bookmark median
Teachers	27.0	33.0
Administrators	NA	NA
Other	31.5	33.0
All online panelists	28.0	33.0
IQR	9.25–43.25	16.0–60.25

Table D.3.5 Bookmark Medians and IQR, Grade 7

Panelist Category	ELA	Math
	Level 3 bookmark median	Level 3 bookmark median
Teachers	25.5	35.0
Administrators	NA	NA
Other	31.5	36.0
All online panelists	26.5	35.0
IQR	9.0-54.75	25.0-54.0

Table D.3.6 Bookmark Medians and IQR, Grade 8

Panelist Category	ELA	Math
	Level 3 bookmark median	Level 3 bookmark median
Teachers	40.0	47.0
Administrators	45.0	47.0
Other	34.0	39.0
All online panelists	40.0	47.0
IQR	14.0-56.0	26.0-52.0

Table D.3.7 Bookmark Medians and IQR, Grade 11

Panelist Category	ELA	Math
	Level 3 bookmark median	Level 3 bookmark median
Teachers	35.0	50.0
Higher Education	34.5	50.0
Administrators	42.0	48.0
Other	31.0	49.5
All online panelists	35.0	50.0
IQR	11.5-58.0	30.0-59.0

Appendix F – In-Person Workshop Documentation

• F01_Staff Assignments	658
• F02_Generic Agenda	659
• F03_Final Amplified Agenda	661
• F04_List of Scripts and Presentations	667
• F05_In-Person Panel Participant Demographics	669
• F06_Facilitator Script	675
• F07_Facilitator Script Supplement.....	735
• F08_Observer List.....	740
• F09_Observer Script.....	741
• F10_Table Leader Training	748
• F11_In-Person Intro	778
• F12_OIB Training	813
• F13_Bookmark Training.....	839
• F14_Readiness and Evaluation Form	856
• F15_In-Person Results by Round.....	865
• F16_ELA03 Bookmark Charts.....	877
• F17_ELA04 Bookmark Charts.....	880
• F18_ELA05 Bookmark Charts.....	883
• F19_ELA06 Bookmark Charts.....	886
• F20_ELA07 Bookmark Charts.....	889
• F21_ELA08 Bookmark Charts.....	892
• F22_ELA11 A Bookmark Charts	895
• F23_ELA11B Bookmark Charts	898
• F24_Math03 Bookmark Charts	901
• F25_Math04 Bookmark Charts	904
• F26_Math05 Bookmark Charts	907
• F27_Math06 Bookmark Charts	910
• F28_Math07 Bookmark Charts	913
• F29_Math08 Bookmark Charts	916
• F30_Math11A Bookmark Charts	919
• F31_Math11B Bookmark Charts.....	922

October 13-15	11-A ELA	11-B ELA	11-A Math	11-B Math
Facilitator	Ann Wilder	Gretchen Schultz	Maud Eno	Judy Hickman
Co-Facilitator	Craig Deville	Rick Mercado	Winnie Reid	Jennifer Lord-Bessen
Leadership	Mike Bunch, Dan Lewis			
Data Analysis	Adele Brandstrom, Anne Woods, Irene Hendrawan, Ari Wibowo			
IT	Nelson Androes, Eric Lewis, Ben Weber, Anthony Jordan, Ira Henderson			
Project Mgmt.	Barbara Scherich, David Breen, Mandy Hunter, Kathleen Stapleton			
Auditor	Greg Cizek			

October 15-17	8 ELA	8 Math	7 ELA	7 Math	6 ELA	6 Math
Facilitator	Amy Griswold	John Upchurch	Molly Buck	Lisa Johnson	Chris Dunbar	Kelly Bolton
Co-Facilitator	Dan Bowen	Juan D'Brot	Sarah Hagge	Jennie Bowen	Corey Palermo	Rick Mercado
Leadership	Mike Bunch, Dan Lewis					
Data Analysis	Adele Brandstrom, Anne Woods, Irene Hendrawan, Ari Wibowo					
IT	Nelson Androes, Eric Lewis, Anthony Jordan, Ira Henderson					
Project Mgmt.	Barbara Scherich, David Breen, Mandy Hunter, Kathleen Stapleton					
Auditor	Greg Cizek					

October 17-19	5 ELA	5 Math	4 ELA	4 Math	3 ELA	3 Math
Facilitator	Ruth Hargis	Judy Hickman	Kelly Connelly	Lisa Johnson	Sheryl Grady	Heather Farina
Co-Facilitator	Joe McClintock	Jennifer Lord-Bessen	Sarah Hagge	Winnie Reid	Craig Deville	Rick Mercado
Leadership	Mike Bunch, Dan Lewis					
Data Analysis	Adele Brandstrom, Anne Woods, Irene Hendrawan, Ari Wibowo					
IT	Nelson Androes, Eric Lewis, Anthony Jordan, Ira Henderson					
Project Mgmt.	Barbara Scherich, David Breen, Mandy Hunter, Kathleen Stapleton					
Auditor	Greg Cizek, Heather Koons					

Achievement Level Setting Agenda

Day	Time	Event
Day 1	7:30 a.m.	Registration opens (Ballroom Lobby)/Continental breakfast
	8:30 a.m.	Opening Session (Carpenter Ballroom).
	9:00 a.m.	Overview of schedule of events for the in-Person Panel
	9:30 a.m.	Panelists are dismissed to breakout rooms
	9:35 a.m.	Orientation to hardware and software
	9:55 a.m.	Discussion of CCSS and ALDs
	10:30 a.m.	Break
	10:45 a.m.	Introduction to the test
	10:55 a.m.	Practice Test administration
	11:55 a.m.	Table-wide discussion of tests
	12:20 p.m.	Closing comments on test construction; completion of Readiness Form
	12:30 p.m.	Lunch (Ballroom Lobby)
	1:20 p.m.	Introduction to the Ordered Item Booklet
	1:50 p.m.	Review of Ordered Item Booklet
	5:00 p.m.	Adjourn
Day 2	7:30 a.m.	Registration opens (Ballroom Lobby)/Continental breakfast
	8:30 a.m.	Orientation to the Bookmark Procedure (Ballroom)

Day	Time	Event
Day 2	8:50 a.m.	Reconvene in breakout rooms
	9:00 a.m.	Group practice exercise/Completion of Readiness Form
	9:45 a.m.	Bookmark Round 1
	12:00 Noon	Lunch (Ballroom Lobby)
	1:30 p.m.	Review of Round 1
	2:30 p.m.	Completion of Readiness Form
	2:45 p.m.	Bookmark Round 2
	4:50 p.m.	Wrap-up for Day 2
		Panelists log out; computers secured
	5:00 p.m.	Adjourn
Day 3	7:30 a.m.	Registration opens (Ballroom Lobby)/Continental breakfast
	8:30 a.m.	Convene in breakout rooms
		Review Round 2 results/Completion of Readiness Form
	10:00 a.m.	Bookmark Round 3
	11:30 a.m.	Wrap-up for Day 3
	11:45 a.m.	Completion of Evaluation Forms
	12:00 Noon	Adjourn for lunch (Ballroom Lobby)

Day	Time	New Event
Arrival	4:00 p.m.	Facilitator Training - Oct. 12th, 14th, 16th
	6:00 p.m.	Table Leader Training - Oct. 12th, 14th, 16th
Day 1	7:30 a.m.	Registration opens (Ballroom Lobby) Badges, place cards, and room and table assignments
		Continental breakfast
		Opening Session (Carpenter Ballroom). Mike opens and introduces Smarter Balanced staff. Joe follows.
	8:30 a.m.	
		Keynote address and charge
		Q&A; introduction of the parking lot concept; Joe gives Mike the floor.
		Overview of schedule of events for the in-Person Panel (Mike)
	9:00 a.m.	
	9:30 a.m.	Panelists are dismissed to breakout rooms
	9:35 a.m.	Introductions
	9:45 a.m.	Orientation to hardware and software
	9:55 a.m.	Discussion of CCSS and ALDs
	10:30 a.m.	Break
	10:45 a.m.	Introduction to the test
		Overview of test development and scoring
	10:55 a.m.	Practice Test administration
	11:55 a.m.	Table-wide discussion of tests
	12:20 p.m.	Closing comments on test construction
	12:25 p.m.	Completion of Readiness Form Part 1
		Lunch (Ballroom Lobby). Joe answers Parking Lot questions last 20 minutes.
12:30 p.m.	Introduction to the Ordered Item Booklet (Ballroom: Dan)	
1:20 p.m.		
1:50 p.m.	OIB training and review	
	Wrap-up for Day 1. Panelists log out; computers secured	
4:50 p.m.		
	Adjourn	
5:00 p.m.		
	Smarter Balanced, MI, CTB, and auditor debriefing	
5:15 p.m.		
6:00 p.m.	Secure meeting rooms and depart	
Day 2	7:30 a.m.	Registration opens (Ballroom Lobby) Badges, place cards, and room and table assignments
		Continental breakfast
		Orientation to the Bookmark Procedure (Ballroom)
8:30 a.m.		

- 8:50 a.m. Reconvene in breakout rooms
- 9:00 a.m. **Orientation round**
- 9:40 a.m. Completion of Readiness Form Part 3
- 9:45 a.m. Round 1
- 11:45 Reflection questionnaire
- Lunch (Ballroom Lobby). Joe answers Parking Lot
- 12:00 Noon questions last 20 minutes.
- Confer with Smarter Balanced staff to approve Round 1 analyses and reports**
- 1:00 p.m. Review of Round 1
- Grade group facilitators review results of Round 1; lead discussion of bookmark placements, impact data, and rationales for bookmark placements
- 2:30 p.m. Completion of Readiness Form Part 4
- 2:35 p.m. Begin Round 2
- Panelists work in small groups as in Round 1
- 4:40 Reflection questionnaire
- 4:50 p.m. Wrap-up for Day 2
- Panelists log out; computers secured
- 5:00 p.m. Adjourn
- MI/CTB staff analyze data and prepare reports.
- Meet with Vertical Articulation Committee panelists.
- 5:30 p.m. Smarter Balanced, MI, CTB, and auditor debriefing
- 6:00 p.m. Secure meeting rooms and depart
- 6:00 p.m. Table Leader Training - Oct. 12th, 14th, 16th
- Facilitator Training

- Day 3**
- 7:30 a.m. Registration opens (Ballroom Lobby)
 - Badges, place cards, and room and table assignments
 - Continental breakfast
 - 8:30 a.m. Convene in breakout rooms
 - Review Round 2 results; grade group facilitators review results of Round 1; lead discussion of bookmark placements, impact data, and rationales for bookmark placements
 - 9:45 a.m. Completion of Readiness Form Part 5
 - 10:00 a.m. Begin Round 3
 - Panelists work in small groups as in Rounds 1 and 2.
 - 11:30 a.m. Wrap-up for Day 3

MI/CTB staff process and share Round 3 results for this group plus others that have completed previously (not applicable to Grade 11; grade 6–8 groups will see only their results and those for Grade 11; grade 3–5 groups may see all results).

11:45 a.m. Completion of Evaluation Forms; log out; all computers secured

Adjourn for lunch (Ballroom Lobby). On Days 3 and 5, outgoing panelists will be joined by incoming panelists. The Parking Lot concept should still work, but there will be different kinds of questions.

12:00 Noon

Panelists may stay for lunch or leave if they have successfully logged out.

MI/CTB staff discuss Round 3 data r with Smarter Balanced staff and auditor, update achievement level setting report.

Comments

Rick Mercado- (time may change on 14 & 16)

Jennifer Lord-Bessen in breakout room

Need checklist for Mandy

Meet in main ballroom for one hour.

PowerPoint for Joe (10 min)

Additional slides for Joe's PPT

Facilitator-led discussion

Facilitator-led discussion

Facilitator-led discussion

PPT

Facilitator led (need talking points)

Reconvene in main ballroom

Facilitator checks all computers; need checklist for facilitators and data analysts.

Need checklists, protocols, discussion topics; brief auditor/SB on expectations.

Need checklist

Dan presents the rest of the Bookmark PPT

Jennifer Lord-Bessen in breakout room
Rick Mercado

Discuss at a Prep meeting. What to include or exclude or emphasize when incoming and outgoing panelists are in the same room.

Smarter Balanced Achievement Level Setting Scripts and PowerPoints

#	Title	Presenter	File	Notes
1	Opening Session Keynote	J. Willhoft	In Person Intro v1 071814.pptx	PPT; Joe welcomes and gives charge
2	Overview/Housekeeping	M. Bunch	In Person Intro v1 071814.pptx	PPT; Mike explains schedule, procedures, shows online supports, room assignments
3	Common Core State Standards	Facilitators	CCSS Script v1 071814.docx	Script; introduction to CCSS; Q&A
4	ALDs	Facilitators	CCSS Script v1 071814.docx	Script; Introduction to ALDs, Q&A, Readiness Part 1
4a	ELA ALDs	Facilitators	Common Core ALDs\ELA 11 ALDs v1 071814.dotx	
4b	Math ALDs	Facilitators		Threshold ALDs for Math
5	Smarter Balanced Tests	Facilitators	Test Development Script v1 071814.d	Script; overview of development; practice test; Readiness Part 2
6	Bookmark Procedure	M. Bunch	Bookmark.pptx	PPT; introduction to the Bookmark procedure
7	Bookmark Practice	Facilitators	Bookmark Script.docx	Script; administration of 6-item practice round; discussion; Readiness Part 3
8	Introduction to Round 1	Facilitators	Bookmark Script.docx	Script; familiarize panelists with program; instructions for Round 1
9	Wrap-Up for Day 1	Facilitators	Bookmark Script.docx	Script; enter 1 cut score; log out
10	Resume Round 1	Facilitators	Bookmark Script.docx	Script; review of Day 1; Q&A; resume
11	Wrap-Up for Round 1	Facilitators	Bookmark Script.docx	Script; enter 3 cut scores; check; log out
12	Review of Round 1	Facilitators	Bookmark Script.docx	Script; presentation of R1 data, OP data; Q&A; Readiness Part 4
13	Introduction to Round 2	Facilitators	Bookmark Script.docx	Script; same directions as Round 1
14	Wrap-Up for Round 2	Facilitators	Bookmark Script.docx	Script; same directions as Round 1
15	Review of Round 2	Facilitators	Bookmark Script.docx	Script; presentation of R2 data; Q&A; Readiness Part 5
16	Introduction to Round 3	Facilitators	Bookmark Script.docx	Script; same directions as Round 2
17	Wrap-Up for Round 3	Facilitators	Bookmark Script.docx	Script; enter 3 cut scores; check; complete evaluation; log out
18	Introduction to Vertical Articulation	M. Bunch	Vertical Articulation.pptx	PPT; Readiness Form

19	Conduct of Vertical Articulation	Bunch/Lewis	VAC SScript v1 071814.docx	Script; start with obvious deviations; follow parliamentary procedure; ad lib
20	Readiness Form	Facilitators	Readiness Form v1 071814.docx	Online Form in 5 parts
21	Presenting External Data	Facilitators	Presenting External Data v1 071814.docx	

In-Person Workshop Participant Demographics

Table A.1.1 In-Person Participants by State/Territory, ELA

State/Territory	3	4	5	6	7	8	11	Total
California	4	3	5	3	6	4	12	37
Connecticut	2	1	3	3	1	2	3	15
Delaware	2	2	1	1	2	2	3	13
Hawaii	1	0	0	1	1	1	3	7
Iowa	1	1	1	0	1	0	3	7
Idaho	2	1	1	2	0	2	1	9
Maine	2	2	2	3	3	3	5	20
Michigan	1	1	1	1	0	1	3	8
Missouri	0	1	0	0	1	1	1	4
Montana	0	1	2	2	1	2	3	11
North Carolina	0	0	1	0	0	0	3	4
North Dakota	0	2	2	1	0	1	3	9
New Hampshire	0	2	1	2	1	0	2	8
Nevada	1	1	0	0	0	1	2	5
Oregon	0	1	2	0	1	1	5	10
South Dakota	0	0	1	0	0	0	2	3
Virgin Islands	1	2	0	3	2	3	1	12
Vermont	1	0	1	0	0	1	0	3
Washington	2	1	1	2	2	1	6	15
Wisconsin	3	2	1	2	3	1	1	13
West Virginia	2	3	1	3	1	2	2	14
Wyoming	1	0	0	1	1	1	4	8
Total	26	27	27	30	27	30	68	235

Table A.1.2 In-Person Participants by State/Territory, Math

State/Territory	3	4	5	6	7	8	11	Total
California	4	0	4	5	3	2	11	29
Connecticut	1	2	1	1	1	3	2	11
Delaware	3	1	1	2	0	2	2	11
Hawaii	1	3	3	1	3	2	2	15
Iowa	2	1	1	2	1	1	2	10
Idaho	1	1	3	1	1	1	3	11
Maine	2	1	2	1	1	1	1	9
Michigan	1	0	0	1	1	0	3	6
Missouri	1	2	1	2	0	1	3	10
Montana	1	2	0	2	1	2	4	12
North Carolina	1	0	1	0	0	1	1	4
North Dakota	1	1	0	3	1	2	3	11
New Hampshire	2	2	1	1	1	2	3	12
Nevada	2	1	1	1	2	0	3	10
Oregon	0	1	1	2	1	2	4	11
South Dakota	0	2	0	0	2	0	3	7
Virgin Islands	1	1	1	1	1	1	0	6
Vermont	1	3	1	0	1	0	1	7
Washington	2	1	2	1	2	1	8	17
Wisconsin	2	2	3	1	2	2	3	15
West Virginia	1	1	1	2	4	2	6	17
Wyoming	0	1	1	0	1	1	2	6
Total	30	29	29	30	30	29	70	247

Table A.1.3 In-Person Participants by Role, ELA

Role	3	4	5	6	7	8	11	Total
Educator: ELL Experience	1	4	3	3	2	2	4	19
Educator: General	19	15	15	11	14	19	23	116
Educator: Non-Teaching	3	3	7	11	5	4	10	43
Educator: SWD Experience	1	2	1	2	3	3	4	16
General Public	2	3	1	3	3	2	1	15
Higher Education	0	0	0	0	0	0	26	26
Total	26	27	27	30	27	30	68	235

Table A.1.4 In-Person Participants by Role, Math

Role	3	4	5	6	7	8	11	Total
Educator: ELL Experience	2	1	4	2	2	3	3	17
Educator: General	14	19	17	20	17	14	25	126
Educator: Non-Teaching	9	5	5	5	6	10	7	46
Educator: SWD Experience	2	3	2	2	2	1	3	15
General Public	3	1	1	1	3	1	3	13
Higher Education	0	0	0	0	0	0	29	30
Total	30	29	29	30	30	29	70	247

Table A.1.5 In-Person Participants by State/Territory and Role

State/Territory	ELA								Math								Grand Total
	3	4	5	6	7	8	11	Total	3	4	5	6	7	8	11	Total	
CA	4	3	5	3	6	4	12	37	4	0	4	5	3	2	11	29	66
Educator: ELL experience	1	1	2	1	0	2	2	9	2	0	3	2	1	2	2	12	21
Educator: General	2	1	2	0	4	1	4	14	0	0	1	2	0	0	1	4	18
Educator: Non-Teaching	0	0	1	1	1	0	2	5	1	0	0	0	1	0	2	3	8
Educator: SWD experience	1	0	0	1	0	1	1	4	1	0	0	1	1	0	0	3	7
General Public	0	1	0	0	1	0	0	2	0	0	0	0	0	0	2	2	4
Higher Education	0	0	0	0	0	0	3	3	0	0	0	0	0	0	4	5	8
CT	2	1	3	3	1	2	3	15	1	2	1	1	1	3	2	11	26
Educator: General	2	1	0	2	0	1	1	7	1	2	1	1	1	1	1	8	15
Educator: Non-Teaching	0	0	2	1	0	0	1	4	0	0	0	0	0	2	0	2	6
Educator: SWD experience	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1
General Public	0	0	1	0	0	1	0	2	0	0	0	0	0	0	0	0	2
Higher Education	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	2
DE	2	2	1	1	2	2	3	13	3	1	1	2	0	2	2	11	24
Educator: ELL experience	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	1	2
Educator: General	1	1	1	0	1	1	2	7	1	1	0	2	0	2	1	7	14
Educator: Non-Teaching	0	0	0	1	0	0	1	2	1	0	0	0	0	0	0	1	3
Educator: SWD experience	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1
General Public	1	0	0	0	1	0	0	2	1	0	0	0	0	0	0	1	3
Higher Education	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
HI	1	0	0	1	1	1	3	7	1	3	3	1	3	2	2	15	22
Educator: ELL experience	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1
Educator: General	0	0	0	0	0	1	0	1	1	1	3	1	3	1	1	11	12
Educator: Non-Teaching	1	0	0	0	1	0	1	3	0	0	0	0	0	0	0	0	3
Educator: SWD experience	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1
General Public	0	0	0	1	0	0	0	1	0	1	0	0	0	1	0	2	3
Higher Education	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	2
IA	1	1	1	0	1	0	3	7	2	1	1	2	1	1	2	10	17
Educator: General	1	1	1	0	0	0	1	4	1	0	0	1	1	0	1	4	8
Educator: Non-Teaching	0	0	0	0	1	0	0	1	1	1	1	1	0	1	0	5	6
Higher Education	0	0	0	0	0	0	2	2	0	0	0	0	0	0	1	1	3
ID	2	1	1	2	0	2	1	9	1	1	3	1	1	1	3	11	20
Educator: ELL experience	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
Educator: General	2	0	1	2	0	2	0	7	0	1	1	1	1	1	1	6	13
Educator: Non-Teaching	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	3	3
General Public	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1
Higher Education	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	2
ME	2	2	2	3	3	3	5	20	2	1	2	1	1	1	1	9	29
Educator: ELL experience	0	0	1	0	1	0	1	3	0	0	0	0	0	0	0	0	3
Educator: General	2	1	1	2	2	3	2	13	2	1	0	1	1	1	0	6	19
Educator: Non-Teaching	0	1	0	1	0	0	1	3	0	0	1	0	0	0	0	1	4
Educator: SWD experience	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1
Higher Education	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	2
MI	1	1	1	1	0	1	3	8	1	0	0	1	1	0	3	6	14
Educator: General	1	1	1	1	0	1	0	5	1	0	0	1	1	0	1	4	9
Higher Education	0	0	0	0	0	0	3	3	0	0	0	0	0	0	2	2	5
MO	0	1	0	0	1	1	1	4	1	2	1	2	0	1	3	10	14
Educator: General	0	1	0	0	1	0	0	2	1	2	1	1	0	1	1	7	9
Educator: Non-Teaching	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1
Educator: SWD experience	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1
General Public	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1
Higher Education	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	2

State/Territory	ELA								Math								Grand Total
	3	4	5	6	7	8	11	Total	3	4	5	6	7	8	11	Total	
MT	0	1	2	2	1	2	3	11	1	2	0	2	1	2	4	12	23
Educator: ELL experience	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
Educator: General	0	1	1	1	0	2	1	6	1	1	0	1	0	0	2	5	11
Educator: Non-Teaching	0	0	0	1	1	0	1	3	0	0	0	1	0	2	0	3	6
Educator: SWD experience	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	1	2
General Public	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1
Higher Education	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	2
NC	0	0	1	0	0	0	3	4	1	0	1	0	0	1	1	4	8
Educator: General	0	0	1	0	0	0	2	3	0	0	1	0	0	1	0	2	5
Educator: Non-Teaching	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1
Higher Education	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	2
ND	0	2	2	1	0	1	3	9	1	1	0	3	1	2	3	11	20
Educator: ELL experience	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1
Educator: General	0	2	2	0	0	1	2	7	1	1	0	3	1	2	2	10	17
Higher Education	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	2
NH	0	2	1	2	1	0	2	8	2	2	1	1	1	2	3	12	20
Educator: General	0	2	0	1	0	0	1	4	1	1	1	0	1	1	2	7	11
Educator: Non-Teaching	0	0	1	0	1	0	0	2	1	1	0	1	0	1	0	4	6
Educator: SWD experience	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1
General Public	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1
Higher Education	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
NV	1	1	0	0	0	1	2	5	2	1	1	1	2	0	3	10	15
Educator: General	1	1	0	0	0	0	0	2	0	1	1	1	0	0	1	4	6
Educator: Non-Teaching	0	0	0	0	0	0	1	1	1	0	0	0	2	0	1	4	5
Educator: SWD experience	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1
General Public	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1
Higher Education	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	2
OR	0	1	2	0	1	1	5	10	0	1	1	2	1	2	4	11	21
Educator: ELL experience	0	1	0	0	0	0	1	2	0	0	0	0	0	0	0	0	2
Educator: General	0	0	2	0	1	1	0	4	0	1	1	1	0	1	0	4	8
Educator: Non-Teaching	0	0	0	0	0	0	1	1	0	0	0	0	1	1	2	4	5
General Public	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1
Higher Education	0	0	0	0	0	0	3	3	0	0	0	0	0	0	2	2	5
SD	0	0	1	0	0	0	2	3	0	2	0	0	2	0	3	7	10
Educator: ELL experience	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1
Educator: General	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	3	3
Educator: Non-Teaching	0	0	1	0	0	0	1	2	0	1	0	0	0	0	1	2	4
Higher Education	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	2
VI	1	2	0	3	2	3	1	12	1	1	1	1	1	1	0	6	18
Educator: ELL experience	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1
Educator: General	1	1	0	0	1	2	1	6	0	1	0	1	0	0	0	2	8
Educator: Non-Teaching	0	1	0	3	0	1	0	5	1	0	1	0	0	1	0	3	8
General Public	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1
VT	1	0	1	0	0	1	0	3	1	3	1	0	1	0	1	7	10
Educator: General	1	0	0	0	0	0	0	1	1	2	1	0	1	0	0	5	6
Educator: Non-Teaching	0	0	1	0	0	1	0	2	0	1	0	0	0	0	0	1	3
Higher Education	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
WA	2	1	1	2	2	1	6	15	2	1	2	1	2	1	8	17	32
Educator: General	1	0	1	1	2	0	2	7	1	0	1	1	2	1	2	8	15
Educator: Non-Teaching	0	1	0	1	0	1	0	3	0	0	1	0	0	0	0	1	4
Educator: SWD experience	0	0	0	0	0	0	0	0	1	1	0	0	0	0	2	4	4
General Public	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
Higher Education	0	0	0	0	0	0	4	4	0	0	0	0	0	0	4	4	8
WI	3	2	1	2	3	1	1	13	2	2	3	1	2	2	3	15	28
Educator: General	3	0	0	0	1	1	0	5	0	1	2	1	0	0	2	6	11

State/Territory	ELA								Math								Grand Total
	3	4	5	6	7	8	11	Total	3	4	5	6	7	8	11	Total	
Educator: Non-Teaching	0	0	1	0	0	0	0	1	1	0	0	0	1	1	0	3	4
Educator: SWD experience	0	2	0	1	2	0	1	6	0	1	1	0	1	1	1	5	11
General Public	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	1	2
WV	2	3	1	3	1	2	2	14	1	1	1	2	4	2	6	17	31
Educator: ELL experience	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	2
Educator: General	0	1	1	1	0	2	1	6	1	0	1	0	2	1	4	9	15
Educator: Non-Teaching	2	0	0	1	0	0	0	3	0	1	0	2	1	0	0	4	7
General Public	0	2	0	0	1	0	0	3	0	0	0	0	1	0	1	2	5
Higher Education	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	2
WY	1	0	0	1	1	1	4	8	0	1	1	0	1	1	2	6	14
Educator: General	1	0	0	0	1	0	3	5	0	1	1	0	1	0	1	4	9
Educator: Non-Teaching	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	2
General Public	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1
Higher Education	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	2
Grand Total	26	27	27	30	27	30	68	235	30	29	29	30	30	29	70	247	482



Achievement Level Setting

Dallas, Texas
October 13–20, 2014

Facilitator Materials

TABLE OF CONTENTS

Facilitator Script Overviews

Day 1 AM Content Standards Review, ALDs Review, and Practice Test	678
Day 1 PM Software and Ordered Item Booklet Review	679
Day 2 AM Bookmark Placement – Orientation and Round 1	680
Day 2 PM Round 2 Bookmark Placement.....	682
Day 3 AM Round 3 Bookmark Placement.....	683

Full Facilitator Script

Day 1 AM Content Standards Review, ALDs Review, and Practice Test	684
The Smarter Balanced Assessment Framework.....	686
CCSS Review	690
Achievement Level Descriptor Review.....	691
Practice Test Administration	692
Completion of Readiness Form Part 1.....	693
Security Issues	694
Day 1 PM Software and Ordered Item Booklet Review	695
Item Map Navigation.....	696
OIB Navigation.....	698
Presenting External Data (Grades 4, 8, and 11 Only).	701
Presenting Performance Tasks	702
Reviewing the OIB	703
Day 2 AM Bookmark Placement– Orientation and Round 1	705
Day 2 PM Round 2 Bookmark Placement.....	711
Pre-Session Table Leader Meeting.....	711
Item Map Review.....	711
Round 2	714
Round 2 Wrap-up	714
Day 3 AM Round 3 Bookmark Placement.....	715
Review of Round 2	715
Readiness Form and Round 3	719
Review of Round 3.....	719
Evaluation Form	720
Wrap-Up and Closeout.....	721
Appendix A Presenting External Data to Achievement Level Setting Panelists	722
Background	722

The Bookmark Procedure	722
The Benchmark Method	723
The Briefing Book Method	724
Selecting, Evaluating, and Presenting Data	725
Selecting data.	725
Evaluating Data.....	725
Presenting Data.....	727
Feedback and Discussion	730
Conclusion	730
References	732
Appendix B Presentation of Sample NAEP and PISA Items.....	733
Sample 8 th Grade NAEP Reading Item with Data.....	733
Sample PISA Mathematics Item with Data.....	734

FACILITATOR SCRIPT OVERVIEW

Day 1 AM

Content Standards Review, ALDs Review, and Practice Test

Goal: Review content standards and threshold ALDs. Allow panelists to experience the assessment online so they can see how the various item types operate with the technology and empathize with the students.

What you need: A panelist login that has never been used for this grade and content for computer attached to projector. Paper copies of ALDs. Smarter Balanced Assessment Framework slides.

Outline of session: Facilitator leads group introductions. Everyone logs in as a panelist. Facilitator walks through content standards, leads brief discussion, and provides time for individual review and repeats the process for ALDs. Panelists access Smarter Balanced website and take practice test. Panelists complete Readiness Form Part 1.

Highlights

- *Make sure you have the correct panelists signed in and at the correct tables.*
- *Model logging in with your assigned panelist login, explaining where to find their event code (name tag) and table number (card on table).*
- *Show them how to access the CCSS. Review material mentioned in morning session with them. Time spent here should be minimal because they should already be familiar with the content standards and the focus of their efforts should be on the next section.*
- *Show them how to access the ALDs and discuss threshold ALDs. The more they discuss the ALDs now, the easier it will be to use them later on. Give them the paper copies of the ALDs.*
- *Grades 3-8: Level 3 is considered On Target. Levels 3 & 4 combined represent all the students who are considered ready for the next phase of education or career.*
- *Grade 11: Level 3 is Conditionally Content-Ready/Exempt from Developmental Course Work. Only Level 4 is considered Content-Ready. This distinction is critical in the setting of the grade 11 cut score for Levels 3 and 4.*
- *Encourage empathy with students and exploration of new item types.*
- *Have panelists log out before leaving for lunch. Otherwise, software appears to be logged in but is not recording information after lunch.*

Schedule

9:35 a.m.	Introductions
9:45	Orientation to hardware and software
10:00	Discussion of CCSS and ALDs
10:30	Break
10:45	Practice test administration
11:55	Discussion of practice test
12:20 p.m.	Closing comments on tests
12:25	Completion of Readiness Form Part 1
12:30	Lunch

Full script: Page 684

FACILITATOR SCRIPT OVERVIEW

Day 1 PM Software and Ordered Item Booklet Review

Goal: First, introduce navigating the ordered item booklet with the software and train panelists on how to use the two focus questions to analyze each item. Then, review the entire OIB using those two questions.

What you need: Panelist login for computer attached to projector. Slide showing the two questions.

Outline of session: Everyone logs back in as a panelist. Facilitator leads them to OIB in the software and walks through explanation of item map and ordered item booklet navigation and introduces two focus questions. Facilitator leads discussion through the first few items, modeling how to use the focus questions and add their answers to these questions in the Comments section. Identify items from NAEP and PISA and indicate they are not on the tests students will take but are provided for context. Panelists spend the remaining time working through the remaining items in the OIB.

Focus Questions

1. *What do you know about a student who responds successfully to this item; that is, what skills must a student have in order to know the correct answer?*
2. *What makes this item more difficult than preceding items?*

Highlights

- *Make sure the panelists are entering the appropriate full OIB and not the 6-item Orientation Round OIB.*
- *Avoid discussion of bookmarks. Focus should solely be on the items and the two questions. They will receive bookmark training later.*
- *Grades, 4, 8 and 11: Note NAEP (4, 8, 11) and PISA (11) items.*
- *All grades: note placement and connectedness of performance task components.*
- *Keep working through the two focus questions until it appears that the group understands how to use them. Then have the table leaders reinforce.*
- *Remind the panelists of bookmark training the next morning. Emphasize being on time.*

Schedule

1:50 p.m.	Introduce the OIB and item map
2:00	Lead panelists through focus questions
2:20	Model how to use focus with sample of items
2:40	Panelists review items in OIB and enter comments on item map
4:50	Wrap up
5:00	Adjourn

Full Script: Page 695

FACILITATOR SCRIPT OVERVIEW

Day 2 AM

Bookmark Placement – Orientation and Round 1

Goal: Practice setting the Level 3 bookmark in the 6-item orientation round and then set all three Round 1 bookmarks.

What you need: Panelist login for computer attached to projector. Second computer with facilitator login. The two focus question slide and the Bookmark placement slide. Excel spreadsheet to calculate medians based on room's frequency distribution of bookmark placements.

Outline of session: Everyone logs back in as a panelist. Facilitator reviews how to set a Level 3 bookmark with the software. Panelists discuss 6 items using 2 focus questions. Facilitator discusses the Bookmark placement slide and then panelists set their Level 3 bookmark. Facilitator collects bookmark placements into Excel spreadsheet and leads discussion into reasoning behind panelists' bookmark placements. Panelists complete Readiness Form and then Facilitator directs panelists to set Round 1 bookmarks.

Focus Questions

1. *What do you know about a student who responds successfully to this item; that is, what skills must a student have in order to know the correct answer?*
2. *What makes this item more difficult than preceding items?*
3. *Would a student at the threshold of the achievement level have at least a 50% chance of responding successfully to this item (or score point)? If they would, then move on to the next item; if not, place a bookmark on this page.*

Highlights for Orientation Round

- *Do not actually submit any bookmarks when modeling the process for panelists as they can affect the data collected. If you accidentally submit, please notify the operations room.*
- *Ask them why the item before their bookmark fits the Level 3 threshold descriptor and why the bookmarked page does not.*
- *Reinforce we don't expect everyone to always agree.*
- *Keep focus on Bookmark Placement poster and ALDs.*

Highlights for Round 1

- *Remind the panelists to look at their comments from yesterday regarding the answers to the two questions for each item.*
- *Remind panelists of nature and purpose of NAEP/PISA items (grades 4, 8, and 11 only) and performance tasks.*
- *Keep focus on Bookmark Placement poster and ALDs.*
- *Make sure to keep an eye on completion status in facilitator login, close Round 1 when everyone is finished, and notify Operations Room.*
- *Arrange for table leaders to return 15 minutes early from lunch.*

Schedule

9:00 a.m.	Orientation round exercise
9:40	Completion of Readiness Form Part 2
9:45	Round 1
11:45	Completion of Reflection Questionnaire Part 1
12:00 Noon	Lunch

Full Script: Page 705

FACILITATOR SCRIPT OVERVIEW

Day 2 PM Round 2 Bookmark Placement

Goal: Review Round 1 results. Set Round 2 bookmarks.

What you need: Panelist login for computer attached to projector. Second computer with facilitator login. Bookmark Placement slide.

Outline of session: Facilitator meets with table leaders before session to explain how to run table discussion. Everyone else returns and logs back in as a panelist. Facilitator reviews additions (thumbtacks, location, ACT) to item map. Table leaders lead discussion of all panelist bookmarks, starting at Level 3 and direct discussion on rationale behind different placements at a table level. Discuss Online Panel median bookmark slide. Table leaders finish table discussion including new information. Panelists complete Round 2 readiness form. Facilitator displays Bookmark Placement slide and panelists set their Round 2 bookmarks.

Highlights

- *Review with table leaders how to lead discussion at their table.*
- *New items: thumbtacks, location, ACT.*
- *Direct tables to share their bookmarks at their tables and note low, high, and median BMs.*
- *Display online panelist median bookmarks as a point of reference and provide context.*
- *Review setting the bookmark on item after which they stop having at least a 50% chance of correctly answering the item.*
- *Make sure to keep an eye on completion status in facilitator login, close Round 2 when everyone is finished, and notify Operations Room.*

Schedule

1:00 p.m.	Review of Round 1
2:30	Completion of Readiness Form Part 3
2:35	Round 2
4:50	Completion of Reflection Questionnaire Part 2
5:00	Adjourn

Full Script: Page 711

FACILITATOR SCRIPT OVERVIEW

Day 3 AM Round 3 Bookmark Placement

Goal: Review Round 2 results including table and room medians and impact data. Share online panel results again, now with impact data and complete readiness form. Set Round 3 bookmarks. Break. Review Round 3 results and complete evaluation form.

What you need: Panelist login for computer attached to projector. Second computer with facilitator login. Results of Round 2 slides. Online impact data slides. Bookmark placement slide.

Outline of session: Everyone logs back in as a panelist. Facilitator walks through at a room level the results of Round 2 for the room (histogram), each table median, the room median, and the associated impact data with a discussion of the meaning of it and how it can be used (30 minutes). Facilitator reviews the range of bookmarks at each Level, using the different table medians as a focus, and leads discussion on rationale behind different placements at a room level (60 minutes). Panelists complete Round 3 readiness form. Facilitator displays Bookmark Placement slide and panelists set their Round 3 bookmarks. Participants break and return for review of final recommendations. Panelists complete final evaluation form.

Highlights

- *You are explaining impact data and how they can use it. Remind them that they are the ones to decide if they will use it.*
- *Discussion is at a room level in this round.*
- *Review setting the bookmark on first item on which they stop having at least a 50% chance of correctly answering the item.*
- *Make sure to keep an eye on completion status in facilitator login, close Round 3 when everyone is finished, and notify Operations Room.*

Schedule

8:30 a.m.	Convene in breakout rooms; review Round 2 results
9:45	Completion of Readiness Form Part 4
9:50	Round 3
11:30	Review of Round 3 Results
11:40	Completion of Evaluation Forms and Reflection Questionnaire Part 3
12:00	Adjourn/Lunch

Full Script: Page 715

FACILITATOR SCRIPT

Note: This script contains screen shots, data displays, and other information that will be updated as they are available. Discussions of data are also illustrative of what might be said during Bookmark training and will also be updated to reflect actual data as they become available.

Day 1 AM Content Standards Review, ALDs Review, and Practice Test

Make sure all panelists have signed the sign-in sheet and are seated at their correct tables (see your seating chart). Welcome the panelists, beginning with introductions. Start with yourself, then allow each panelist to introduce him or herself, providing name, state, and role. This should take approximately 15 seconds per panelist, and it will begin the community building that will develop during the discussions that will take place later. Let panelists know that they were seated at specific tables to achieve a balance of expertise.

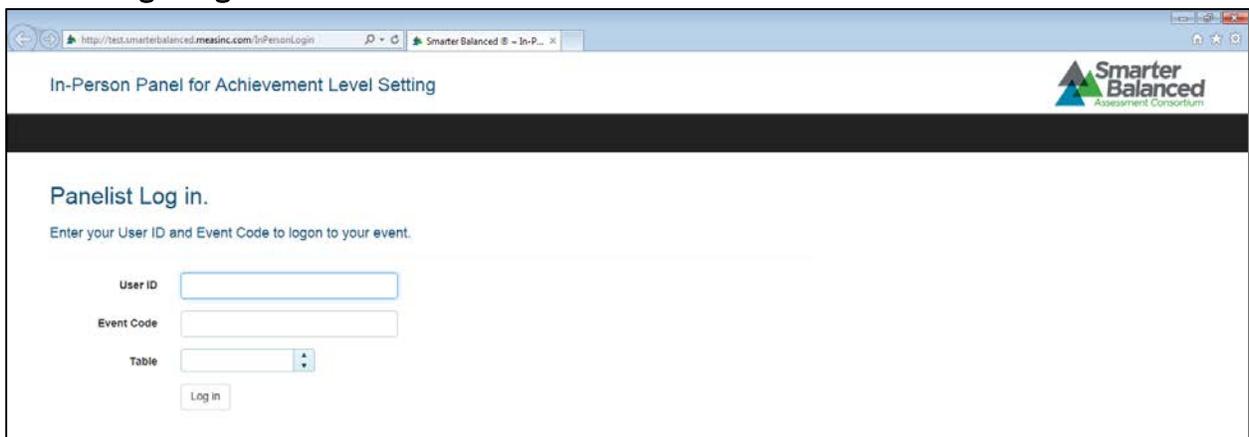
Next, instruct panelists to sign and complete their Non-Disclosure Agreement and Multimedia Release Form. Both of these forms were previously distributed to panelists electronically, but they will now fill out a paper copy.

Please remove the ivory Non-Disclosure Agreement and blue Multimedia Release Form from your registration packet. You will need to sign both forms before we can proceed. Please pass to your table leader upon completion.

After all completed forms have been collected and put into the envelope provided to you, then say,

For the next several minutes we are going to take a look at the software program you will be using to carry out the tasks you just learned about. Please make sure your computers are turned on and that you see the login screen [refers to session preceding first breakout]. To login, enter your user id without any spaces, the event code listed on the back of your name tag, and your table number. [Model this on the projector with a new panelist login. Your event code will be different than the panelists so remind them to use their event code not yours.]

Panelist Login Page

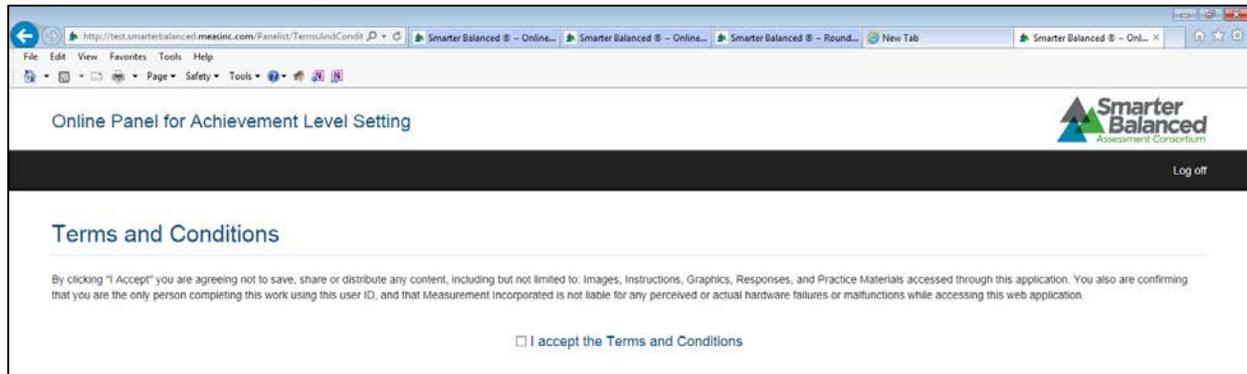


The screenshot shows a web browser window with the URL <http://test.smarterbalanced.measinc.com/InPersonLogin>. The page title is "In-Person Panel for Achievement Level Setting" and features the Smarter Balanced Assessment Consortium logo. The main heading is "Panelist Log in." followed by the instruction "Enter your User ID and Event Code to logon to your event." The form contains three input fields: "User ID" (text), "Event Code" (text), and "Table" (dropdown menu). A "Log in" button is located below the "Table" field.

Verify that all panelists were able to log in before proceeding.

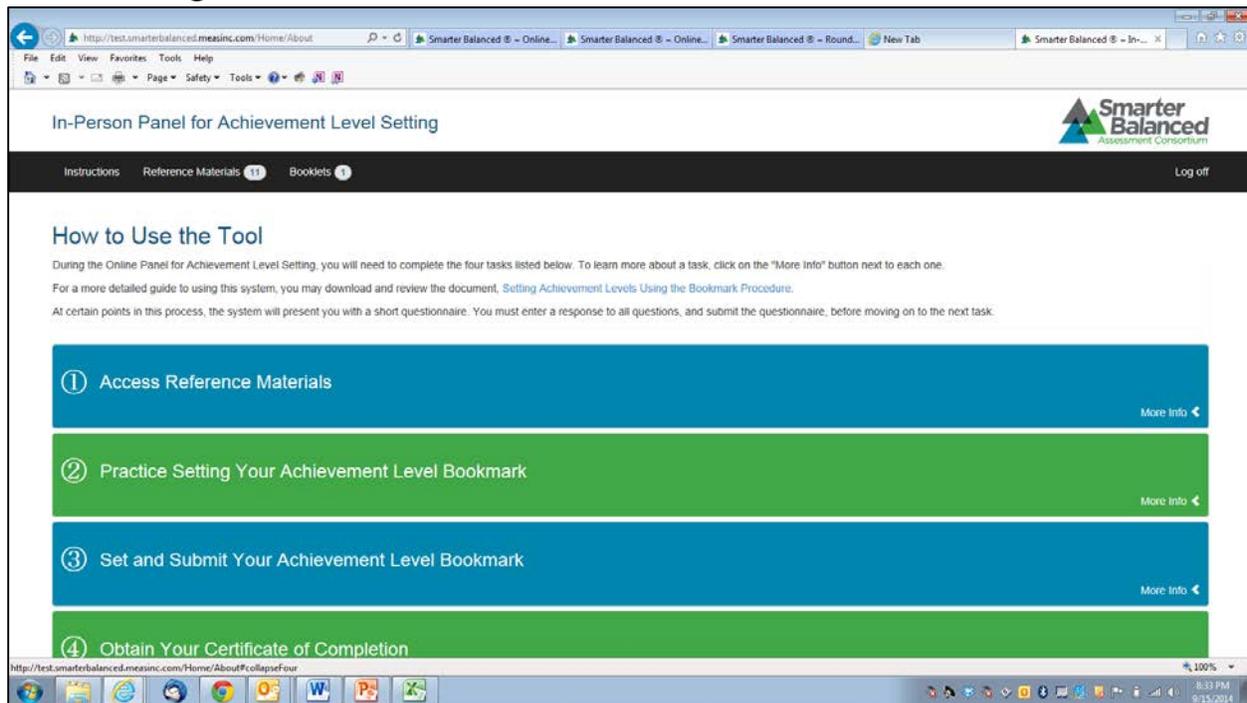
The first page that you see should be the “Terms and Conditions” page. Please read the page and then click the check box when you are ready.

Terms and Conditions Page



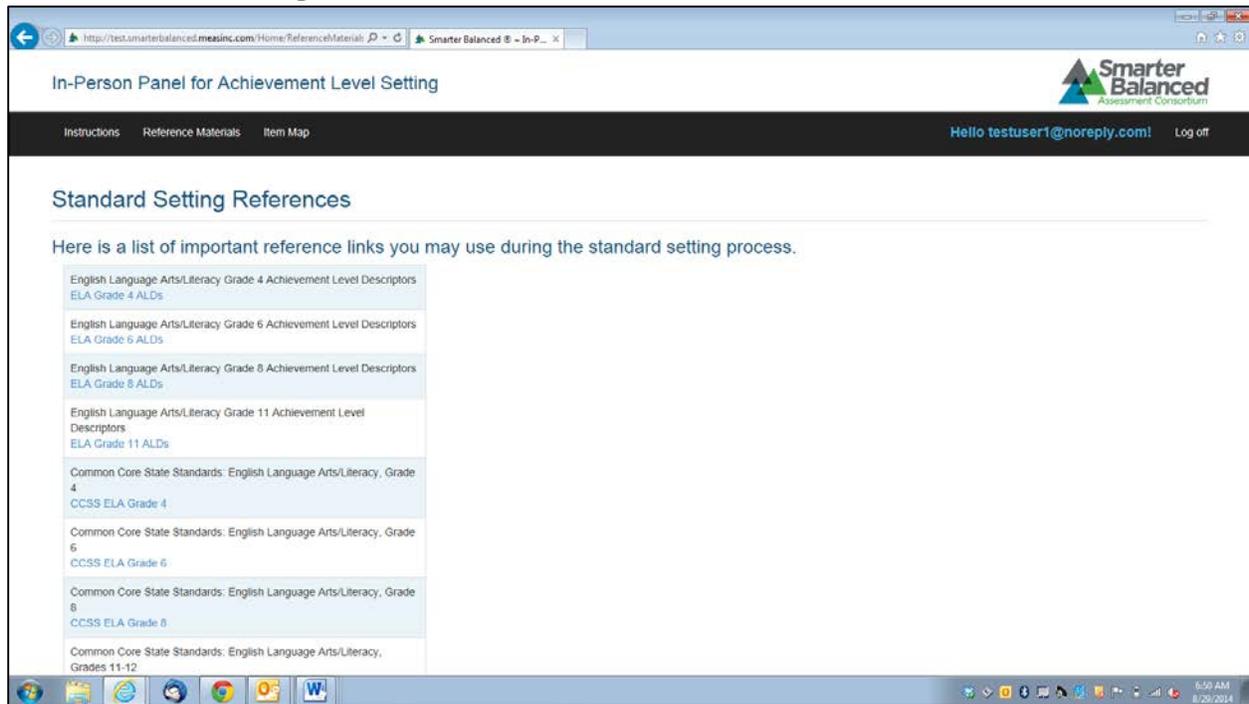
The next page you will see is “How to Use the Tool.” It describes how the software operates. I will be walking you through demonstrations of the steps discussed here, but should you want to see this page again, simply click on the “Instructions” link in the black bar at the top. There are four tasks to be completed over the next three days. We will start with the first.

Instructions Page



Click on the first blue bar “Access Reference Materials” and you will see a description of the first task. It shows you how to access all the reference materials that will be used at this standard setting. This morning we will be focusing on the standards and the ALDs for [insert appropriate grade and content].

Reference Materials Page



In-Person Panel for Achievement Level Setting

Standard Setting References

Here is a list of important reference links you may use during the standard setting process.

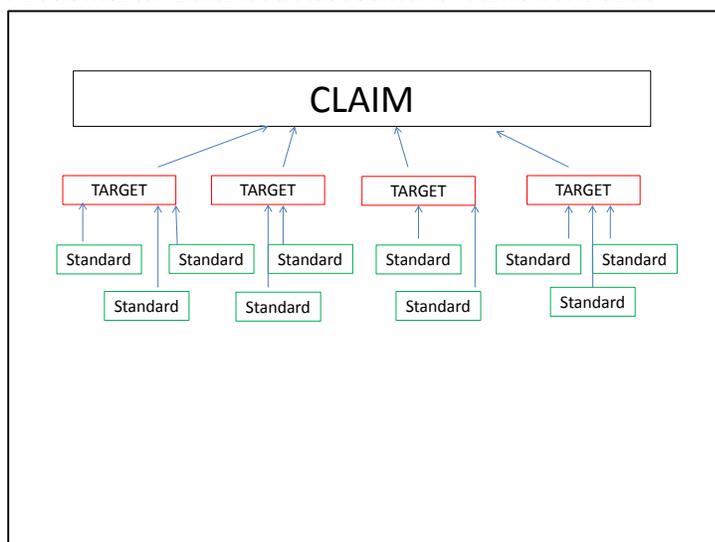
- English Language Arts/Literacy Grade 4 Achievement Level Descriptors
ELA Grade 4 ALDs
- English Language Arts/Literacy Grade 6 Achievement Level Descriptors
ELA Grade 6 ALDs
- English Language Arts/Literacy Grade 8 Achievement Level Descriptors
ELA Grade 8 ALDs
- English Language Arts/Literacy Grade 11 Achievement Level Descriptors
ELA Grade 11 ALDs
- Common Core State Standards: English Language Arts/Literacy, Grade 4
CCSS ELA Grade 4
- Common Core State Standards: English Language Arts/Literacy, Grade 6
CCSS ELA Grade 6
- Common Core State Standards: English Language Arts/Literacy, Grade 8
CCSS ELA Grade 8
- Common Core State Standards: English Language Arts/Literacy, Grades 11-12

The Smarter Balanced Assessment Framework

There is some animation in the next PowerPoint slides. Within this section, the * represents pressing the Enter key to advance the animation.

Display first assessment slide on the projector.

First Smarter Balanced Assessment Framework Slide



As with any framework, there is an organizational hierarchy. In the Smarter Balanced assessment framework, the claims and the targets are the most important hierarchical components of the framework.

The claims are overarching statements within each content area; the assessments are written to elicit evidence so that inferences can be made to support the claims. Each claim—and there are four for ELA and four for math—focuses on one domain within a content area, such as reading for ELA or concepts and procedures for math. The specific wording of the reading claim states that “Students can read closely and analytically to comprehend a range of increasingly complex literary and informational texts.” Claims are global statements of expected academic performance.

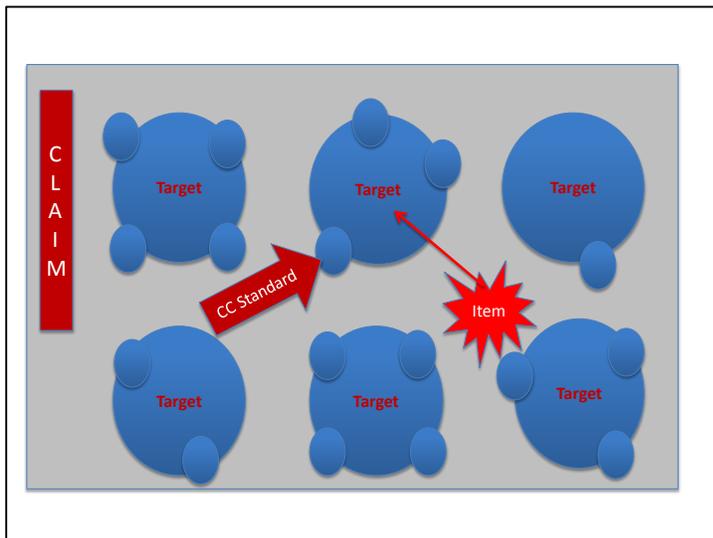
The next level of the Smarter Balanced Framework is that of the Targets. Targets further categorize and organize areas of focus relative to each claim. For example, one Target of the Reading claim focuses on “Central Ideas.” For each claim there may be as few as 1 or 2 targets and as many as 14.

All of you are very familiar with the Common Core standards, and you know that assessments have been written to measure student performance relative to those standards. So, where do the Common Core standards fit into the Smarter Balanced framework?

The Targets serve to organize the Common Core standards. Each Target is a cluster of related Common Core standards. The standards help define the content that is assessable within each target. All Smarter Balanced items are written to the Target; the standards are part of the item specifications that item writers use when they write their “targeted” items for the Smarter Balanced assessments.

Display second assessment slide on the projector.

Second Smarter Balanced Assessment Framework Slide



Here’s another way to look at the components of the Smarter Balanced framework.

Imagine you’re at a conference where there are multiple meeting rooms. Each room is focused on one Smarter Balanced claim. There’s a room focused on reading, another on writing, another on concepts and procedures, etc. You get the idea.

In each meeting room are tables, and each table represents a target. Each “participant” at a table represents one of the Common Core standards. Some tables have a couple of standards present, and some have more.

The participants at the target tables have been tasked with writing assessment items together. They will collaborate, and their end result will be items that are written to measure their target. In some cases, then, one participant may contribute more than another. So, an item may actually measure all or part of one standard or parts of more than one standard, but each item will be written to elicit evidence about what a student knows and can do relative to the target.

Display third assessment slide on the projector.

Third Smarter Balanced Assessment Framework Slide

Achievement Level Descriptors

- **Policy ALDs**
 - general descriptors that articulate the goals and rigor for the final performance standards
- **Range ALDs**
 - written to describe expectations for each TARGET at each proficiency level
 - grade- and content-specific descriptors that may be used by test developers to guide item writing
 - describe the cognitive and content rigor that is encompassed within particular achievement levels
- **Threshold ALDs**
 - developed in conjunction with or following range ALDs and are used to guide standard setting
 - are a SUBSET of the range ALDs
 - use only the information from the range ALDs that defines the minimum performance required for meeting a particular achievement-level expectation

Let’s talk now about how the standard setting process relates to the Smarter Balanced framework.

*The achievement level descriptors have their own hierarchy as well. The three levels we will focus on here are the * policy ALDs, * the range ALDs, and the *threshold ALDs. The later will become well known to you during the standard setting process.*

*The policy ALDS are * those descriptors that articulate the goals and rigor for the final performance standards. They are written at a broad level for each content area.*

*The range ALDs have been * written to describe the expectations for each target—the ones we just discussed—and for each of the four proficiency levels. So, what can a student performing at level 3 do that a student at level 1 cannot? The range ALDS are also *grade- and content-specific descriptors that may be used by test developers to guide item writing, and they *describe the cognitive and content rigor that is encompassed within particular achievement levels. In other words, the range ALDs describe the kind of performance, or evidence that a student will show, for EACH target and EACH proficiency level. The educator groups who developed and wrote the Range ALDs relied heavily on the language of each target as well as the language and expectations of the standards that were aggregated within each target. All of the ALDs are public documents, and you will find copies of the full documents in each meeting room to use as reference.*

Finally, and most importantly for your work, are the threshold ALDs. These were developed immediately after the range ALDs were and will be used to guide this week's standard setting decisions. They are a SUBSET of the range ALDs that use only the information from the range ALDs that defines the minimum performance required for meeting a particular achievement-level expectation. They describe the performance of the student who just crosses the line into a proficiency level.

Display fourth assessment slide on the projector.

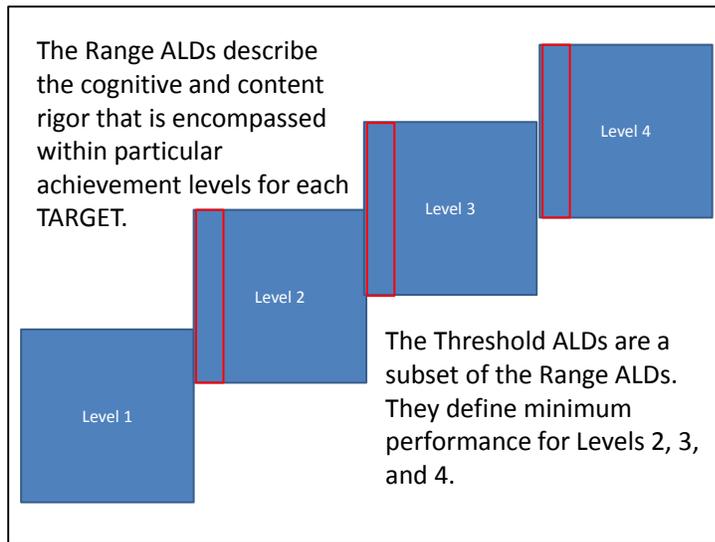
Fourth Smarter Balanced Assessment Framework Slide



So the Smarter Balanced assessment framework was created to guide what is to be assessed. The Achievement Level Descriptors were developed to describe performance levels and expectations. Extremely clear and detailed item specifications further guided item development. And then the items were field tested.

Display fifth assessment slide on the projector.

Fifth Smarter Balanced Assessment Framework Slide



*This week we are using the data from the field test to set cut scores for each proficiency level. *****

*Once again, * the Range ALDs describe the cognitive and content rigor that is encompassed within particular achievement levels. So, for each TARGET in the Smarter Balanced assessment framework there are descriptions of all the things a student is expected to do relative to that target (and its aggregated standards) for each proficiency level.*

**The Threshold ALDs are a subset of the Range ALDs. They define minimum performance for Levels 2*, 3*, and 4*. It is these areas in red that will be your primary focus when setting your book marks. These areas describe minimal performance at that level.*

CCSS Review

This morning, only the CCSS and ALDs applications are active. We will activate the others over the next two days. Take a moment now to click on CCSS. You will see 14 different sets of content standards. Click on the one for [grade/subject].

Pause to make sure everyone has found the appropriate set of content standards; then say:

As you heard in the introductory session, the questions on the Smarter Balanced Assessments are derived from the Common Core State Standards. A few weeks ago, you received an e-mail with the [grade/subject] portion of the standards attached. I hope you spent some time looking those over. Now, we want you to take another look at these standards, along with some sample items, and discuss them with others at your table. Then we will discuss them as a group.

Check to make sure all panelists have the content standards on their screens; then say:

Review the content standards for [grade/subject], noting the structure of domains and then content standards within each domain. Discuss with others at your table what these content standards mean to you in terms of what you believe students are expected to know and be able to do within each domain.

At the end of your content standards, you will find a section about Depth of Knowledge (DOK). Most of the content standards are broad enough to require students to perform at two or more cognitive levels; for example recall and application. Review the DOK and discuss with others at your table. I will call time on your conversations in about 15 minutes and ask each table to give a brief summary of their discussion. Are there any questions about what you need to do for the next 15 minutes?

Answer any questions about the process.

Please begin.

Circulate through the room to get a sense of the direction of each table's conversation. Answer questions about the process and goal of the activity, but do not offer interpretations of the standards. Instead, save questions about interpretation for the room-wide discussion. After 15 minutes (or whenever conversation appears to be winding down, whichever comes first), call time. Then ask for someone at each table to summarize the conversation for that table. Many of the summaries will be quite similar. If questions regarding interpretation of specific standards have arisen, share these with the whole group. It may be helpful to ask for an interpretation from the panel before offering one yourself. If no one volunteers quickly, answer the question yourself. Remember that this activity has two objectives: understanding the content standards and building a group.

Achievement Level Descriptor Review

Direct panelists to close the CCSS application and open the ALD application. Pause to make sure everyone has opened the application. Key points to cover:

- Distribute hard copy ALDs
- Point out that ALDs will serve as the principal guidance in setting cut scores.
- ALDs describe what students just entering Level 2, 3, or 4 know or can do.

*As with the content standards, I am going to ask you to review the achievement level descriptors with others at your table and talk about what they mean in terms not only of what students should know and be able to do relative to the content standards but **how much** they should know and be able to do when they are just entering Level 2, Level 3, and Level 4. I want you to take about 10-15 minutes to review the descriptors and discuss them with others at your table. If you have questions, pose them to the others at your table. I will be circulating through the room; if there are questions others at your table cannot answer, ask me. I won't answer them immediately, but I will collect them and share the answers with the rest of the room. Are there any questions about what you need to do for the next 10-15 minutes?*

Answer questions about process, then say:

Please begin.

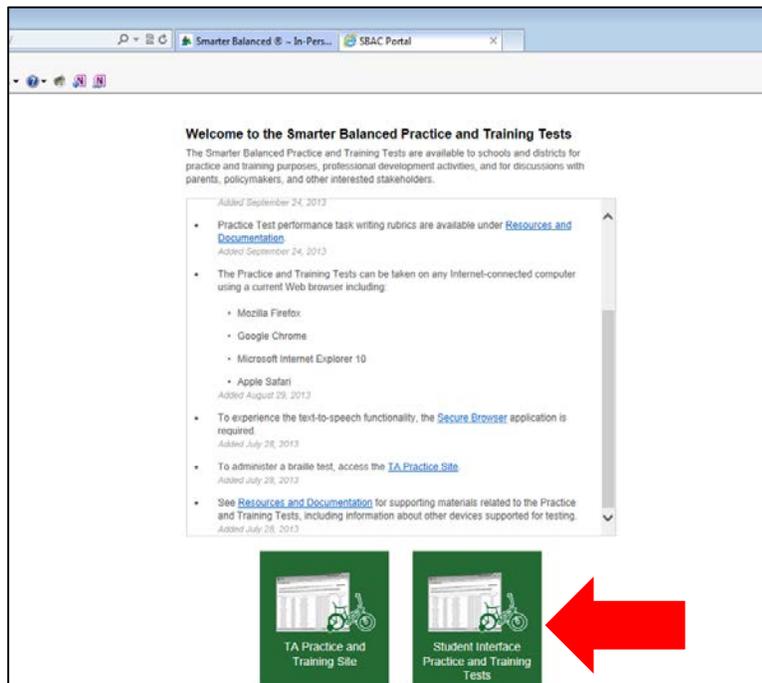
Circulate through the room to get a sense of the direction of each table's conversation. Answer questions about the process and goal of the activity, but do not offer interpretations of the ALDs. Instead, save questions about interpretation for the room-wide discussion. After 10-15 minutes (whenever conversation appears to be winding down, whichever comes first), call time. Then ask for someone at each table to summarize the conversation for that table. Many of the summaries will be quite similar. If questions regarding interpretation of specific ALDs have arisen, share these with the

whole group. It may be helpful to ask for an interpretation from the panel before offering one yourself. If no one volunteers quickly, answer the question yourself. Remember that this activity has two objectives: understanding the achievement level descriptors and building a group.

Practice Test Administration

From the Reference Materials page, direct panelists to click the “Smarter Balanced Practice and Training Tests” link. This should bring up the launch page for the Practice Test.

Smarter Balanced Launch Screen for Practice Tests



Have them sign in as a guest. At the next screen, they should choose their grade from the dropdown box and click “Yes.” There will be a list of grade appropriate test. They should select their content and choose the “Practice” not the “Training” test. They should continue to follow the prompts and begin the test. Check to make sure everyone has the correct page on their screens; then say:

You may have taken a practice test at some point during the past few months, and if you did, that is fine. This morning, I would like you to take this practice test and think about the students who will take a test like this one next spring. You should keep a couple of things in mind as you do so:

1. *This test is a fixed-form test. That means each of you will be taking exactly the same items as everyone else in the room. The tests students will take next spring will be computer adaptive, which means that there will be differences from one student to another. Students who answer a lot of test items correctly will get progressively more difficult items to answer, while students who answer a lot of test items incorrectly will get progressively easier items to answer.*

2. *Students will complete performance tasks that may take multiple class periods over several days as part of their assessments next spring. We will not be doing that here, but we will talk briefly about what is included in those performance tasks.*

There are directions to take you through the practice test and entering your responses to the items. Please follow those directions carefully, and do your best to answer as many questions as you can within the next hour. When everyone at your table has completed the practice test, you may begin to discuss the tests among yourselves, taking care not to interfere with others at nearby tables who are still taking the tests. I will call time later and encourage everyone to discuss their experience with the practice test with others at their table only. Are there any questions about what you need to do before we begin?

Scan the room for questions. Answer procedural questions and defer content or global questions and comments until the room-wide discussion at the end of the activity; then say:

Please begin.

Walk through the room to observe panelists taking the tests. If there are mechanical or internet issues, determine if they are specific to one panelist or room-wide. Technical staff will be on site to deal with these issues and will be on call. Call the Work Room at once if a problem arises. If a panelist has a question about the content of a particular item, do not attempt to explain or defend the item. Make a note of the item and include it in the room-wide discussion at the end of the session. Allow about an hour for this activity unless it becomes clear that nearly all panelists have completed the Practice Test sooner than that.

Start the discussion with a general question about the Practice Test experience:

Now that you have had an opportunity to take the practice test, what are your general impressions of its contents and the way it is presented?

This is an opportunity for panelists to discuss their impressions, but there may be many questions. It is not necessary for the facilitator to answer all questions, as other panelists may answer them. That is actually preferable, as it adds to community building and gives those panelists more ownership of the process. However, if responses are decidedly off target, gently guide the panel back to a more accurate understanding. During this discussion, remind panelists of the comments made earlier in the day that while the items on the Practice Test are not exactly the ones that will be on the 2015 tests, they are similar in content and difficulty and have been through several rounds of review. Now ask:

I promised to answer remaining questions when we began this session, and I would be happy to do so now. I would like to start with a few that have come up already and then open the floor for additional questions.

Answer content questions that panelists asked during the earlier activities, relying on the FAQ and other Smarter Balanced published materials. Then open the floor for additional questions, following these same guidelines. Save about 10 minutes at the end of the session for completion of the Readiness Form. Remind panelists that there will also be an opportunity at the end of lunch to ask general questions about the tests.

Completion of Readiness Form Part 1

Draw the room-wide discussion to a close and direct panelists to their computers and Part 1 of the Readiness Form:

If you would, please click on Readiness Form on your computer screens and turn to Part 1. Please respond to the questions on that form. When you have finished, submit your responses and log out. When you have completed this activity, this session will be over, and it will be time for lunch. Thank you for your attention. After lunch, we will meet back in the ballroom where we met earlier this morning for an introduction to the Bookmark procedure you will be using for the rest of your time here.

Security Issues

MI has produced a notetaking pad that will be included in panelists' packets. This pad will be the only paper panelists will be allowed to use to take notes on during the workshop. Remind panelists, however, that there are places in the item map to enter notes about the items. The notetaking pad is just a backup. Facilitators will collect all notetaking pads at the end of each day and instruct panelists to leave them at their work tables during lunch and breaks. Say to the panelists regarding the notetaking guides, ALDs, and CCSS:

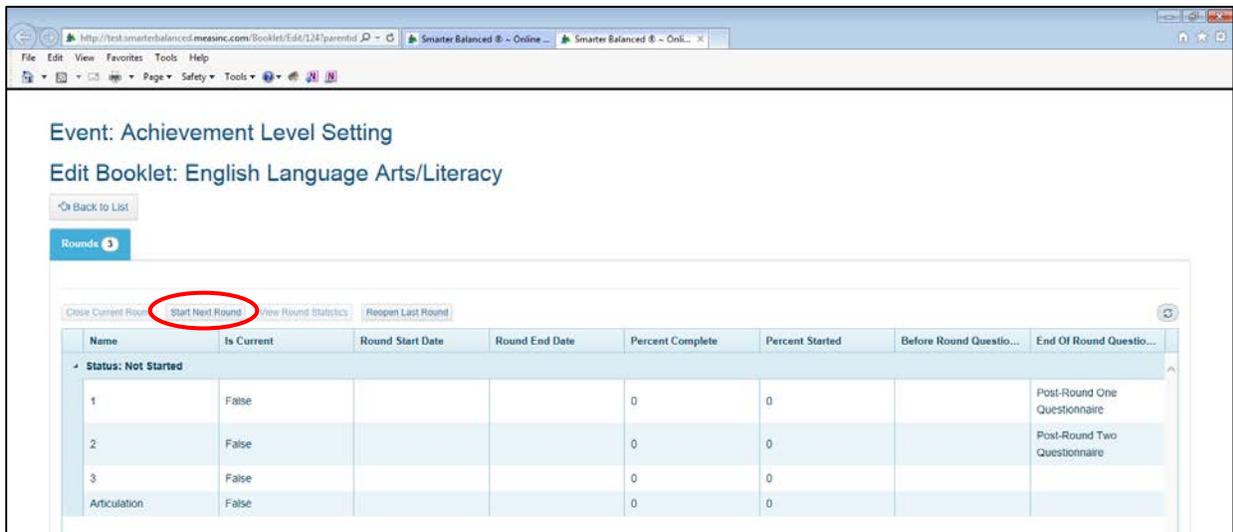
- You can't take any of this with you when you leave.
- You can't talk to anyone about any test content, now or ever.
- You can talk to others about the process but not the data or test items.
- You can't call another person (e.g., your supervisor, colleague, friend, spouse) to ask for advice.
- You can't take this back to your room to study tonight or to lunch to discuss with colleagues

Have panelists log out before going to lunch.

Day 1 PM Software and Ordered Item Booklet Review

The first purpose of this session is to familiarize the panelists with navigating the software with regard to the ordered item book. The next purpose is for panelists to answer two focus questions for every item in the OIB in preparation for bookmark placement tomorrow morning. From the Facilitator’s screen for the booklet, click “Start Next Round” to open Round 1.

Facilitator View



The screenshot shows a web browser window with the URL <http://test.smarterbalanced.com/measinc.com/Booklet/Edit/1241parentid>. The page title is "Event: Achievement Level Setting" and the subtitle is "Edit Booklet: English Language Arts/Literacy". There is a "Back to List" button and a "Rounds 3" indicator. Below this, there are buttons for "Choose Current Round", "Start Next Round" (circled in red), "View Round Statistics", and "Reopen Last Round". A table below shows the status of rounds:

Name	Is Current	Round Start Date	Round End Date	Percent Complete	Percent Started	Before Round Questio...	End Of Round Questio...
Status: Not Started							
1	False			0	0		Post-Round One Questionnaire
2	False			0	0		Post-Round Two Questionnaire
3	False			0	0		
Articulation	False			0	0		

Have panelists login after lunch and direct them to the Booklet for the correct grade and content area by clicking on the Booklet link in the black bar.

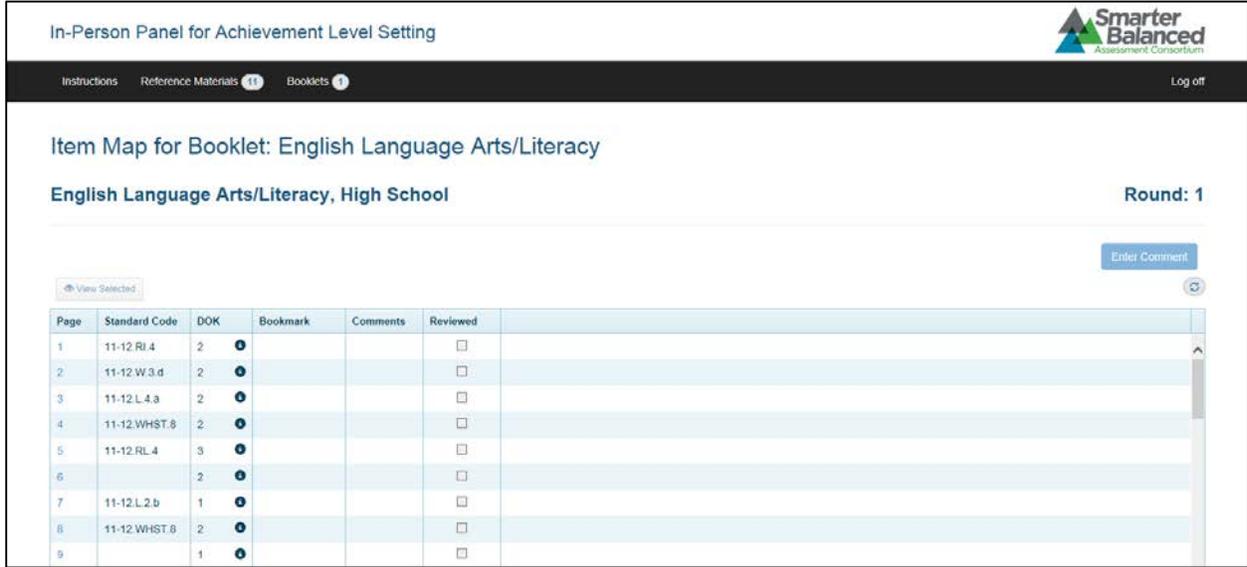
Booklet Link



The screenshot shows the "In-Person Panel for Achievement Level Setting" interface. The navigation bar at the bottom contains links for "Instructions", "Reference Materials", and "Booklets 1" (circled in red). The Smarter Balanced Assessment Consortium logo is visible in the top right corner, and a "Log off" link is in the bottom right corner.

Be aware that they should have a second Booklet they can access for the 6-item Orientation Round. Tell them they will access that later but that they should access the full OIB now by double clicking on it. Model this for them.

Sample Grade 4 ELA Item Map



In-Person Panel for Achievement Level Setting

Instructions Reference Materials 11 Booklets 1 Log off

Item Map for Booklet: English Language Arts/Literacy

English Language Arts/Literacy, High School Round: 1

View Selected Enter Comment

Page	Standard Code	DOK	Bookmark	Comments	Reviewed
1	11-12.RI.4	2	<input checked="" type="checkbox"/>		<input type="checkbox"/>
2	11-12.W.3.d	2	<input checked="" type="checkbox"/>		<input type="checkbox"/>
3	11-12.L.4.a	2	<input checked="" type="checkbox"/>		<input type="checkbox"/>
4	11-12.WHST.8	2	<input checked="" type="checkbox"/>		<input type="checkbox"/>
5	11-12.RL.4	3	<input checked="" type="checkbox"/>		<input type="checkbox"/>
6		2	<input checked="" type="checkbox"/>		<input type="checkbox"/>
7	11-12.L.2.b	1	<input checked="" type="checkbox"/>		<input type="checkbox"/>
8	11-12.WHST.8	2	<input checked="" type="checkbox"/>		<input type="checkbox"/>
9		1	<input checked="" type="checkbox"/>		<input type="checkbox"/>

You should see the item map on your computer that is projected from my screen.

(The sample item map above and the following OIB pictures are for ELA, Grade 4. Your item map will be similar, but will be for the grade and content area you are facilitating.)

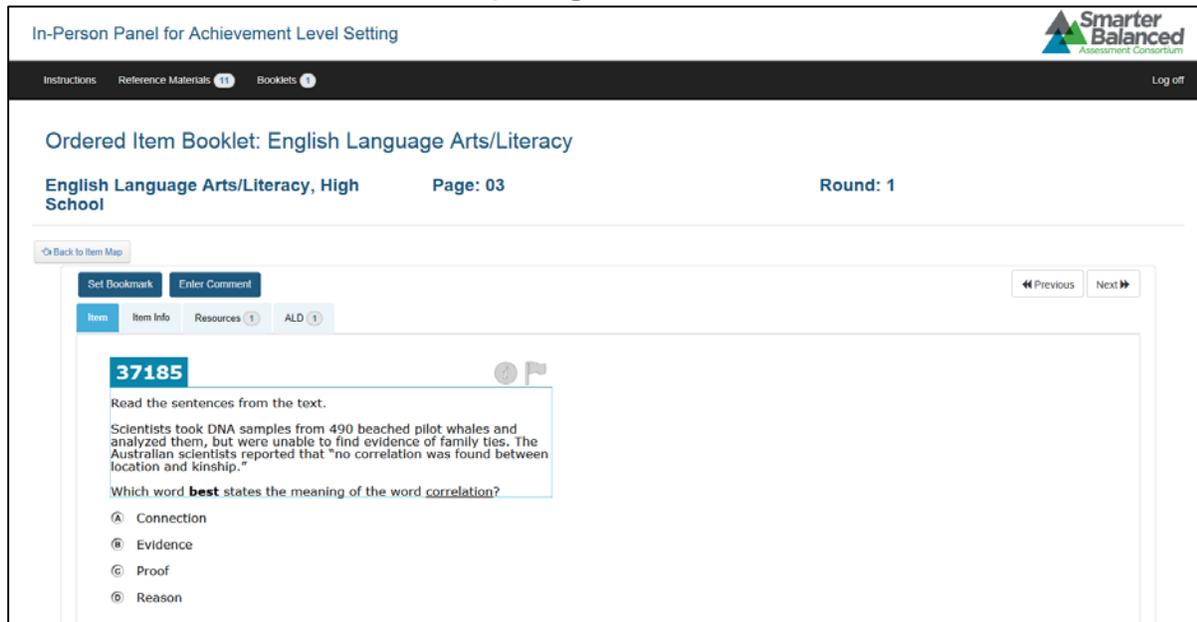
Item Map Navigation

Take a moment to make sure everyone has the proper image on screen. Assist those who might be having difficulty. Then say,

*Let's take a few minutes to familiarize ourselves with the navigation within the item map and from the item map to the ordered item booklet (OIB). Let's start with **Page**. Click on page 3 and see what happens. (Wait a few seconds to allow panelists to attempt this independently, then model it using your computer.) That should take you to page 3 in the OIB. You should see the image I now have on the screen. On the screen you see the item, or question, that students answered. You know you are on page 3 of the OIB by observing the number "03" after "Page:" toward the top of the page in the OIB (point to the Page: 03 on the screen)—this is where the page number is displayed in the OIB. Is anyone not seeing page 3 of the OIB as displayed on my screen?*

Make sure everyone has gotten to page 3 of the OIB before moving on. If there are technical problems, technical staff is onsite for support.

Grade 4 ELA Ordered Item Booklet Sample Page



In-Person Panel for Achievement Level Setting

Instructions Reference Materials 11 Booklets 1 Log off

Ordered Item Booklet: English Language Arts/Literacy

English Language Arts/Literacy, High School Page: 03 Round: 1

Back to Item Map

Set Bookmark Enter Comment Previous Next

Item	Item Info	Resources 1	ALD 1
37185	<p>Read the sentences from the text.</p> <p>Scientists took DNA samples from 490 beached pilot whales and analyzed them, but were unable to find evidence of family ties. The Australian scientists reported that “no correlation was found between location and kinship.”</p> <p>Which word best states the meaning of the word <u>correlation</u>?</p> <p>(A) Connection</p> <p>(B) Evidence</p> <p>(C) Proof</p> <p>(D) Reason</p>		

The page number refers to the relative difficulty of the item. In this case, page 3 shows the 3rd easiest item in the OIB. We will learn more about the information about the ordered item booklet shortly. For now, click one of the links labeled “Back to Item Map,” that appear to the left, both above and below the item, so we can complete our understanding of the item map first. (Facilitator models by clicking “Back to Item Map”).

Make sure everyone has navigated back to the item map.

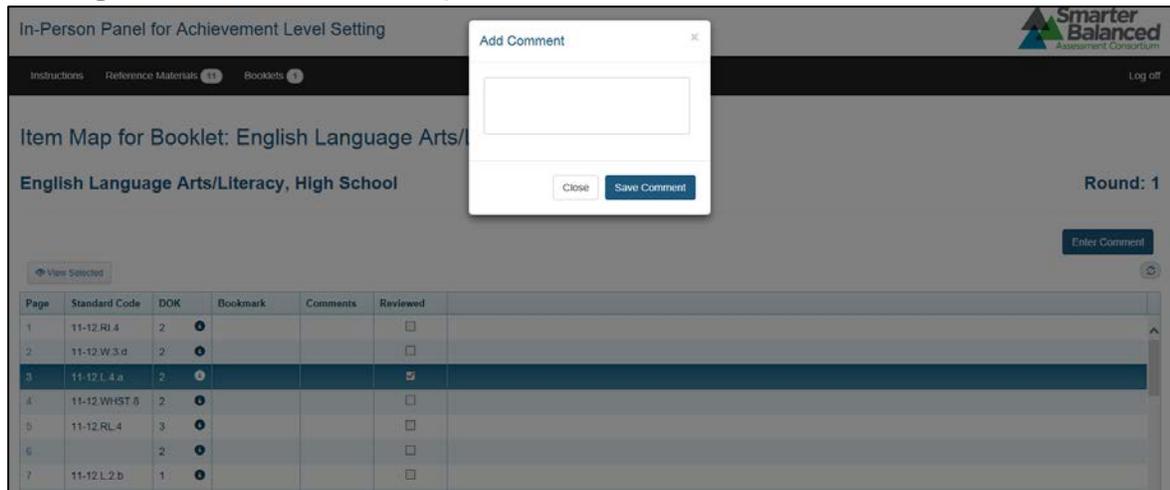
So, clicking the page number on the item map moves you to the OIB and clicking on the “Back to Item Map” link in the OIB moves you to the item map.

Now use your cursor to hover over the small, dark circle next to the standard for the item on page 1, which is in the column with the header “Claim and Target.” You will see [Claim and Target code] for that item. (Facilitator models.) Now hover over the circle (model). You should see the full text of the content standard underlying that item, as I have on my screen. Does anyone not see the full text for the content standard after hovering? (Make sure everyone sees the text for the content standard.)

Now use your cursor to hover over the small, dark circle in the DOK column (model). DOK refers to depth of knowledge. If you hover over the circle in the DOK column of the item, you will get a definition of that DOK level —1, 2, 3, or 4 (model). Notice that there are items with DOK levels of 1, 2, 3, and 4 throughout the item map (point to items with DOK levels 1, 2, 3, and 4 on your item map on the screen).

We will learn about the bookmark columns tomorrow. For now, we want to focus on the next column which is for comments. Let’s click once on the row for Page 3 to highlight the entire row (model). Then click the “Enter Comment” box at the top right of your screen and type in “I have a comment about this item.” (Model. The following graphic is illustrative). If you want to save your comment, click “Save Comment.” You can then view your comment by hovering over that column.

Entering Comments on the Item Map



Take a moment to make sure panelists understand how to comment.

When you review an item in your OIB it will register as “reviewed” on the item map. (As seen below.) That way, if you have to take a break and then come back, you can quickly see where you left off in the OIB.

Reviewed Items

Page	Standard Code	DOK	Bookmark	Comments	Reviewed
1	11-12.RI.4	2			<input type="checkbox"/>
2	11-12.W.3.d	2			<input type="checkbox"/>
3	11-12.L.4.a	2			<input checked="" type="checkbox"/>
4	11-12.WHST.8	2			<input type="checkbox"/>
5	11-12.RL.4	3			<input type="checkbox"/>
6		2			<input type="checkbox"/>
7	11-12.L.2.b	1			<input type="checkbox"/>
8	11-12.WHST.8	2			<input type="checkbox"/>
9		1			<input type="checkbox"/>
10	11-12.RI.4	2			<input type="checkbox"/>

And that wraps up our review of the item map. Are there any questions?

Answer any questions that may arise. Some may require reiteration of portions of the script; others may be better answered through the next experience in navigating the OIB. Make sure all questions are answered before moving on. If necessary, feel free to contact Michael Bunch or Daniel Lewis.

OIB Navigation

Ask panelists to double-click on page 1 of the item map (model).

Sample Page 1 from OIB

In-Person Panel for Achievement Level Setting

Instructions Reference Materials 11 Booklets 1

Ordered Item Booklet: English Language Arts/Literacy

English Language Arts/Literacy, High School Page: 01 Round: 1

Back to Item Map

Set Bookmark Enter Comment

Item Item Info Resources 1 ALD 1

42315

First, read the dictionary definition. Then, complete the task.

(adv.) **1.** existing or happening at the same time

Click on the word in the paragraph that **most closely** matches the definition provided.

The game, called NeuroRacer, involves doing two things simultaneously. A player uses a joystick to guide a car along a hilly, twisting road, steering it and controlling its speed. At the same time, a series of signs - actually colored shapes - appears on the screen. The player is supposed to push a button only when a particular kind of sign appears. Players were scored on how quickly and accurately they reacted to the right signs.

You should see page 1 of the OIB. The page number appears above the item. Let me know if at any time you do not see on your screen what I show on mine. First, notice the five-digit number in blue. This is the item number which is just a reference number and holds no other meaning for our purposes. Next, there are tabs above the item number labeled “Item,” “Item Info,” “Resources,” and “ALD.”

Item Tabs

Item Item Info Resources 1 ALD 1

42315

Initially you will be on the “Item” Tab, displaying the item. If you click the “Item Info” Tab you will view additional information about the item, such as standard alignment. If the item is based on a passage, or has an audio or video file associated with the item, it will appear in the “Resource(s)” Tab with a number behind it signifying how many resources there are.

The “Resource(s)” Tab

Set Bookmark Enter Comment

Item Item Info Resources 1 ALD 1

Passage
58.JPG

Find an item with a passage and then have panelists go to that item's Resource tab and open the passage which will show up as a new tab. Try to keep panelists from randomly clicking on too many items so that they don't have a lot of items marked "Reviewed" before they actually start reviewing items.

Passage Sample



Read the text and answer the questions.

Why Do Whales Beach Themselves?

by Stuart Landers

In early June of 2009, a group of more than fifty whales swam onto a beach in South Africa and could not return to the water. A team of scientists, along with hundreds of local volunteers, were able to move all of them back into the ocean, but the whales just turned around and beached themselves for a second time. Despite the heroic efforts of the people on the beach, within two days all 55 whales were dead or dying. To this day scientists have been unable to figure out exactly what caused this group of cetaceans to lose their lives in this way. In fact, the causes of whale beachings, which have been observed regularly for more than 2,000 years, are not fully understood by marine biologists and other scientists. Although many whales swim onto beaches and die because they are sick or starving or for other obvious reasons, at least half of all whale beachings cannot be explained at all.

Whales, dolphins, porpoises, and other cetaceans beach themselves both individually and in groups, and these groups can range in size from three or four whales to close to 100 of these animals. The reasons for individual beachings are often the easiest for scientists to understand. According to an international group called Whale and Dolphin Conservation, or WDC, sickness is one clear reason why cetaceans come ashore and die. Examinations of beached whales frequently reveal that the animals suffered from serious illnesses. Another easily understood reason is old age; scientists have been able to determine that some beached whales were quite old and therefore at the end of their natural lifespans. Injury, finally, is also a common cause of whale beachings. According to WDC, many of the dead whales found on beaches have serious injuries caused by being struck by boats or getting tangled in fishing nets. Nonetheless, every year there are numerous individual whale beachings that cannot be easily explained by any of these obvious causes.

It is often even more difficult for marine scientists to determine why whales beach themselves in groups. The WDC offers several unproven theories as to why cetaceans come ashore to die in groups. First, they suggest that a food shortage in the area where a whale or group of whales usually feeds can cause these whales to move into unfamiliar areas. Being in unfamiliar parts of the ocean might cause whales to swim onto shore by mistake. There is also some evidence that whales use Earth's magnetic fields to help them navigate when they migrate. Some scientists have suggested that natural phenomena such as earthquakes can cause magnetic fields to change. When this happens, it is possible that whales become confused and swim too close to shore, where they become stranded. Another theory holds that some groups of whales are driven ashore by large boats or heavy boat traffic. Navy training exercises involving large numbers of warships may be a cause as well. The WDC also suggests that the Navy's use of sonar is causing whale beachings. Whales also navigate using a method called echolocation. Whales are able to make loud sounds underwater, and when these sound waves bounce off of land or solid objects in the water, the whales are able to hear this and therefore determine where they are. Naval sonar uses powerful underwater sound transmissions to locate enemy ships, and it is possible that these loud sounds damage the whales' hearing and interfere with their ability to echolocate. They therefore become confused and swim onto shore by mistake, or so some scientists believe.

One of the more controversial explanations for whale beachings that WDC offers is the role of family relationships in whale beachings. According to this theory, whales often travel in family groups and are led by a leader. If that leader makes a mistake and swims out of the ocean, the other members of its family group will follow it and become beached as well. WDC reports that a certain type of cetacean called a pilot whale sometimes beaches in groups for this reason. However, a 2013 study of pilot whales in the oceans around Australia rejects this family leader theory. Scientists took DNA samples from 490 beached pilot whales and analyzed them, but were unable to find evidence of family ties. The Australian scientists reported that "no correlation was found between location and lineages." The pilot whales that they studied came from several different family groups, and were not closely related at all.

Even though the reasons for whale beaching are not fully understood, marine scientists around the world will certainly continue to study this phenomenon closely. Understanding why whales beach themselves is important because it is something of a scientific mystery. Whales are a very important part of the world's ocean life, and many cetacean species are endangered because of hunting and water pollution. Fully understanding the causes of whale beaching is a critical part of making sure that whales continue to survive in the future.

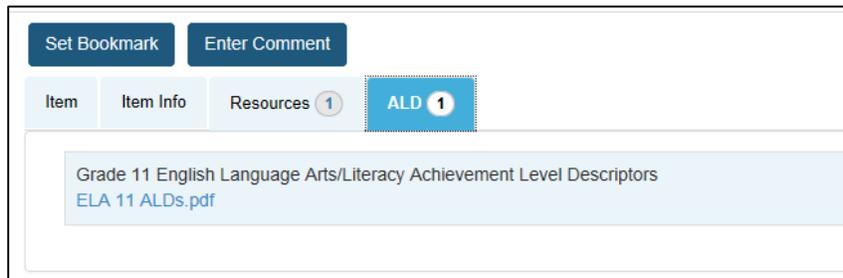
You will see multiple items in the booklet based on the same passage. On the tests the students take they are arranged together but they will not be together in the OIBs because the OIBs have the items arranged in order of difficulty. There will be other items, based on other passages, in between the set of items associated with a single passage.

You can leave the passage by clicking the "X."

There are also constructed response items and performance task components included in the OIB. Remember, you will see these multiple times, once for each positive score point. So for a CR item scores 0, 1, 2, or 3 you will see the item 3 times—once for a score point or 1, once for 2, and once for a score of 3. We saw this in the presentation this morning and we learned how to interpret each score point and we will review this later when we begin studying the OIB.

If you click on the ALD Tab, you will be able to access the Achievement Level Descriptors that guide your conceptualization of what students just entering each achievement level should know and be able to do. If you click on the link, the ALDs will appear in a new window (model).

The ALD Tab



Above the tabs you will see two buttons—“Set Bookmark” and “Enter Comment.” We will learn about the “Set Bookmark” button tomorrow morning. For now, I want you to see how to use the “Enter Comment” button in case you want take notes about some of the items as you review them. You have also received a paper-based notetaking guide if you prefer to use that instead of or in addition to these electronic notes. We will be taking up your paper notes each day, but we will give them back to you each morning. As we saw with the item map, you can keep your notes by entering them as a comment. This is the item page version of creating comments. If you want to enter a comment about the item, as a reminder for your own use later or to include in discussion with others at your table, click “Enter Comment.” You will see that you get a comment box in which you may enter your comment. Does everyone see that? You will be able to view that comment on the item’s page or in the comment column for that item on the item map.

Wait to make sure everyone has opened the comment box; help those who have not found it.

Now click the “X” on the comment box to close it and return to the item.

Suppose you have reviewed the item and want to review the next item in the OIB. To move on, click “Next” at the right of the page. That will take you to the next page. On subsequent pages, you will find both “Next” and “Previous” buttons at the right of the page so you can go in either direction, or you can click the “Return to Item Map” at the left to return to the item map, and from there, you can go to any page in the booklet simply by clicking the page number. For example, you may have entered a comment on page 12 and want to go back and look at it when you get to page 17. Then, you would click “Back to the Item Map” at the left of the page which will return you to the item map, and then you can hover over the comment box on line 12. The comment you entered for the item on page 12 will open for you. Are there any questions? After this we will move on to examining the ordered item booklet.

Answer any navigation questions panelists may have, using your own OIB and item map projected onto the screen at the front of the room. Once you have completed the demonstration and answered questions, the panel should be ready to review the OIB.

Presenting External Data (Grades 4, 8, and 11 Only).

If you are leading a group for grade 4, 8, or 11, point out the external items embedded in the OIBs and item maps. Otherwise, skip this section.

Before allowing panelists to begin their review of the OIBs, point out that some of the items in the booklets are not Smarter Balanced items but actually come from other tests such as the National Assessment of Educational Progress (NAEP) and the Programme for International Student

Assessment (PISA). These items were embedded in the spring 2014 field test because stakeholders and the public may draw comparisons between the performances of students on the Smarter Balanced tests with the performance of students on other tests.

Go to the first page where an external item appears and point it out to the panelists. Show it on screen and have panelists go to that page. You can find the item by looking in the Standards column on the item map for the first item marked “NAEP” or “PISA.” Note first the item itself, then its source (NAEP, PISA), and finally the item location associated with it. [We will have covered the meaning of NAEP and PISA scale scores and proficiency levels in Facilitator Training.] Then say:

This item and others like it from other tests are included to give you an idea where in the overall order of difficulty of the Smarter Balanced items the items on other well-known tests lie. They are taken from among the easy, moderately difficult, and very difficult items of those tests. Note their locations relative to Smarter Balanced items, particularly if they are in the vicinity of where you may be considering placing a bookmark. Remember that the Smarter Balanced ALDs should guide your cut score recommendations. Finally, the cut scores for the other tests were set for purposes which might be very different than those for Smarter Balanced: for example, the cut scores established for a national or international test may have different uses or aspirational uses than those for this test.

Because of these important differences, caution panelists against using any external data as their sole or even primary reason for recommending a cut score. Instead, note that the external items comprise a helpful, supplementary set of information to complement the Smarter Balanced items and test data they study. Encourage panelists to ground their cut score recommendations in the Smarter Balanced items and ALDs, and to use the external items for additional context. It may be necessary to remind panelists of the nature and purpose of these items during each round.

One helpful way for participants to use this information is to place the Smarter Balanced content as a primary source for their judgments, then to step back and compare their potential bookmark placements with the items from other tests as secondary sources for their judgments.

Once panelists have an opportunity to view external data, they will evaluate it in a very straightforward way:

1. Where do these items and their data suggest that a cut score should be set?
2. Consider the items in the ordered item booklet in the vicinity of that cut score. Does the content of those items, given the relevant achievement level descriptor, support placement of a cut score here? Why or why not?
3. If the cut score suggested by the external items and their data and the contents of the relevant items seem not to agree, to which will you give the greater weight, and why?

Presenting Performance Tasks

All OIBs will have elements of one performance task. Find the first element of the PT in your OIB, show it on the screen at the front of the room, and direct panelists to that page in their OIBs. Then say:

Part of each student's score will come from a performance task. Each performance task has multiple parts, some of which are selected-response items, and some of which are short answer or extended-response items. The various components of the performance task will appear in the ordered item booklet just like any other item. To get an idea of what the whole

performance task looks like, simply click on “Resources,” and you will see the entire performance task with all its parts.

Click on “Resources” to demonstrate the performance task.

Reviewing the OIB

Next, you will go through all of the items in the OIB, answering the following two questions:

- *What do you know about a student who responds successfully to this item or score point? Here we are asking you to think about what a student needs to know and be able to do in order to earn this point.*
- *What makes this item more difficult than preceding items?*

You can write down what you have learned about the item by answering these questions in the comments field. As you review the items, also discuss your answers with others at your table. You will notice that by working together, and listening to your fellow panelists’ thinking, you will learn much more about the item than you would working alone. Let’s work through the first few items together.

Have panelists open the OIB and solicit responses to the first question for the first item. You cannot address the second question until you get to the second item. Encourage discussion and then have panelists put their own responses to that question in the comments section. Move to the second item and repeat the process for the first questions. When you get to the second question, clarify:

We are not asking you to compare the difficulty of an item only with the item immediately before it—items may not vary in difficulty much from item to item. However, you will see an increase in difficulty in general as you move from earlier items to later items. Make your observations about item difficulty in comparison to any or all of the items prior to the item of focus.

Continue this procedure for the next few questions until panelists seem to understand the process. Then ask,

Are there any questions about how to analyze each item? [Answer any questions that arise at a group level.] Now you will continue this process on your own for the remaining items in the OIB. However, if you wish to discuss specific items at your table, feel free to do so. The intention is for you to complete your review of these items by the end of today so be mindful of the time. It is very important for you to put the answers to these questions for each item in the Comments section as you will need to reflect upon this information during the bookmark placement rounds.

Circulate through the room to make sure everyone is progressing through the OIB. Give assistance when asked, but do not intrude. Some panelists will catch on more quickly than others, and the pace will vary from table to table. Do not rush anyone, but do caution anyone who seems to be racing through the OIB without reading the items carefully; likewise prompt those who seem to be lingering too long on items to keep a reasonable pace. Throughout this process, give time checks indicating, for example, that it is one hour in, and panelist should be about [x] way through the OIB. At these time checks, have table leaders check in with their table to encourage those lagging behind to keep pace and slow down those who are racing through. Remind panelist that they will have many more opportunities to review the items in the bookmark placement rounds.

When panelists have completed their review of the OIB, they should log out of the software and are dismissed for the day. Remind them that the first session of the next day will be Bookmark training which will be crucial to their remaining time there and they should make every effort to be on time. Gather your frequently asked questions to take to the facilitator debriefing.

Day 2 AM

Bookmark Placement– Orientation and Round 1

Panelists will have had an evening to think about what they did the previous day and may have several questions about the CCSS, the ALDs, or the tests themselves. They will also have just completed bookmark training. The morning of the second day will be a good time to answer those questions.

The morning of Day 2 has 2 goals. The first goal is for panelists to practice the information learned during the Bookmark presentation that occurred immediately preceding this session with the software. This will happen with the use of the 6-item orientation OIB. The second goal is to actually set the Round 1 bookmarks in the appropriate grade/content OIB.

Panelists should login to the software. Then model the following steps,

Select “Booklet” in the header. You will see two booklets. Yesterday we got to know every item in the full OIB. We will be returning to the booklet later this morning, but at this time you should double click the other booklet, the Orientation Round. This will bring up a 6-item item map linked to a 6-item OIB. The purpose of this orientation round booklet is to give you an opportunity to try out the navigation buttons and set one cut score – Level 3. Does everyone see what I have on my screen?

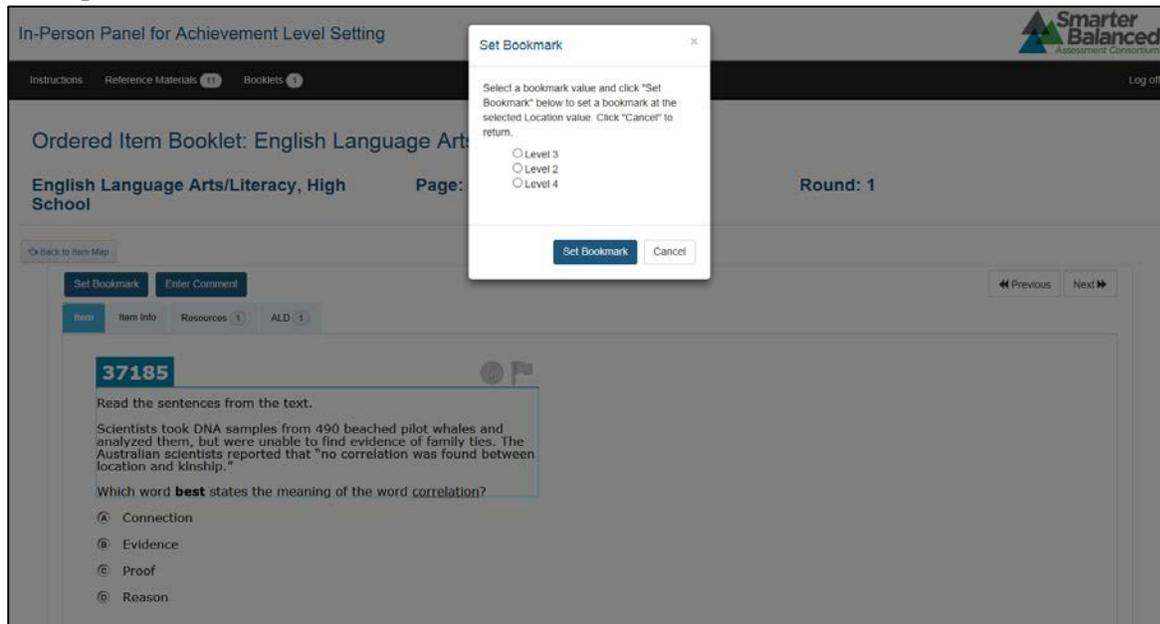
Orientation Round Item Map

[Add screenshot of Orientation Round OIB Item Map]

Next, you see the Bookmark column. When you are working through the OIB and you set a bookmark on a particular page, it will mark that page on the item map. You can also set a bookmark from here on the item map by highlighting the desired item row and clicking on the “Set Bookmark” button.

Go to Page 1 in the Orientation Round. Next to the “Enter Comment” button you will see “Set Bookmark.” To enter a bookmark to indicate your cut score recommendation, you would click “Set Bookmark” and be prompted to enter a bookmark and designate it for Level 3. Go ahead and click the “Set Bookmark” box.

Setting Bookmarks



NOTE: Because the facilitator logins use a different event code, you can set a bookmark without affecting the panelists' data. However, you should model canceling out of the "Set Bookmark" box so that the panelists do not set bookmarks prematurely.

Wait to make sure everyone has opened the "Set Bookmark" box; help those who have not found it.

OK, let's cancel out of the "Set Bookmark" box for now.

When you have set your one bookmark for this orientation round or all three bookmarks for the full booklet, you will click "Submit Bookmarks."

That will submit your bookmarks into the system.

We're going to put all of this information together now in a Bookmark orientation session and take a look at a very abbreviated Bookmark orientation OIB. It has only 6 items. These items should look familiar because they are six of the items in the sample test you took yesterday morning. To get started, click on Page 1 to get to the OIB. Read that item and discuss the following two questions with the panelists at your table in order to help you better understand the nature of the item [Show 2-question slide]:

- *What do you know about a student who responds successfully to this item; that is, what does a student need to know and be able to do in order to know the correct response?*
- *What makes this item more difficult than preceding items?*

Panelists should be familiar with the process of using the 2 questions.

Ask panelists what the correct answer is for the first item. Then, ask the first question, "What does a student need to know and be able to do in order to earn this point?" Wait for panelists to call out answers. Ask them to write these down in their item map comments. Push panelists to go beyond the obvious. Move on to the second item, asking panelists for the correct answer and skills students need to know and be able to do in order to answer the item correctly. Then, ask panelist the second question, "What makes this item more difficult than preceding items?"

Ask panelists to write down their answers in their item map comments. Then, direct panelists to take about 10 minutes to consider the remaining 4 items in their tables. They may discuss with their fellow table panelists.

Circulate through the room to make sure everyone is on task. Answer questions as necessary and appropriate, making note of any that arise more than one or two times. For those, it may be helpful to make a general announcement explaining or clarifying the issue that precipitated those questions. At the end of about 10 minutes or when it appears that everyone has completed the task, call time.

Now we're going to practice setting one bookmark for Level 3. Make sure you have your Achievement Level Descriptors handy. When you place your Level 3 bookmark, think about a student at the threshold of Level 3 based on the Achievement Level Descriptors. Place your Level 3 bookmark at the point in the OIB such that

- *A threshold Level 3 student would have at least a 50% likelihood of success on each of the items before the bookmark, with about a 50% chance on the item just before the bookmark*
- *A threshold Level 3 student would have less than a 50% likelihood of success on the items from the bookmark on.*
- *For each of the items, think of a large group of students at the cut score for this level. Would at least 50% of them be expected to respond successfully to this and prior items in the OIB? If you answer Yes on page 1, go on to page 2. If you answer Yes on page 2, go on to page 3, and so on until you can no longer say Yes. You will place a bookmark on the item AFTER the last item for which you were able to say "yes," or equivalently, ON the first item for which you said "no."*

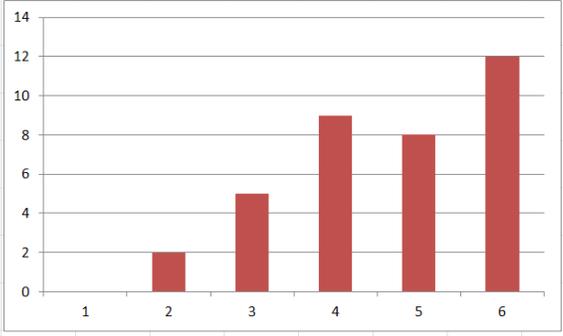
Let's take about 5 minutes to enter your Level 3 bookmark. I will circulate through the room to see how everyone is doing. Remember, when you enter your bookmark, it should be based on your own judgment, not necessarily the consensus of everyone at your table.

*Remember, your bookmark is on the item **after** the last item you expected Level 3 students to master (just like a bookmark in a physical book separates what you have read (mastered) from what you have not read (not mastered). In this orientation round, you will only have the option of setting the Level 3 bookmark, but in the actual OIB you will be entering bookmarks for Level 2, 3, and 4.*

Circulate through the room to make sure everyone is on task. Answer questions as necessary and appropriate, making note of any that arise more than one or two times. For those, it may be helpful to make a general announcement explaining or clarifying the issue that precipitated those questions. When it appears that everyone has completed the task, call time. To get a quick calculation of the group median cut score recommendation, it will be best to get a show of hands and do the calculation using the Excel spreadsheet set up on your computer so you can show the tally on the projector.

Sample Excel Spreadsheet for Medians

	A	B	C	D	E	F	G	H	I	J	K	L	M
	Page	Votes	Cum. Freq.	Cum. %									
2													
3	1	0	0	0%									
4	2	2	2	6%									
5	3	5	7	19%									
6	4	9	16	44%									
7	5	8	24	67%									
8	6	12	36	100%									
9	Total	36											
10	Median Page			5									



First of all, did anyone have any difficulty finding an item that you thought was the most difficult item students just barely at Level 3 would be expected to answer successfully?

Take time to allow panelists to respond to this question. Discussion may ensue. If so, keep it focused on the task at hand and draw it to a close after a few minutes or if it appears to be drifting away from the task.

Raise your hand if you placed your bookmark on page 1.

Enter the count on the tally sheet and on your spreadsheet. Repeat for Pages 2, 3, 4, 5, and 6. The median will calculate as you enter the last tally.

Now draw their attention to the tally sheet. Starting with the lowest bookmark placement, the panelist(s) who placed their bookmark here should share his or her rationale behind their bookmark placement. Be sure to refer to the item map, the ALDs, and the target students. Allow 1 or 2 panelists to explain why that item fits the Level 3 threshold descriptor. After 1 or 2 people have shared his or her rationale, move on to the next highest bookmark.

Now, let's hear from some of you who placed your bookmark on page [y]. How do the items before page [y] fit the descriptor and why do the items after not fit the descriptor?

Again, allow 1 or 2 panelists to explain why that item fits the Level 3 threshold descriptor. Follow this process for each of the bookmark placements.

You've just had your first experience applying the Bookmark Procedure, and we've learned a couple of things:

- *Even though you all had the same Achievement Level Descriptors and the same set of items, your judgments were different—not all of you placed your bookmark on the same page. That's fine; we don't expect everyone to always agree.*
- *Some of you may have wanted to set your bookmark between, rather than on these pages because there were not enough items to really capture the skills represented in the ALD. However, as you've already seen, the full booklet you worked with yesterday has many more items. You will see that additional items better fulfill the definition not only of Level 3 but of Levels 2 and 4 as well.*

When you begin the bookmark procedure with the actual OIBs, you will follow this same process. You will begin by setting your Level 3 bookmark, followed by your Level 4 and Level 2 bookmark recommendations. Before we break out the booklet, however, we need to do two things. First, do you have any questions about this process or how you will apply it?

Answer questions panelists may still have. Refer to the script or the PowerPoint notes from the earlier sessions for CCSS, ALDs, and Bookmark. Also remind panelists about the inclusion of NAEP and PISA items (grades 4, 8, and 11) and the placement of the performance tasks. Then continue while modeling:

The last thing we need to do before we start Round 1 is to ask you if you are ready. Close your Orientation Round screen and go the Menu to call up the Readiness Form Part 2. Please take a few minutes to read and complete and submit it. When everyone has successfully completed it, we can begin.

Circulate through the room to make sure everyone is completing the Readiness Form Part 2. Answer questions as necessary and appropriate. For questions that come up several times, it may be necessary to make room-wide clarifications. Note these questions for the debriefing at the end of the day. When everyone has completed and submitted the Readiness Form, have them close the screen, click on the Booklets link, and return to the full OIB.

Allow all panelists time to open the booklet. It should say “Round 1” at the top; then begin:

Yesterday you did a thorough review of all the items in the OIB using the two focus questions shown here. Now you are going to set three bookmarks in that OIB, one each for Levels 2, 3, and 4. Since you’ve already been working with the Level 3 threshold ALD in the orientation round, the first bookmark you should set is the Level 3 bookmark.

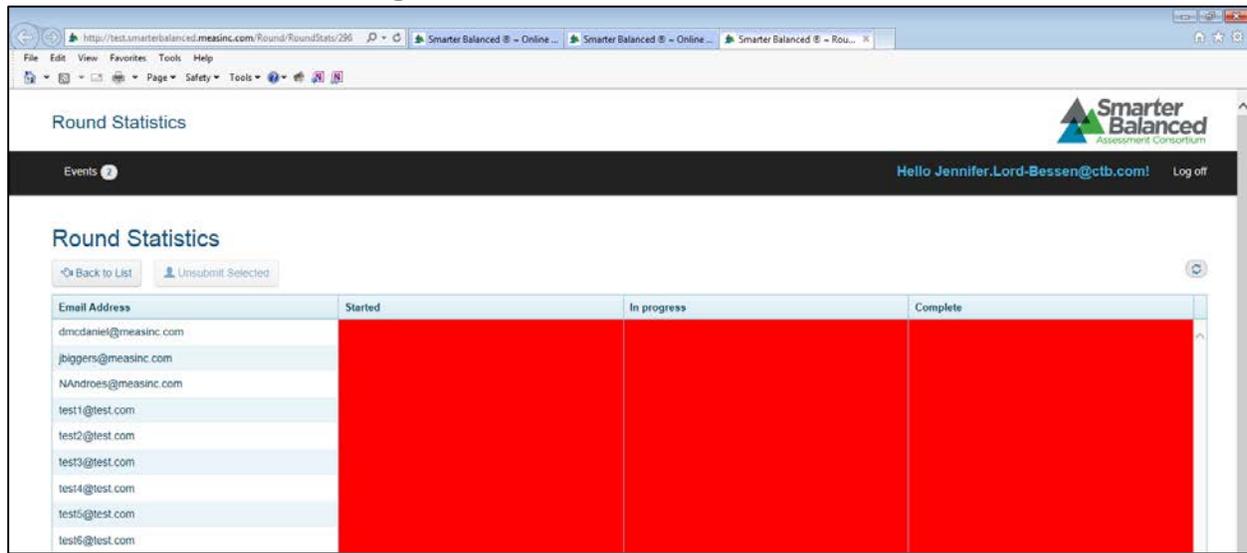
When setting your bookmarks, remember to consider a third question [display Bookmark Placement slide], namely “Would a student at the threshold of the achievement level have at least a 50% chance of responding successfully to this item (or score point)? If they would, then move on to the next item; if not, place bookmark on this page.” Keep referring to your ALDs as your work through this process.

Point out to panelists the posters with the third question in the room in case they want to reference it later.

In Round 1, you will decide individually where the bookmarks should be placed. That is, you will make your cut score recommendation individually, without discussion. Remember, you will set your Level 3 bookmark, followed by your bookmarks for Level 4 and Level 2. In Round 1 it is not uncommon to find it challenging to find an exact spot that seems “perfect” for your bookmark placement. Don’t be overly concerned. Find the best location you can, perhaps by finding a range of possible bookmark placements, say you think it belongs between page x and page y, then pick the best place, knowing that you will have two more rounds in which you can change your mind based on discussion with your fellow panelists and with additional information provided to you. We will finish this task before lunch and then see the results of the group’s work after lunch.

Panelists proceed to place their bookmarks. Walk around the room, reinforcing the instructions about bookmark placement. Panelists should take their time and look at each item with respect to the third question before placing their bookmarks. Monitor the panelists’ progress using the facilitator login by clicking on the “Round Statistics” button. Green bars will show progress by the panelists through the round.

Facilitators View of Panelist Progress



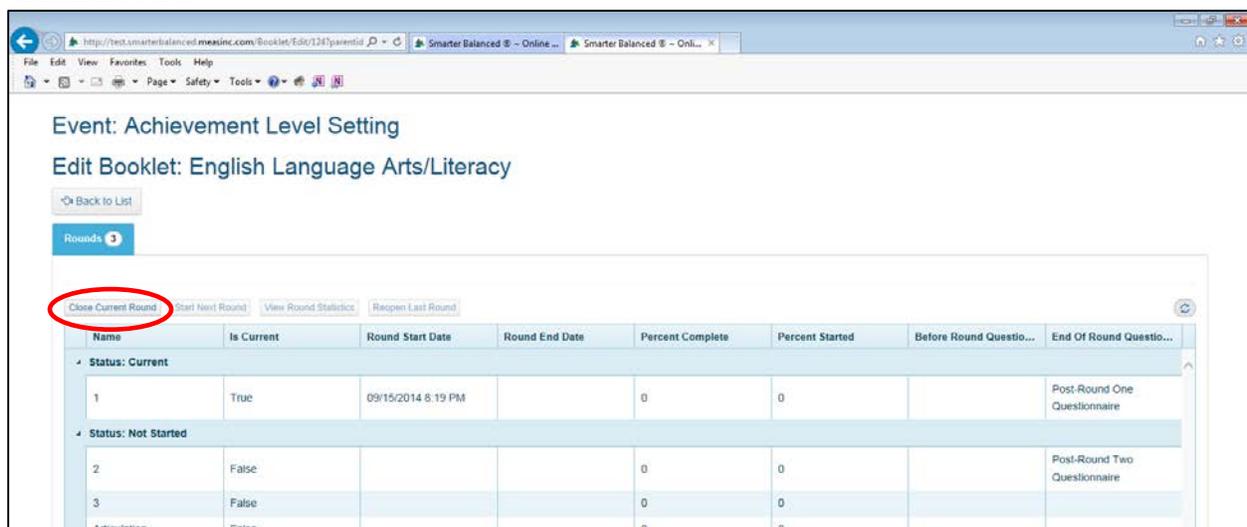
Email Address	Started	In progress	Complete
dmcDaniel@measinc.com			
jbiggers@measinc.com			
NAndrees@measinc.com			
test1@test.com			
test2@test.com			
test3@test.com			
test4@test.com			
test5@test.com			
test6@test.com			

There will be a 90 minute lunch break today to allow for data analysis and reporting. Panelists will need to complete and submit their Round 1 bookmarks before being dismissed for lunch. Panelists will complete Round 1 at various times, so be alert and make sure each panelist submits and logs out before leaving for lunch. Remind them that we cannot tally their bookmarks until they submit, get a “Bookmarks Submitted” message, and log out. Check screens to make sure panelists are following this important procedure, and dismiss panelists individually.

Remind table leaders that they are to return 15 minutes prior to the start time to receive their table’s median bookmarks and to review the procedures for the afternoon.

When all participants have completed the bookmark placement, close out the round from the facilitator screen and notify the Operations Room that your Round 1 is complete.

Facilitator View



Name	Is Current	Round Start Date	Round End Date	Percent Complete	Percent Started	Before Round Questionnaire	End Of Round Questionnaire
Status: Current							
1	True	09/15/2014 8:19 PM		0	0		Post-Round One Questionnaire
Status: Not Started							
2	False			0	0		Post-Round Two Questionnaire
3	False			0	0		
Articulation	False			0	0		

Day 2 PM

Round 2 Bookmark Placement

After lunch, panelists will reassemble in their respective rooms to discuss at their tables the results of Round 1. Prepare table leaders prior to the session for the table discussions. A review of the additions to the item map should occur before the table discussions. After table discussions are completed, panelists should complete the Round 2 Readiness Form (Part 3) and then will proceed to set the Round 2 bookmarks.

Pre-Session Table Leader Meeting

You should have arranged to meet with the table leaders previous to the start of the session to remind them that they will need to facilitate the sharing and noting of their table's panelists' bookmark locations, writing them down for each panelist and identifying the lowest, highest, and median (middle) score and to remind them of how to lead their table discussions of the range of bookmark placements. They will have had training on this, but reinforcement is useful at this time point.

The purpose of this discussion is to give all panelists at your table an opportunity to share where they placed their bookmark and the content rationales. The purpose of this discussion is not to come to consensus; bookmark placement is always an individual activity.

Start with the Level 3 bookmarks and have each panelist quickly go around and state where they placed it. Each panelist can write them down, and note the low, high, and median (middle) bookmark. Note that the table did not have perfect agreement as to where the bookmarks should be placed and that's fine. But before reconsidering our own bookmark placements for Round 2, it is helpful to understand the thinking behind our fellow table panelists' bookmark placements. You want to get some feedback on everyone's content rationales for placing their bookmarks where they did so that others can have the benefit of their point of view.

Lead a discussion of the content rationales for bookmark placements, beginning with the lowest and highest bookmarks (those that are most discrepant with the "consensus." Make sure to bring back the discussion of the placements to the ALDs.

Dismiss table leaders to return to their tables as panelists enter the room.

Item Map Review

From the Facilitator's login, open Round 2. Once all panelists have returned, have panelists log in and go to the item map.

Welcome back. In a few moments, your table leaders will be conducting a discussion of the bookmark placements at your table. After that, we will complete the readiness form and then begin the placement of bookmarks for Round 2. However, I'd first like to bring your attention to some new features on the item map. [Model the following.]

You will see the page numbers on which you placed your bookmarks in Round 1 represented by the thumbtack icons. You will refer to these in your table discussions starting in the next

few minutes. In Round 2, you are free to keep your bookmarks on these same pages, or you may move them to different pages.

Another new column on the item map is the Location column. The location of an item provides greater insight into the relative difficulty of an item. Items don't increase in difficulty in equal increments. Location provides a more accurate indicator of difficulty. An item whose location is the same as the item before it only differed by a decimal and rounding resulted in equal locations. The greater the difference in location, the greater the difference in difficulty.

Use the following script for panels for grades 4, 8, and 11 ONLY.

To support your content based decisions, we also provided some contextual information that illustrates the connection between your recommended cut scores for the Smarter Balanced tests and other testing programs with which you may be familiar. This can support the validity of your judgments and provides useful information that may be communicated to stakeholders who will interpret the results of the standard setting.

Recall that some NAEP and PISA items were embedded in your OIBs. The field tests we administered last spring in grades 4 and 8 had items from the National Assessment of Educational Progress (NAEP) embedded in them. We have placed those items on the same scales as the Smarter Balanced items and have included some of them in your OIB. We noted them when we studied the OIBs yesterday. Notice, for example, that the item on page [x] is a NAEP (PISA also, for HS) item that has a location value of [xyz]. If you go to page [x] in the OIB, you will see where that item fits into the overall hierarchy of items in the National Assessment and an estimate of the average ability of students who took the National Assessment and answered that item correctly. (Find an example in your OIB). Items like these will help you gauge student achievement on Smarter Balanced tests against that of students who performed at or near the Basic, Proficient, and Advanced levels on the National Assessment. You may use this contextual information as you place your Round 2 bookmarks.

Use the following script for grade 11 panels only.

The final column of the item map has the header "ACT".

ACT conducted a comprehensive study to establish benchmark scores on its assessments, which provide guidance with respect to the setting of the score needed on the ACT exams for a student to be considered college and career ready. They operationalized readiness to mean that a student who attains the benchmark score on the ACT subject-area tests would have a 50% chance of obtaining a B or higher or about a 75% chance of obtaining a C or higher in corresponding credit-bearing first-year college courses. These college courses include English Composition and College Algebra (among others). The benchmark scores are 18 for English Composition and 22 for College Algebra.

We matched students' ACT Math and ELA scores with their corresponding scores from the Smarter Balanced field test (SBFT) forms for a large Midwestern state that administers the ACT to all high school students. We then calculated the Smarter Balanced scale scores that correspond to the ACT CCR cut scores and have included those scores on the item maps. Because they are projections, we have shown them as ranges rather than as discrete scores.

Emphasize to panelists that this is meant to provide them with additional, contextual information. Panelists should not be overly influenced by this information because the ACT and Smarter Balanced assessments are different in scope and focus with respect to the Common Core and are administered very differently. Moreover this comparison between the

two assessments is based on data from only one state. The data, and resulting projections on the Smarter scale, may not generalize to other states.

Round 2 Item Map.

[Insert Round 2 Item Map]

Are there any questions about these additions to the item map? [Answer any remaining questions.] For the next 45 minutes, we will be having table level discussions concerning the content rationales for bookmark placements at your table. In particular, when directed by the table leader you will share your bookmark placements only with the panelists at your table. You will look at the highest, lowest, and middle bookmark for your table and discuss the reasons for you bookmark placement. Do not feel that you have to defend your bookmark placement as if it is your final judgment—in Round 1 it is not uncommon to find a range of possible bookmark placements and to pick the best place, but you may not have had very strong opinions at that time. We only want you to discuss your thinking and listen to what others were thinking so that you can reconsider and make a Round 2 bookmark placement informed by that discussion. It is important for all table members to be involved in this discussion as everyone’s viewpoint has the possibility of providing a perspective that other table members may have missed in their consideration.

Allow table discussions to begin. Encourage tables to use the full 45 minutes, but if they finish early allow them to take a short break and return to the room at the end of the 45 minutes. After 45 minutes, interrupt any remaining discussion to look at an additional piece of contextual information.

Note that there was a period of time that other educators from across the Smarter states were able to participate in a similar bookmark activity, but only online, individually (without discussion with other educators), and only for the Level 3 threshold student.

Display the results of the online panel, so that they may discuss this at their tables in light of their own bookmarks and previous discussion.

Sample Online Panel Results for Level 3

Role	Median	Range
Teachers	44	38-50
Non-Teacher Educators	49	45-53
Administrators	52	49-55
Other	47	41-53

When all the discussion is complete, they can complete the Readiness Form Part 3.

Go to the Menu to call up the Readiness Form Part 3. Please take a few minutes to read and complete Part 3 and submit it. When everyone has successfully completed the form, we can begin.

Circulate through the room to make sure everyone is completing the Readiness Form. Answer questions as necessary and appropriate.

Round 2

When everyone has completed and submitted the Readiness Form, have them close the screen and call up Round 2 from the Menu.

As you begin Round 2, make sure the bookmarks you enter are yours—you are not required to reach consensus with others at your table or in the room but you may discuss any questions you have with your table. As you did in Round 1, you will set your Level 3 bookmark first, followed by your bookmarks for Level 4 and Level 2. You will have until about 4:45 to finish this round. Many of you will finish sooner than that. If you finish early, make sure you have entered 3 valid bookmarks and pressed “Submit Bookmarks.” When you get verification that you have submitted your bookmarks and log out, you are free to leave.

Place the Bookmark Placement slide on the projector.

Panelists begin Round 2, working individually, though they may consult with other panelists at their tables. Circulate through the room to make sure everyone is on task and not having any difficulty with the task or the navigation of the system.

Round 2 Wrap-up

At about 4:30, let panelists who are still working that they will need to wrap up in about 15 minutes. If there are any who will not be able to finish in 15 minutes, speak to them individually to find out how much time they will need. Be prepared to stay late but not past 5:30 P.M. to help them finish. At 4:45, invite panelists to wrap up, enter their bookmarks, and log out. Stay to help any who need more time. If the panelist needs to stay past 5:30, arrange for another staff member to stay while you attend the debriefing. When all panelists are done, close out the round from the facilitator’s login and notify the Operations Room.

Day 3 AM Round 3 Bookmark Placement

Review of Round 2

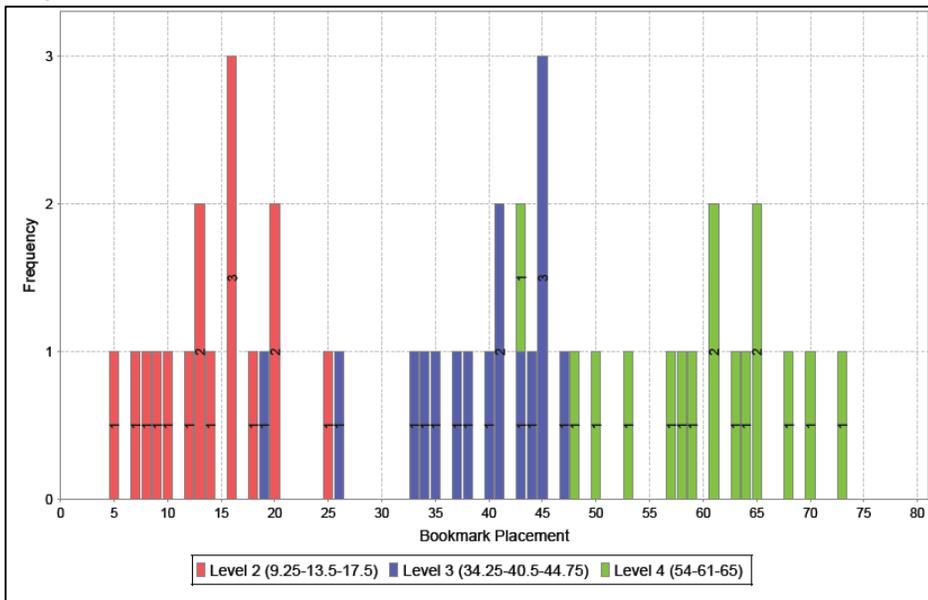
From the Facilitator’s login, open Round 3. When panelists enter on the morning of Day 3, MI and CTB staff will have analyzed Round 2 data and prepared reports. Present these reports on the projector. Direct panelists to call up the Round 2 item maps while you put the first Round 2 report on the screen in the front of the room.

Welcome back. This morning, we are going to review the results of Round 2 and then complete Round 3. Similar to our table discussions yesterday, we will now engage in the Round 2 discussion as a room. The purpose of this discussion is to give all panelists an opportunity to share where they placed their bookmark and the reason why. The purpose of this discussion is not to come to consensus; bookmark placement is always an individual activity.

I am displaying the distribution of bookmarks for panelists in our room.

Display the distribution of room bookmarks, as in the following sample graphic.

Sample Distribution of Bookmarks



You will notice that we did not have perfect agreement as to where the bookmarks should be placed. That’s fine. In the legend, the three numbers shown next to the level name are the 25th percentile, the median, and the 75th percentile. For level 3, about 50% of you placed your bookmarks between [34.25] and [44.75] (replace with appropriate level 3 numbers), with the median being [40]. But before we reconsider our own bookmark placements for Round 3, it will be helpful to understand the thinking behind our fellow panelists’ bookmark placements. We want to get some feedback on your content rationales for placing your bookmarks where you did so that others can have the benefit of your point of view.

Lead a discussion of the rationales for bookmark placements, beginning with the lowest and highest bookmarks (those that are most discrepant with the “consensus.”

Begin with Level 3, move on to Level 4 and to Level 2.

We’ve seen where you placed your bookmarks. We’ve taken your median bookmark recommendation and found the cut score associated with that median bookmark recommendation (the score where a student would have a 50% chance of responding successfully to the item before the bookmark). If we take the median location for the rooms’ bookmarks then the page numbers associated with each table and the whole room “consensus” recommendations are indicated in the following table. In addition, we are also showing the interquartile range (25th-75th percentile) of each of the three levels for the room. This is the range of the middle 50% of the bookmark placements for each level and will give you an idea of how spread out placements were.

Sample Table and Room Median Summary

Round 2 Median Bookmark Summary			
Table	Level 2	Level 3	Level 4
1	6.5	12	38
2	15	29.5	47
3	13	32	43
4	7	22	31
Overall	9	27	42

If this was the final round, the Overall medians would be the group recommendations. We understand that everyone does not agree—there is no consensus—which is why we take the median or middle score. Let’s begin with the Level 3 median. The table indicates the group would recommend page XX (27 in the example). This means that we would expect students to have at least a 50% chance of success on the content reflected by items 1 through XX (26; remember you placed your bookmarks AFTER the last item meeting the mastery criterion) to qualify as a threshold Level 3 student. Would anyone care to make a case for a higher or lower bookmark?

Facilitate conversation, allowing panelists to voice their opinions on the Level 3 bookmark. Before you conclude the conversation about Level 3, have the panelists enter a comment so they remember their current thinking about Level 3. Then move on to facilitate group discussion about bookmark differences for Levels 4 and 2.

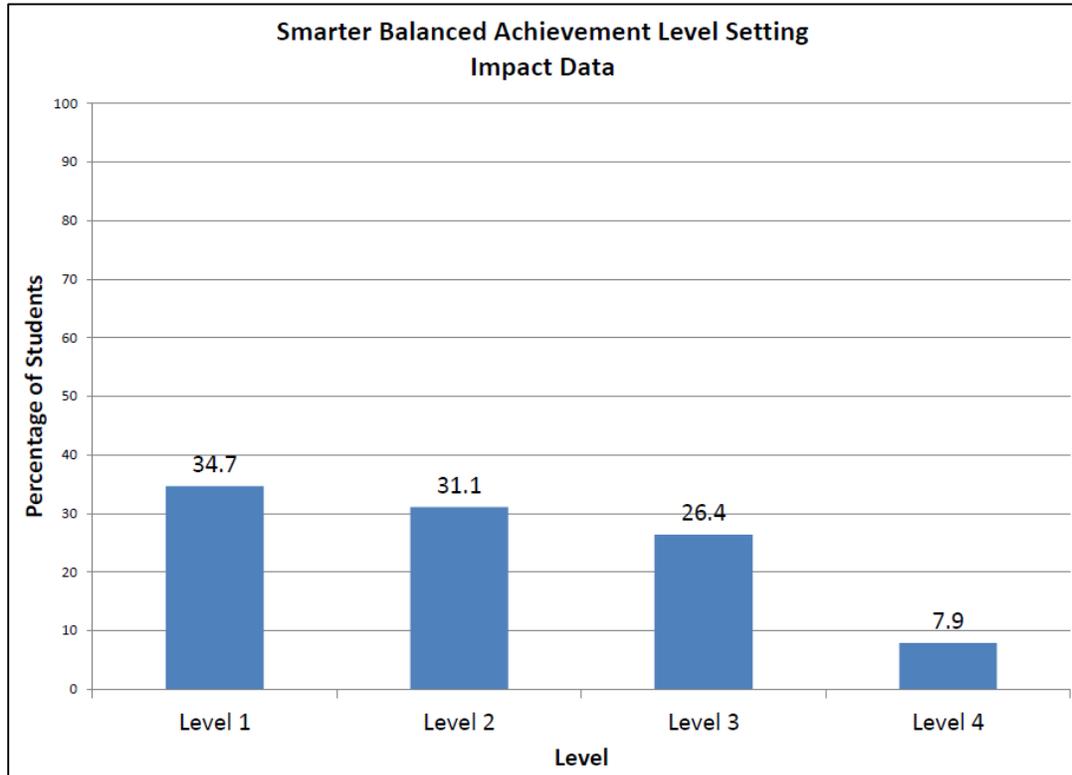
Now, before we break off conversation to make our Round 3 bookmark placements, let’s take a look at one more piece of information—how many students actually would be placed into each Level based on the room median bookmarks. This is provided as another point of reference—it is a reality check. We do not expect you to try to target a specific proportion of students in each level. However, the data you see should not be very surprising to you. If it is, that may be cause to reconsider your bookmark placement, but you must consider the changes to your expectations with respect to content in doing so. Remember, this data is from actual student work from the Spring 2014 Field Test.

Show impact data on the screen in the front of the room.

Sample Impact Data Table

Impact Data				
	Level 1	Level 2	Level 3	Level 4
Overall	34.7	31.1	26.4	7.9

Sample Impact Data Graph



Using the cut scores you as a group came up with in Round 2, these are the percentages of students tested last spring who would be in each of the four achievement levels. Let's take a few minutes to talk about how these results compare to what you expected.

Panelists will need some time to absorb the impact of this information. It may help to remind them that Level 3 is considered Conditionally College/Career Ready or On Target. At grade 11, only Levels 4 students are considered to be College or Career Ready. Then ask,

Does this percentage seem low, high, or about right?

Once panelists have digested these impact numbers, reshare the Online Panel results for Level 3, now with impact data. Note whether this panel's cut score for Level 3 is lower, higher, or about the same as those of the Online Panel subgroups. Note similarities and differences among Online Panel subgroups. Get reactions from the panelists.

Online Panel Median Bookmarks with Impact Data

Role	Median	Range
Teachers	44	38-50
Non-Teacher Educators	49	45-53
Administrators	52	49-55
Other	47	41-53

Impact Data		
	Levels 1 & 2	Levels 3 & 4
Overall	59.7	40.4

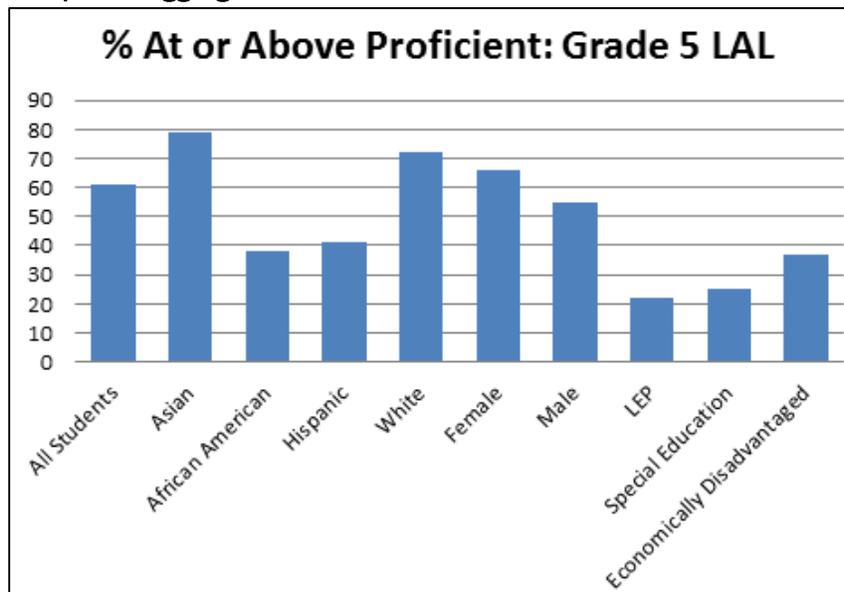
Just as we examined items in the vicinity of your cut scores, let's take a look at the items in the vicinity of the cut scores suggested by these teachers, administrators, higher education faculty, business leaders, parents, and others to see how they interpreted Level 3.

Find pages with locations close to each of the Online Panel subgroup mean cut scores. Discuss these items relative to the Level 3 ALD. Ask the panel,

Do any of these cuts seem out of sync with the ALDs? Do any of them seem close to it?

Next, provide panelists with the disaggregated impact data **based on the overall median bookmarks for the room** such as the sample below. Note the range of percents at or above Level 3. Provide a few minutes for discussion at a room level.

Sample Disaggregated Data



Elicit responses from the panel and note any insights you think should be shared at the daily debriefing.

Readiness Form and Round 3

Draw the Round 2 review to a close and ask panelists to open the Readiness Form Part 4 and complete it. Once they have completed this Readiness Form and submitted it, they will be able to start Round 3 item review. Tell the panelists,

Go the Menu to call up the Readiness Form Part 4. Please take a few minutes to read and complete it. When everyone has successfully completed Part 4, we can begin.

Circulate through the room to make sure everyone is completing the Readiness Form. Answer questions as necessary and appropriate. When everyone has completed and submitted the Readiness Form, have them close the screen and call up Round 3 from the Menu.

I will be available to answer any questions you may have during this final round. Also, as in Rounds 1 and 2, feel free to discuss any aspect of this process with or ask questions of others at your table. However, keep in mind that setting your bookmark is an individual activity. As you did in Rounds 1 and 2, you will set your Level 3 bookmark first, followed by your bookmarks for Level 4 and Level 2. Once you have submitted your final cut scores, do not log out.

Answer any questions that panelists may have, again referring questions about bookmark placements to the alignment of the ALD to the item's content. By this time, there are not likely to be many questions about process, and, indeed, there may be few questions of any type. After answering panelists' questions, direct them to begin Round 3, followed by a break, filling out the Evaluation Form, review of results, and lunch.

Display Bookmark Placement slide. Once panelists have all submitted their Round 3 bookmark recommendations, direct them to take a break until 11:30 while results are tabulated. Close out the round when all panelists have finished and notify the Operations Room.

Review of Round 3

Once panelists return, first present the Round 3 median bookmark recommendations.

Sample Round 3 Median Bookmarks

Round 3 Median Bookmark Summary			
	Level 2	Level 3	Level 4
Overall	9	27	42

Now, move on to Round 3 impact data.

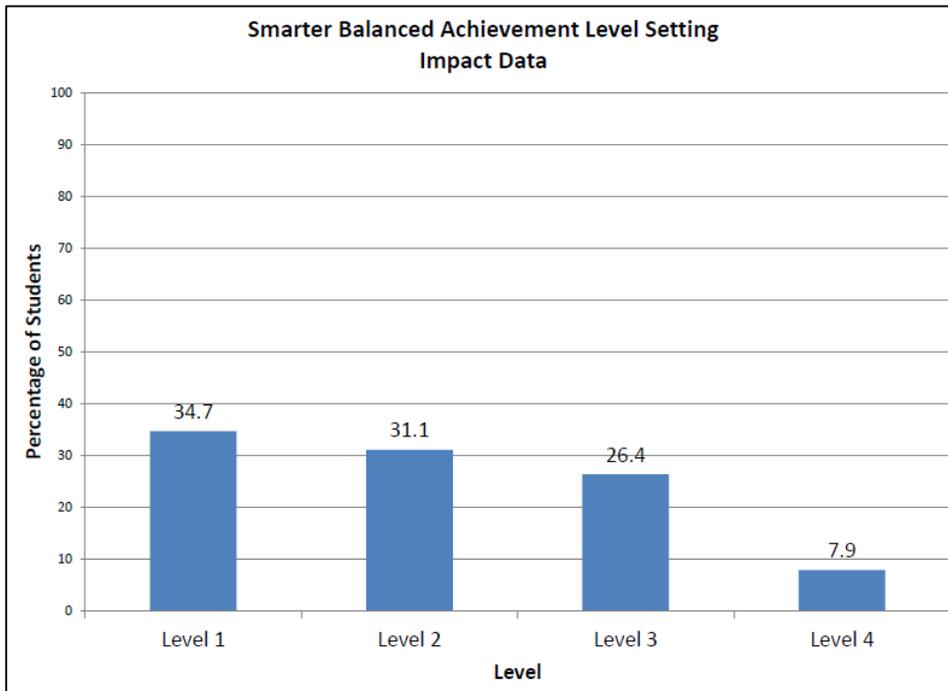
Now that we know where we as a group set our median bookmark recommendations, let's take a look at how many students actually perform at those levels.

Show impact data on the screen in the front of the room.

Sample Round 3 Impact Data Table

Impact Data				
	Level 1	Level 2	Level 3	Level 4
Overall	34.7	31.1	26.4	7.9

Sample Round 3 Impact Data Graph



Using the bookmark recommendations you as a group came up with in Round 3, these are the percentages of students tested last spring who would be in each of the four achievement levels. Let's take a few minutes to talk about how these results compare to what you expected.

Evaluation Form

Once panelists have finished discussing the Round 3 results, explain the Evaluation Form.

Go to your Menu and open the Evaluation Form. It contains questions and statements about the process you used to enter your bookmarks to obtain cut scores. We will need to summarize and present your anonymous responses, along with the final recommended cut scores to Smarter Balanced leadership, external review groups, and the Governing States.

As they finish the Evaluation Form, to the extent possible, thank panelists individually for their participation. For those who will be returning for Vertical Articulation, remind them of that activity and let them know you are looking forward to seeing them again on October 20. Once panelists have



finished their Evaluation Form, thank them for their participation and remind them about the confidentiality of the items and results of the process.

On behalf of Smarter Balanced, I thank you for your contribution to this important project. Lunch will be served at noon, and you are welcome to stay. If you need to leave, the shuttles will be ready to take you back to the airport starting at noon.

Wrap-Up and Closeout

As panelists prepare to leave, check to make sure they have correctly logged out. When the last panelist completes and submits his or her Evaluation Form, notify the Operations Room. When you have verified that all panelists have successfully submitted their Round 3 bookmarks and Evaluation Forms, gather your materials and secure the room for lunch.

Appendix A

Presenting External Data to Achievement Level Setting Panelists

In the course of achievement level setting (ALS) for Smarter Balanced assessments, panelists will review one or more sources of data external to the tests themselves:

- Embedded NAEP and PISA items for grades 4 and 8 (NAEP), and grade 11 (PISA)
- Results of matched-data studies involving ACT and SAT data for grade 11
- Recommendations from the Online Panel for all grades

There is a great deal of concern about how achievement level setting facilitators will present the external data to panelists and how panelists will use those data in determining where to place bookmarks. However, there is much more to the issue than presenting and using the data. In this brief, we also address how data will be selected and evaluated before they are presented to and used by panelists.

Background

Traditionally, achievement level setting has been a closed system. Panelists considered items on a test, the content standards on which those items were based, characteristics of examinees, and score distributions of those examinees on those tests. Advances in the 1990s and early 2000s brought formal achievement level descriptors (ALDs; Reckase, 2001; Cizek & Bunch, 2007, Chapter 3; Reckase & Chen, 2012), which panelists reviewed and applied to the test items in methods such as Bookmark (Mitzel, Lewis, Patz, & Green, 2001; Cizek & Bunch, 2007, Chapter 10; Lewis, Mitzel, Mercado, & Schulz, 2012) and Item Descriptor Match (Ferrara & Lewis, 2012). Performance on the tests in question was typically considered strictly within the framework of those parameters. External criteria were considered in methods such as contrasting or borderline groups (Cizek & Bunch, 2007, Chapter 8), and even then, the external criteria were typically the judgments of teachers or others familiar with the performances of the examinees. Information outside this system was rarely considered.

When claims are made about examinees regarding their readiness for college and career or the next phase of education (e.g., readiness of rising fifth graders for actual fifth grade work), there is a clear implication that performance on the test at hand is predictive of some future performance, if only another test similar to the one at hand, or to performance on some other relevant, parallel measure. This presumption of comparable performance on parallel or future measures has given rise to alternate ways of viewing achievement level setting.

The Bookmark Procedure

Lewis, Mitzel, Mercado & Schulz (2012) describe the use of external benchmarks in Bookmark standard settings conducted since 2002 for various purposes. Lewis, Mitzel, Mercado & Schulz also analyze and describe the results of 11 Bookmark standard settings using external benchmarks with respect to the influence of the benchmarks on the standard setting panelists. In the described Bookmark standard settings, external benchmarks are translated to locations on the test scale so that they can be viewed in terms of a location in the ordered item booklet. The approach described by Lewis, Mitzel, Mercado & Schulz provides an external frame of reference to support the content-

based judgments of standard setting panelists, and is the foundation of the two methods described next—The Benchmark Method (Philips, 2012) and the Briefing Book Method (Haertel, 2002; Haertel, Beimers, & Miles, 2012; O’Malley, Keng, & Miles, 2012).

Lewis, Mitzel, Mercado & Schulz indicate that “(b)enchmarking adds an additional—and perhaps the most persuasive—form of feedback to the standard setting process. As such, the use of benchmarks should be carefully considered with respect to when the benchmarks are introduced, and how panelists are expected to use them to support the judgment task...The policy goals must be explained to panelists and different presentations will alter the use of benchmark information by participants. Instructions accompanying the presentation of benchmarks may be of two types:

Agnostic

- “We provide this information to provide additional context for your judgments; use this information to inform your judgments...,” or
- “Use this information and your own perspective to find the place in the OIB...,” or

Assertive and moderating

- Policymakers believe the cut scores for this assessment should be similar to the benchmark cut scores...”
- “These benchmarks represent optimal cuts in terms of the consistency of percents of students, but you may want to adjust them based on a review of the content students should master...,” or
- “Start with the benchmarks, and realize that the further you move from them, the further you move from the goal of...”

Benchmarks are provided to support and perhaps moderate panelists’ decision making and unless the information is carefully communicated, panelists may believe the message is “Your cuts should be similar to these.” Therefore, the purpose of the benchmarks should be carefully considered by policymakers, and the timing and nature of their communication to panelists should be well-considered and well-delivered.”

The Benchmark Method

The Benchmark method (Phillips, 2012) uses external, national or international benchmark exams such as the National Assessment of Educational Progress (NAEP) and Trends in International Mathematics and Science Study (TIMSS) to inform local (typically state-level) decision making on high-stakes exams. Starting with the premise that state-level performance standards should be based on state-level content standards, Phillips argues that the process is relatively impervious to empirical data. States have perpetuated performance standards that intrinsically seemed adequate or even rigorous when, in reality, they were far below national or international standards.

Phillips’ solution to the problem was to link national or international tests to state tests statistically. Once the tests are statistically linked, scores on one can be compared to scores on the other. Thus, for example, the cut score on a state test that indicates proficiency can be translated onto the scale for NAEP allowing decision makers to see where that score falls relative to the NAEP cut scores for Basic, Proficient, and Advanced.

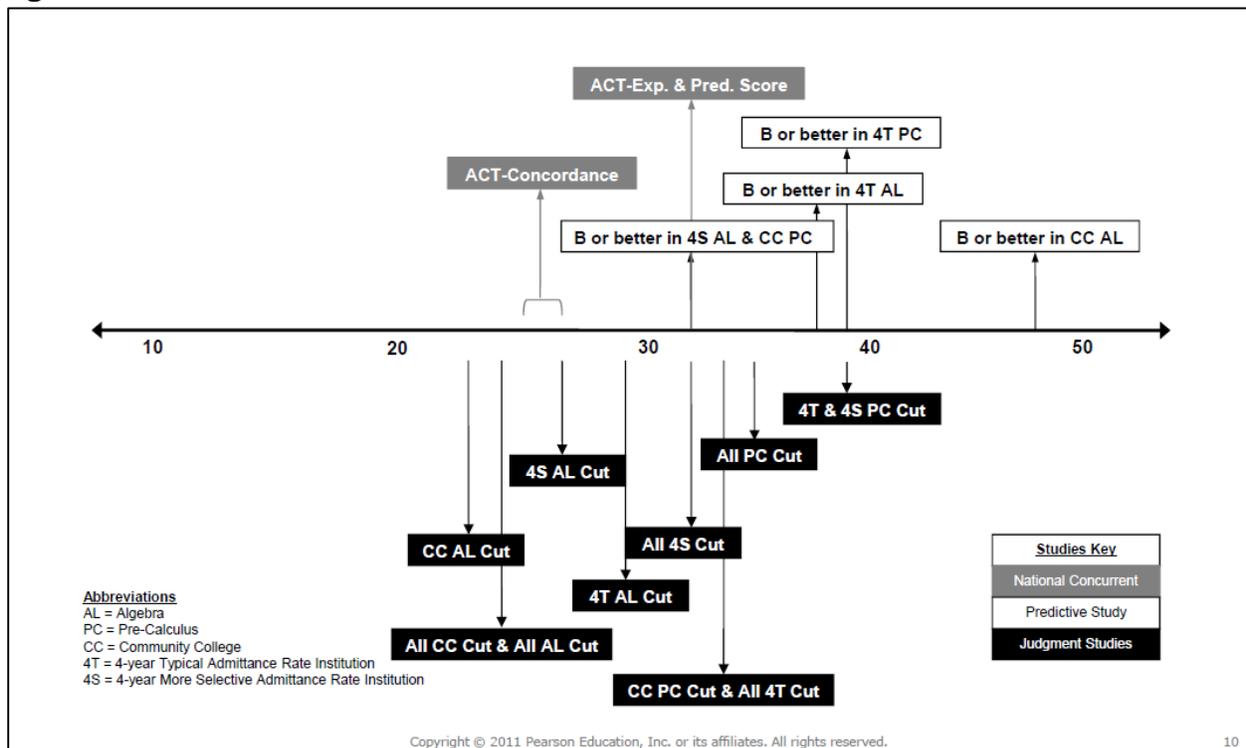
In 2009, West Virginia reset cut scores on some of its statewide tests (Phillips, 2012). Panelists were given ordered item booklets (OIBs) with item maps showing not only the difficulty of each item but

the NAEP-equivalent score for the state scale score associated with that item. NAEP indicators of Advanced, Proficient, and Basic were clearly indicated on the item map. For each item, panelists were able to see the percentage of students obtaining the associated West Virginia scale score, the associated NAEP scale score, and the percentage of students in the NAEP sample reaching that score or higher. The item map for grade 8 Mathematics, for example, showed that the existing cut score for Mastery (state scale score = 624) corresponded to the lower end of the NAEP Basic scale, and that 61 percent of students nationally would reach that score, while only 53% of West Virginia students would reach that score. While panelists were free to set their cut scores anywhere they wished, they were keenly aware that the state desired to raise standards, and the NAEP cut scores were ever before them.

The Briefing Book Method

The Briefing Book method (Haertel, 2002; Haertel, Beimers, & Miles, 2012; O'Malley, Keng, & Miles, 2012) is perhaps the most ambitious application of external data to achievement level setting to date. With this method, panelists receive a wealth of data from concurrent, cross-sectional, and judgmental studies, along with a graphic indicating possible cut scores or “cut score neighborhoods,” as shown in Figure 1 (taken from Way, Miles, & McClarty, 2013).

Figure 1



All the studies are presented, along with an introduction and summary, in a briefing book, shared with panelists in advance of achievement level setting and reviewed again on site. This method was used in setting cut scores for the Algebra II test of the American Diploma Project in 2009. Given the extensive nature of the information to be presented and explained, the agenda called for three

discussion periods lasting a total of three hours and 15 minutes (Miles, Beimers, & Way, 2010, Appendix E).

Selecting, Evaluating, and Presenting Data

The relevance of these methods to our present situation is that each has derived a way of selecting, evaluating, and presenting external data to achievement level setting panelists. While our situation is not identical to those described in the previous section, there are enough common elements to consider these approaches, particularly with respect to the presentation and explanation of the external data.

Selecting data.

As noted above, three forms of external data will be available to panelists – embedded NAEP and PISA items, results of regression studies involving ACT scores, and recommendations from the Online Panel. These data are in the mix for different reasons and will be used in different ways.

- NAEP and PISA items – during field testing, thousands of students responded to forms containing embedded NAEP or PISA items. Those items will be calibrated on Smarter Balanced scales, and some of them will be placed in the ordered item booklets, interspersed with Smarter Balanced items in difficulty order.
- ACT data – the studies currently underway will match students' ACT scores with their scores on Smarter Balanced field tests. We will employ Phillips' (2012) approach to linking the scales, using Phillips' (2012) equation 4 to calculate ACT equivalents of Smarter Balanced scores and his equation 5 to calculate the error variance.
- Online Panel recommendations – Online Panelists will review ordered item booklets and enter a single cut score (for Level 3) between October 6 and October 17. Panelists reviewing Round 1 results on October 14, 16, and 18 will have access to all relevant Online Panel data available up to that point.

Evaluating Data

The windows for ACT data and Online Panel responses are very short. In anticipation of receipt of these data, we have set forth specific evaluation questions and methods, summarized in Table 1. Details follow Table 1.

Table 1. External Data Evaluation Questions and Methods.

External Data Source	Evaluation Question	Evaluation Method
NAEP/PISA Items	Are the items content relevant?	MI/CTB Content specialists will evaluate methodology for initial selection.
	Can items be used to link Smarter Balanced scores to NAEP/PISA scores?	MI/CTB psychometricians will obtain from Smarter Balanced contractors and staff original scale values of NAEP and PISA items used in the OIBs and compare original scale values to Smarter Balanced scale values.
ACT Data	How are tests related?	MI/CTB will do a simple regression of Smarter thetas and ACT scores, and then indicate +/- 1 SE at the ACT Benchmarks for College and Career Readiness.
	Are tests comparable?	MI/CTB psychometricians will employ correlation techniques; minimum expected correlation = .71 (50% shared variance).
Online Panel Recommendations	Are there enough?	Set a minimum of 50 observations per subgroup per test. Failing this minimum, responses will be aggregated across groups.
	Are recommendations stable?	Calculate standard error of the mean for each group mean, and report ranges rather than point estimates; use box and whisker plots or other methods of showing range.
	Are recommendations reasonable and responsible?	MI and CTB have established minimum engagement times and other measures to minimize the risk of an Online Panelist simply entering a cut score without following most of the prescribed process. Responses failing these minima will be discarded.

MI and CTB staff will evaluate NAEP and PISA items in August prior to constructing dummy OIBs for the field test. Preference will be given to items with known NAEP or PISA scale values that place them in the vicinity of critical scores, such as NAEP Basic and Proficient. Items selected will be presented in the OIBs, along with their original national or international statistics and statistical summaries from the Smarter Balanced field test. Appendix B shows two sample items, one from NAEP and one from PISA.

MI/CTB staff will evaluate ACT data in August, using raw scores on Smarter Balanced selected-response items as a surrogate for scale scores (which will become available on September 12) to perform a preliminary linking to Smarter Balanced field tests using Phillips' (2012) benchmarking approach,

MI/CTB staff will evaluate Online Panel data on a daily basis, starting October 6. Using the criteria given in Table 1, they will calculate means and standard deviations for each subgroup for each test. These data will be made available during the Round 1 reviews on October 14 (grade 11), October 16 (grades 6-8), and October 18 (grades 3-5).

Presenting Data.

It is first necessary to note that the external data will be presented in a very well defined context. Facilitators will carefully prepare panelists by building a case around the Common Core State Standards, the Achievement Level Descriptors, the practice test, and the ordered item booklets, emphasizing the rigor of the content standards, the goals of the policy ALDs (and indeed the goals of the Smarter Balanced Assessment Consortium), and the translation of the threshold ALDs into statements of what students at Levels 2, 3, and 4 must know and be able to do, all before they see any external data. Bunch (2012) has described a process for creating such a context by focusing on four goals:

- Gain and maintain panelist trust and respect – we actually begin this process during the recruitment stage and continue it through the summer with travel plans and check-in, greet everyone cheerfully and make sure their lodging and other needs are met, and treat them as professionals and colleagues throughout the workshop.
- Establish rapport with panelists – we also initiate rapport during the recruitment and planning process, make ourselves available for discussions during breaks and off-hour times, and learn their names. Facilitators work to create a friendly, collegial atmosphere during the workshop.
- Elicit feedback – perhaps one of the most important aspects of the three rounds of OIB review, making sure every panelist gets a chance to speak and be heard goes a long way toward making sure every panelist listens.
- Promote equity and transparency – from the outset, parents, teachers, administrators, higher education faculty, and others will be treated as equals, partners in creating something all can point to with pride. Every process will be thoroughly explained, and every question will be honestly and completely answered.

All activities are being planned around these four goals.

According to the agenda for the In-Person Workshop, the first opportunity to view external data will be the afternoon of the first day, in the context of the item maps, during bookmark training. During their initial orientation, panelists will receive their charge from Dr. Willhoft who will elaborate on the goal of every student leaving high school ready for college and career. In general, attainment of that goal requires more rigorous content standards and higher performance standards. Later in the morning and early afternoon, panelists will review and discuss the CCSS, the ALDs, and the practice test. By the time they view any external data, they will be thoroughly steeped in the content and aims of the Smarter Balanced tests and the foundation upon which they are built.

Panelists will be introduced to the purpose of the external benchmarks. Namely, stakeholders and the public may draw comparisons between the performance of students on the Smarter tests with the performance of students on other tests.

The items or benchmarks from NAEP/PISA/ACT/online panels that panelists see on the item maps provide contextual information in the form of items associated with performance levels (NAEP and PISA), scale locations associated with performance levels such as College and Career Ready (ACT), or cut score recommendations of other committees who examined Smarter test content (online panel recommendations).

There are a few things to consider when panelists think about these benchmarks that must be communicated to panelists during training. First, NAEP, PISA, and ACT are not designed to measure the Common Core State Standards. That is, they have test items and item formats which are different—and in some cases, very different—than the Smarter Balanced tests. Second, the cut scores for those tests were not set using the Smarter Balanced achievement level descriptors: remember that the Smarter Balanced ALDs should guide your cut score recommendations. Third, the cut scores for the other tests were set for purposes which might be very different than those for Smarter Balanced: for example, the cut scores established for a national or international test may have different uses or aspirational uses than those for this test.

Because of these important differences, panelists will be cautioned against using any external data as their sole determinative factor in recommending a cut score. Instead, panelists will be instructed that the benchmarks comprise a helpful, supplementary set of information to complement the Smarter Balanced items and test data they study. Facilitators will encourage panelists to ground their cut score recommendations in the Smarter Balanced items and ALDs, and to use the benchmarks for additional context. To assist panelists in the appropriate use of the benchmarks, this information will be scaffolded to the panels throughout the process, starting with training and continuing during and after each round.

One helpful way for participants to use this information is to place the Smarter Balanced content as a primary source for their judgments, then to step back and compare their potential bookmark placements with the items from other tests as secondary sources for their judgments.

Once panelists have an opportunity to view external data, they will evaluate it in a very straightforward way:

- Consider the items in the ordered item booklet in the vicinity of that cut score. Does the content of those items, given the relevant achievement level descriptor, support placement of a cut score here? Why or why not?

With this preface, we can now move on to presentation of data.

Embedded NAEP and PISA items will be included in the OIB, along with Smarter Balanced items. The primary difference in their appearance in the item map will be that NAEP and PISA items may not have a CCSS designation. The NAEP or PISA scale score associated with the item will be presented in the appropriate column of the item map and again on the appropriate page in the OIB.

Figure 2 shows how MI and CTB plan to present ACT data for the high school panels. As noted above, critical scores (in this case the college/career ready scores for each test) will be shown as bands, rather than point estimates.

Figure 2. Sample Item Map.

Page	ID	CCSS	DOK	Ability	Book mark	Comment/ Question	Reviewed	ACT
23	7734	L6.1c	2	4.29			<input checked="" type="checkbox"/>	
24	4089	L7.1c	2	4.55			<input checked="" type="checkbox"/>	
25	1776	L6.1d	3	4.63			<input checked="" type="checkbox"/>	18
26	4789	L7.3a	3	4.67			<input checked="" type="checkbox"/>	
27	1882	L6.1e	2	4.84			<input checked="" type="checkbox"/>	
28	1443-3	L6.3b	3	4.99	Level 3		<input checked="" type="checkbox"/>	
29	4072	L8.1d	3	5.11			<input checked="" type="checkbox"/>	
30	9104	L6.1c	2	5.63			<input checked="" type="checkbox"/>	
31	2345	L4.3b	3	6.07			<input checked="" type="checkbox"/>	
32	1949	L8.1d	3	6.67			<input checked="" type="checkbox"/>	

This particular section of the item map shows ACT scores corresponding to Smarter Balanced Ability scores of 4.63 through 5.11 on pages 25 through 29 of the OIB. At each page, panelists will be confronted with the questions listed above. They will discuss these questions and their answers with other panelists at their table. During the inter-round discussions, they will discuss their answers to these questions with the entire panel. As with other issues, the table-wide and room-wide discussions are expected to have a moderating effect on panelists' responses to the data.

Online Panel data will be shared with In-Person panelists after Round 1, as shown in Figure 3. Panelists will consider these recommended cut scores in the context of room-wide discussion, once again employing the three questions listed above in the same manner that these three questions will have been applied during discussions of their own cut score recommendations prior to viewing the Online Panel recommendations.

Figure 3. Presentation of Online Panel Data.

Role	Median	Range
Teachers	44	38-50
Non-Teacher Educators	49	45-53
Administrators	52	49-55
Other	47	41-53

Feedback and Discussion

Achievement level setting occurs over multiple rounds for a reason. Panelists consider possible cut scores, discuss them, revise their thinking, and enter different ones. Along the way, they consider different information and reasons for setting a cut score in one place versus another. These reasons emerge during the small-group and large-group discussions, but until recently, they were not formally recorded or analyzed. The Bookmark method was the first to include this feedback formally in the standard setting process (Lewis, Mitzel, Mercado, & Schulz, 2012) and Mercado, Barton, & Brandstrom (2013) used “reflection” forms to capture panelists’ thoughts about which data or scenarios they considered most important in placing their cut scores.

During this reflection process, panelists will be reminded that achievement level setting incorporates both content-based information (e.g., test items, achievement level descriptors) and policy-based information (e.g., percentages of students classified in each achievement level, connections to other tests). Panelists will be instructed that the judgments made during the Bookmark Procedure are largely content-based, but policy-based information is presented to add context to their judgments. In the reflection process, panelists will be asked to describe the elements of content- and policy-based information that they considered and how they made a decision. Facilitators will use this information in the discussions during subsequent rounds.

MI and CTB plan to incorporate reflections into the Comments section of the item map/OIB system. When they enter their Round 1 cut scores, they will be prompted to comment on the factors that influenced their decisions, including the NAEP/PISA items and the ACT data. Facilitators will analyze these comments between rounds and include them in the inter-round discussion. When panelists enter their Round 2 cut scores, they will be prompted with specific reference to the Online Panel data as well as the other external data. Once again, facilitators will analyze the comments and include them in the inter-round discussion.

When panelists enter their Round 3 cut scores and complete the Process Evaluation form, they will be prompted to enter their impressions of the external data contained in the OIB and item map, the external data from the Online Panel, and the feedback from their panel as well as the panels who met before them (i.e., final results from the grade 11 panels for the grades 6-8 panels, and final results from grade 11 and grades 6-8 for the grades 3-5 panels). Open-ended questions will prompt panelists to tell which information they used and how they used it.

Conclusion

MI and CTB will carefully select, evaluate, and present relevant external data to panelists during the In-Person Workshop. Similar data, in condensed form, will also be presented to Online Panelists. The process by which these data are selected, evaluated, presented, and used will be established on the



four goals spelled out by Bunch (2012) and honed by application of techniques adapted from the benchmark and briefing book procedures.

References

- Bunch, M. B. (2012). Practical issues in standard setting. In G. J. Cizek (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*. New York: Routledge.
- Cizek, G. J., & Bunch, M. B. (2007). *Standard Setting: Establishing and Evaluating Performance Standards on Tests*. Thousand Oaks, CA: Sage, Chapters 3, 8, and 10.
- Ferrara, S. & Lewis, D. M. (2012). The item-descriptor (ID) matching method. In G. J. Cizek (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*. New York: Routledge.
- Haertel, E. H., Beimers, J., & Miles, J. (2012). The briefing book method. In G. J. Cizek (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*. New York: Routledge.
- Lewis, D. M., Mitzel, H. C., Mercado, R. L., & Schulz, M. (2012). The bookmark standard setting procedure. In G. J. Cizek (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*. New York: Routledge.
- Mercado, R., Barton, K., & Brandstrom, A. (2013). Design for Standard Setting DC CAS Grades 3–8 and 10 Mathematics. Report to District of Columbia Public Schools. Monterey, CA: CTB/McGraw-Hill.
- Miles, J., Beimers, J., & Way, D. (2010). The Modified Briefing Book Standard Setting Method: Using Validity Data as a Basis for Setting Cut Scores. Paper presented at the annual meeting of the American Educational Research Association (AERA) and the National Council on Measurement in Education (NCME).
- Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting Performance Standards: Concepts, Methods, and Perspectives*. Mahwah, NJ: Lawrence Erlbaum.
- O'Malley, K., Keng, L., & Miles, J. (2012). From Z to A: Using validity evidence to set performance standards. In G. J. Cizek (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*. New York: Routledge.
- Phillips, G. (2012). The benchmark method of standard setting. In G. J. Cizek (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*. New York: Routledge.
- Reckase, M. D. (2001). Innovative methods for helping standard-setting participants perform their task: The role of feedback. In G. J. Cizek (Ed.), *Setting Performance Standards: Concepts, Methods, and Perspectives*. Mahwah, NJ: Lawrence Erlbaum.
- Reckase, M. D., & Chen, J. (2012). The role, format, and impact of feedback to standard setting panelists. In G. J. Cizek (Ed.), *Setting Performance Standards: Concepts, Methods, and Perspectives*. Mahwah, NJ: Lawrence Erlbaum.
- Way, D., Miles, J. & McClarty, K. (2013). Using Construct Maps in Evidence Based Standard Setting. Paper presentation at the National Conference on Student Assessment.

Appendix B Presentation of Sample NAEP and PISA Items

These are some sample items for your review only and need not be shared with panelists. Note the types of metadata included, and check your OIB in advance to see what metadata are available.

Sample 8th Grade NAEP Reading Item with Data

Passage about E. B. White

How do you think E. B. White feels by the end of his visit to the zoo? Support your answer with reference to the essay.

Score & Description

Acceptable

Responses at this level provide an appropriate feeling inferred from the essay. Responses may make specific reference to White's experience at the zoo or may provide a more general statement that reflects White's experience.

- He feels exhilarated by watching the deer give birth.
- He feels happy that he got to scratch the fawns behind the ears.
- He feels that he has witnessed something special and it was a once in a lifetime deal.

Unacceptable

Responses at this level may provide irrelevant details or personal opinions or a feeling that is so general it could apply to any visit to a zoo.

- I think he feels fulfilled and happy because his description uses lots of details.
- He feels he has seen too much for one day.
- He was happy to see the animals.

Scale Score: 278 (Basic = 243; Proficient = 281)

Percentage of NAEP 8th grade sample answering correctly in 2009: 39%

Smarter Balanced Field Test Sample: XX%

Sample PISA Mathematics Item with Data

HEIGHT

There are 25 girls in a class. The average height of the girls is 130 cm.

Question 1: HEIGHT M421Q01 - 019

Explain how the average height is calculated.

Scoring

Correct

Explanations that include:

- Sum the individual heights and divide by 25.
- You add together every girl's height and divide by the number of girls.
- Take all the girls' heights, add them up, and divide by the amount of girls, in this case 25.
- The sum of all heights in the same unit divided by the number of girls.

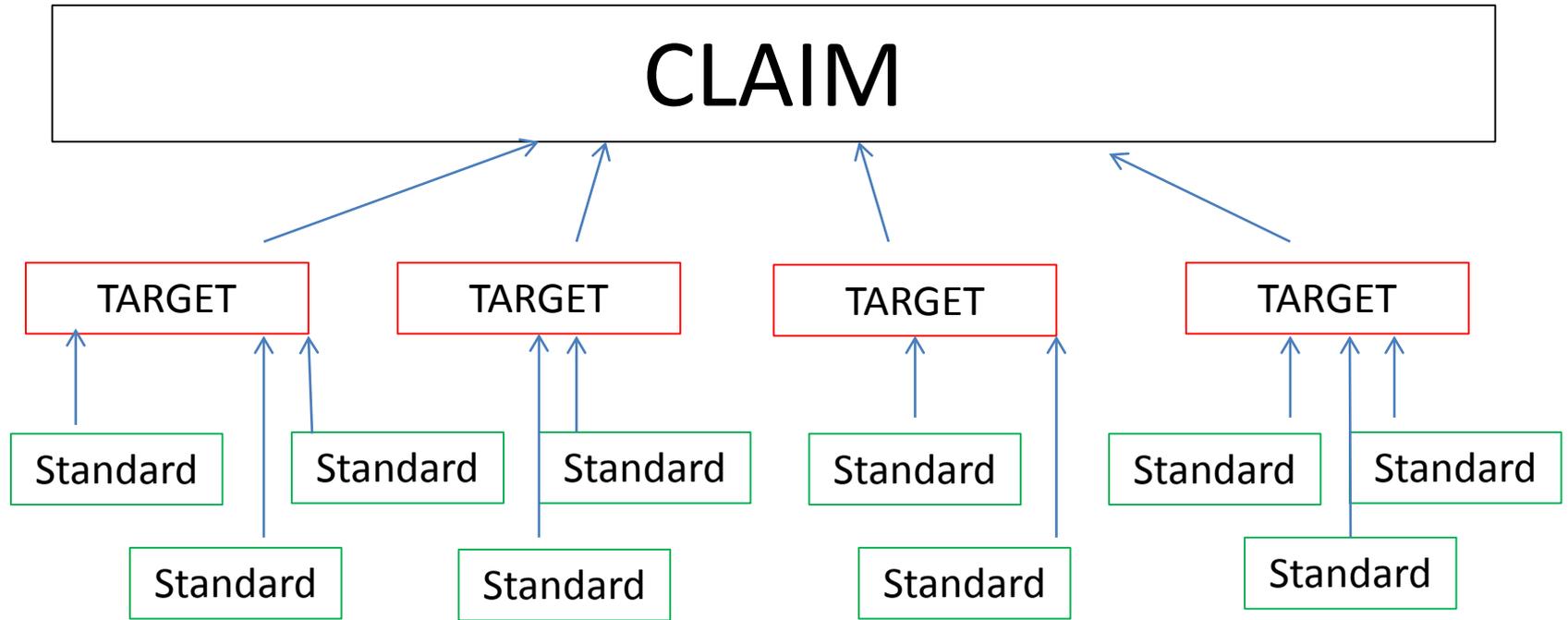
Incorrect

Other responses.

Percentage of students answering correctly in PISA 2006:

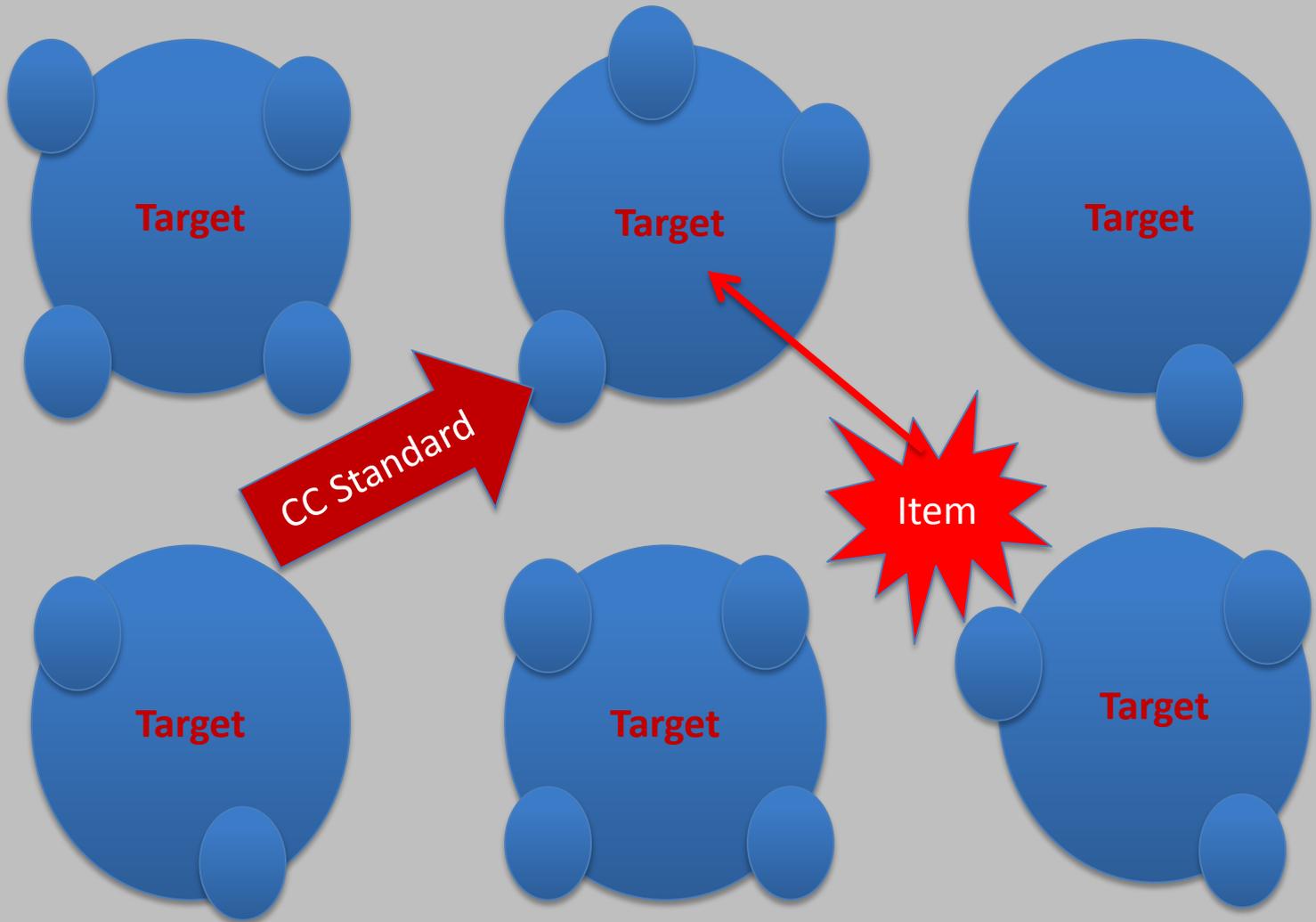
OECD average: 63%

U.S.: 60%





**C
L
A
I
M**





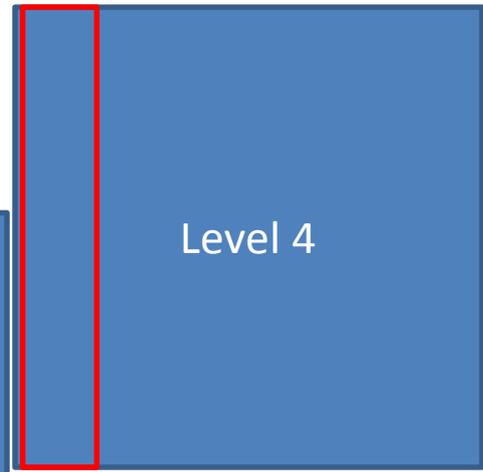
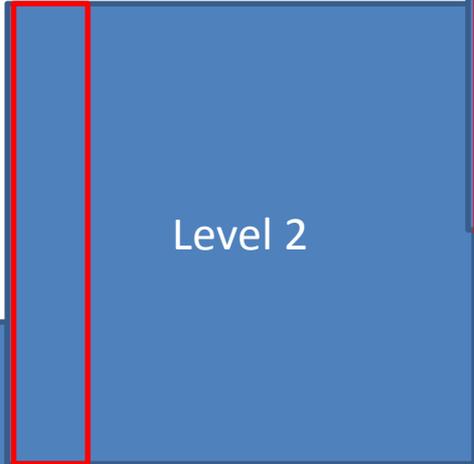
Achievement Level Descriptors

- **Policy ALDs**
 - general descriptors that articulate the goals and rigor for the final performance standards
- **Range ALDs**
 - written to describe expectations for each TARGET at each proficiency level
 - grade- and content-specific descriptors that may be used by test developers to guide item writing
 - describe the cognitive and content rigor that is encompassed within particular achievement levels
- **Threshold ALDs**
 - developed in conjunction with or following range ALDs and are used to guide standard setting
 - are a SUBSET of the range ALDs
 - use only the information from the range ALDs that defines the minimum performance required for meeting a particular achievement-level expectation





The Range ALDs describe the cognitive and content rigor that is encompassed within particular achievement levels for each TARGET.



The Threshold ALDs are a subset of the Range ALDs. They define minimum performance for Levels 2, 3, and 4.

State	Lead	Observer	Nominee	Nominee E-Mail	Info Sent	HS	6-8	3-5	VAC
CA	Diane Hernandez	Yes	Eric Zilbert	ezielbert@cde.ca.gov	12-Sep			✓	✓
CT	Gail Pagano	Yes	Gail Pagano	gail.pagano@ct.gov	12-Sep	✓			
DE	Brian Touchette	Yes	Carolyn Lazar	carolyn.lazar@doe.k12.de.us	12-Sep	✓	✓	✓	✓
HI	Patricia Reiss	Yes	Patricia Reiss	patricia_reiss@notes.k12.hi.us	12-Sep	✓	✓	✓	✓
ID	Angela Hemingway	Yes	Luci Willits	lbwillits@sde.idaho.gov	12-Sep				✓
ID	Angela Hemingway	Yes	Toni Wheeler	tcwheeler@sde.idaho.gov	12-Sep	✓			
ID	Angela Hemingway	Yes	Nichole Hall	nhall@sde.idaho.gov	12-Sep		✓		
ID	Angela Hemingway	Yes	Angela Hemingway	ahemingway@sde.idaho.gov	12-Sep			✓	
IA	Tom Deeter								
ME	Rachelle Tome	Yes	Charlene Tucker	charlene.tucker@maine.gov	16-Sep	✓	✓		
MI	Andrew Middlestead	Yes	Linda Howley	howleyl@michigan.gov	12-Sep		✓		
MO	Michael Muenks	Yes	Michael Muenks	michael.muenks@dese.mo.gov	12-Sep			✓	✓
MT	Judy Snow	No							
NV	Cindy Sharp	No							
NH	Scott Mantie	Yes	Nicole Heimarck	nheimarck@sprise.com	12-Sep		✓	✓	✓
NC	Tammy Howard	Yes	Kinge Mbella	kinge.mbella@dpi.nc.gov	16-Sep	✓	✓	✓	
ND	Robert Bauer	Yes	Justin Wageman	justin.wageman@ndsu.edu	12-Sep	✓	✓	✓	
OR	Doug Kosty	No							
SD	Jan Martin	Yes	Jan Martin	jan.martin@state.sd.us	12-Sep			✓	✓
VI	Alexandria Baltimore-Hookfin	No							
VT	Frank Gerdman	No							
WA	Robin Munson	Yes	Linda Drake	linda.drake@k12.wa.us	12-Sep	✓			
WV	Vaughn Rhudy	Yes	Timothy Butcher	tbutcher@k12.wv.us	12-Sep			✓	✓
WI	Troy Couillard	Yes	Viji Somasundaram	visalakshi.somasundaram@dpi.wi.gov	16-Sep			✓	
WY	Deb Lindsey	No							
Total On Site						8	8	11	8



TABLE OF CONTENTS

Facilitator Script Overviews

Day 1 AM Content Standards Review, ALDs Review, and Practice Test	742
Day 1 PM Software and Ordered Item Booklet Review	743
Day 2 AM Bookmark Placement – Orientation and Round 1	744
Day 2 PM Round 2 Bookmark Placement.....	746
Day 3 AM Round 3 Bookmark Placement.....	747

FACILITATOR SCRIPT OVERVIEW

Day 1 AM

Content Standards Review, ALDs Review, and Practice Test

Goal: Review content standards and threshold ALDs. Allow panelists to experience the assessment online so they can see how the various item types operate with the technology and empathize with the students.

What you need: A panelist login that has never been used for this grade and content for computer attached to projector. Paper copies of ALDs. Smarter Balanced Assessment Framework slides.

Outline of session: Facilitator leads group introductions. Everyone logs in as a panelist. Facilitator walks through content standards, leads brief discussion, and provides time for individual review and repeats the process for ALDs. Panelists access Smarter Balanced website and take practice test. Panelists complete Readiness Form Part 1.

Highlights

- *Make sure you have the correct panelists signed in and at the correct tables.*
- *Model logging in with your assigned panelist login, explaining where to find their event code (name tag) and table number (card on table).*
- *Show them how to access the CCSS. Review material mentioned in morning session with them. Time spent here should be minimal because they should already be familiar with the content standards and the focus of their efforts should be on the next section.*
- *Show them how to access the ALDs and discuss threshold ALDs. The more they discuss the ALDs now, the easier it will be to use them later on. Give them the paper copies of the ALDs.*
- *Grades 3-8: Level 3 is considered On Target. Levels 3 & 4 combined represent all the students who are considered ready for the next phase of education or career.*
- *Grade 11: Level 3 is Conditionally Content-Ready/Exempt from Developmental Course Work. Only Level 4 is considered Content-Ready. This distinction is critical in the setting of the grade 11 cut score for Levels 3 and 4.*
- *Encourage empathy with students and exploration of new item types.*
- *Have panelists log out before leaving for lunch. Otherwise, software appears to be logged in but is not recording information after lunch.*

Schedule

9:35 a.m.	Introductions
9:45	Orientation to hardware and software
10:00	Discussion of CCSS and ALDs
10:30	Break
10:45	Practice test administration
11:55	Discussion of practice test
12:20 p.m.	Closing comments on tests
12:25	Completion of Readiness Form Part 1
12:30	Lunch

FACILITATOR SCRIPT OVERVIEW

Day 1 PM Software and Ordered Item Booklet Review

Goal: First, introduce navigating the ordered item booklet with the software and train panelists on how to use the two focus questions to analyze each item. Then, review the entire OIB using those two questions.

What you need: Panelist login for computer attached to projector. Slide showing the two questions.

Outline of session: Everyone logs back in as a panelist. Facilitator leads them to OIB in the software and walks through explanation of item map and ordered item booklet navigation and introduces two focus questions. Facilitator leads discussion through the first few items, modeling how to use the focus questions and add their answers to these questions in the Comments section. Identify items from NAEP and PISA and indicate they are not on the tests students will take but are provided for context. Panelists spend the remaining time working through the remaining items in the OIB.

Focus Questions

1. *What do you know about a student who responds successfully to this item; that is, what skills must a student have in order to know the correct answer?*
2. *What makes this item more difficult than preceding items?*

Highlights

- *Make sure the panelists are entering the appropriate full OIB and not the 6-item Orientation Round OIB.*
- *Avoid discussion of bookmarks. Focus should solely be on the items and the two questions. They will receive bookmark training later.*
- *Grades, 4, 8 and 11: Note NAEP (4, 8, 11) and PISA (11) items.*
- *All grades: note placement and connectedness of performance task components.*
- *Keep working through the two focus questions until it appears that the group understands how to use them. Then have the table leaders reinforce.*
- *Remind the panelists of bookmark training the next morning. Emphasize being on time.*

Schedule

1:50 p.m.	Introduce the OIB and item map
2:00	Lead panelists through focus questions
2:20	Model how to use focus with sample of items
2:40	Panelists review items in OIB and enter comments on item map
4:50	Wrap up
5:00	Adjourn

FACILITATOR SCRIPT OVERVIEW

Day 2 AM

Bookmark Placement – Orientation and Round 1

Goal: Practice setting the Level 3 bookmark in the 6-item orientation round and then set all three Round 1 bookmarks.

What you need: Panelist login for computer attached to projector. Second computer with facilitator login. The two focus question slide and the Bookmark placement slide. Excel spreadsheet to calculate medians based on room's frequency distribution of bookmark placements.

Outline of session: Everyone logs back in as a panelist. Facilitator reviews how to set a Level 3 bookmark with the software. Panelists discuss 6 items using 2 focus questions. Facilitator discusses the Bookmark placement slide and then panelists set their Level 3 bookmark. Facilitator collects bookmark placements into Excel spreadsheet and leads discussion into reasoning behind panelists' bookmark placements. Panelists complete Readiness Form and then Facilitator directs panelists to set Round 1 bookmarks.

Focus Questions

1. *What do you know about a student who responds successfully to this item; that is, what skills must a student have in order to know the correct answer?*
2. *What makes this item more difficult than preceding items?*
3. ***Would a student at the threshold of the achievement level have at least a 50% chance of responding successfully to this item (or score point)? If they would, then move on to the next item; if not, place a bookmark on this page.***

Highlights for Orientation Round

- *Do not actually submit any bookmarks when modeling the process for panelists as they can affect the data collected. If you accidentally submit, please notify the operations room.*
- *Ask them why the item before their bookmark fits the Level 3 threshold descriptor and why the bookmarked page does not.*
- *Reinforce we don't expect everyone to always agree.*
- *Keep focus on Bookmark Placement poster and ALDs.*

Highlights for Round 1

- *Remind the panelists to look at their comments from yesterday regarding the answers to the two questions for each item.*
- *Remind panelists of nature and purpose of NAEP/PISA items (grades 4, 8, and 11 only) and performance tasks.*
- *Keep focus on Bookmark Placement poster and ALDs.*
- *Make sure to keep an eye on completion status in facilitator login, close Round 1 when everyone is finished, and notify Operations Room.*
- *Arrange for table leaders to return 15 minutes early from lunch.*

Schedule

9:00 a.m.	Orientation round exercise
9:40	Completion of Readiness Form Part 2
9:45	Round 1
11:45	Completion of Reflection Questionnaire Part 1
12:00 Noon	Lunch

FACILITATOR SCRIPT OVERVIEW

Day 2 PM Round 2 Bookmark Placement

Goal: Review Round 1 results. Set Round 2 bookmarks.

What you need: Panelist login for computer attached to projector. Second computer with facilitator login. Bookmark Placement slide.

Outline of session: Facilitator meets with table leaders before session to explain how to run table discussion. Everyone else returns and logs back in as a panelist. Facilitator reviews additions (thumbtacks, location, ACT) to item map. Table leaders lead discussion of all panelist bookmarks, starting at Level 3 and direct discussion on rationale behind different placements at a table level. Discuss Online Panel median bookmark slide. Table leaders finish table discussion including new information. Panelists complete Round 2 readiness form. Facilitator displays Bookmark Placement slide and panelists set their Round 2 bookmarks.

Highlights

- *Review with table leaders how to lead discussion at their table.*
- *New items: thumbtacks, location, ACT.*
- *Direct tables to share their bookmarks at their tables and note low, high, and median BMs.*
- *Display online panelist median bookmarks as a point of reference and provide context.*
- *Review setting the bookmark on item after which they stop having at least a 50% chance of correctly answering the item.*
- *Make sure to keep an eye on completion status in facilitator login, close Round 2 when everyone is finished, and notify Operations Room.*

Schedule

1:00 p.m.	Review of Round 1
2:30	Completion of Readiness Form Part 3
2:35	Round 2
4:50	Completion of Reflection Questionnaire Part 2
5:00	Adjourn

FACILITATOR SCRIPT OVERVIEW

Day 3 AM Round 3 Bookmark Placement

Goal: Review Round 2 results including table and room medians and impact data. Share online panel results again, now with impact data and complete readiness form. Set Round 3 bookmarks. Break. Review Round 3 results and complete evaluation form.

What you need: Panelist login for computer attached to projector. Second computer with facilitator login. Results of Round 2 slides. Online impact data slides. Bookmark placement slide.

Outline of session: Everyone logs back in as a panelist. Facilitator walks through at a room level the results of Round 2 for the room (histogram), each table median, the room median, and the associated impact data with a discussion of the meaning of it and how it can be used (30 minutes). Facilitator reviews the range of bookmarks at each Level, using the different table medians as a focus, and leads discussion on rationale behind different placements at a room level (60 minutes). Panelists complete Round 3 readiness form. Facilitator displays Bookmark Placement slide and panelists set their Round 3 bookmarks. Participants break and return for review of final recommendations. Panelists complete final evaluation form.

Highlights

- *You are explaining impact data and how they can use it. Remind them that they are the ones to decide if they will use it.*
- *Discussion is at a room level in this round.*
- *Review setting the bookmark on first item on which they stop having at least a 50% chance of correctly answering the item.*
- *Make sure to keep an eye on completion status in facilitator login, close Round 3 when everyone is finished, and notify Operations Room.*

Schedule

8:30 a.m.	Convene in breakout rooms; review Round 2 results
9:45	Completion of Readiness Form Part 4
9:50	Round 3
11:30	Review of Round 3 Results
11:40	Completion of Evaluation Forms and Reflection Questionnaire Part 3
12:00	Adjourn/Lunch



Table Leader Training



Smarter Balanced Achievement Level Setting
Dallas, TX, October 2014





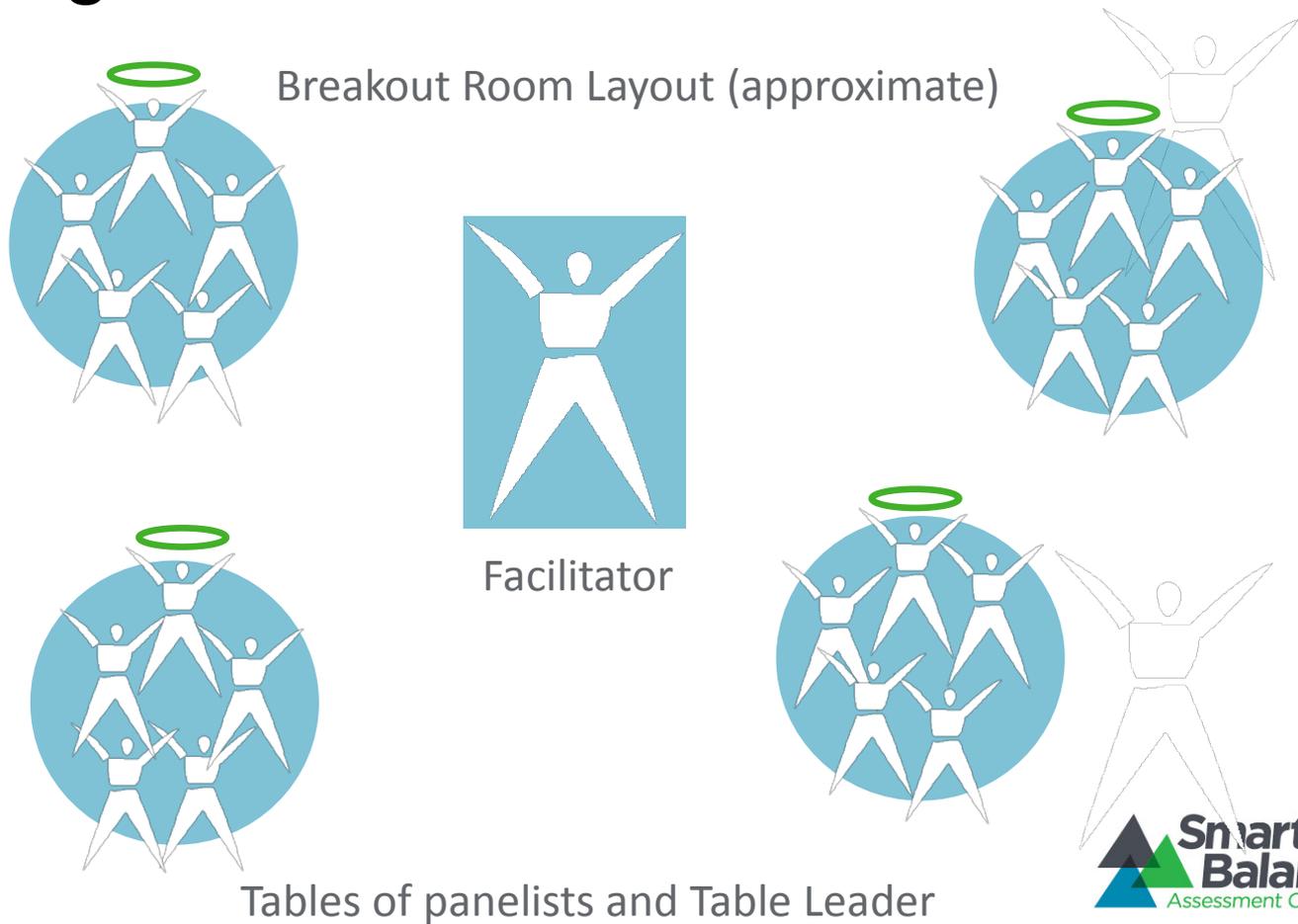
Overview of Table Leader Training

- Understanding the achievement level setting process
- Understanding the Bookmark Procedure
- Understanding your role



Facilitator Names

- Name, grade, content





Understanding the Achievement Level Setting Process





What Is an Achievement Level Setting?

- A process that lets educators recommend threshold test scores that separate students into achievement levels.
- Recommendation is based on the content that students should know in each of those levels.



Smarter Balanced Policy Achievement Level Descriptors

Level 1

- Demonstrates minimal understanding of and ability to apply the knowledge and skills associated with college content-readiness.

Level 2

- Demonstrates partial understanding of and ability to apply the knowledge and skills associated with college content-readiness.

Level 3*

- Demonstrates adequate understanding of and ability to apply the knowledge and skills associated with college content-readiness.

Level 4

- Demonstrates thorough understanding of and ability to apply the knowledge and skills associated with college content-readiness.

* Level 3 is used for federal NCLB purposes.



Types of ALDs

- Policy
- Range
- Threshold
- Reporting



Types of ALDs

- Policy
- Range
- **Threshold – students just entering Level 2, 3, 4**
- Reporting

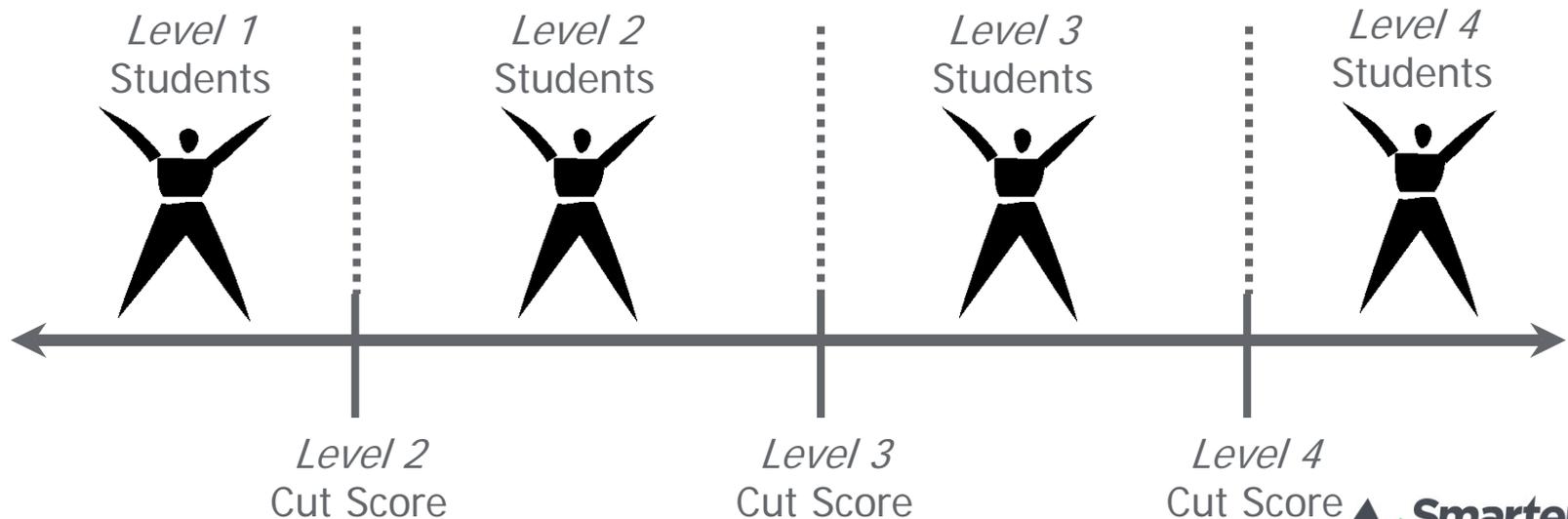


How Do We Set Achievement Levels?

- ~~Percentages~~
 - Arbitrary
 - Test-specific
 - Do not consider content
- Content
 - Uses pre-established content standards
 - Considers educational objectives
- Bookmark Procedure

Purpose of the Achievement Level Setting

- Allows cut scores to be set on the test scale
- The test scale represents the ability of students





Purpose of the Achievement Level Setting

- You will set three cut scores on the test scale.
- Students who meet or exceed the cut score will have enough knowledge, skills and abilities to be classified as *Level 3*.
 - Also *Level 1, Level 2, and Level 4*.
- Decisions will be based on the Achievement Level Descriptors based on the Common Core State Standards.



Developing the Expertise to Make Content-Based Achievement Level Recommendations

- Activities that must be completed:
 - Study content levels
 - Study achievement level descriptors
 - Study test items
 - Take practice test
 - Study ordered item booklets using item maps

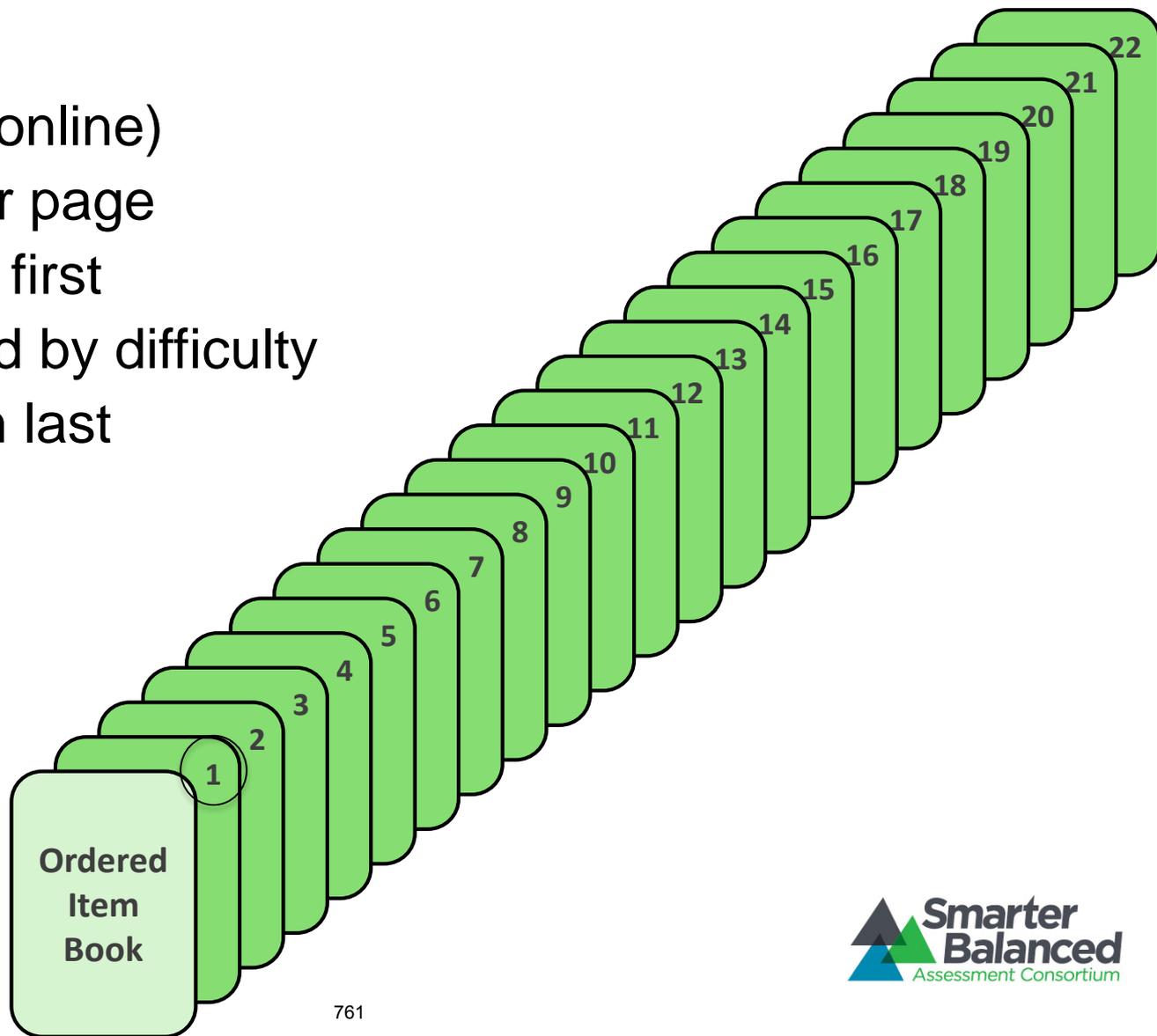


Understanding the Bookmark Procedure



Ordered Item Books (OIBs)

- Virtual OIB (online)
- One item per page
- Easiest item first
- Items ascend by difficulty
- Hardest item last





Item Map

Item Map for Booklet: Achievement Level Setting

English Language Arts/Literacy, 8

Round: 1

Enter Comment

Review each question in the ordered item booklet by clicking on the page number below, starting with page 1.

View Selected



Page	Location	Claim and Target	Depth of Knowledge	Bookmark	Comments	Reviewed
1	362	Claim: 3-L,Target: 4-8	2			<input type="checkbox"/>
2	368	Claim: 3-L,Target: 4-8	1			<input type="checkbox"/>
3	397	Claim: 2-W,Target: 8-8	2			<input type="checkbox"/>
4	409	Claim: 3-L,Target: 4-8	1			<input type="checkbox"/>
5	420	Claim: 1-LT,Target: 7-8	3			<input type="checkbox"/>
6	433	Claim: 1-LT,Target: 1-8	2			<input type="checkbox"/>
7	438	Claim: 3-L,Target: 4-8	3			<input type="checkbox"/>
8	443	Claim: 2-W,Target: 8-8	1			<input type="checkbox"/>
9	444	Claim: 2-W,Target: 8-8	2			<input type="checkbox"/>
10	452	Claim: 2-W,Target: 9-8	1			<input type="checkbox"/>
11	455	Claim: 1-LT,Target: 3-8	2			<input type="checkbox"/>
12	457	Claim: 3-L,Target: 4-8	3			<input type="checkbox"/>
13	458	Claim: 4-CR,Target: 3-8	2			<input type="checkbox"/>

Set Bookmark

Submit Bookmarks



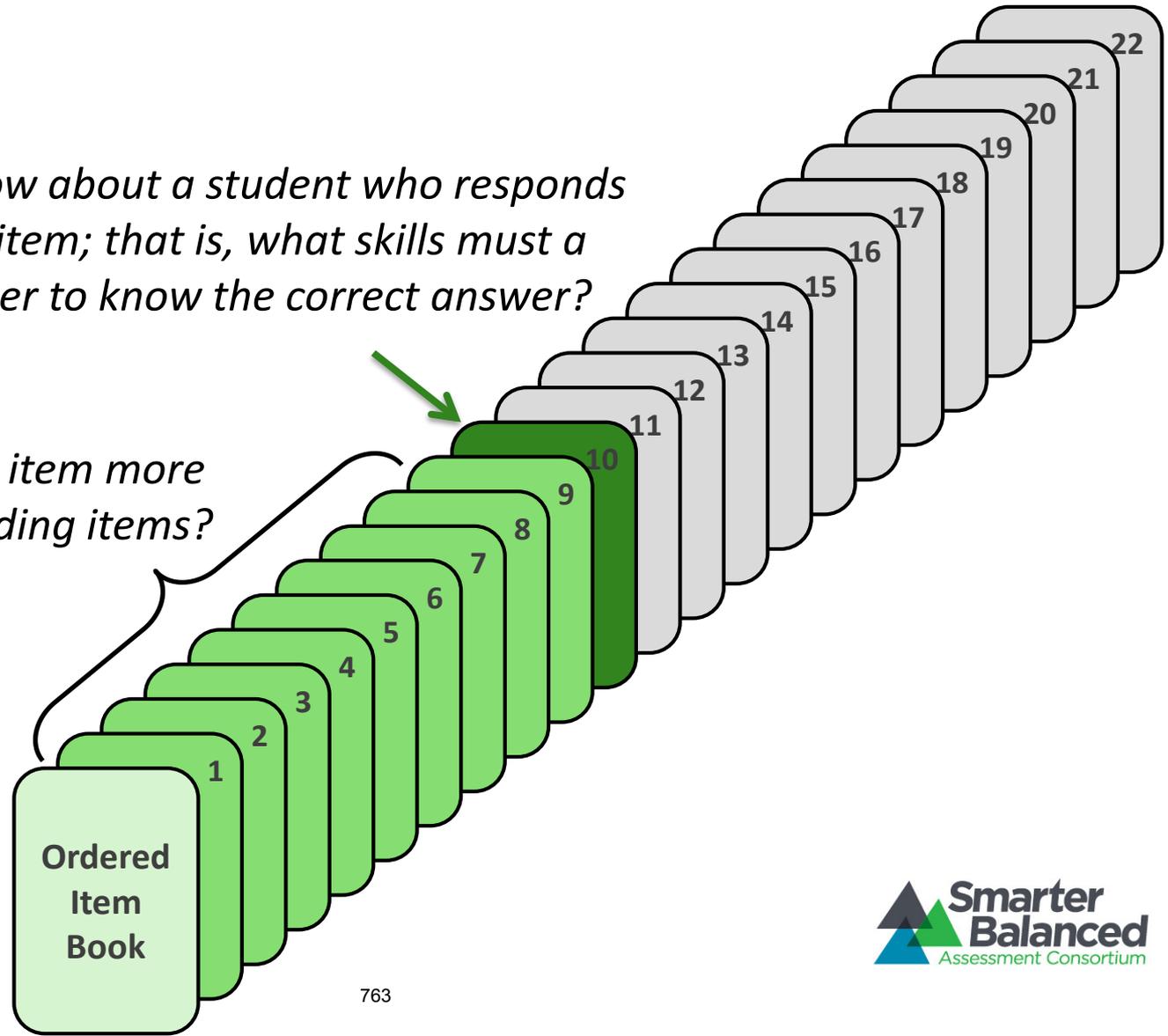
Study the Items in the OIB

As you study each item in the OIB, discuss two questions with your fellow panelists:

1. *What do you know about a student who responds successfully to this item; that is, what skills must a student have in order to know the correct answer?*

2. *What makes this item more difficult than preceding items?*

This activity helps you acquire the knowledge to make content-based cut score recommendations.



Learning With Your Fellow Panelists

- You will study the OIB tomorrow afternoon
- You may and should discuss the OIB items with your fellow panelists
- You may study the OIB individually, if this is more efficient for you, but may want to confer with your fellow panelists from time to time



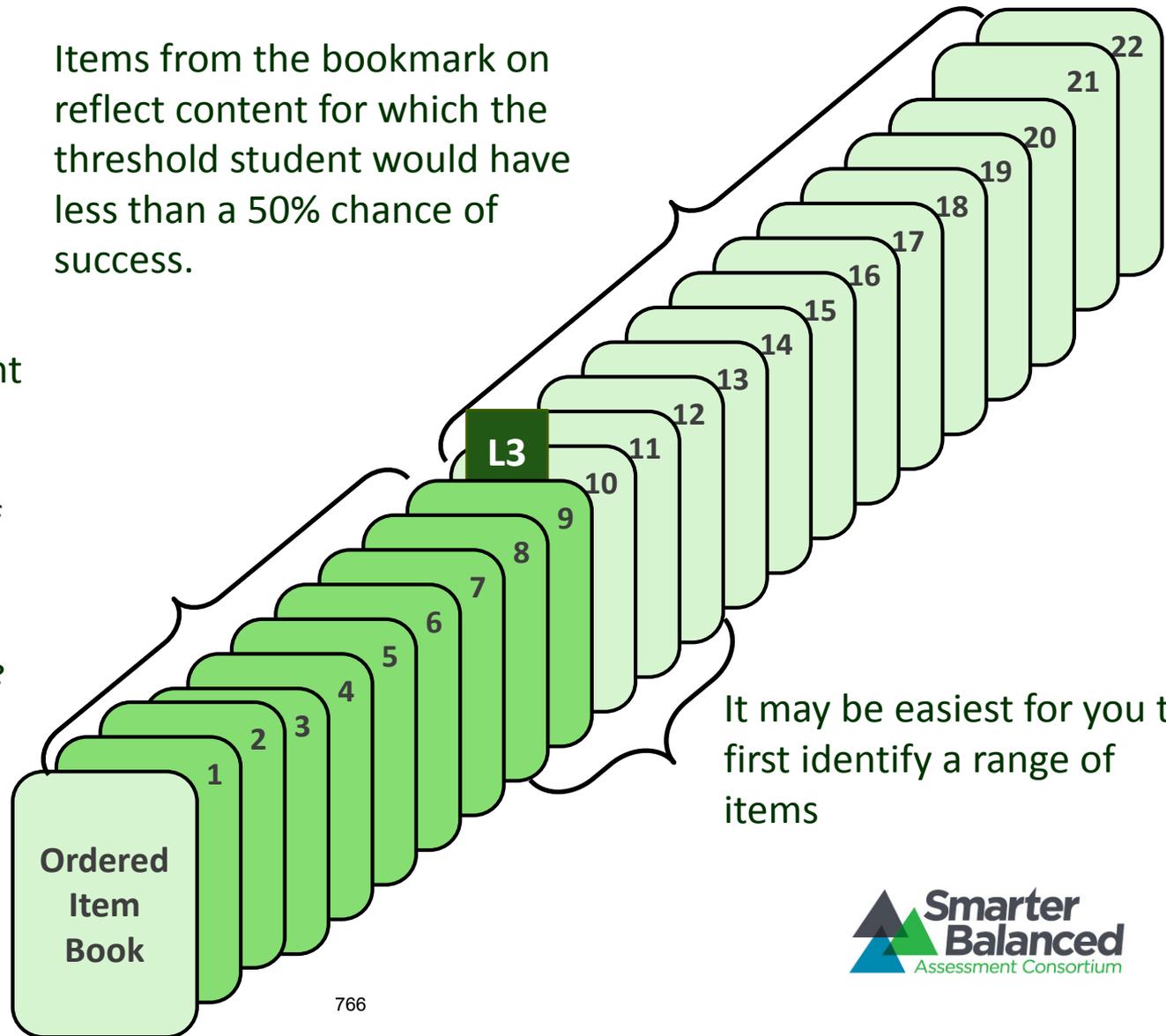
Bookmark Overview

- The Bookmark Procedure is a process used to set cut scores that define achievement levels by comparing and aligning items and content to achievement level descriptors
- It is so named because you place bookmarks in an ordered item booklet of test items to make your cut score recommendations



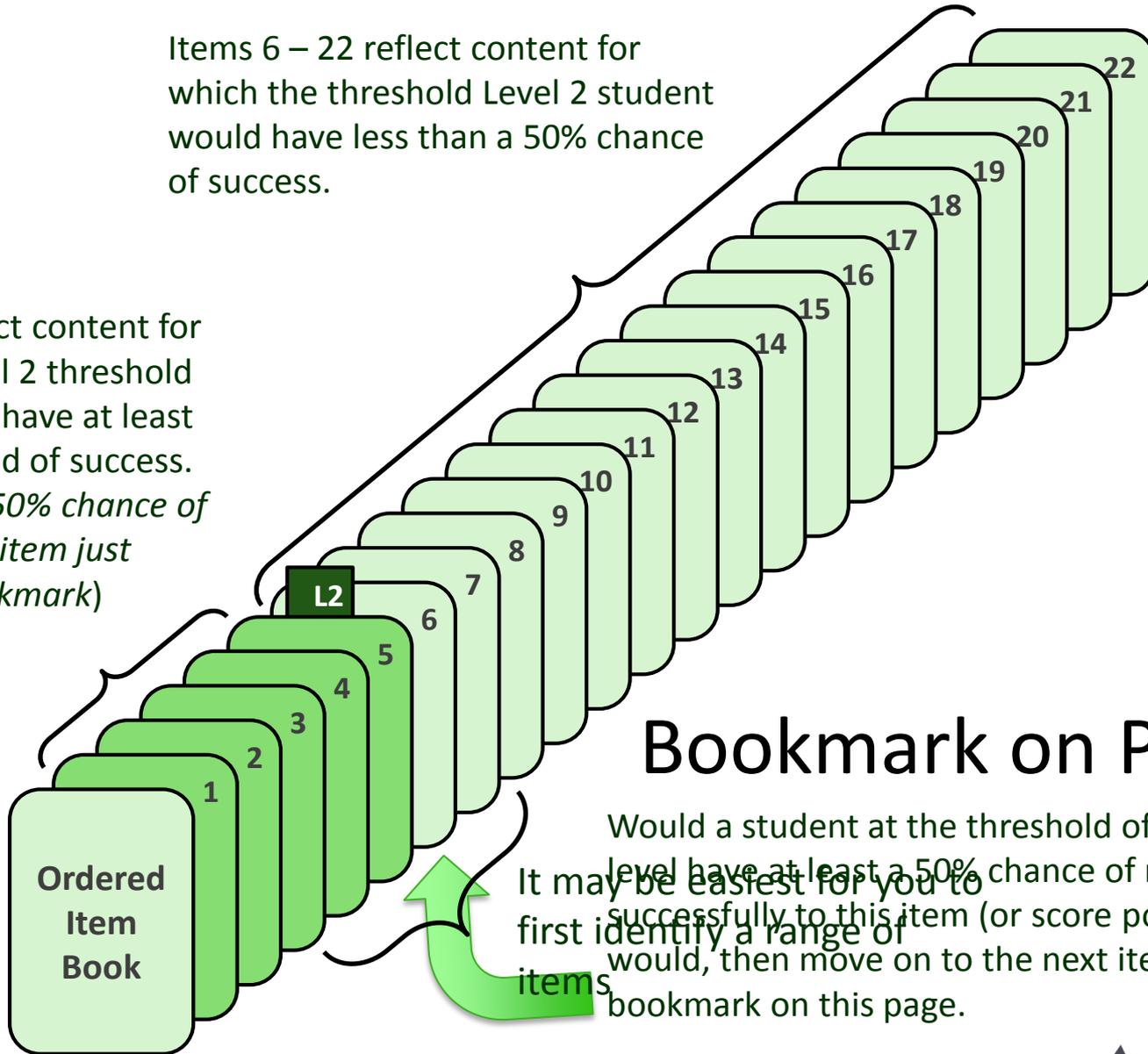
The items before the bookmark reflect content for which the threshold student would have at least a 50% likelihood of success (with about a 50% chance of success on the item just before the bookmark).

Items from the bookmark on reflect content for which the threshold student would have less than a 50% chance of success.



Items 6 – 22 reflect content for which the threshold Level 2 student would have less than a 50% chance of success.

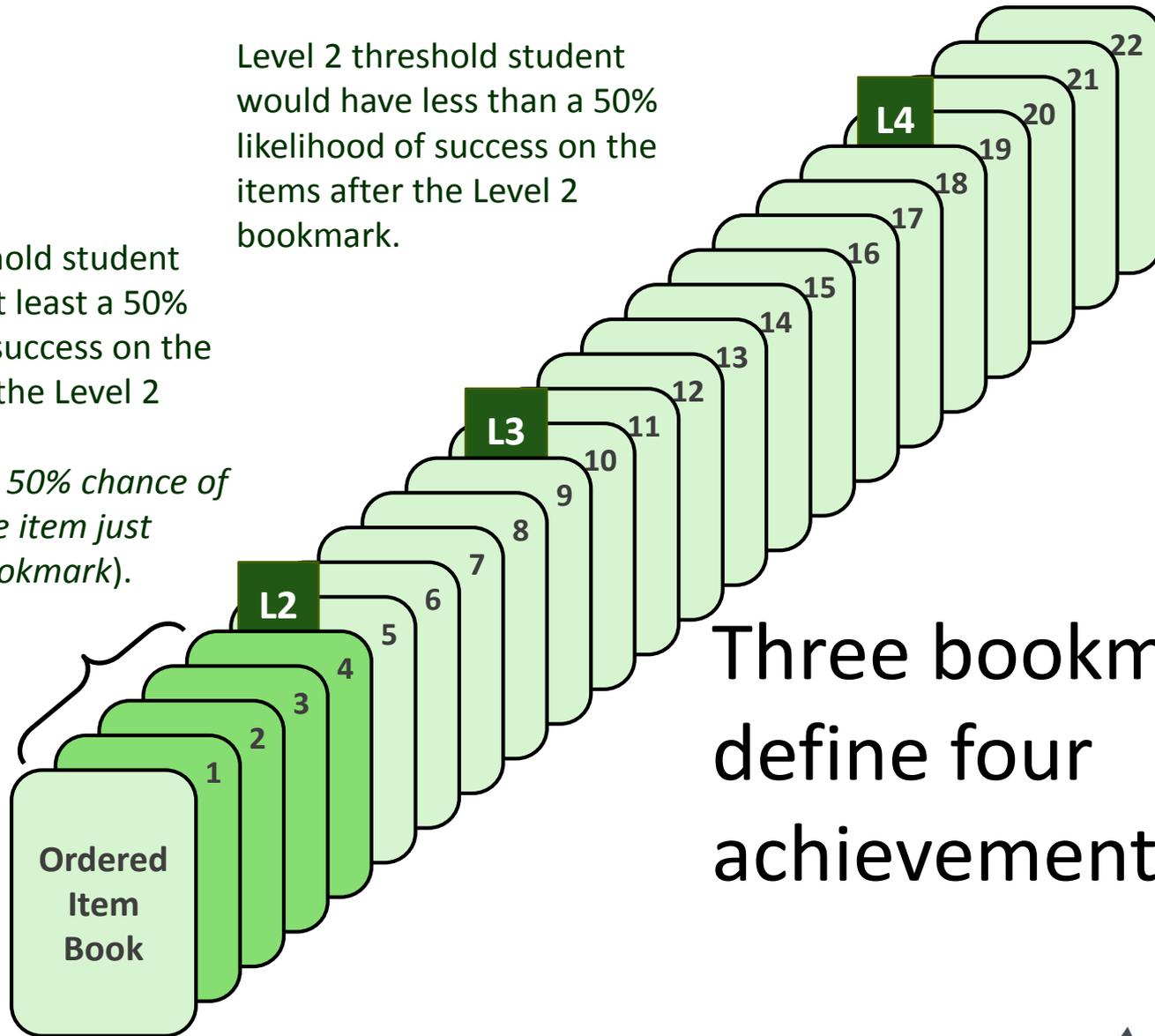
Items 1-5 reflect content for which the Level 2 threshold student would have at least a 50% likelihood of success. (with about a 50% chance of success on the item just before the bookmark)





Level 2 threshold student would have at least a 50% likelihood of success on the items before the Level 2 bookmark (with about a 50% chance of success on the item just before the bookmark).

Level 2 threshold student would have less than a 50% likelihood of success on the items after the Level 2 bookmark.



Three bookmarks define four achievement levels.



Understanding Your Role as Table Leader



Table Leader Responsibilities

- Facilitate discussion at the table level
 - Day 1 Studying the OIB
 - Day 2 discussion after lunch
- Bring important questions to the attention of the room facilitators
- Check understanding at table
- Keep an eye on time
- Monitoring and collection of secure materials at your table

Table Leader Responsibilities

- Facilitate discussion at the table level
 - Day 1 Studying the OIB
 - Facilitate conversation around the two questions
 - 1. What do you know about a student who responds successfully to this item; that is, what skills must a student have in order to know the correct answer?
 - 2. What makes this item more difficult than preceding items?
 - Remind panelists to put thoughts into comments
 - Monitor time for this task
 - Discourage sidebar conversation
 - Try to get all panelists to participate in discussion

Table Leader Responsibilities

- Facilitate discussion at the table level
 - Day 2 discussion after lunch
 - At the direction of the facilitator, write down the bookmark locations for each panelist at your table
 - Levels 2, 3, 4
 - At the direction of the facilitator, discuss the reasons for panelist differences one level at a time, beginning with Level 3.
 - Begin with the first bookmark, then the highest, to spur conversation.
 - When conversation is complete for Level 3, you should understand the rationale for the differences, and move on to the next level's bookmarks

Table Leader Responsibilities

- Monitoring and collection of secure materials at your table
 - At the end of each day we will collect materials
 - The facilitator will call out specific materials (say, ALDs) and panelists should place that material in front of them, for you to see. You can give the thumbs up to the facilitator when each panelist has the specified material on the table, so the next piece can be called out.
 - This allows us to have an efficient and orderly collection of materials to support security.



Agenda: Day 1

- Morning:
 - Opening session
 - Software orientation
 - Discussion of content standards and ALDs
 - Practice test
- Afternoon:
 - Introduction to the OIB
 - Study the OIB

Agenda: Day 2

- Morning: Round 1
 - Learn Bookmark Procedure
 - Practice in breakout rooms
 - Make bookmark recommendations individually
- Afternoon: Round 2
 - Discuss Round 1 bookmark placements at each table
 - Table leaders will be asked to return 15 minutes prior to start of session to review instructions for discussion
 - Place Round 2 bookmarks individually

Agenda: Day 3

- Morning: Round 3
 - Discuss Round 2 bookmark placements for the entire room
 - View supporting data based on Round 2 bookmarks
 - Place Round 3 bookmarks individually
 - Review final recommendations
 - Evaluate the process

Questions?

- Thank you for your participation!



Achievement Level Setting

Joe Willhoft
Mike Bunch



In-Person Workshop
October 13, 2014



Welcome!

- Purpose
 - Provide tools that teachers and schools need to improve teaching and learning so that all students can leave high school prepared for postsecondary success in college or a career
- Your Part
 - Place 4 separate performance levels for each test



We Are a Diverse Group

- You Were Nominated by Your State
 - All Smarter Balanced states invited to send panelists
 - Set achievement levels for all, not just a few
- You Represent an Important Stakeholder Group
 - Teachers, administrators, higher education, public
 - All points of view carefully considered
 - Important to public acceptance



Process/Expectations

- Small Groups – Panels
- Chance for Everyone to be Heard
- Chance for Everyone to Listen
- We Recommend – States Set



Online Panel

- Crowd Sourcing
- Focusing on Level 3
- Review Results Here



Your Job

- Review Support Materials
 - Smarter Balanced tests
 - Common Core State Standards
 - Achievement Level Descriptors
 - External Data

- Learn an Achievement Level Setting Procedure
 - Receive instruction
 - Practice
 - Apply



Forum

- Questions and Concerns
- Parking Lot



Thank You!



Overview





Goals for This Workshop

- Deepen Understanding of Content of Tests
- Deepen Understanding Achievement Level Descriptors
- Learn Bookmark Procedure
- Recommend Cut Scores



Cut Score

- A Dividing Line
- A Meaningful Distinction
- Varies with Test Difficulty
- Can Be Adjusted



Activities

- Review Common Core State Standards
- Review Achievement Level Descriptors
- Take Practice Tests
- Receive Bookmark Instruction
- Practice Using Bookmark Procedure
- Apply Bookmarks
 - 3 Rounds
 - Discussion Between Rounds



Agenda

Day - Time	Event(s)
Day 1 A.M.	Welcome; overview, training on CCSS, ALDs, tests
Day 1 P.M.	Review of Ordered Item Booklet
Day 2 A.M.	Orientation to the Bookmark Procedure; complete Round 1
Day 2 P.M.	Review Round 1; complete Round 2
Day 3 A.M.	Review Round 2; complete Round 3; evaluate process



Common Core State Standards





Key Shifts

English Language Arts/Literacy	Mathematics
Complex texts	Greater focus on fewer topics
Use of evidence from texts	Coherence – linking across grades
Building knowledge – nonfiction texts	Rigor <ul style="list-style-type: none">• Conceptual understanding• Procedural skills and fluency• Application



Claims

English Language Arts/Literacy	Mathematics
Reading	Concepts & Procedures
Writing	Problem Solving/Modeling & Data Analysis
Speaking/Listening	Communicating/Reasoning
Research/Inquiry	



Achievement Level Descriptors





Policy Framework for Grade 11 Assessment Results

Level 1

- Not Yet Content-Ready - Substantial Support Needed
- K-12 & higher education may offer interventions

Level 2

- Not Yet Content-Ready – Support Needed
- Transition courses or other supports for Grade 12, retesting option for states

Level 3*

- Conditionally Content-Ready/Exempt from Developmental
- In each state, K-12 and higher ed must jointly develop Grade 12 requirements for students to earn exemption

Level 4

- Content-Ready/Exempt from Developmental
- K-12 and higher education may jointly set Grade 12 requirements to retain exemption (optional for states)

* Level 3 is used for federal NCLB purposes.



Types of ALDs

- Policy
- Range
- Threshold
- Reporting



Types of ALDs

- Policy
- Range
- **Threshold – students just entering Level 2, 3, 4**
- Reporting



Smarter Balanced Assessments





Review Tests

- Become familiar with practice tests
- Discuss contents



Question and Test Development

- Common Core State Standards
- Development of Blueprints
- Development of Templates
- Question Development
- Question Review
- Field Testing
- Statistical Review
- Question Retention and Banking



Question Types

- Traditional Selected-Response Questions
- Traditional Constructed-Response Questions
- Technology Enhanced Questions
- Viewing/Listening Questions



Tools and Supports

Universal Tools	Designated Supports
Calculator	Color Contrast
Digital Notepad	Highlighter
English Dictionary	Magnification
English Glossary	Masking
Expandable Passages	Text to Speech
Math Tools	Translations (Glossary)
Spell Check	
Writing Tools	
Zoom	



Accommodations

- Braille
- Abacus
- Alternate Response Options
- Calculator
- Multiplication Table
- Print on Demand
- Read Aloud



Getting Started





Recap

- Study the Common Core State Standards
- Study the Achievement Level Descriptors
- Take the Practice Test
- Examine an Ordered Item Booklet
- Learn About the Bookmark Procedure
- Place 3 Bookmarks



Groundrules

- Security/Confidentiality
- Group Process
- All Voices Equal
- Recommend – Not Set

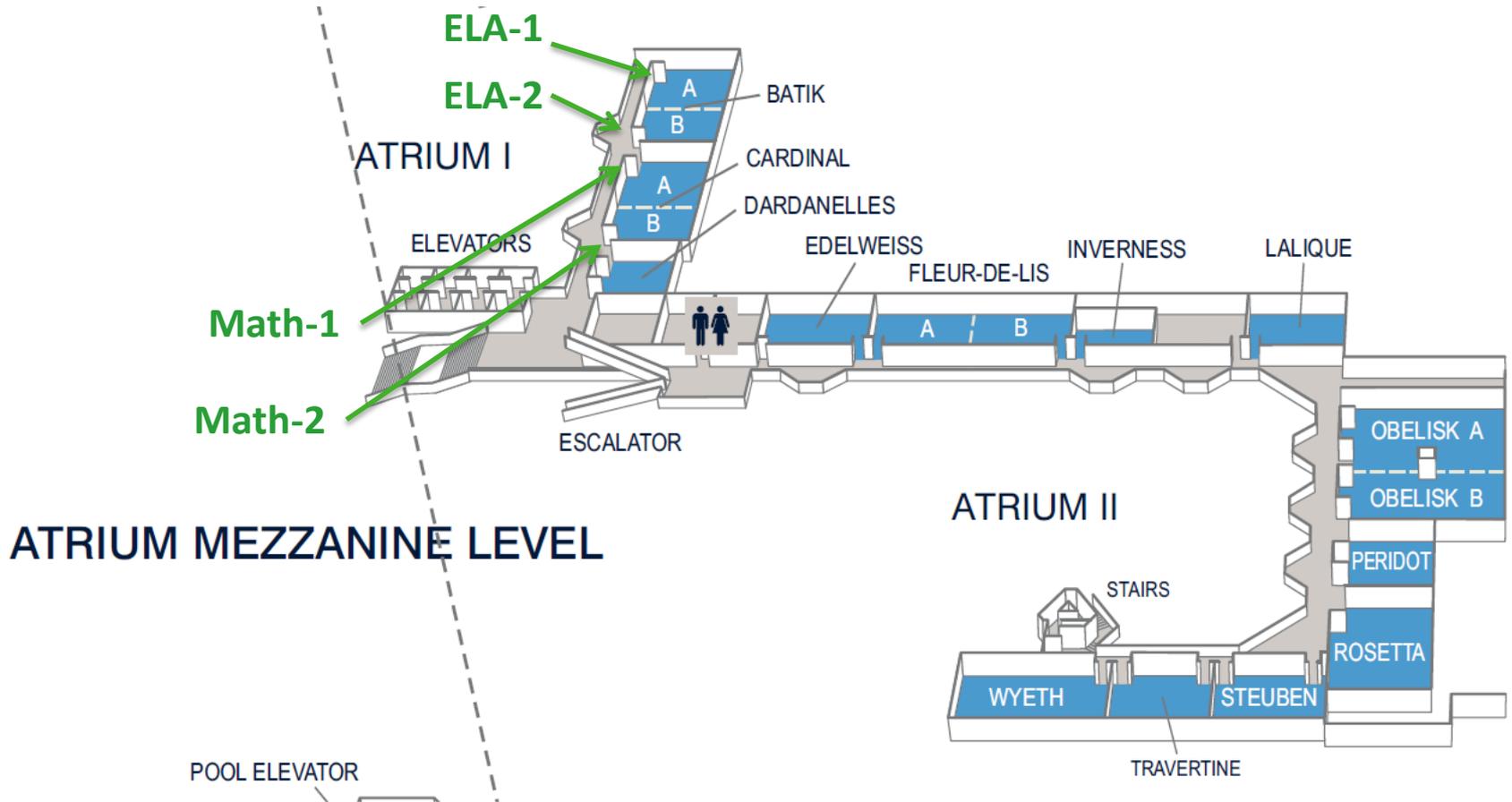


Room Assignments

Panel	Room	Facilitator(s)
High School ELA – 1	Batik A	Wilder/Deville
High School ELA – 2	Batik B	Schultz/Mercado
High School Mathematics – 1	Cardinal A	Reid/Eno
High School Mathematics – 2	Cardinal B	Hickman/Lord-Bessen



Hotel Map



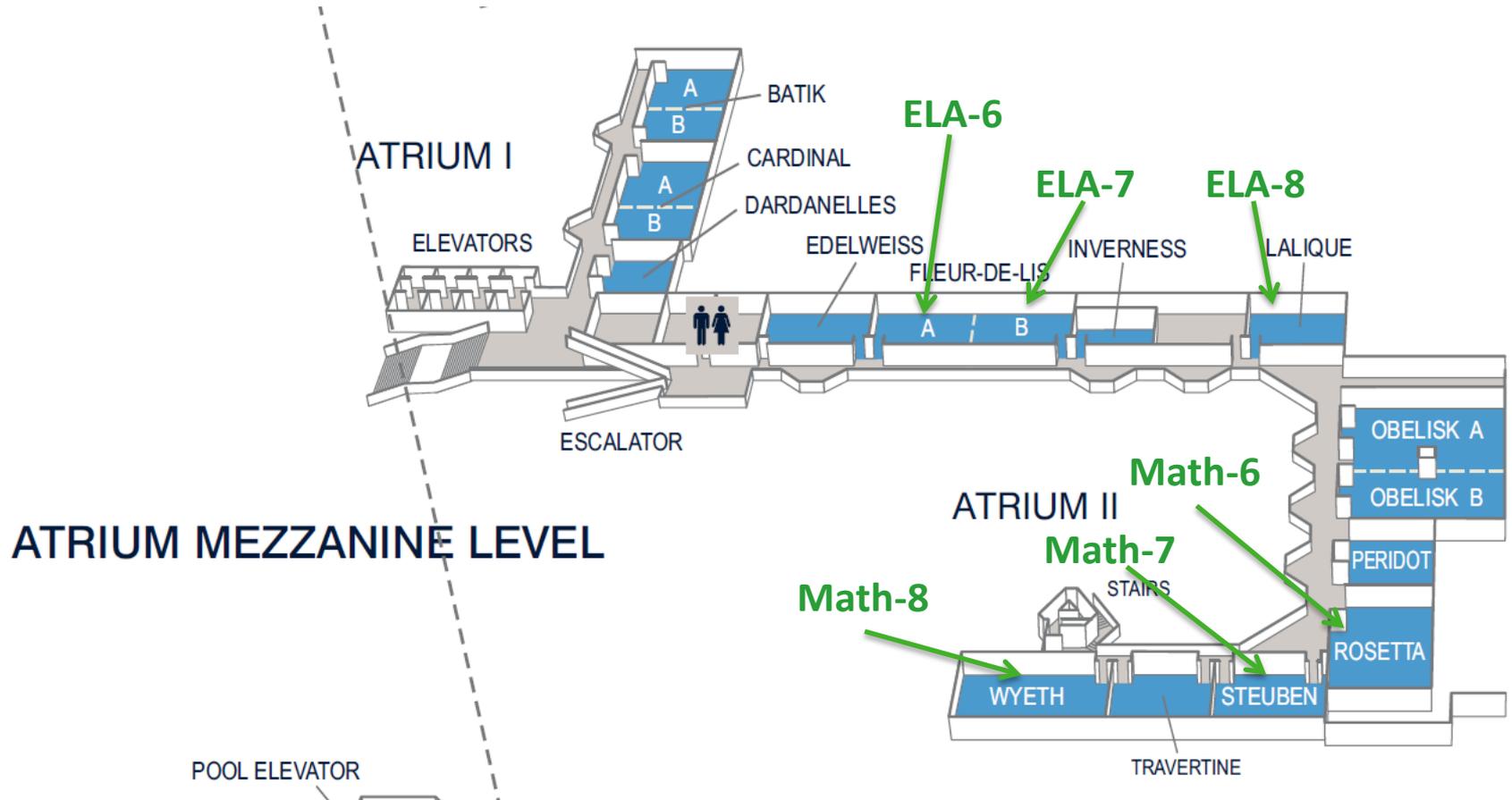


Room Assignments

Panel	Room	Facilitator(s)
Grade 6 ELA	Fleur de Lis A	Dunbar/Palermo
Grade 7 ELA	Fleur de Lis B	Buck/Hagge
Grade 8 ELA	Lalique	Griswold/D. Bowen
Grade 6 Mathematics	Rosetta	Bolton/Mercado
Grade 7 Mathematics	Steuben	J. Bowen/Johnson
Grade 8 Mathematics	Wyeth	Upchurch/Lord-Bessen



Hotel Map



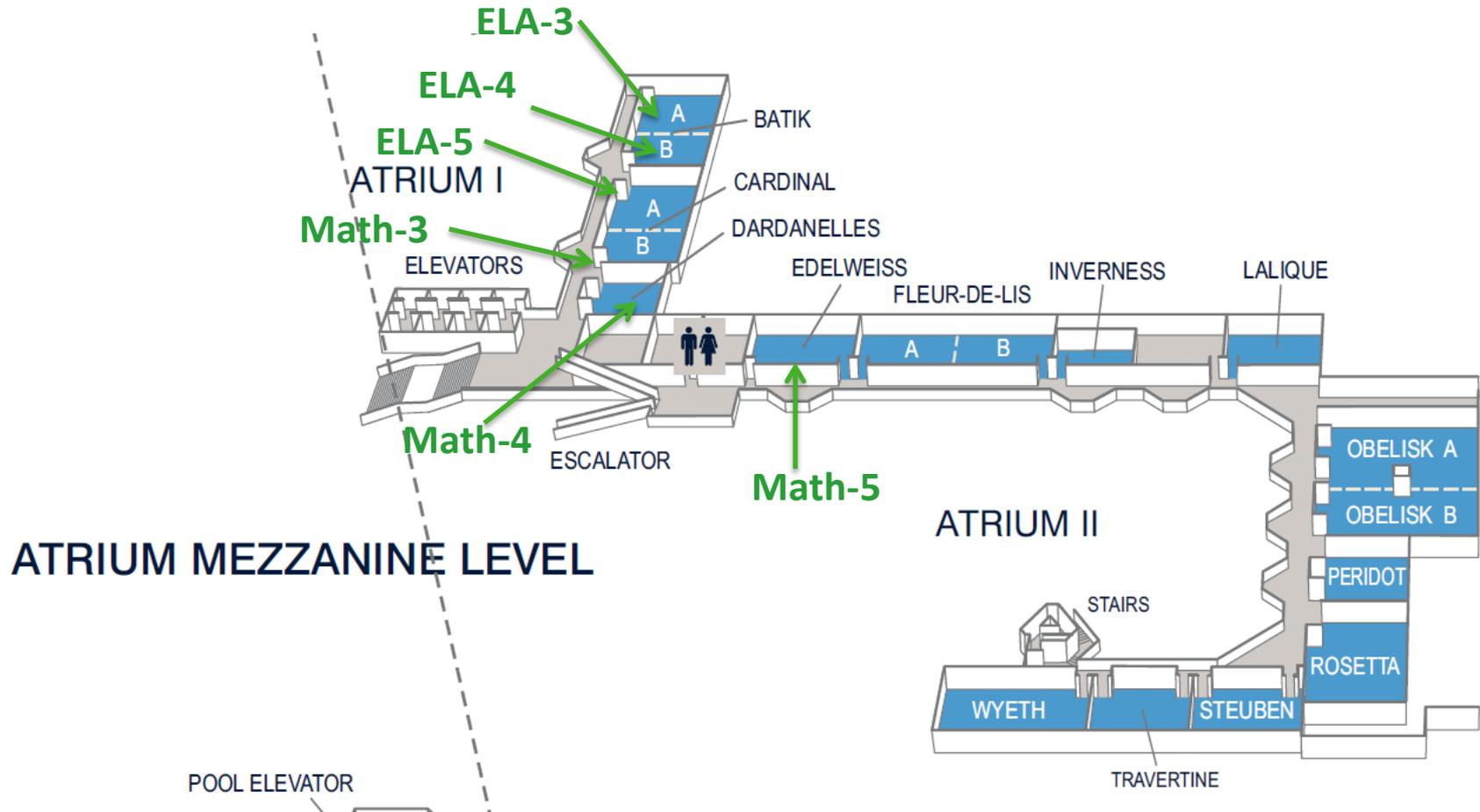


Room Assignments

Panel	Room	Facilitator(s)
Grade 3 ELA	Batik A	Grady/Deville
Grade 4 ELA	Batik B	Connelly/Hagge
Grade 5 ELA	Cardinal A	Hargis/McClintock
Grade 3 Mathematics	Cardinal B	Farina/Mercado
Grade 4 Mathematics	Dardanelles	Reid/Johnson
Grade 5 Mathematics	Edelweiss	Hickman/Lord-Bessen



Hotel Map





Understanding What the Test Measures

Daniel Lewis
Chief Research Advisor
McGraw-Hill Education CTB



Smarter Balanced Achievement Level Setting
Dallas, TX, October 2014

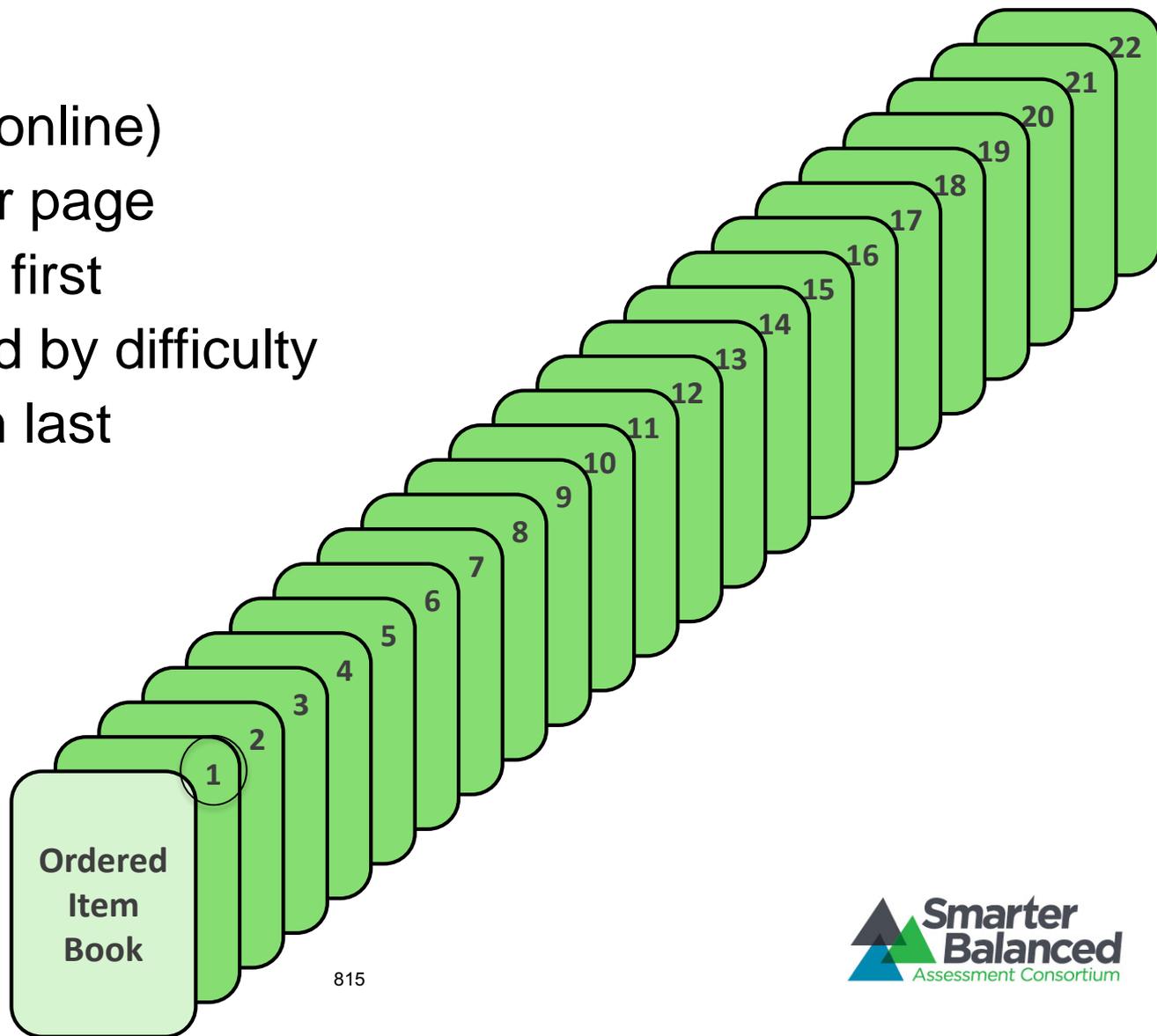


Developing the Expertise to Make Content-Based Achievement Level Recommendations

- Activities that must be completed :
 - Study content standards
 - Study achievement level descriptors
 - Study test items
 - Take practice test
 - Study ordered item booklets using item maps

Ordered Item Books (OIBs)

- Virtual OIB (online)
- One item per page
- Easiest item first
- Items ascend by difficulty
- Hardest item last





Item Map for Booklet: English Language Arts/Literacy, Grade 4

English Language Arts/Literacy, 4

Round: 1

Enter Comment

Page	Standard Code	DOK	Bookmark	Comments	Reviewed
01	4.L.6	1			<input type="checkbox"/>
02	4.L.2.c	1			<input type="checkbox"/>
03	4.RL.4	2			<input type="checkbox"/>
04	4.SL.3	1			<input type="checkbox"/>
05	4.W.3.d	1			<input type="checkbox"/>
06	4.L.1.a	1			<input type="checkbox"/>
07	4.RI.4	2			<input type="checkbox"/>
08	4.RI.5	2			<input type="checkbox"/>
09	4.L.1.d	1			<input type="checkbox"/>
10	4.RL.2	2			<input type="checkbox"/>
11	4.SL.3	2			<input type="checkbox"/>
12	4.SL.3	1			<input type="checkbox"/>
13	4.W.2.c	2			<input type="checkbox"/>

Set Bookmark

Submit Bookmarks

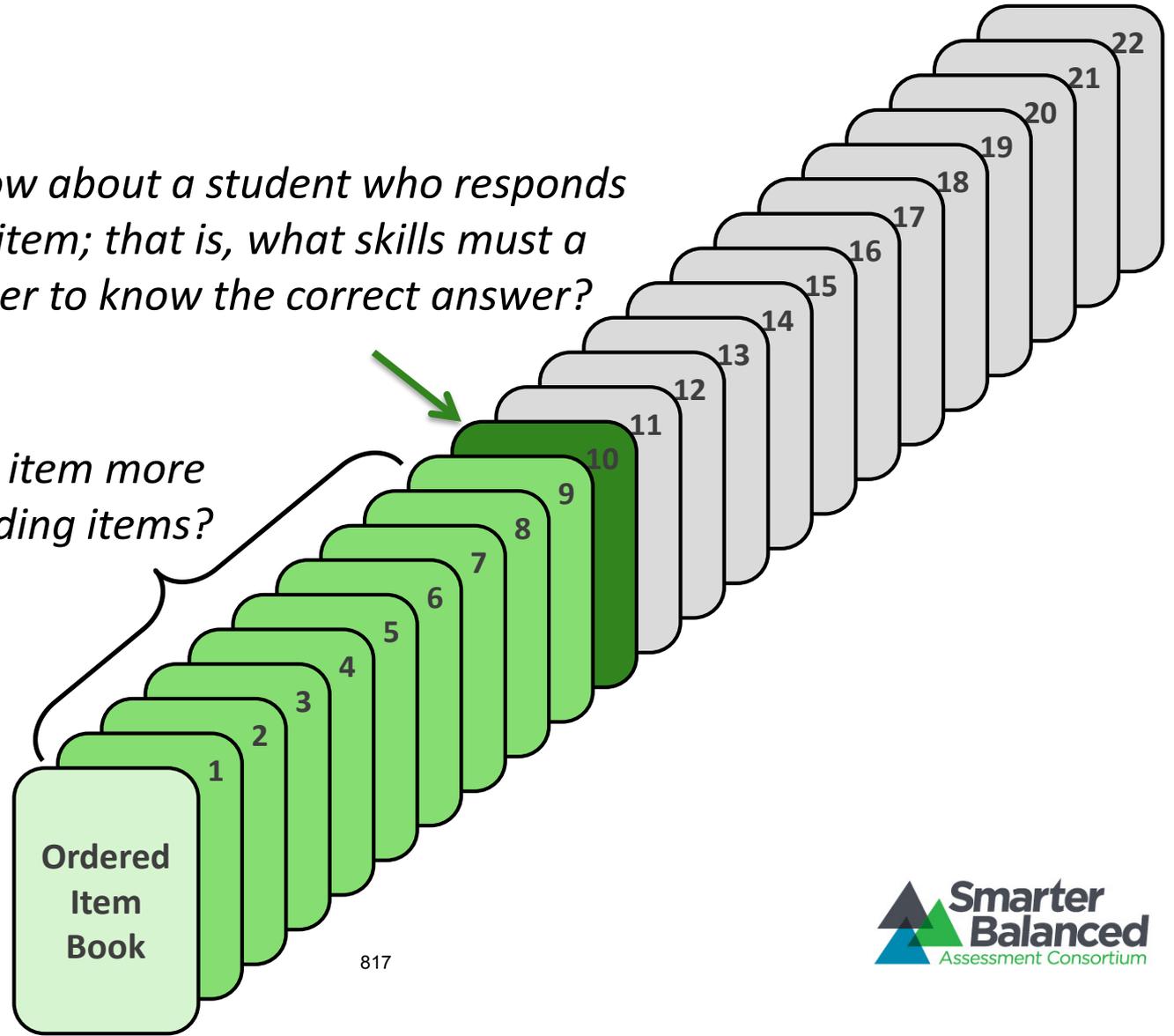
Study the Items in the OIB

As you study each item in the OIB, discuss two questions with your fellow panelists:

1. *What do you know about a student who responds successfully to this item; that is, what skills must a student have in order to know the correct answer?*

2. *What makes this item more difficult than preceding items?*

This activity helps you acquire the knowledge to make content-based cut score recommendations.

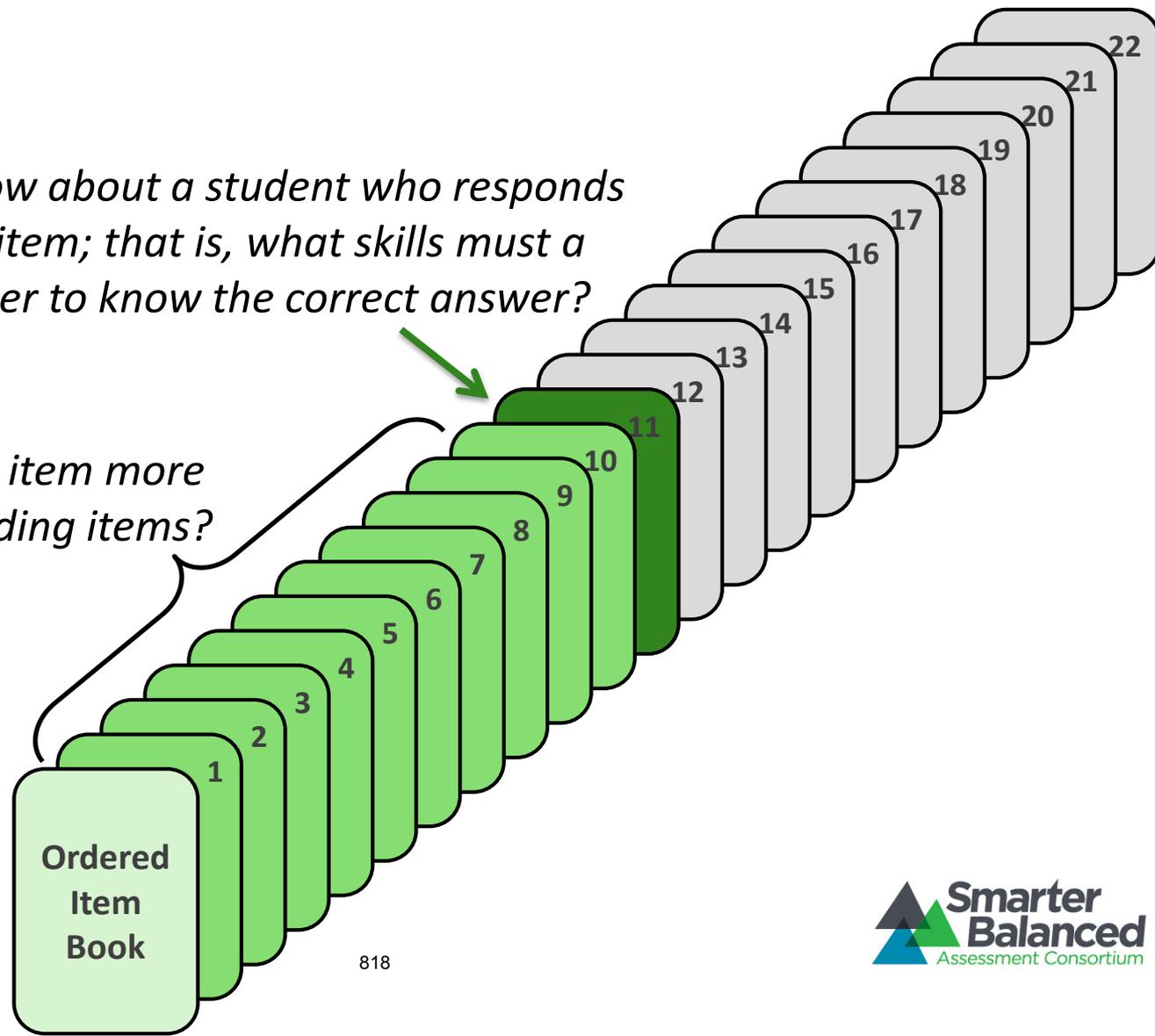


Study the Items in the OIB

As you study each item in the OIB, discuss two questions with your fellow panelists:

1. *What do you know about a student who responds successfully to this item; that is, what skills must a student have in order to know the correct answer?*

2. *What makes this item more difficult than preceding items?*





In-Person Panel for Achievement Level Setting

2. *What makes this item more difficult than preceding items?*

it
e

2

Round: 1

[Set Bookmark](#) [Enter Comment](#)

[Item](#) [Item Info](#) [Resource\(s\)](#) [ALD](#)

22766  

This sentence has an error in grammar usage.
The person who's car is blocking the entrance should move it.
Choose the correct word to replace the underlined word in the sentence.

- (A) that
- (B) which
- (C) who
- (D) whose

[Back to Item Map](#)

2. What makes this item more difficult than preceding items?

◀ Previous

Next ▶

30666



A student is writing a story for language arts about a time she felt proud. Read this part of the story and complete the task that follows.

I volunteer at the local homeless shelter on the weekends. My sister actually got me involved with this because she had to complete a project for school. They could choose to volunteer anywhere they wanted, and she chose the shelter. After the project was over, my sister continued to volunteer because she liked it so much. She talked about it all the time, so I decided to see the shelter for myself.

Volunteering at the shelter was hard work, but it was so rewarding. I enjoyed talking with the people there, listening as they spoke about their daily activities. They were so polite and thanked me for helping, but I was the one who was thankful. I left each day feeling satisfied with a job well done. It felt good to help others in need.

After a few months of volunteering at the shelter, I knew most everyone who came in. I even became friends with some of the younger people—those around my age. One evening, a small child came up to me with a card in his hand. It was a thank-you card. He made it just for me. My eyes teared up and my heart swelled with pride when I read the inside of the card.

In one paragraph, write an ending for the story using details and description.

2. *What makes this item more difficult than preceding items?*

xt
e

Round: 1

42

Set Bookmark Enter Comment

Item Item Info Resource(s) ALD

26433  

Read the sentences and the question that follows.

I remember the time where I lost my wallet at a movie theater. I realized it was gone when I got home. I called the theater right away, and the person who answered said no one had seen it. My mom drove me back to the theater so I could look for it myself. I was so relieved the manager found it in the lost-and-found bin. I asked her why the other worker had not checked there.

Which underlined word in the paragraph is incorrect?

(A) where
(B) when
(C) who
(D) why

2. What makes this item more difficult than preceding items?

◀ Previous

Next ▶

30666



A student is writing a story for language arts about a time she felt proud. Read this part of the story and complete the task that follows.

I volunteer at the local homeless shelter on the weekends. My sister actually got me involved with this because she had to complete a project for school. They could choose to volunteer anywhere they wanted, and she chose the shelter. After the project was over, my sister continued to volunteer because she liked it so much. She talked about it all the time, so I decided to see the shelter for myself.

Volunteering at the shelter was hard work, but it was so rewarding. I enjoyed talking with the people there, listening as they spoke about their daily activities. They were so polite and thanked me for helping, but I was the one who was thankful. I left each day feeling satisfied with a job well done. It felt good to help others in need.

After a few months of volunteering at the shelter, I knew most everyone who came in. I even became friends with some of the younger people—those around my age. One evening, a small child came up to me with a card in his hand. It was a thank-you card. He made it just for me. My eyes teared up and my heart swelled with pride when I read the inside of the card.

In one paragraph, write an ending for the story using details and description.

Learning With Your Fellow Panelists

- You will study the OIB this afternoon
- You may and should discuss the OIB items with your fellow panelists
- You may study the OIB individually, if this is more efficient for you, but may want to confer with your fellow panelists from time to time

The Bookmark Procedure

Daniel Lewis

Chief Research Advisor
McGraw-Hill Education CTB



Smarter Balanced Achievement Level Setting
Dallas, TX, October 2014

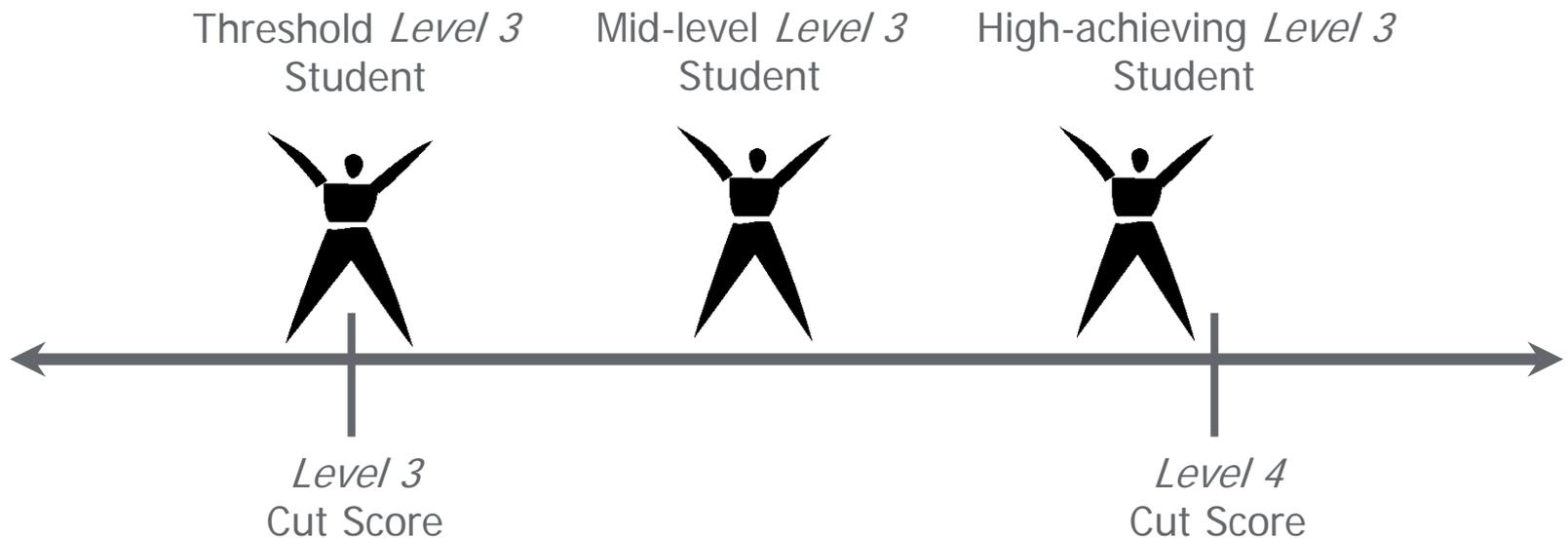


Bookmark Overview

- The Bookmark Procedure is a process used to set cut scores that define achievement levels by comparing and aligning items and content to achievement level descriptors
- It is so named because you place bookmarks in an ordered item booklet of test items to make your cut score recommendations

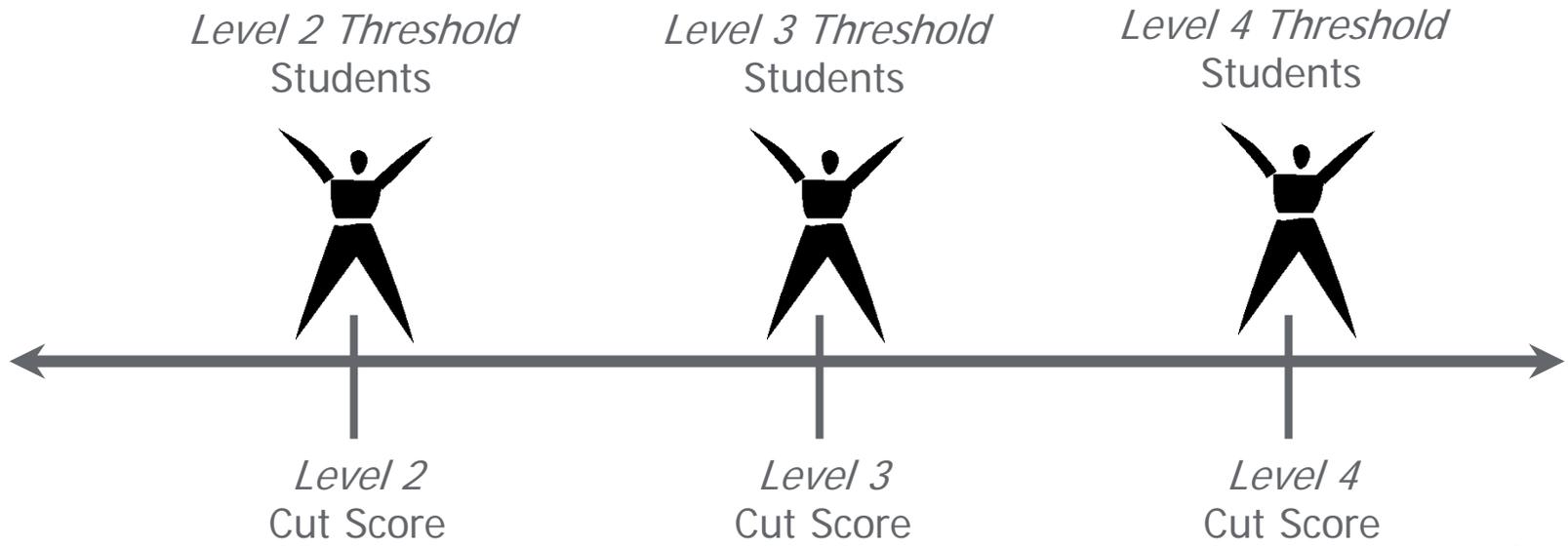
Threshold Student

- We recommend cut scores for Threshold Students—the students with a level of achievement that just barely qualifies them to be in the achievement level
 - These are the skills of the student just entering the achievement level

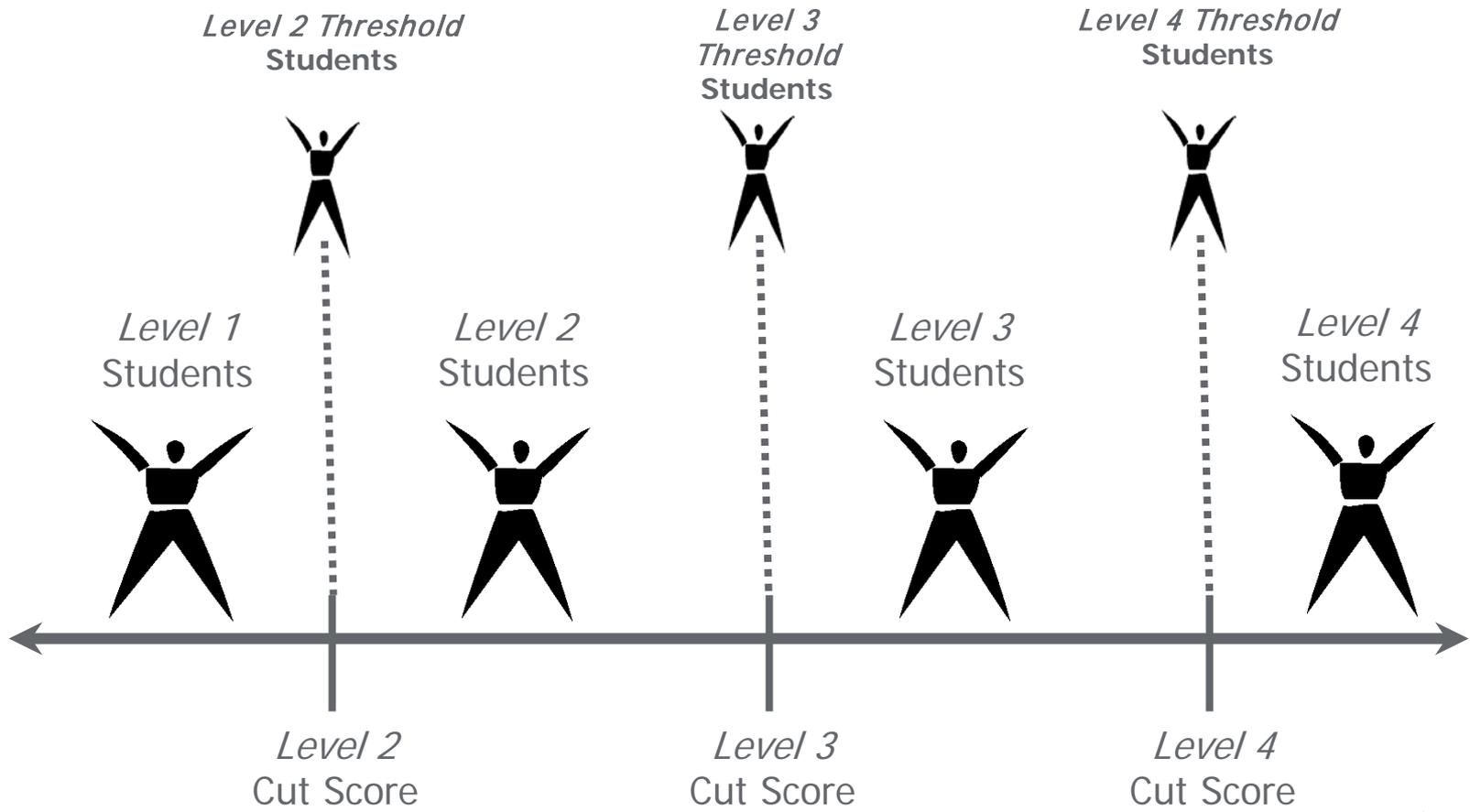




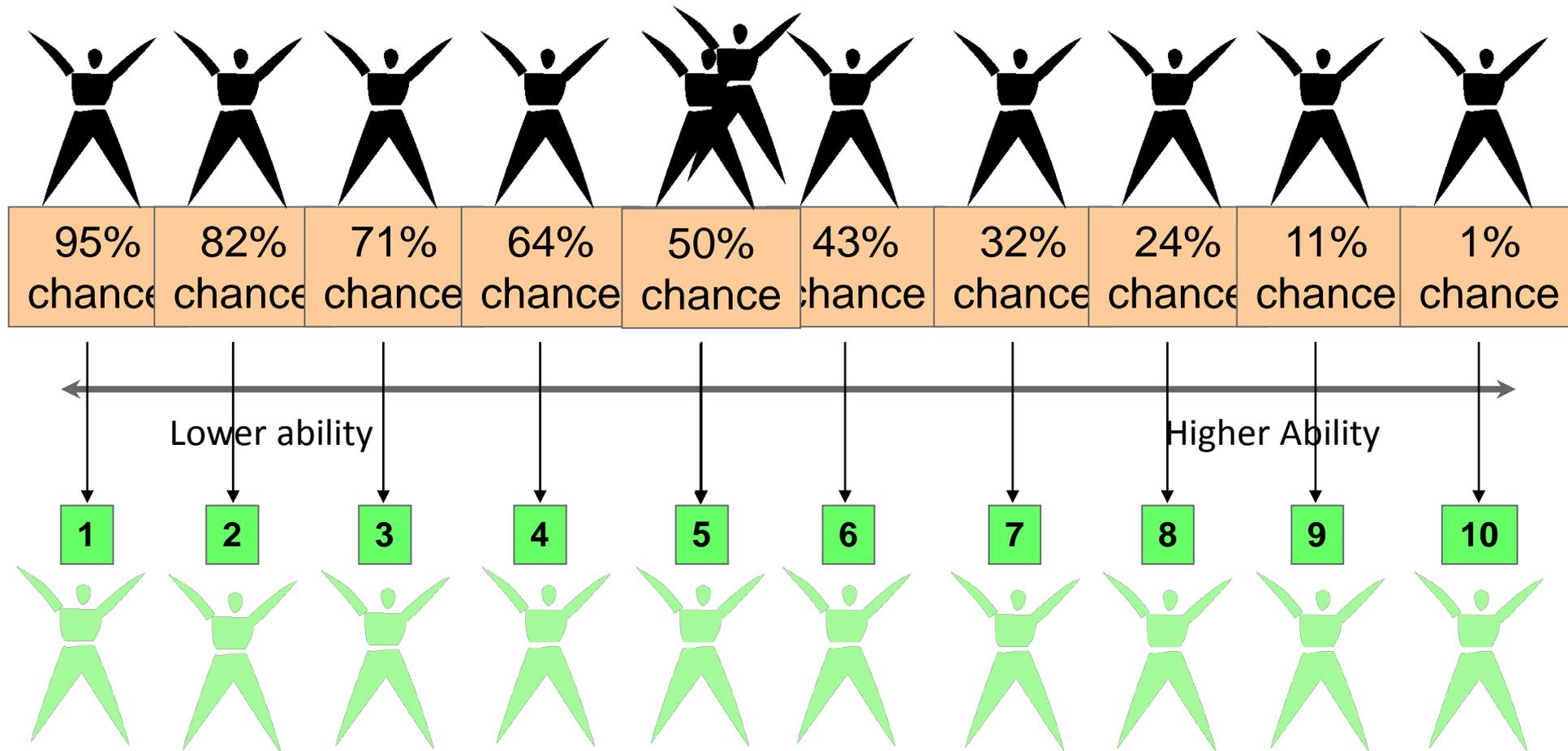
Three Threshold Students



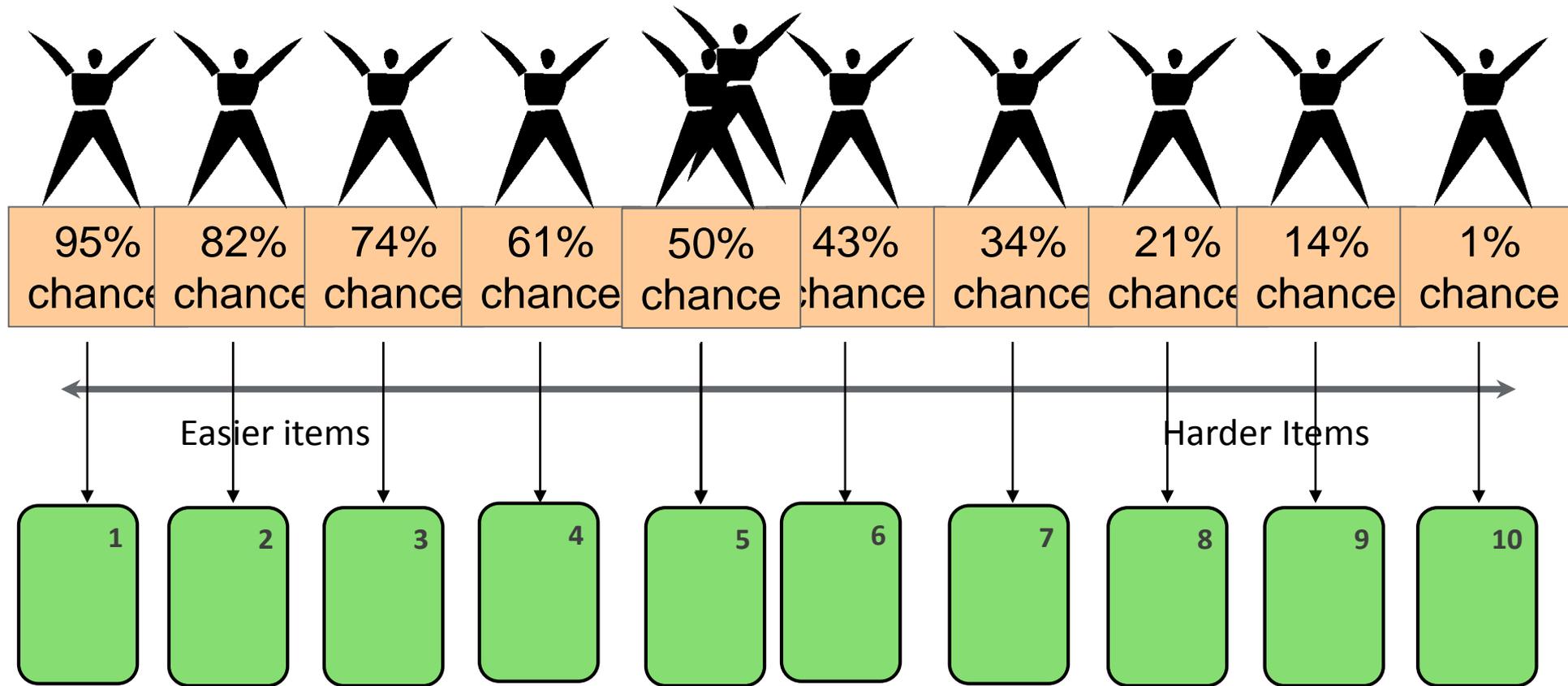
Three Cut Scores Define Four Achievement Levels



Locating the Threshold Student on the Scale

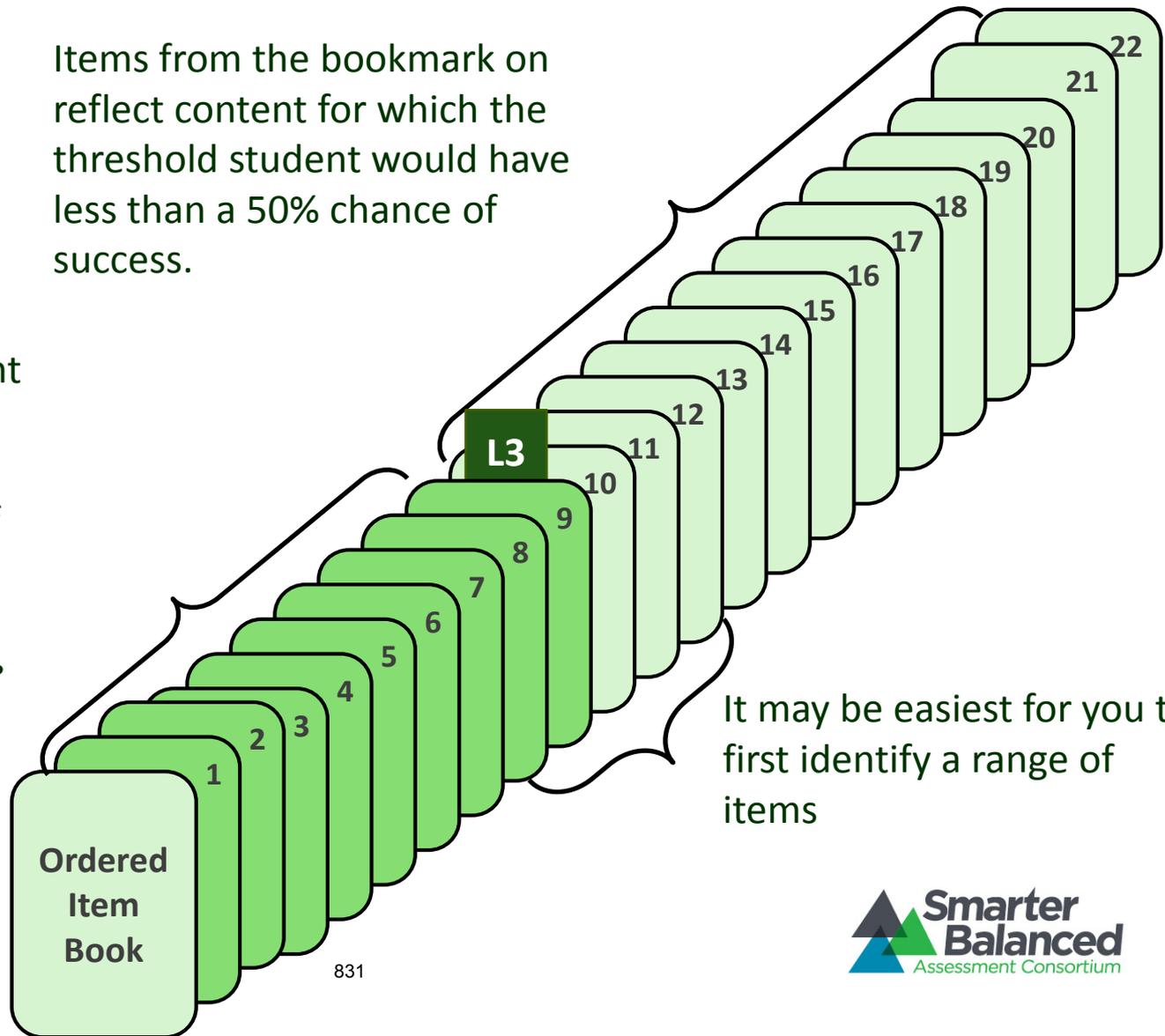


Locating the Threshold Student on the Scale



The items before the bookmark reflect content for which the threshold student would have at least a 50% likelihood of success (with about a 50% chance of success on the item just before the bookmark).

Items from the bookmark on reflect content for which the threshold student would have less than a 50% chance of success.



It may be easiest for you to first identify a range of items

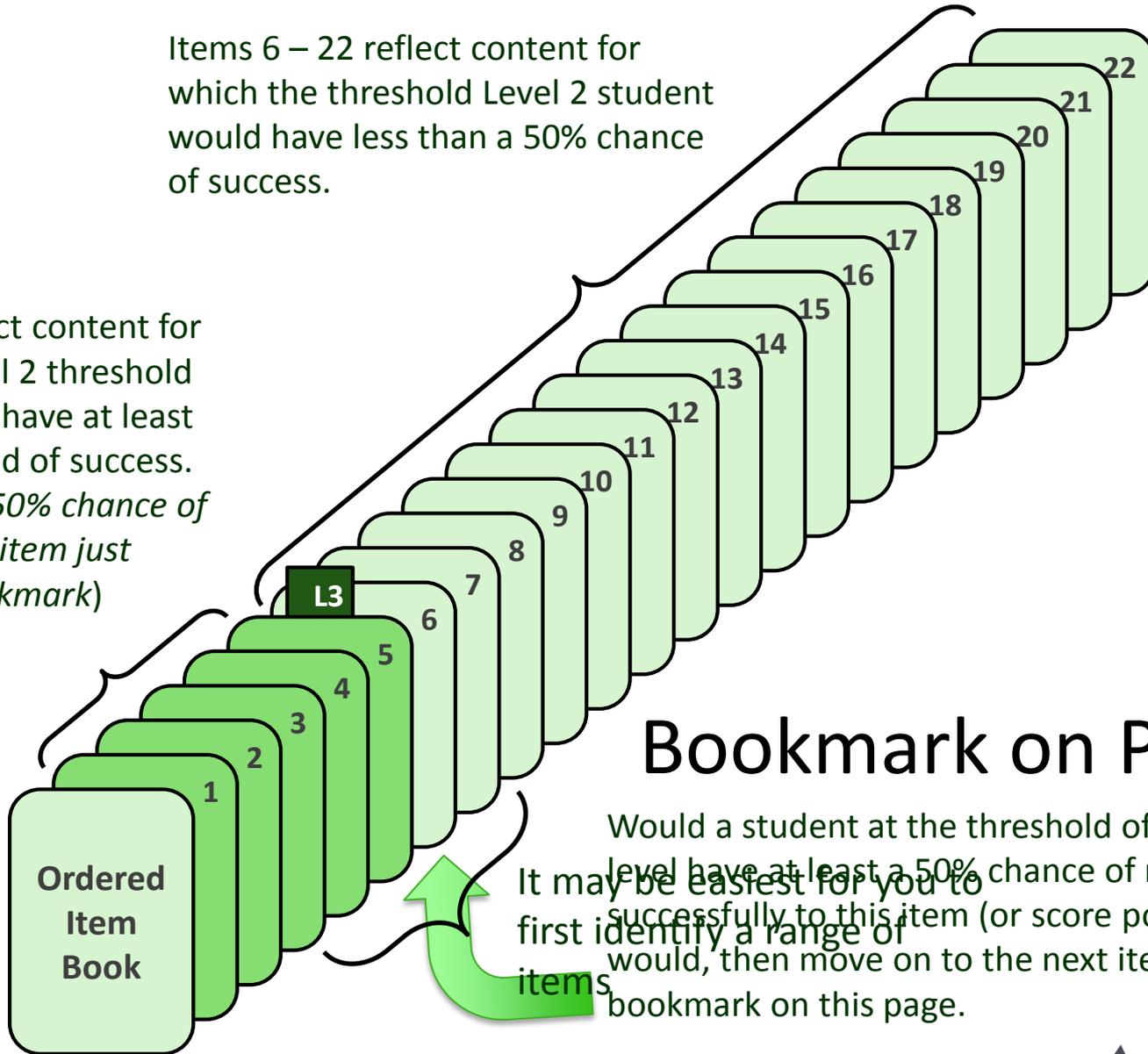


Level 3 Bookmark Placement

- When you place your Level 3 bookmark, think about a student at the threshold of Level 3 based on the achievement level descriptors
 - Place your Level 3 bookmark at the point in the OIB such that
 - A threshold Level 3 student would have at least a 50% likelihood of success on each of the items before the bookmark (and about a 50% chance of success on the item just before the bookmark)
 - A threshold Level 3 student would have less than a 50% likelihood of success on the items from the bookmark on
 - It may be easiest for you to first identify a range of items that begins where threshold Level 3 students begin to be challenged and end where they would definitely be challenged. Then pick the best page in that range.

Items 6 – 22 reflect content for which the threshold Level 2 student would have less than a 50% chance of success.

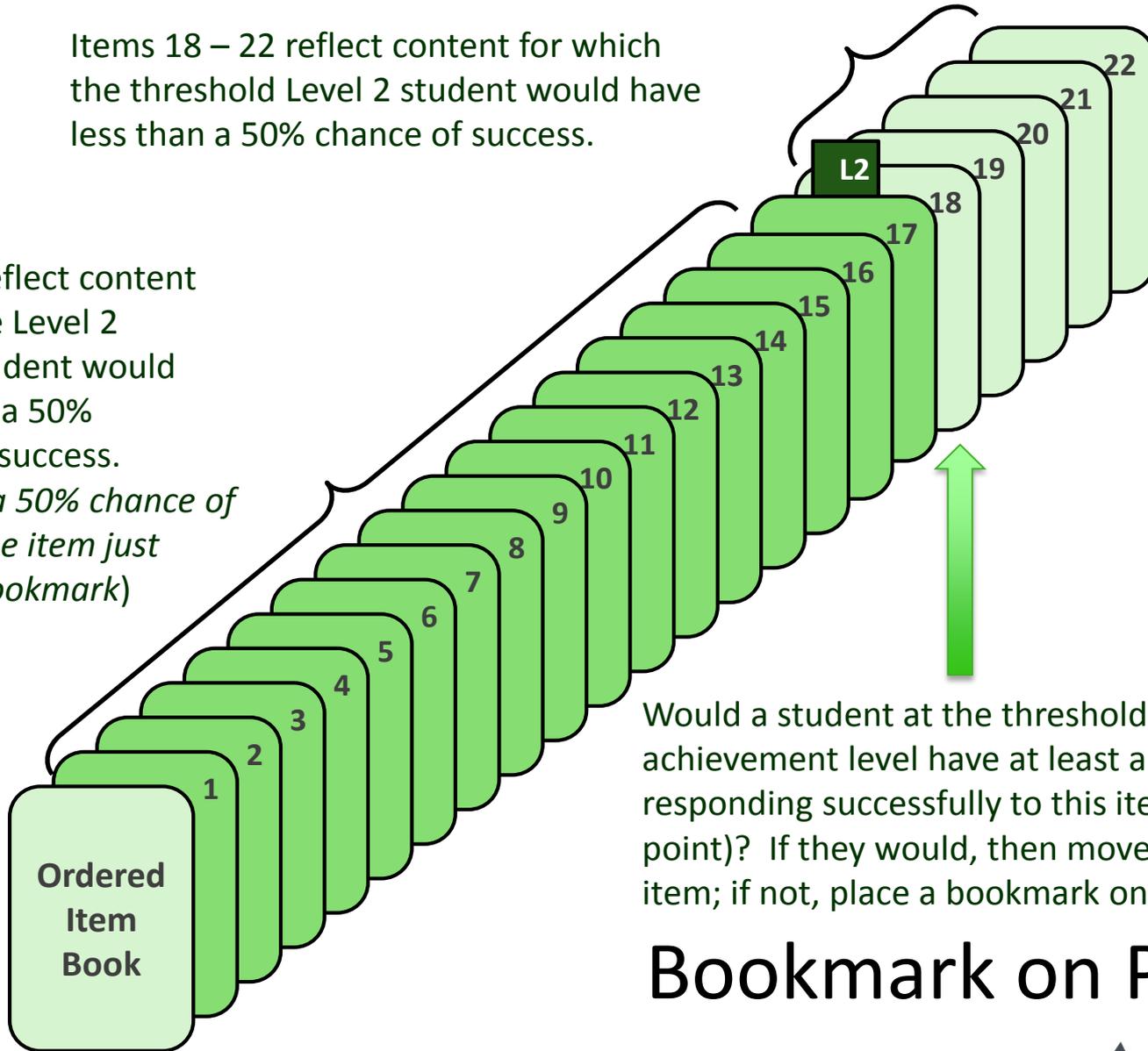
Items 1-5 reflect content for which the Level 2 threshold student would have at least a 50% likelihood of success. (with about a 50% chance of success on the item just before the bookmark)



Would a student at the threshold of the achievement level have at least a 50% chance of responding successfully to this item (or score point)? If they would, then move on to the next item; if not, place a bookmark on this page.

Items 18 – 22 reflect content for which the threshold Level 2 student would have less than a 50% chance of success.

Items 1-17 reflect content for which the Level 2 threshold student would have at least a 50% likelihood of success. (with about a 50% chance of success on the item just before the bookmark)



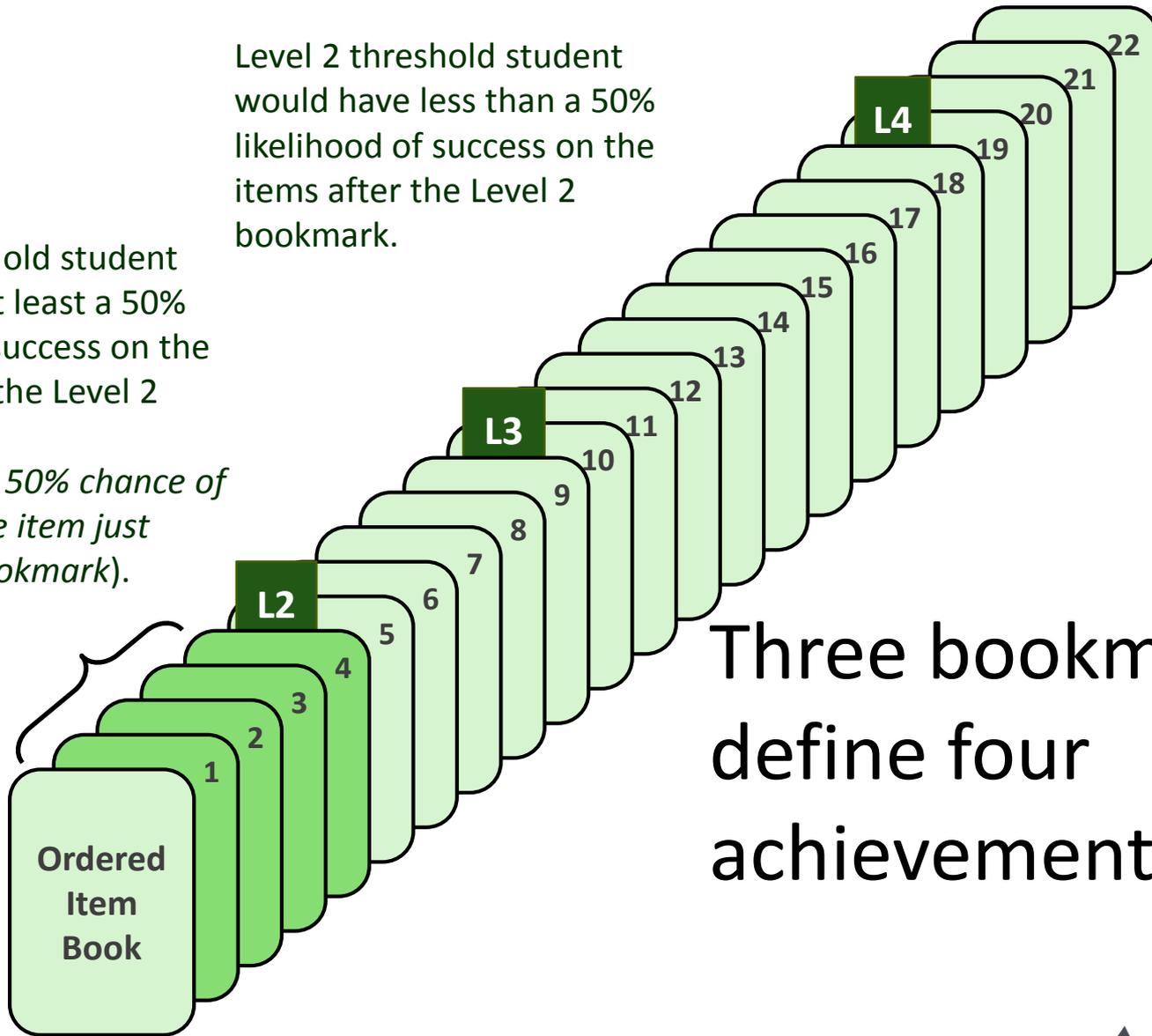
Would a student at the threshold of the achievement level have at least a 50% chance of responding successfully to this item (or score point)? If they would, then move on to the next item; if not, place a bookmark on this page.

Bookmark on Page 18



Level 2 threshold student would have at least a 50% likelihood of success on the items before the Level 2 bookmark (with about a 50% chance of success on the item just before the bookmark).

Level 2 threshold student would have less than a 50% likelihood of success on the items after the Level 2 bookmark.



Three bookmarks define four achievement levels.

Setting Bookmarks

Item Map for Booklet: Achievement

English Language Arts/Literacy, 4

Round: 1

Enter Comment

Review each question in the ordered item booklet by clicking on the page number

View Selected

Page	Claim and Target	Depth of Knowledge	Bookmark	Comments	Reviewed
1	Claim: 2-W, Target: 8-4	1			<input type="checkbox"/>
2	Claim: 2-W, Target: 9-4	1			<input type="checkbox"/>
3	NAEP				<input type="checkbox"/>
4	Claim: 2-W, Target: 1-4	2			<input type="checkbox"/>
5	Claim: 3-L, Target: 4-4	2			<input type="checkbox"/>
6	Claim: 2-W, Target: 2-4	4			<input checked="" type="checkbox"/>
7	Claim: 3-L, Target: 4-4	2			<input type="checkbox"/>
8	Claim: 2-W, Target: 8-4	1			<input type="checkbox"/>
9	Claim: 1-LT, Target: 3-4	2			<input type="checkbox"/>
10	Claim: 2-W, Target: 9-4	1			<input type="checkbox"/>
11	Claim: 3-L, Target: 4-4	2			<input type="checkbox"/>
12	Claim: 1-LT, Target: 7-4	2			<input type="checkbox"/>
13	Claim: 2-W, Target: 3-4	2			<input type="checkbox"/>

Set Bookmark

Select a bookmark value and click "Set Bookmark" below to set a bookmark at the selected Location value. Click "Cancel" to return.

- Level 2
- Level 3
- Level 4



Agenda

Day 2

- Morning: Round 1
 - Practice in breakout rooms
 - Make bookmark recommendations individually
- Afternoon: Round 2
 - Discuss Round 1 bookmark placements at each table
 - Place Round 2 bookmarks individually

Day 3

- Morning: Round 3
 - Discuss Round 2 bookmark placements for the entire room
 - View supporting data based on Round 2 bookmarks
 - Place Round 3 bookmarks individually
 - Review final recommendations
 - Evaluate the process



Questions?

- Thank you for your participation!

The Bookmark Procedure

Daniel Lewis

Chief Research Advisor
McGraw-Hill Education CTB



Smarter Balanced Achievement Level Setting
Dallas, TX, October 2014

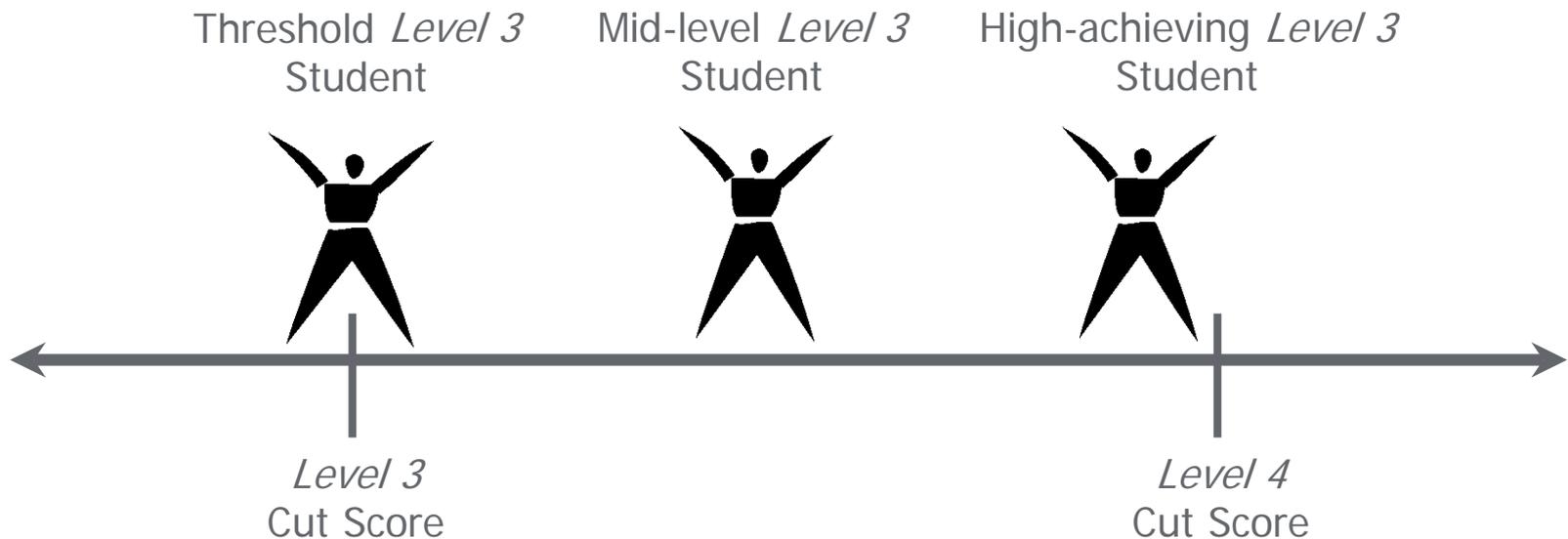


Bookmark Overview

- The Bookmark Procedure is a process used to set cut scores that define achievement levels by comparing and aligning items and content to achievement level descriptors
- It is so named because you place bookmarks in an ordered item booklet of test items to make your cut score recommendations

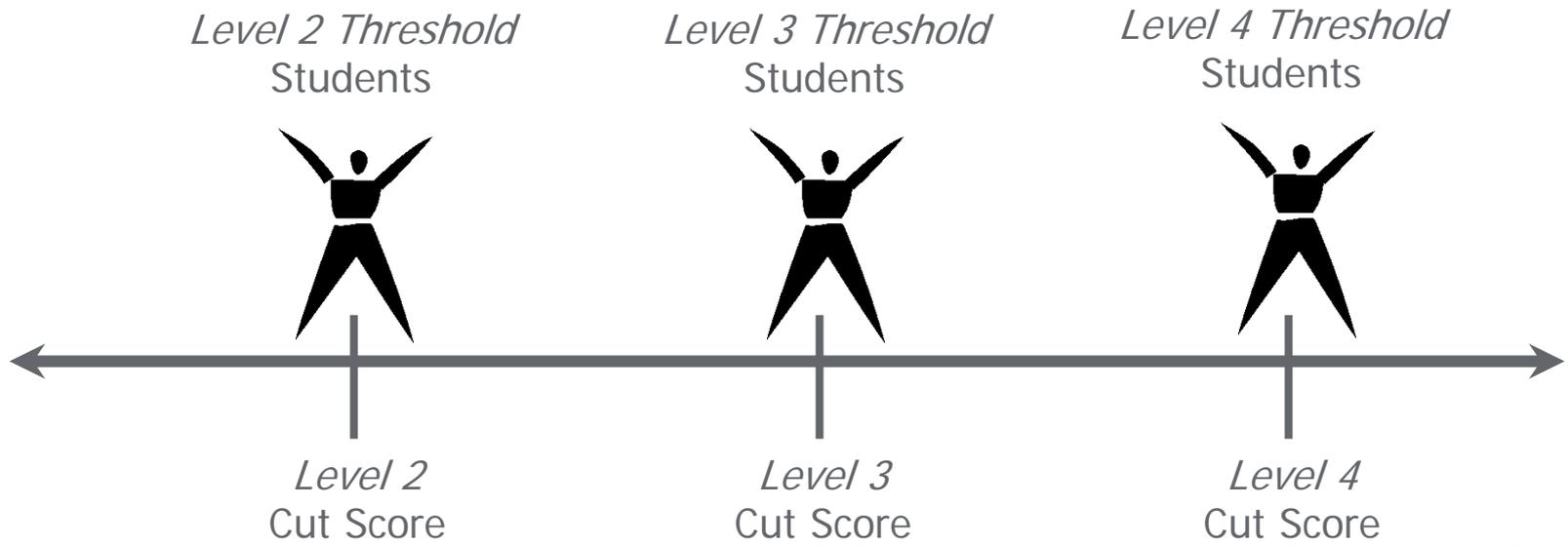
Threshold Student

- We recommend cut scores for Threshold Students—the students with a level of achievement that just barely qualifies them to be in the achievement level
 - These are the skills of the student just entering the achievement level

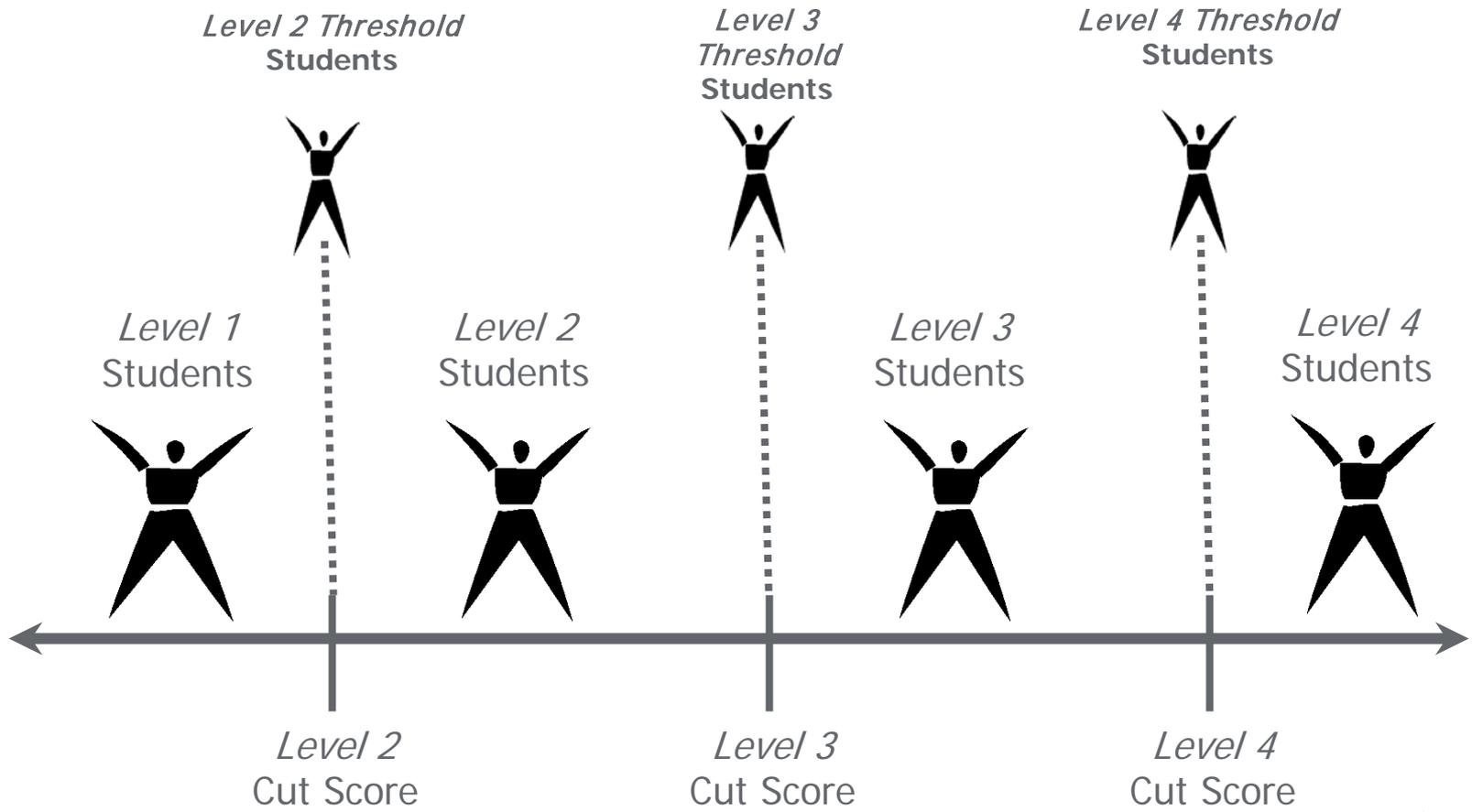




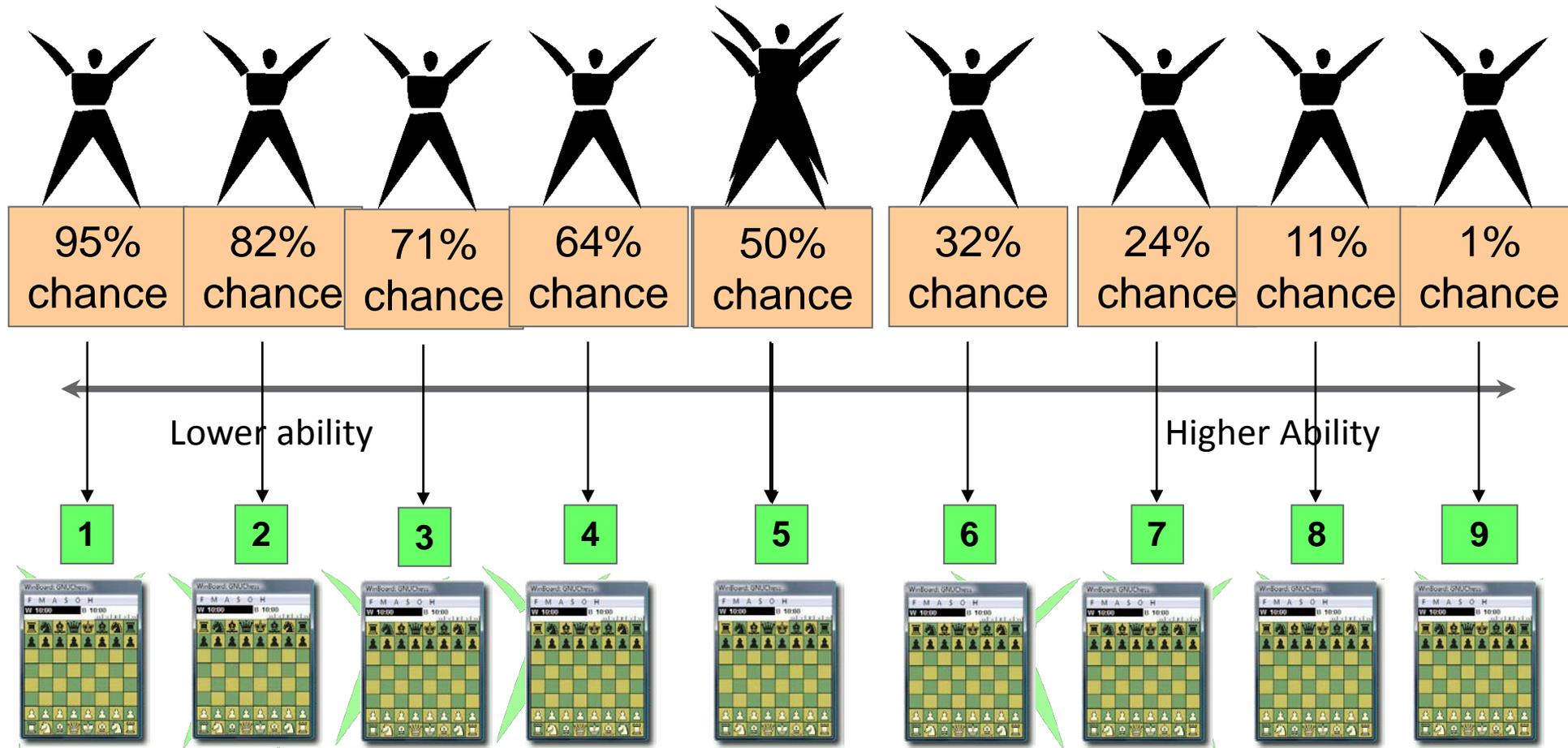
Three Threshold Students



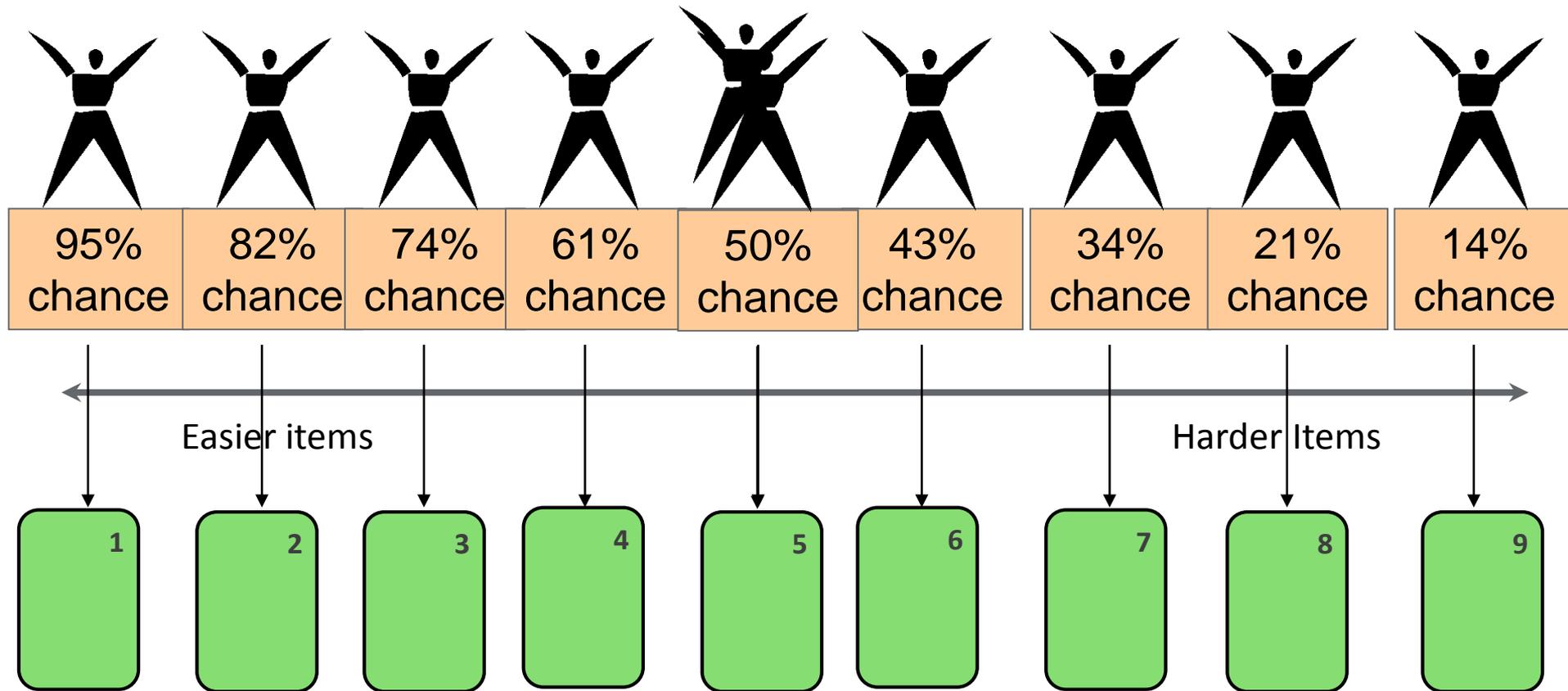
Three Cut Scores Define Four Achievement Levels



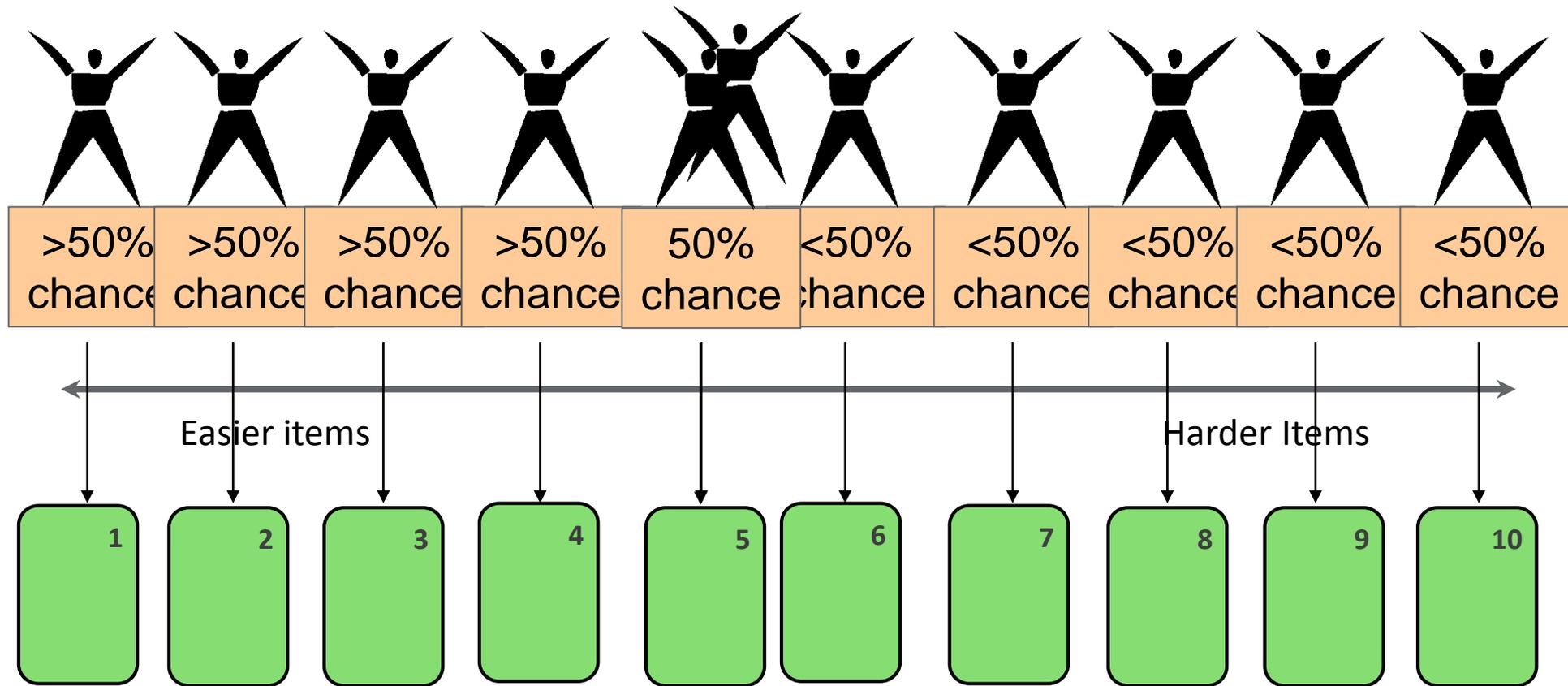
Locating a Chess Player on a Scale



Locating the Threshold Student on the Scale



Locating the Threshold Student on the Scale

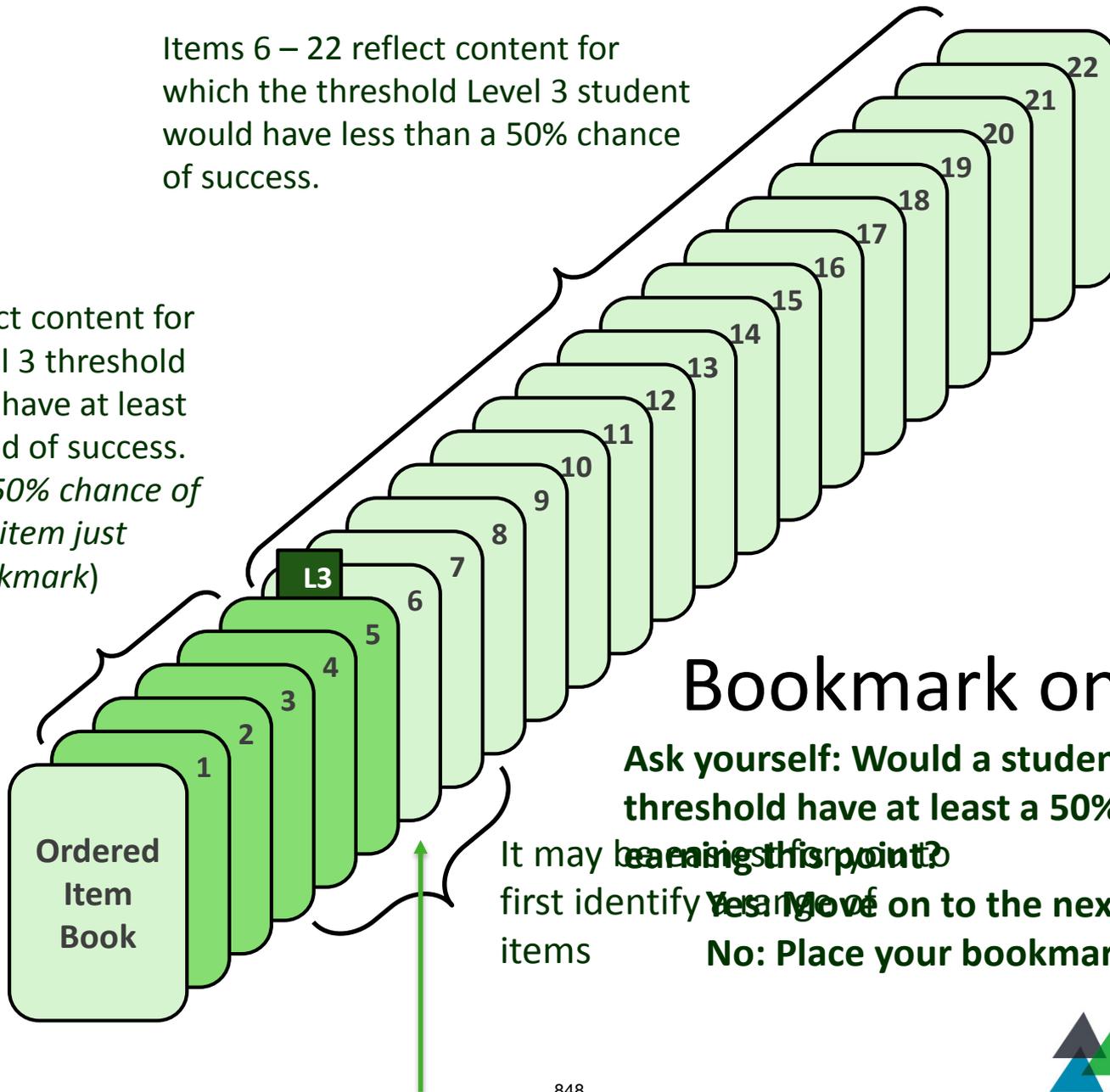


Bookmark Placement Instructions

- Ask yourself: Would a student at the threshold have at least a 50% chance of earning this point?
 - Yes: Move on to the next item.
 - No: Place your bookmark here.

Items 6 – 22 reflect content for which the threshold Level 3 student would have less than a 50% chance of success.

Items 1-5 reflect content for which the Level 3 threshold student would have at least a 50% likelihood of success. (with about a 50% chance of success on the item just before the bookmark)



Bookmark on Page 6

Ask yourself: Would a student at the threshold have at least a 50% chance of

learning this point?

Yes: Move on to the next item.

No: Place your bookmark here.

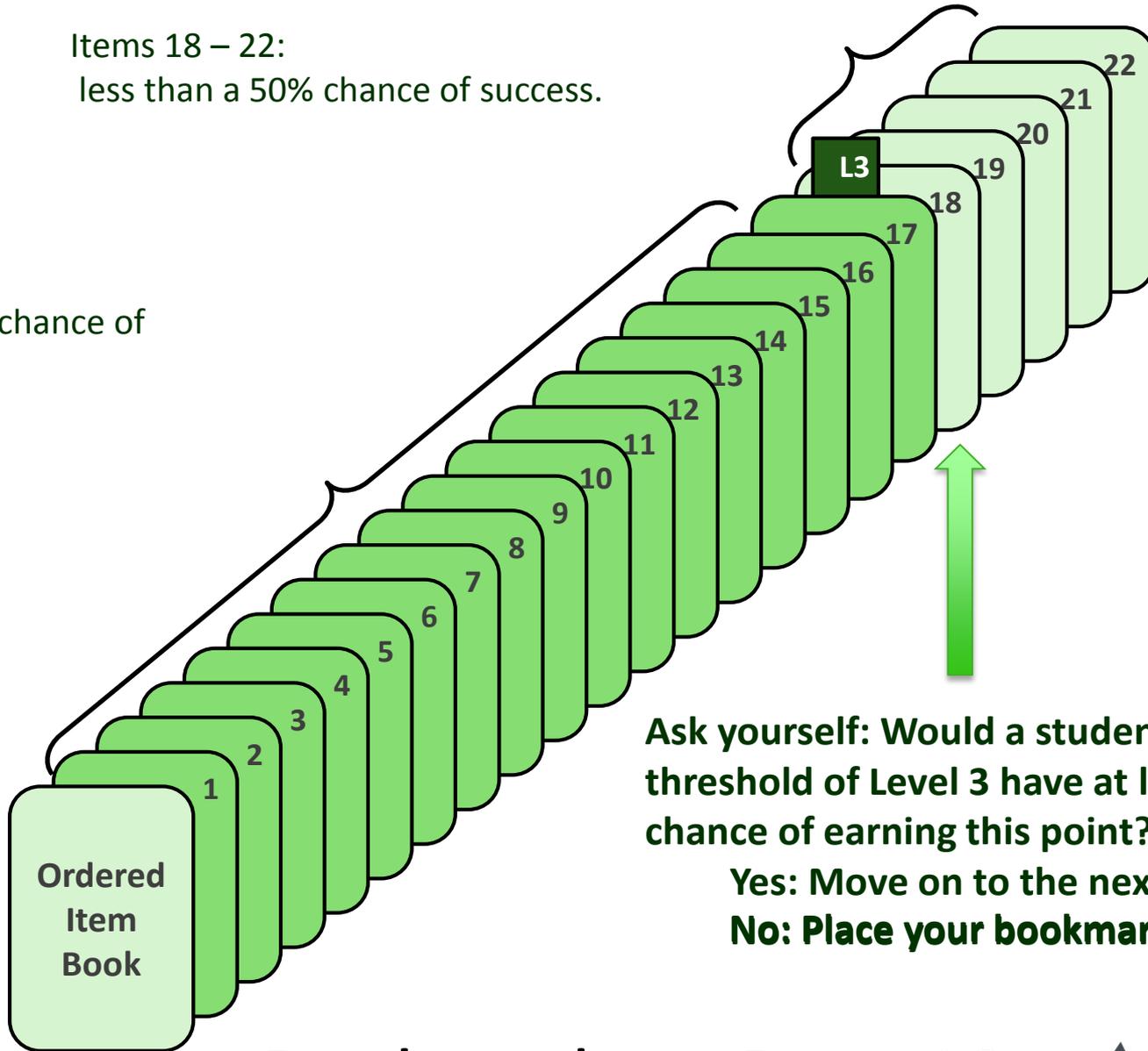


Level 3 Bookmark Placement

- When you place your Level 3 bookmark, think about a student at the threshold of Level 3 based on the achievement level descriptors
 - They consistently demonstrate those skills

Items 18 – 22:
less than a 50% chance of success.

Items 1-17:
At least 50% chance of
success



Ask yourself: Would a student at the
threshold of Level 3 have at least a 50%
chance of earning this point?

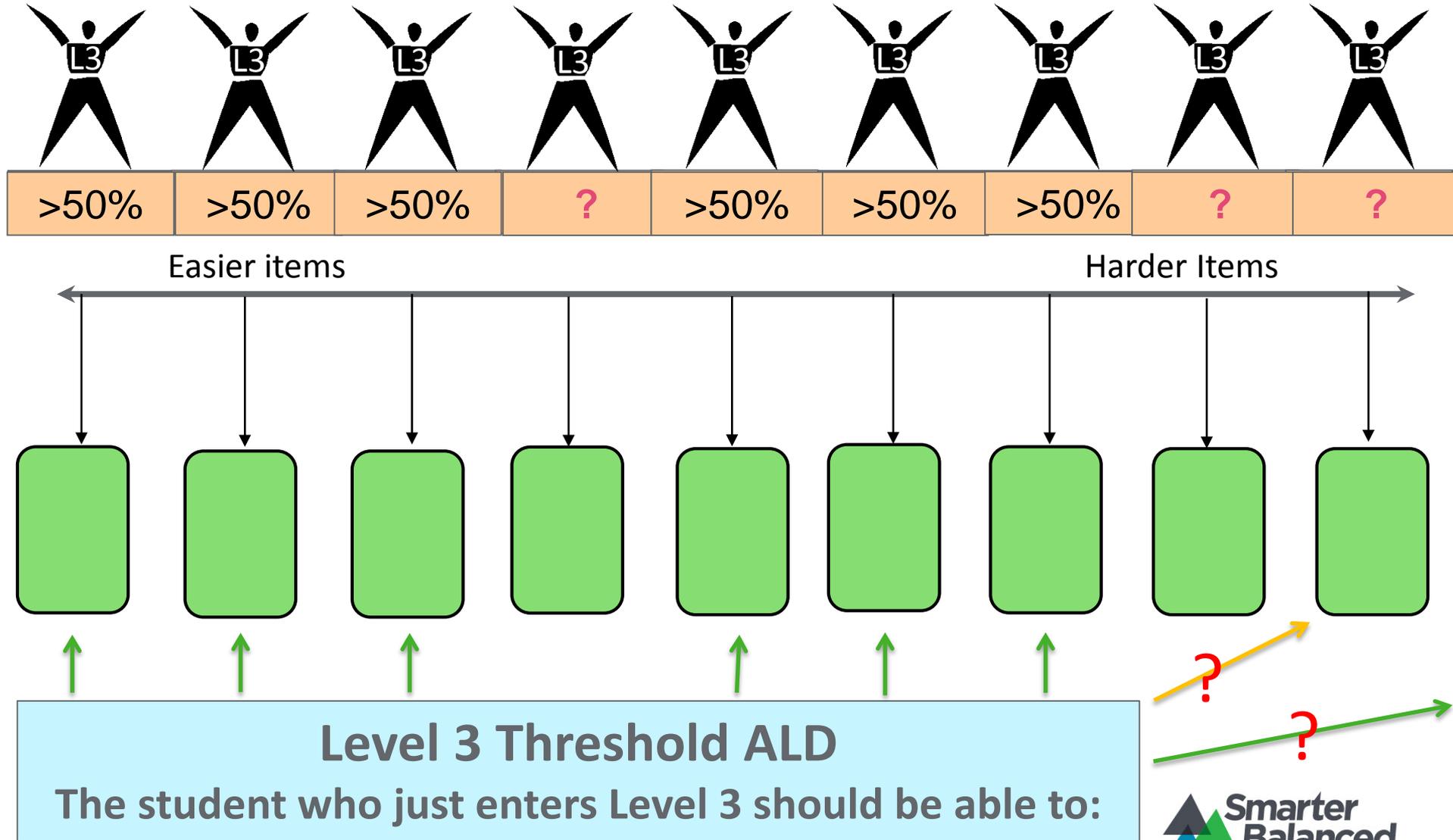
Yes: Move on to the next item.

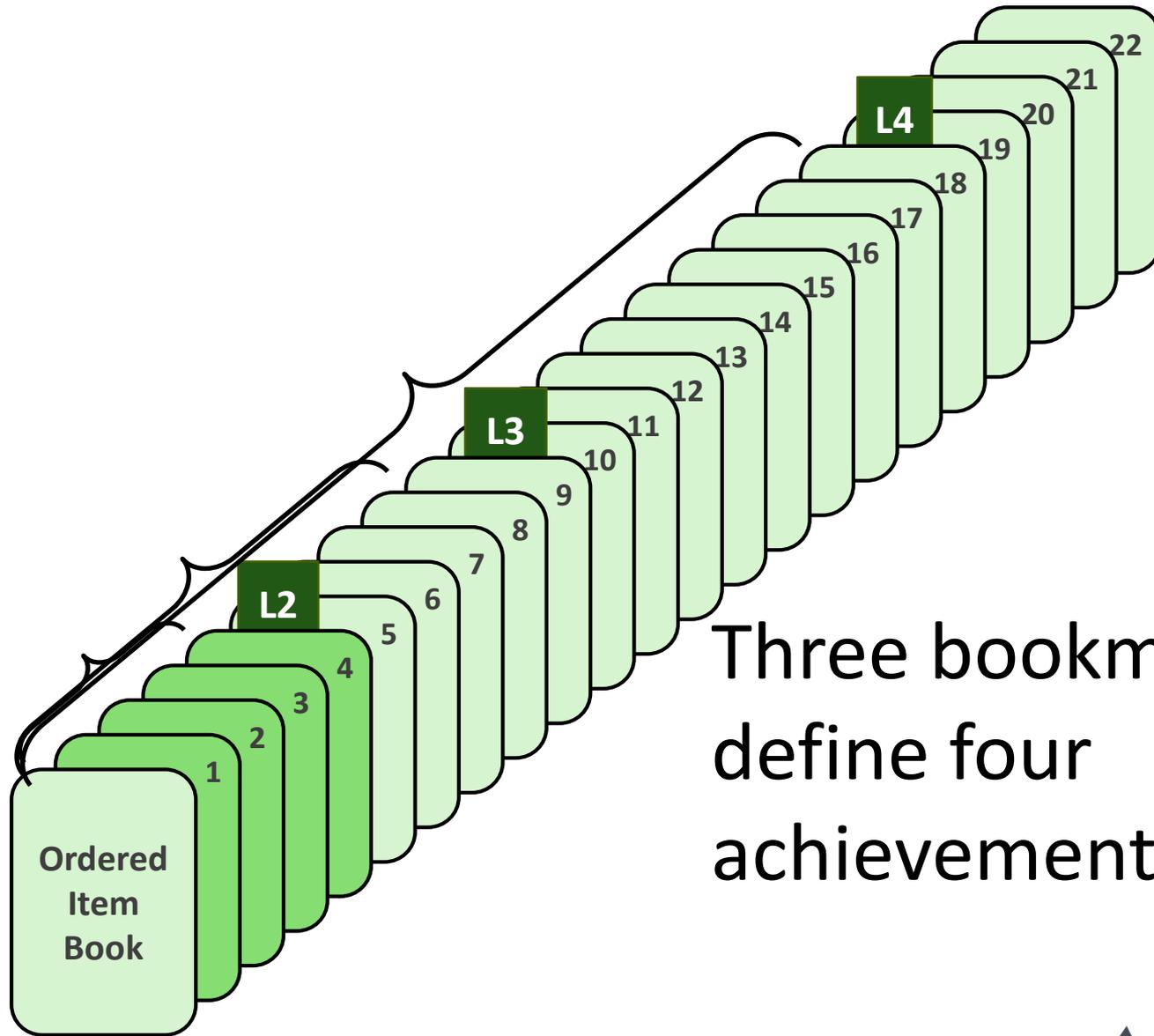
No: Place your bookmark here.

Bookmark on Page 18



How Can the Threshold ALDs Inform Your Bookmark Judgments?





Three bookmarks define four achievement levels.

Setting Bookmarks

Item Map for Booklet: Achievement

English Language Arts/Literacy, 4

Round: 1

Enter Comment

Review each question in the ordered item booklet by clicking on the page number

View Selected

Page	Claim and Target	Depth of Knowledge	Bookmark	Comments	Reviewed
1	Claim: 2-W, Target: 8-4	1			<input type="checkbox"/>
2	Claim: 2-W, Target: 9-4	1			<input type="checkbox"/>
3	NAEP				<input type="checkbox"/>
4	Claim: 2-W, Target: 1-4	2			<input type="checkbox"/>
5	Claim: 3-L, Target: 4-4	2			<input type="checkbox"/>
6	Claim: 2-W, Target: 2-4	4			<input checked="" type="checkbox"/>
7	Claim: 3-L, Target: 4-4	2			<input type="checkbox"/>
8	Claim: 2-W, Target: 8-4	1			<input type="checkbox"/>
9	Claim: 1-LT, Target: 3-4	2			<input type="checkbox"/>
10	Claim: 2-W, Target: 9-4	1			<input type="checkbox"/>
11	Claim: 3-L, Target: 4-4	2			<input type="checkbox"/>
12	Claim: 1-LT, Target: 7-4	2			<input type="checkbox"/>
13	Claim: 2-W, Target: 3-4	2			<input type="checkbox"/>

Set Bookmark

Select a bookmark value and click "Set Bookmark" below to set a bookmark at the selected Location value. Click "Cancel" to return.

- Level 2
- Level 3
- Level 4

Set Bookmark

Bookmark Placement Instructions

- Ask yourself: Would a student at the threshold have at least a 50% chance of earning this point?
 - Yes: Move on to the next item.
 - No: Place your bookmark here.



Questions?

- Thank you for your participation!

Achievement Level Setting Questionnaire

Please review the materials for your assigned subject and grade, and complete this questionnaire

Smarter Balanced Claims and Targets:

ELA Grade 3	Mathematics Grade 3
ELA Grade 4	Mathematics Grade 4
ELA Grade 5	Mathematics Grade 5
ELA Grade 6	Mathematics Grade 6
ELA Grade 7	Mathematics Grade 7
ELA Grade 8	Mathematics Grade 8
ELA Grade 11	Mathematics Grade 11

Achievement Level Descriptors:

ELA Grade 3	Mathematics Grade 3
ELA Grade 4	Mathematics Grade 4
ELA Grade 5	Mathematics Grade 5
ELA Grade 6	Mathematics Grade 6
ELA Grade 7	Mathematics Grade 7
ELA Grade 8	Mathematics Grade 8
ELA Grade 11	Mathematics Grade 11

Part 1: Smarter Balanced Claims and Targets and Achievement Level Descriptors

I have reviewed the Smarter Balanced Claims and Targets for the grade and subject assigned to me.

- Yes
- No

I understand the relationship between the content of the standards and what students are expected to know and be able to do in this subject at this grade level.

- Yes
- No

I understand how items measuring the same content standard can require students to apply different strategies and perform at different cognitive levels.

- Yes
- No

I have reviewed the Achievement Level Descriptors for the grade and subject assigned to me

- Yes
- No

I have discussed with my small group the Achievement Level Descriptors for the grade and subject assigned to me.

- Yes
- No

I understand the progression of student achievement from Achievement Level 1 to Achievement Level 4.

- Yes
- No

Which type of Level 3 student is most relevant to the work you will be doing this week?

- The typical student at a given level
- The very high performing student at a given level
- The student who minimally demonstrates achievement at a given level with supports and/or accommodations

ATTENTION: Please pause here and complete the Smarter Balanced Training Test. Click **Save** below to retain your responses before you begin the Smarter Balanced Training Test. You may access the test by clicking this link: [Smarter Balanced Training Test](#) and logging in as Guest. When you are finished with the training test, you may continue the questionnaire.

Achievement Level Setting Questionnaire

Smarter Balanced Training and Practice Tests

I have taken the Training Test.

- Yes
- No

I have discussed the Training Test with my small group.

- Yes
- No

I have reviewed the Practice Test.

- Yes
- No

I have discussed the Practice Test with my small group.

- Yes
- No

I have reviewed the Performance Task.

- Yes
- No

I have discussed the Performance Task with my small group.

- Yes
- No

Orientation Round Questionnaire

Now that you have completed the orientation round, please complete the following questionnaire.

I have received instruction in the Bookmark Procedure.

- Yes
- No

I have participated in the Orientation Round for the Bookmark Procedure.

- Yes
- No

I placed bookmarks for Levels 2, 3, and 4.

- Yes
- No

I participated in the discussion following the tally of bookmarks.

- Yes
- No

I understand how my bookmark will be combined with the bookmarks of other panelists to derive a cut score.

- Yes
- No

I understand the procedure I am to follow and am ready to begin Round 1.

- Yes
- No

Round 1. Questions for Reflection

Imagine your bookmarks were transformed into cut scores that were used across the nation. About how many students would you expect to be classified as *Level 3* or above, based on *your* Level 3 bookmark placements?

How did the conversation after Round 1 influence your bookmark placements?

Rank the top 3 things that influenced your Round 3 bookmark placement decisions.

①	②	③	Achievement Level Descriptors
①	②	③	Discussions at My Table
①	②	③	My Performance on the Practice Test
①	②	③	Online Panel Results

How confident are you about the three bookmarks you just entered?

Level 2	Level 3	Level 4
<input type="radio"/> Very Confident	<input type="radio"/> Very Confident	<input type="radio"/> Very Confident
<input type="radio"/> Confident	<input type="radio"/> Confident	<input type="radio"/> Confident
<input type="radio"/> Uncertain	<input type="radio"/> Uncertain	<input type="radio"/> Uncertain
<input type="radio"/> Very Uncertain	<input type="radio"/> Very Uncertain	<input type="radio"/> Very Uncertain

I understand the procedure I am to follow and am ready to begin Round 2.

- Yes
- No

Round 2. Questions for Reflection

Imagine your bookmarks were transformed into cut scores that were used across the nation. About how many students would you expect to be classified as *Level 3* or above, based on *your* Level 3 bookmark placements?

How did the conversation after Round 2 influence your bookmark placements?

Rank the top 3 things that influenced your Round 3 bookmark placement decisions.

①	②	③	Achievement Level Descriptors
①	②	③	ACT Data (Grade 11 Only)
①	②	③	Discussions at My Table
①	②	③	Impact Data
①	②	③	NAEP/PISA Items and Data (Grades 4, 8, and 11 only)
①	②	③	My Performance on the Practice Test
①	②	③	Online Panel Results
①	②	③	Roomwide Discussions

How confident are you about the three bookmarks you just entered?

Level 2	Level 3	Level 4
<input type="radio"/> Very Confident	<input type="radio"/> Very Confident	<input type="radio"/> Very Confident
<input type="radio"/> Confident	<input type="radio"/> Confident	<input type="radio"/> Confident
<input type="radio"/> Uncertain	<input type="radio"/> Uncertain	<input type="radio"/> Uncertain
<input type="radio"/> Very Uncertain	<input type="radio"/> Very Uncertain	<input type="radio"/> Very Uncertain

I understand the procedure I am to follow and am ready to begin Round 3.

- Yes
- No

Post-Round Three Questionnaire

Now that you have completed setting your bookmarks, please answer these questions.

I have seen the results of Round 3.

Yes

No

I know where our panel set the Round 3 cut scores for Levels 2, 3, and 4.

Yes

No

Enter those cut scores here:

Level 2 _____

Level 3 _____

Level 4 _____

Based on your panel's Round 3 cut scores, what percentage of students would be classified at Level 3 or above?

Evaluation

Indicate your level of agreement with each of the following statements.

SA = Strongly Agree A = Agree D = Disagree SD = Strongly Disagree

Statement	SA	A	D	SD
1 The orientation provided me with a clear understanding of the purpose of the meeting.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2 The workshop leaders clearly explained the task.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3 The training and practice exercises helped me understand how to perform the task.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4 Taking the practice test helped me to understand the assessment.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5 The Achievement Level Descriptions were clear and useful.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6 The large and small group discussions aided my understanding of the process.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7 The time provided for discussions was appropriate.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8 There was an equal opportunity for everyone in my group to contribute his/her ideas and opinions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9 I was able to follow the instructions and complete the rating tasks accurately.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10 The discussions after the first round of ratings were helpful to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11 The discussions after the second round of ratings were helpful to me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12 The information showing the distribution of student scores was helpful to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13 I am confident about the defensibility and appropriateness of the final recommended cut scores.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14 The facilities and food service helped create a productive and efficient working environment.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

15 Comments

Round 1 Bookmark Results

Table D.4.1 Round 1 Bookmark Medians and IQR, Grade 3

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	17.5	33.5	61.0	21.5	47.5	66.0
Table 2	13.5	33.0	56.0	21.0	46.5	73.0
Table 3	13.0	32.0	57.5	21.5	38.5	70.0
Table 4	28.5	43.5	62.0	23.5	44.0	63.0
Table 5	25.0	46.0	70.0	35.5	52.0	66.0
Total	16.0	38.0	58.5	22.0	47.0	69.5
IQR	12.0–31.0	28.75–51.75	54.0–71.0	21.0–33.0	40.0–54.0	61.0–73.25

Table D.4.2 Round 1 Bookmark Medians and IQR, Grade 4

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	17.5	34.5	59.5	13.0	32.0	68.0
Table 2	23.0	40.0	60.5	12.0	29.0	74.5
Table 3	24.0	54.0	60.0	18.5	46.5	67.0
Table 4	23.5	50.0	65.0	11.0	28.0	67.5
Table 5	18.0	41.5	60.5	10.5	35.0	65.0
Total	20.0	42.0	60.0	12.0	33.0	69.0
IQR	15.0–24.0	38.0–54.0	59.0–65.0	9.5–17.0	27.5–46.0	62.5–73.5

Table D.4.3 Round 1 Bookmark Medians and IQR, Grade 5

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	13.0	27.0	69.0	16.5	50.0	65.0
Table 2	10.5	27.0	53.0	20.5	46.5	65.0
Table 3	17.0	34.0	63.0	25.0	50.0	65.0
Table 4	10.0	24.0	64.0	25.0	50.0	69.5
Table 5	12.5	29.5	59.0	23.5	58.5	71.5
Total	13.0	27.0	63.0	21.5	50.0	65.5
IQR	9.0–17.0	24.0–38.0	52.0–66.0	17.5–26.75	48.0–52.75	64.25–70.75

Table D.4.4 Round 1 Bookmark Medians and IQR, Grade 6

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	24.5	48.0	68.5	21.0	56.5	63.5
Table 2	13.0	33.5	64.5	14.0	30.5	56.5
Table 3	15.0	30.0	59.5	8.0	31.0	52.5
Table 4	15.0	32.5	63.0	17.5	34.0	60.5
Table 5	10.0	31.0	55.5	20.5	50.0	65.5
Total	15.0	35.0	63.0	18.0	37.5	61.0
IQR	10.75–21.75	27.0–46.25	56.5–66.5	8.0–20.25	30.5–55.5	52.5–65.25

Table D.4.5 Round 1 Bookmark Medians and IQR, Grade 7

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	17.0	55.0	74.0	38.0	56.5	63.5
Table 2	22.0	37.0	74.0	23.5	59.5	67.0
Table 3	16.0	49.0	69.0	23.0	41.0	58.0
Table 4	14.0	52.0	70.0	15.5	39.5	56.0
Table 5	16.0	39.5	67.0	13.0	30.5	64.5
Total	16.0	41.0	69.0	21.5	42.5	63.0
IQR	9.0–23.0	37.0–52.0	64.0–74.0	13.0–26.0	34.5–57.0	57.75–67.0

Table D.4.6 Round 1 Bookmark Medians and IQR, Grade 8

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	16.0	40.5	67.5	15.5	48.5	61.5
Table 2	17.0	34.0	67.0	12.0	30.0	54.5
Table 3	19.0	41.5	67.5	26.0	39.0	58.0
Table 4	25.0	47.0	68.0	15.0	27.0	57.5
Table 5	22.0	40.0	67.0	26.0	41.5	59.0
Total	19.0	39.5	68.0	18.0	39.0	58.0
IQR	13.75–26.0	34.0–47.0	60.0–69.0	10.0–25.5	27.0–47.0	53.0–62.5

Table D.4.7 Round 1 Bookmark Medians and IQR, Grade 11¹

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	24.0	42.0	67.0	34.0	58.0	70.0
Table 2	24.0	45.5	64.5	28.5	58.0	69.5
Table 3	20.0	40.0	66.0	23.0	49.5	65.0
Table 4	25.5	58.0	70.0	28.5	59.5	73.0
Table 5	10.5	44.5	67.5	19.0	44.0	56.0
Table 6	29.0	42.0	67.0	20.0	43.0	67.0
Total-A	23.5	45.0	67.0	25.0	56.0	69.0
IQR-A	11.75-27.0	39.75-52.5	64.0-70.0	19.0-30.0	44.0-58.0	63.0-71.0
Table 7	21.0	45.0	69.0	16.0	43.0	73.5
Table 8	16.5	40.0	62.5	16.0	50.0	74.0
Table 9	20.0	45.0	61.0	18.5	44.0	69.0
Table 10	24.0	52.0	66.0	17.0	43.5	59.0
Table 11	23.5	52.5	64.5	23.0	48.0	68.0
Table 12	10.0	46.0	65.0	15.0	49.5	69.5
Total-B	19.5	44.5	65.0	18.0	45.0	69.0
IQR-B	11.75-27.0	36.5-54.0	59.75-67.5	14.0-27.0	38.0-55.0	62.0-73.0
Total	21.5	45.0	66.0	19.0	48.5	69.0
IQR	12.0-27.0	38.25-53.75	61.0-69.75	16.0-29.0	43.0-58.0	62.75-71.25

¹ Grade 11 ALS in ELA and Math was conducted in two panels of six tables each.

Round 2 Bookmark Results

Table D.4.8 Round 2 Bookmark Medians and IQR, Grade 3

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	24.5	40.0	70.0	28.0	50.5	72.0
Table 2	15.0	33.5	55.5	28.0	55.0	74.0
Table 3	11.5	29.0	53.5	25.0	48.5	71.0
Table 4	20.0	36.0	55.5	27.0	46.0	65.0
Table 5	21.5	43.0	69.0	33.0	51.0	66.0
Total	19.0	38.0	57.5	28.0	49.0	70.0
IQR	14.5-22.25	30.0-41.25	54.0-69.0	26.75-29.25	45.0-53.0	66.0-74.0

Table D.4.9 Round 2 Bookmark Medians and IQR, Grade 4

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	20.0	40.5	60.0	13.0	32.0	68.0
Table 2	22.0	44.0	63.0	9.5	30.0	72.5
Table 3	24.0	45.0	60.0	18.0	49.0	73.0
Table 4	22.5	46.0	64.5	9.0	30.0	71.0
Table 5	18.0	43.0	61.5	7.0	36.0	71.0
Total	20.0	44.0	63.0	9.0	32.0	71.0
IQR	18.0-24.0	43.0-48.0	60.0-64.0	9.0-13.0	30.0-37.5	71.0-73.0

Table D.4.10 Round 2 Bookmark Medians and IQR, Grade 5

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	15.0	34.0	68.0	18.5	53.0	64.0
Table 2	11.5	27.0	55.0	19.0	43.0	64.0
Table 3	16.0	33.0	61.0	25.5	50.0	65.0
Table 4	10.0	26.0	63.0	25.0	50.5	64.5
Table 5	11.5	28.5	55.5	20.0	52.5	66.0
Total	14.0	27.0	61.0	20.0	50.5	64.0
IQR	10.0-15.0	26.0-33.0	55.0-63.0	19.0-25.0	29.0-53.0	64.0-66.0

Table D.4.11 Round 2 Bookmark Medians and IQR, Grade 6

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	24.5	47.5	68.0	17.0	50.5	59.0
Table 2	14.0	38.5	64.5	10.0	16.5	51.5
Table 3	15.0	29.5	54.5	12.0	30.0	60.0
Table 4	15.5	33.5	61.5	19.0	37.0	62.0
Table 5	7.5	27.0	52.5	20.0	54.5	64.5
Total	15.0	36.5	63.0	16.5	37.0	60.0
IQR	11.0-24.0	29.0-43.0	54.75-65.0	12.0-20.0	29.75-50.25	59.0-63.0

Table D.4.12 Round 2 Bookmark Medians and IQR, Grade 7

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	10.0	39.0	70.0	23.5	51.0	67.0
Table 2	23.0	37.0	74.0	9.0	51.0	66.0
Table 3	9.0	37.0	63.5	25.0	40.5	58.0
Table 4	16.0	52.0	77.0	10.0	47.0	60.0
Table 5	17.0	36.0	61.0	17.0	39.5	57.5
Total	16.0	37.5	69.5	18.0	46.0	61.0
IQR	9.75-19.0	36.75-41.0	61.75-74.0	10.0-23.5	40.0-51.0	58.0-67.0

Table D.4.13 Round 2 Bookmark Medians and IQR, Grade 8

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	11.5	34.5	53.0	17.5	48.0	62.0
Table 2	17.0	34.0	60.5	17.5	47.0	63.0
Table 3	12.0	42.5	68.0	28.0	40.0	60.0
Table 4	26.0	48.0	74.0	15.5	35.0	57.5
Table 5	21.0	44.5	67.0	12.0	37.0	50.5
Total	17.0	40.5	66.5	17.0	40.0	60.0
IQR	12.0-23.0	34.0-48.0	59.75-70.0	14.0-21.0	35.5-48.0	57.0-62.0

Table D.4.14 Round 2 Bookmark Medians and IQR, Grade 11¹

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	24.0	40.0	67.0	31.0	58.0	69.0
Table 2	23.0	41.0	64.5	30.5	58.0	69.0
Table 3	20.5	42.0	66.5	26.0	50.0	67.0
Table 4	25.5	47.0	66.5	28.5	59.0	74.0
Table 5	15.5	43.0	66.5	19.0	45.0	56.0
Table 6	27.0	42.0	67.0	20.0	50.0	67.0
<i>Total-A</i>	24.0	42.0	66.0	27.0	55.0	69.0
<i>IQR-A</i>	20.0-27.0	40.0-44.5	65.0-67.0	20.0-31.0	49.0-58.0	67.0-69.0
Table 7	18.0	33.0	58.0	19.0	53.0	71.5
Table 8	11.5	37.5	62.5	19.0	51.0	70.0
Table 9	23.0	42.0	60.0	14.5	44.0	61.5
Table 10	24.0	46.0	62.5	18.0	45.0	56.0
Table 11	21.0	40.0	67.5	24.0	49.0	69.0
Table 12	22.0	45.0	65.0	15.0	52.5	68.5
<i>Total-B</i>	21.0	42.0	64.0	19.0	48.0	68.0
<i>IQR-B</i>	14.75-23.0	35.75-44.75	58.0-65.0	15.0-23.0	44.0-58.0	60.0-70.0
Total	22.0	42.0	65.0	20.0	50.0	69.0
IQR	17.0-24.0	38.25-44.0	61.0-67.0	18.0-28.25	45.0-58.0	63.75-70.0

Round 3 Bookmark Results

Table D.4.15 Round 3 Bookmark Medians and IQR, Grade 3

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	14.0	34.5	67.5	28.0	47.5	72.0
Table 2	13.0	34.0	58.0	28.0	56.5	73.0
Table 3	11.5	28.5	53.0	26.0	48.5	71.5
Table 4	11.5	28.0	52.0	26.0	46.0	66.5
Table 5	17.5	38.0	59.5	27.5	47.0	66.0
Total	13.0	33.0	54.0	27.0	47.0	70.0
IQR	11.75-15.5	28.0-38.0	53.0-61.75	26.0-28.0	44.5-53.0	66.0-73.25

Table D.4.16 Round 3 Bookmark Medians and IQR, Grade 4

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	20.0	44.0	61.5	17.0	38.0	71.0
Table 2	20.0	39.0	63.0	16.5	42.5	77.0
Table 3	15.0	37.0	60.0	18.0	49.0	73.0
Table 4	20.0	44.0	63.0	12.0	38.0	71.0
Table 5	17.5	43.0	61.0	13.0	39.0	71.0
Total	19.0	43.0	62.0	15.0	39.0	71.0
IQR	15.0-20.0	37.0-44.0	60.0-63.0	12.0-18.0	38.0-48.0	71.0-73.0

Table D.4.17 Round 3 Bookmark Medians and IQR, Grade 5

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	13.0	33.0	65.0	18.5	52.0	64.0
Table 2	11.0	27.0	63.0	19.0	50.0	64.0
Table 3	15.0	34.0	63.0	20.5	50.0	64.0
Table 4	10.0	27.0	63.0	20.0	50.0	62.0
Table 5	12.5	27.0	61.5	19.0	50.5	63.0
Total	11.0	27.0	63.0	19.0	50.0	64.0
IQR	10.0-15.0	27.0-33.0	61.0-65.0	18.25-21.0	50.0-51.0	62.0-64.0

Table D.4.18 Round 3 Bookmark Medians and IQR, Grade 6

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	21.0	37.5	64.5	18.0	50.0	59.0
Table 2	14.0	34.0	63.0	13.0	31.0	61.0
Table 3	15.0	30.0	54.0	13.0	30.5	60.0
Table 4	14.0	35.0	63.0	19.5	58.0	63.0
Table 5	7.0	28.0	53.0	20.0	51.5	63.0
Total	14.5	34.5	60.5	18.0	45.5	61.5
IQR	7.75-19.0	29.0-40.0	52.0-63.25	13.0-20.0	32.5-53.5	59.0-63.0

Table D.4.19 Round 3 Bookmark Medians and IQR, Grade 7

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	7.0	38.5	67.5	21.5	51.0	61.0
Table 2	16.0	37.0	74.0	16.0	45.5	62.5
Table 3	8.0	36.0	65.0	21.0	42.0	60.5
Table 4	16.0	52.0	77.0	14.0	45.0	59.0
Table 5	16.0	35.0	63.0	17.0	42.5	59.5
Total	16.0	38.0	66.0	17.0	45.0	61.0
IQR	8.0-16.0	34.5-43.5	64.0-74.0	13.25-21.0	42.0-51.0	58.75-64.0

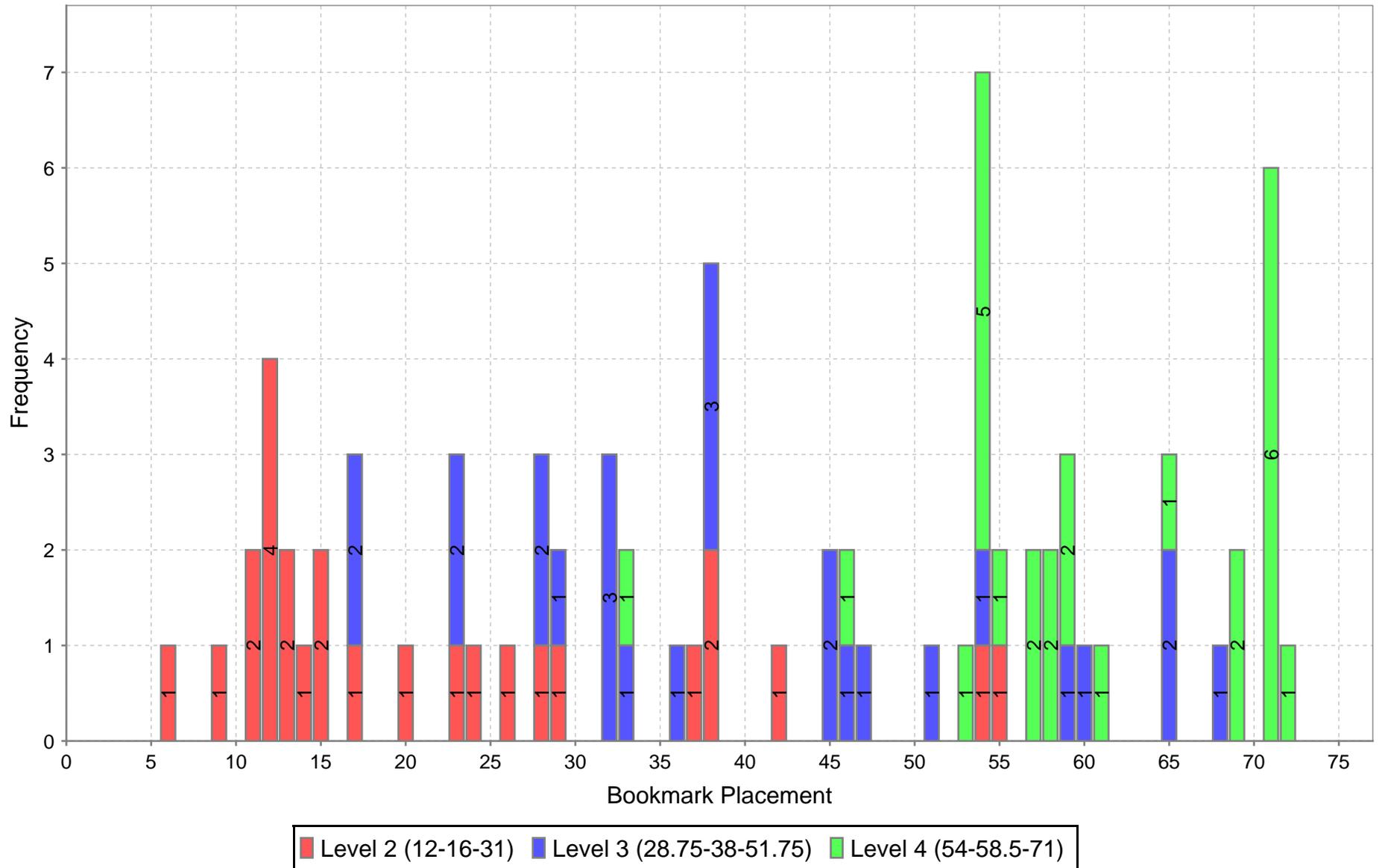
Table D.4.20 Round 3 Bookmark Medians and IQR, Grade 8

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	14.5	35.0	63.0	17.0	48.0	60.0
Table 2	17.0	33.5	60.0	16.0	48.0	63.5
Table 3	16.0	41.0	68.5	18.0	40.0	61.0
Table 4	26.0	51.0	75.0	16.5	37.0	60.5
Table 5	19.0	44.0	68.5	14.0	36.0	56.0
Total	18.0	39.5	68.0	16.0	40.0	60.0
IQR	14.0-21.5	34.0-46.5	60.0-70.0	15.0-18.0	36.5-48.0	57.5-63.0

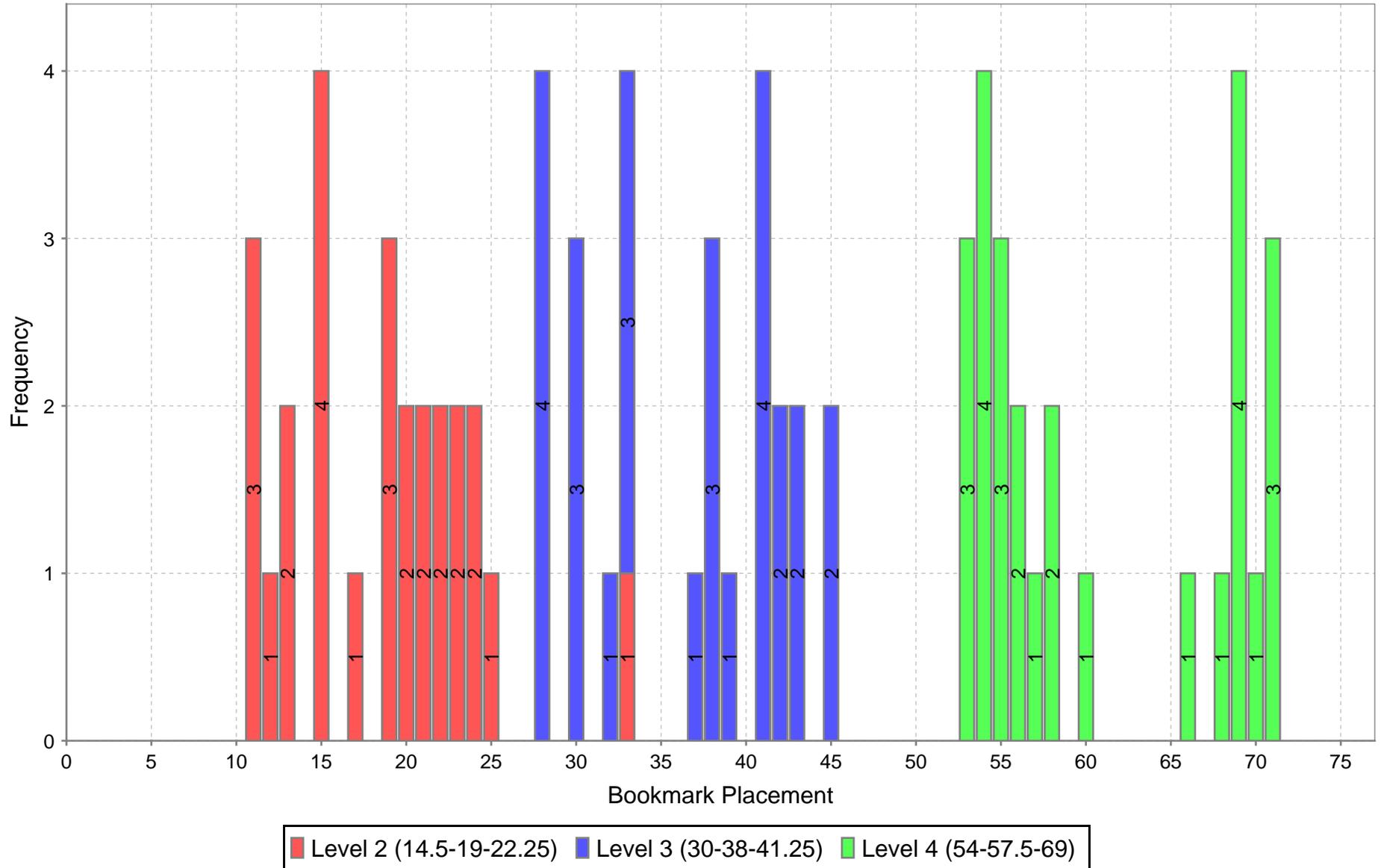
Table D.4.21 Round 3 Bookmark Medians and IQR, Grade 11¹

Table	ELA			Math		
	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median	Level 2 bookmark median	Level 3 bookmark median	Level 4 bookmark median
Table 1	19.0	40.0	67.0	28.0	56.0	69.0
Table 2	16.5	40.0	64.0	20.0	45.5	69.0
Table 3	17.5	40.0	65.0	26.0	50.0	65.5
Table 4	24.0	42.0	66.0	27.0	50.5	70.0
Table 5	14.5	43.0	66.0	20.0	45.0	57.0
Table 6	24.0	41.0	66.0	19.0	50.0	69.0
Total-A	19.0	43.0	66.0	20.0	49.0	69.0
IQR-A	15.75-23.25	39.5-42.0	64.0-67.0	19.0-28.0	41.0-56.0	63.0-69.0
20.0	39.0	63.0	20.0	19.0	52.0	71.0
11.5	38.5	65.0	11.5	19.0	50.0	69.0
22.0	42.0	63.0	22.0	19.0	45.5	69.0
23.0	43.5	63.5	23.0	14.5	44.5	61.0
22.5	44.5	65.0	22.5	21.5	48.0	67.5
17.0	45.0	64.5	17.0	15.5	50.0	69.0
21.0	42.0	64.0	21.0	19.0	48.0	69.0
IQR-B	14.25-23.0	39.75-45.0	62.0-65.0	15.0-20.0	44.0-52.0	64.0-70.0
Total	19.0	42.0	65.0	19.5	48.0	68.0
IQR	15.25-23.0	40.0-44.0	63.0-66.0	17.5-27.0	44.0-55.25	63.75-69.0

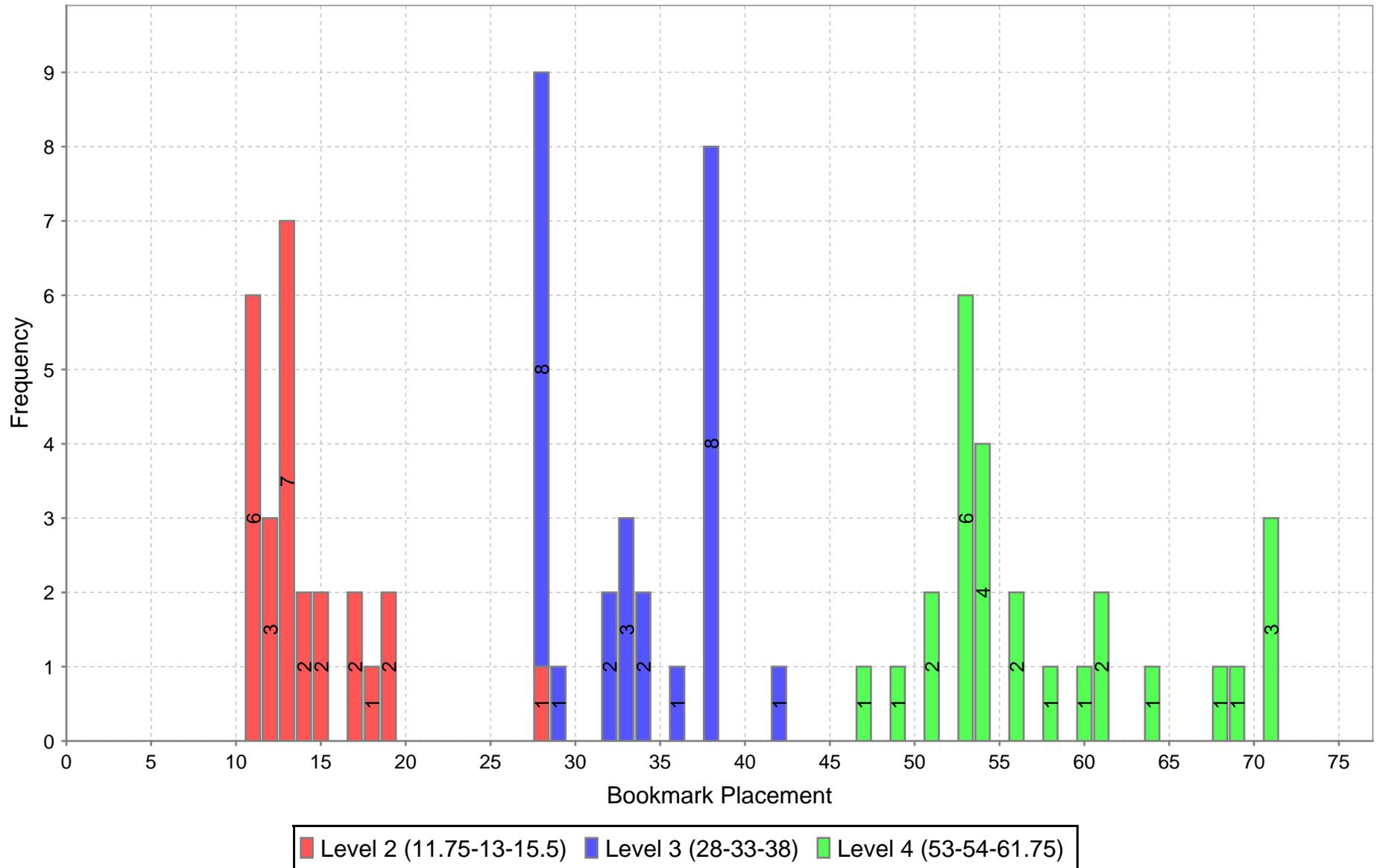
Smarter Balanced In-Person ALS Grade 3 English Language Arts/Literacy Frequency of Bookmark Placements Round 1



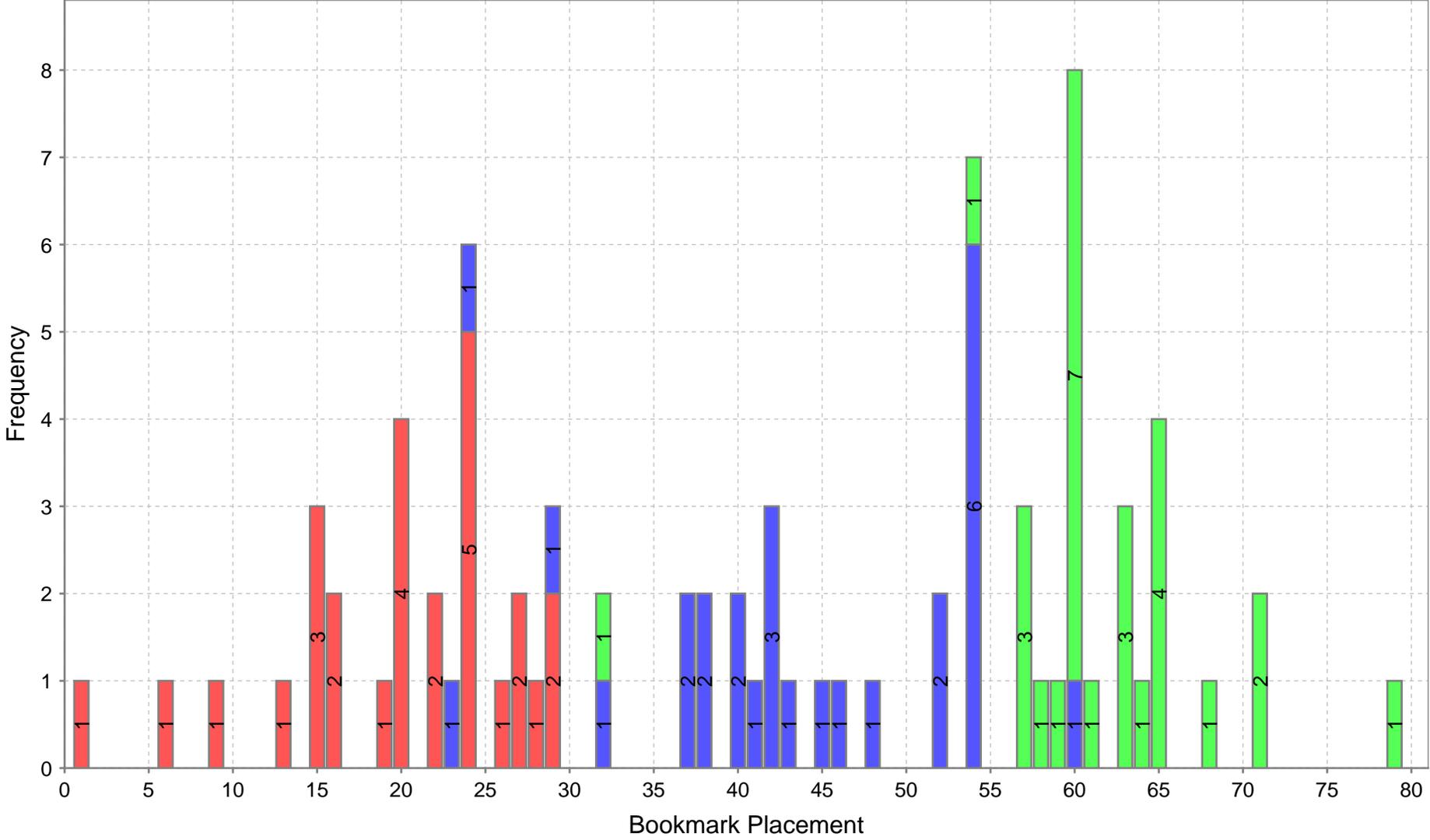
Smarter Balanced In-Person ALS Grade 3 English Language Arts/Literacy Frequency of Bookmark Placements Round 2



Smarter Balanced In-Person ALS Grade 3 English Language Arts/Literacy Frequency of Bookmark Placements Round 3

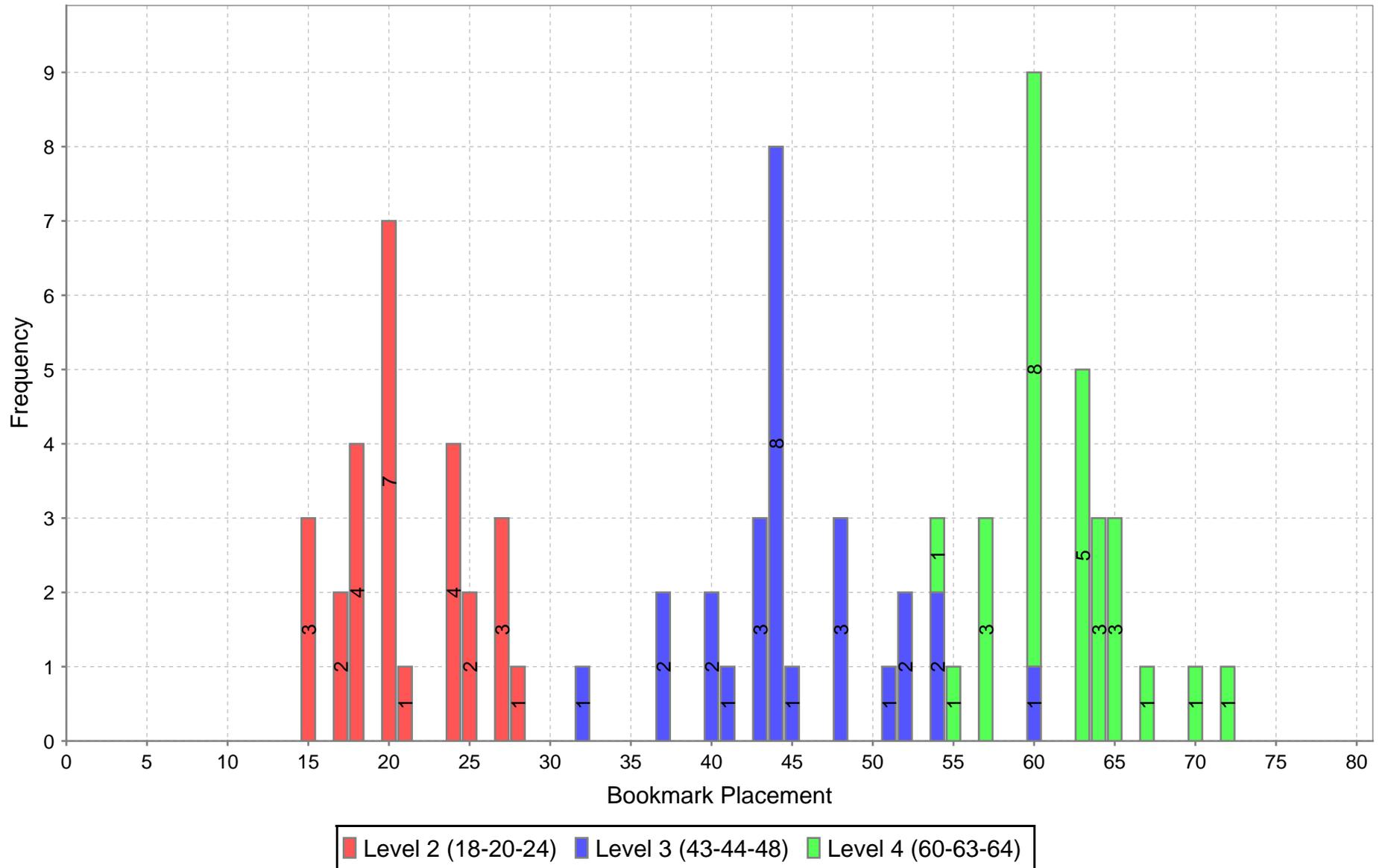


Smarter Balanced In-Person ALS Grade 4 English Language Arts/Literacy Frequency of Bookmark Placements Round 1

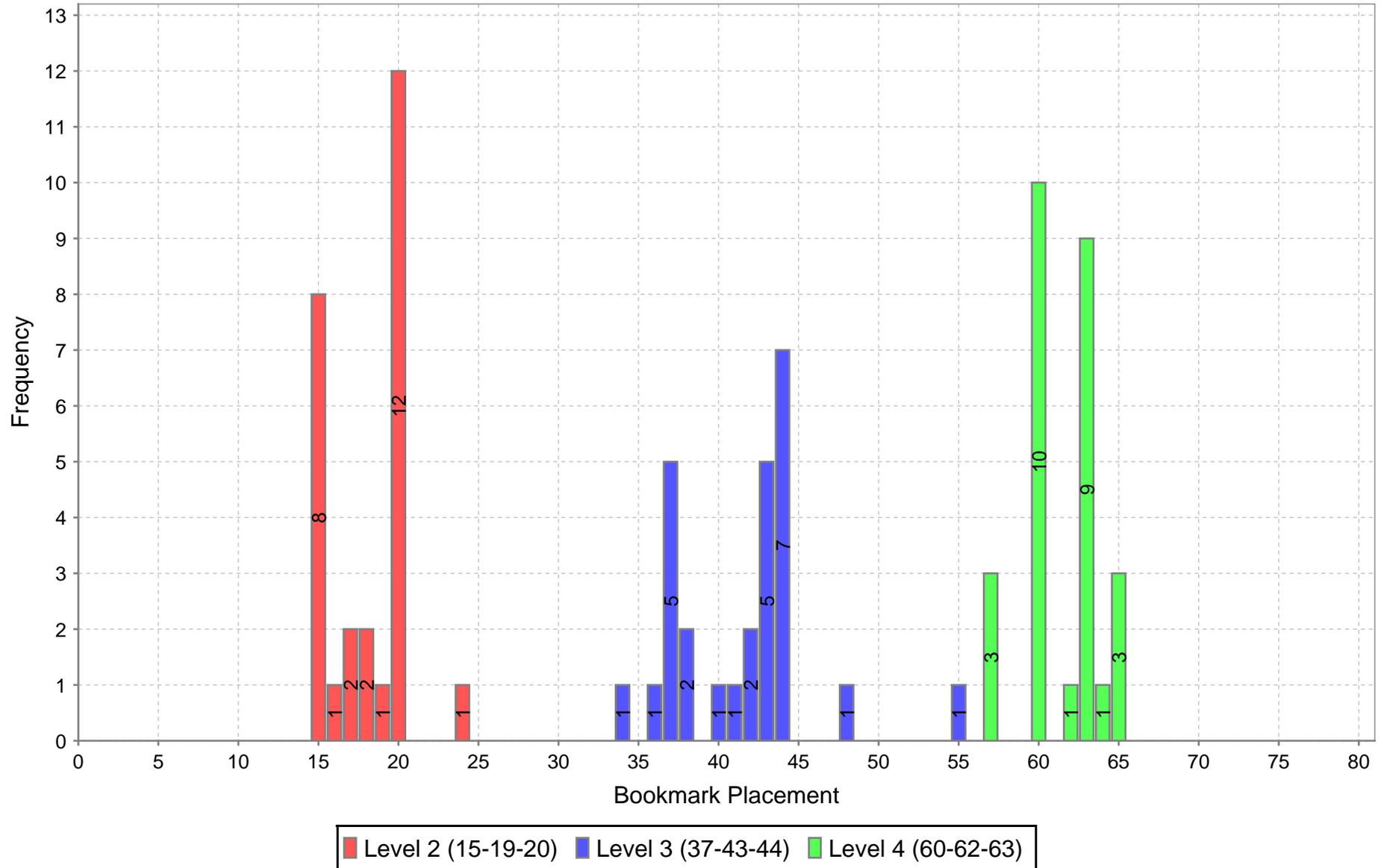


■ Level 2 (15-20-24)
 ■ Level 3 (38-42-54)
 ■ Level 4 (59-60-65)

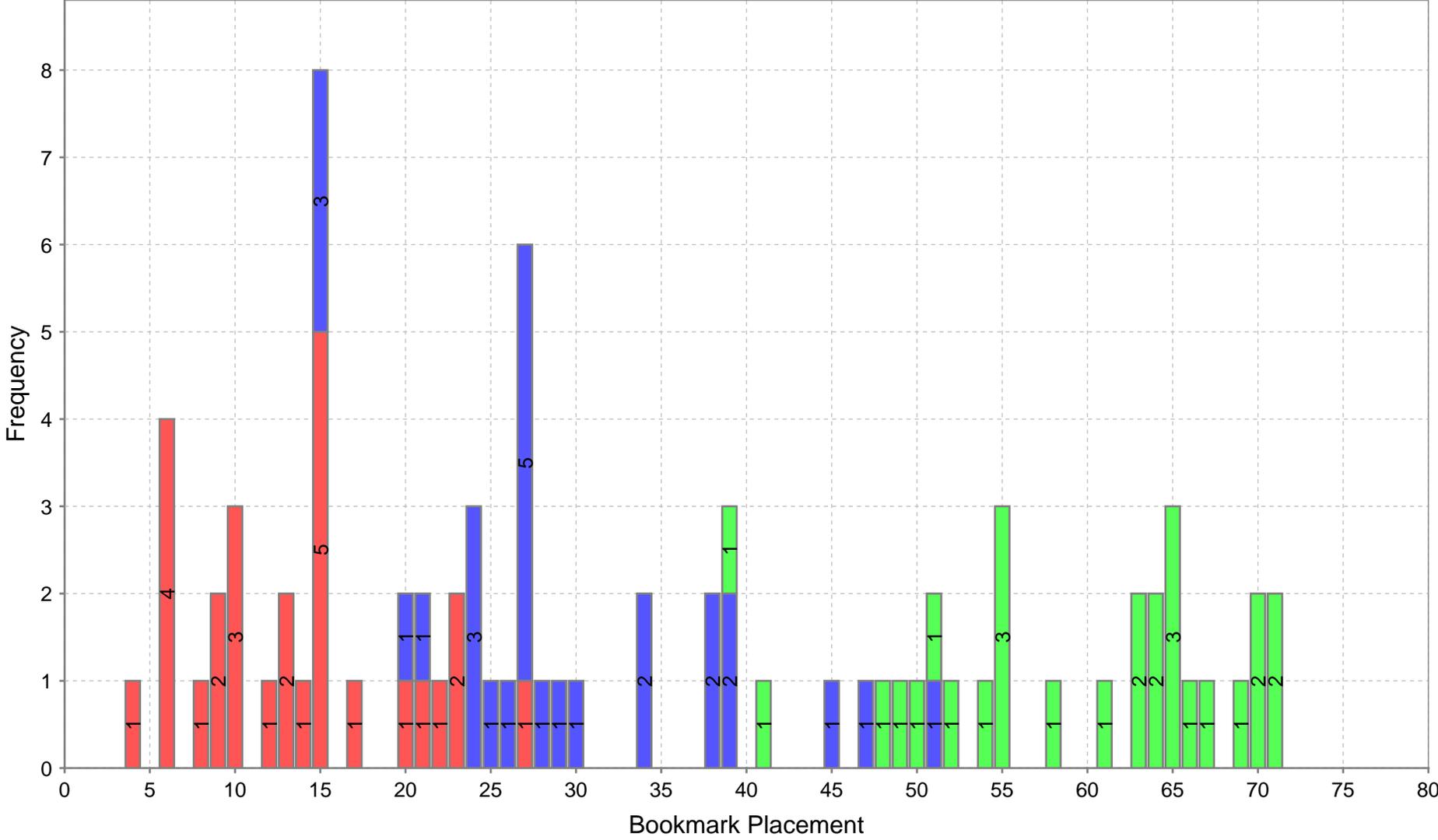
Smarter Balanced In-Person ALS Grade 4 English Language Arts/Literacy Frequency of Bookmark Placements Round 2



Smarter Balanced In-Person ALS Grade 4 English Language Arts/Literacy Frequency of Bookmark Placements Round 3

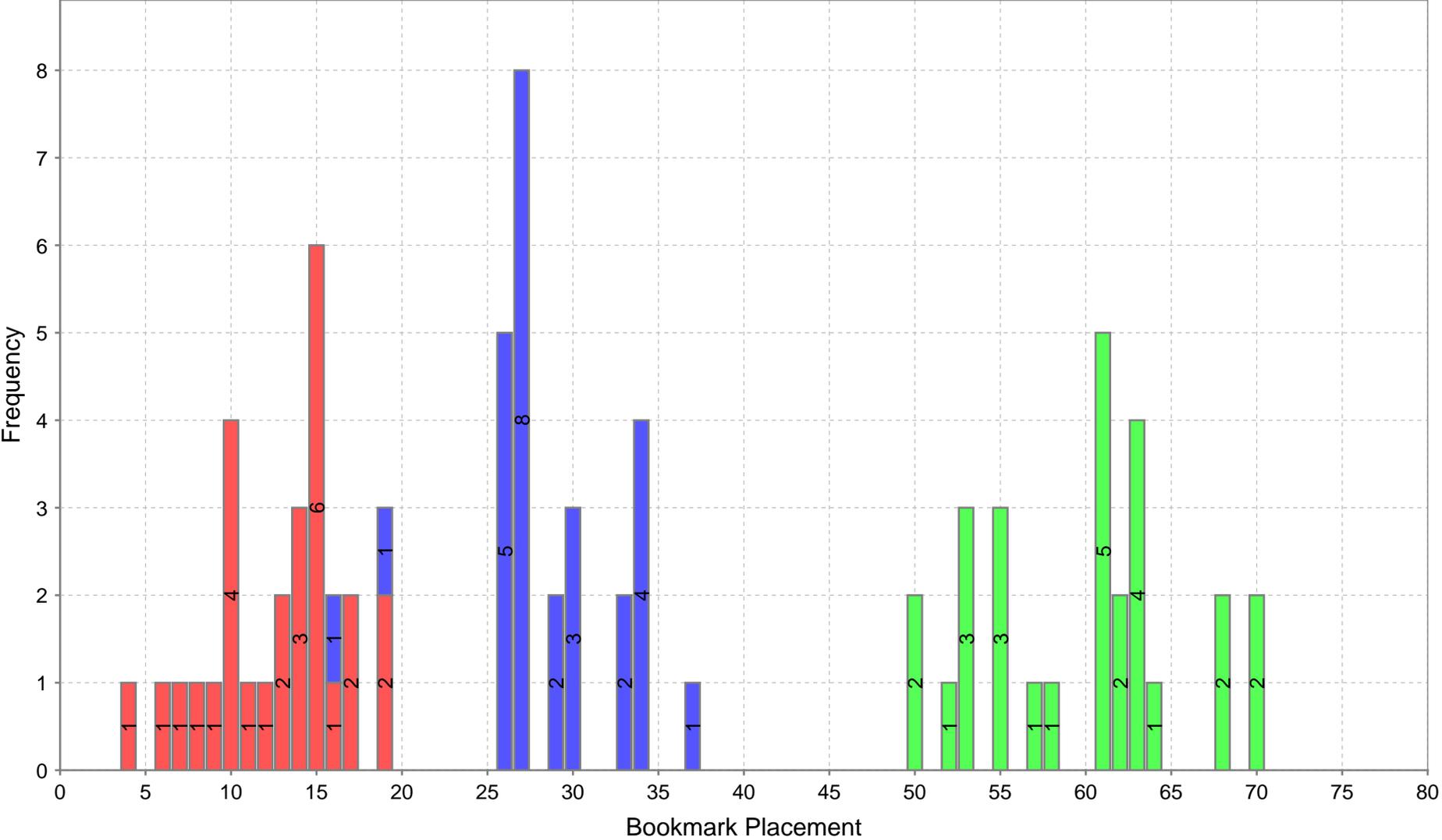


Smarter Balanced In-Person ALS Grade 5 English Language Arts/Literacy Frequency of Bookmark Placements Round 1



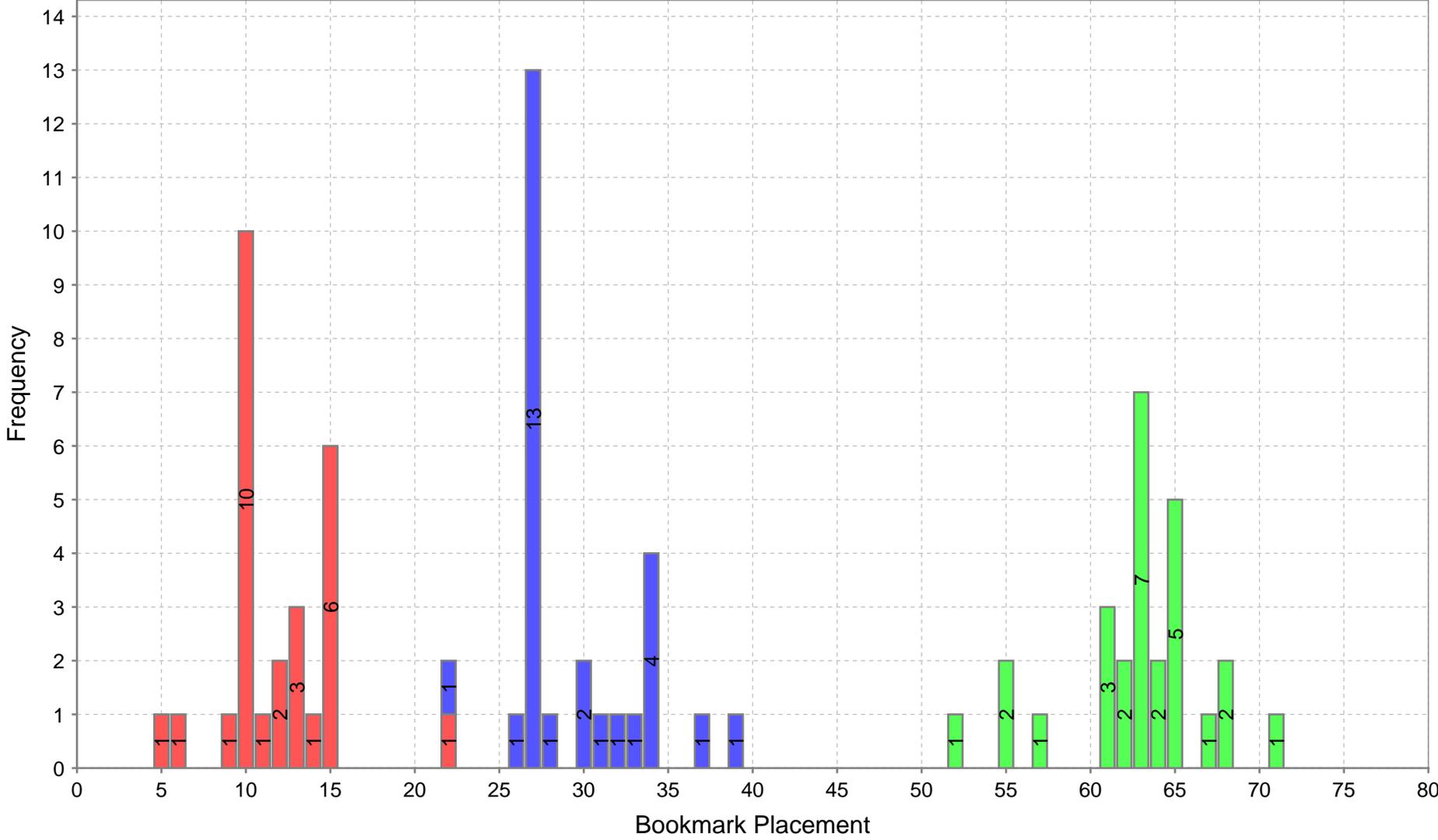
■ Level 2 (9-13-17)
 ■ Level 3 (24-27-38)
 ■ Level 4 (52-63-66)

Smarter Balanced In-Person ALS Grade 5 English Language Arts/Literacy Frequency of Bookmark Placements Round 2



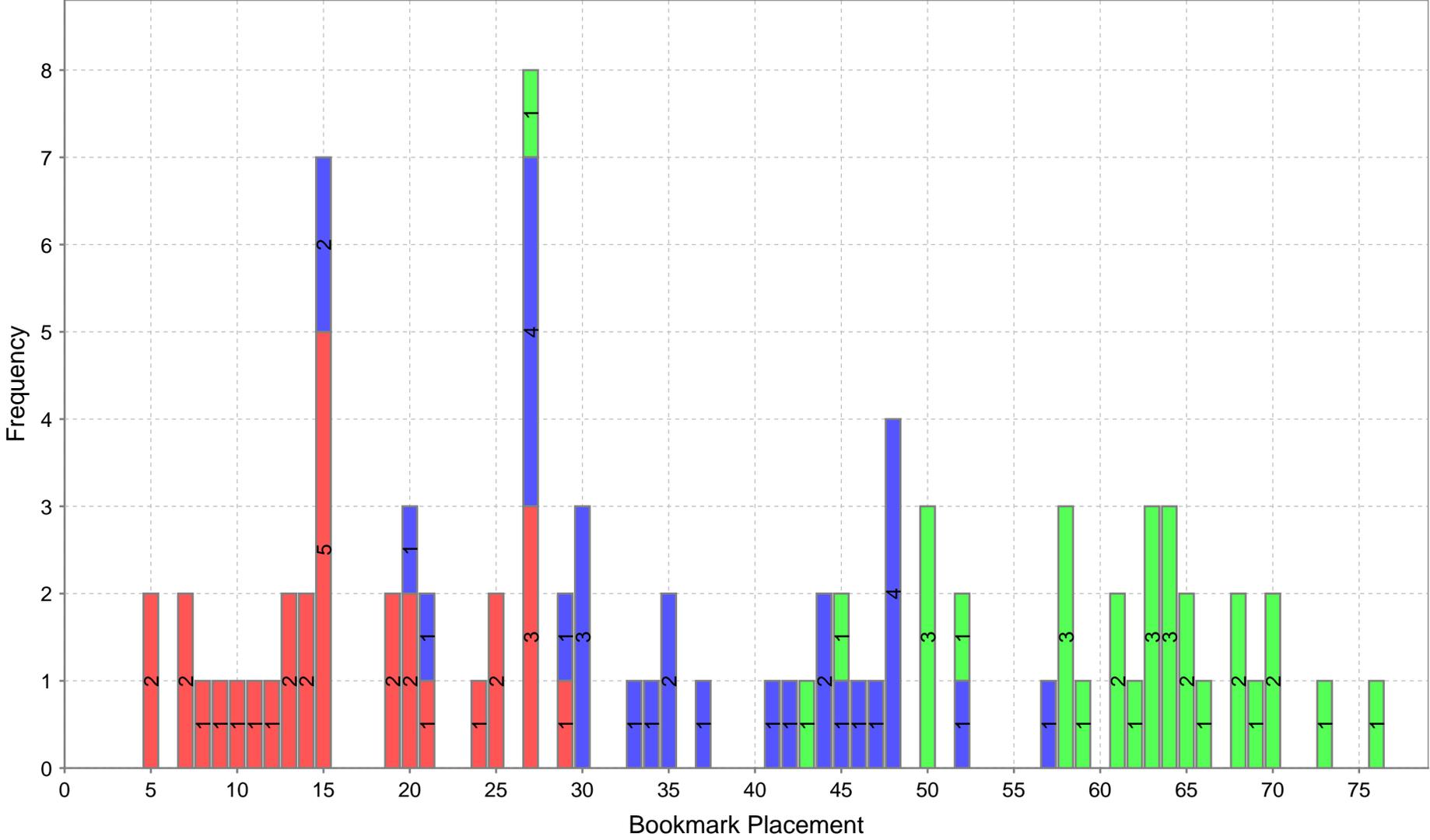
■ Level 2 (10-14-15)
 ■ Level 3 (26-27-33)
 ■ Level 4 (55-61-63)

Smarter Balanced In-Person ALS Grade 5 English Language Arts/Literacy Frequency of Bookmark Placements Round 3



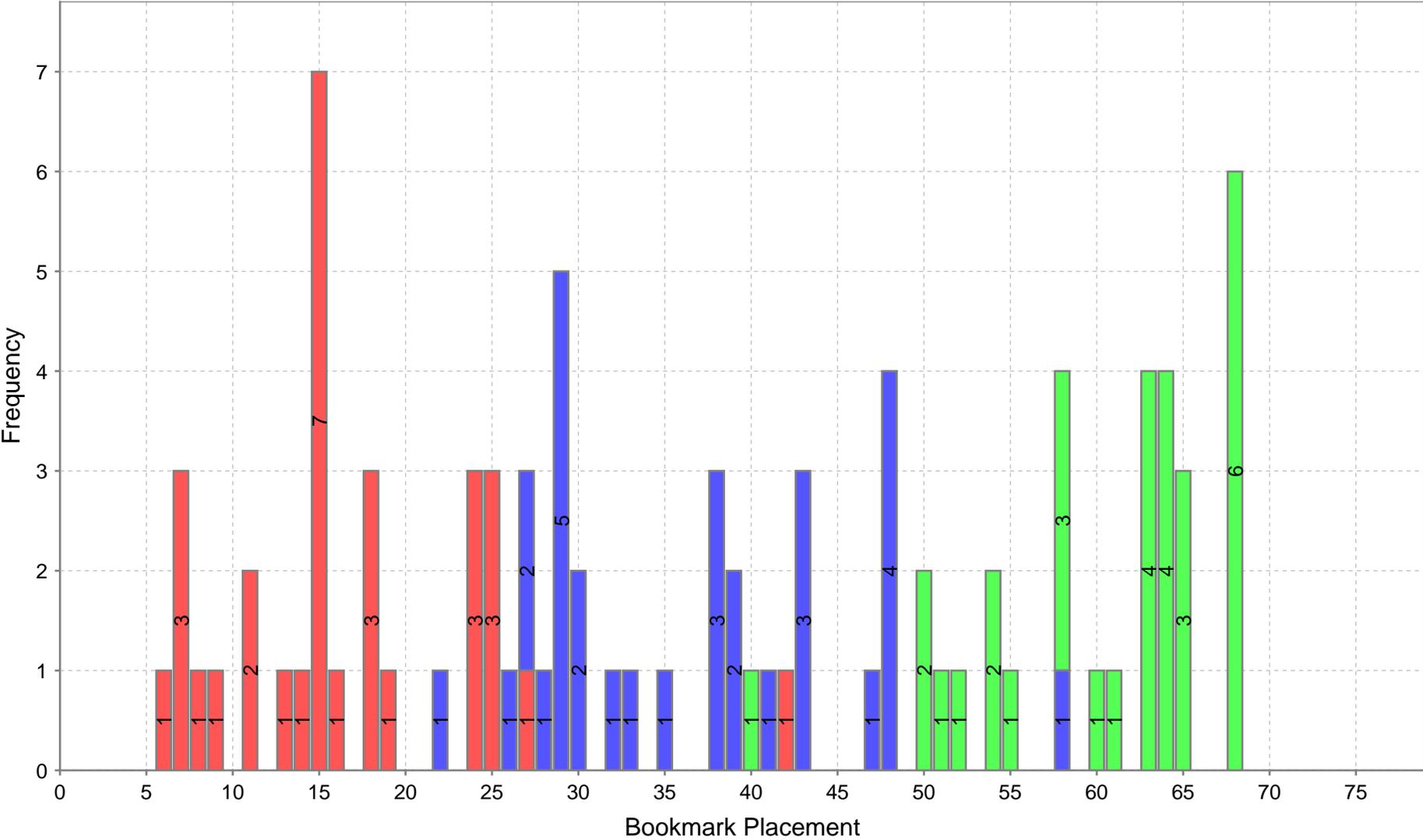
■ Level 2 (10-11-15)
 ■ Level 3 (27-27-33)
 ■ Level 4 (61-63-65)

Smarter Balanced In-Person ALS Grade 6 English Language Arts/Literacy Frequency of Bookmark Placements Round 1



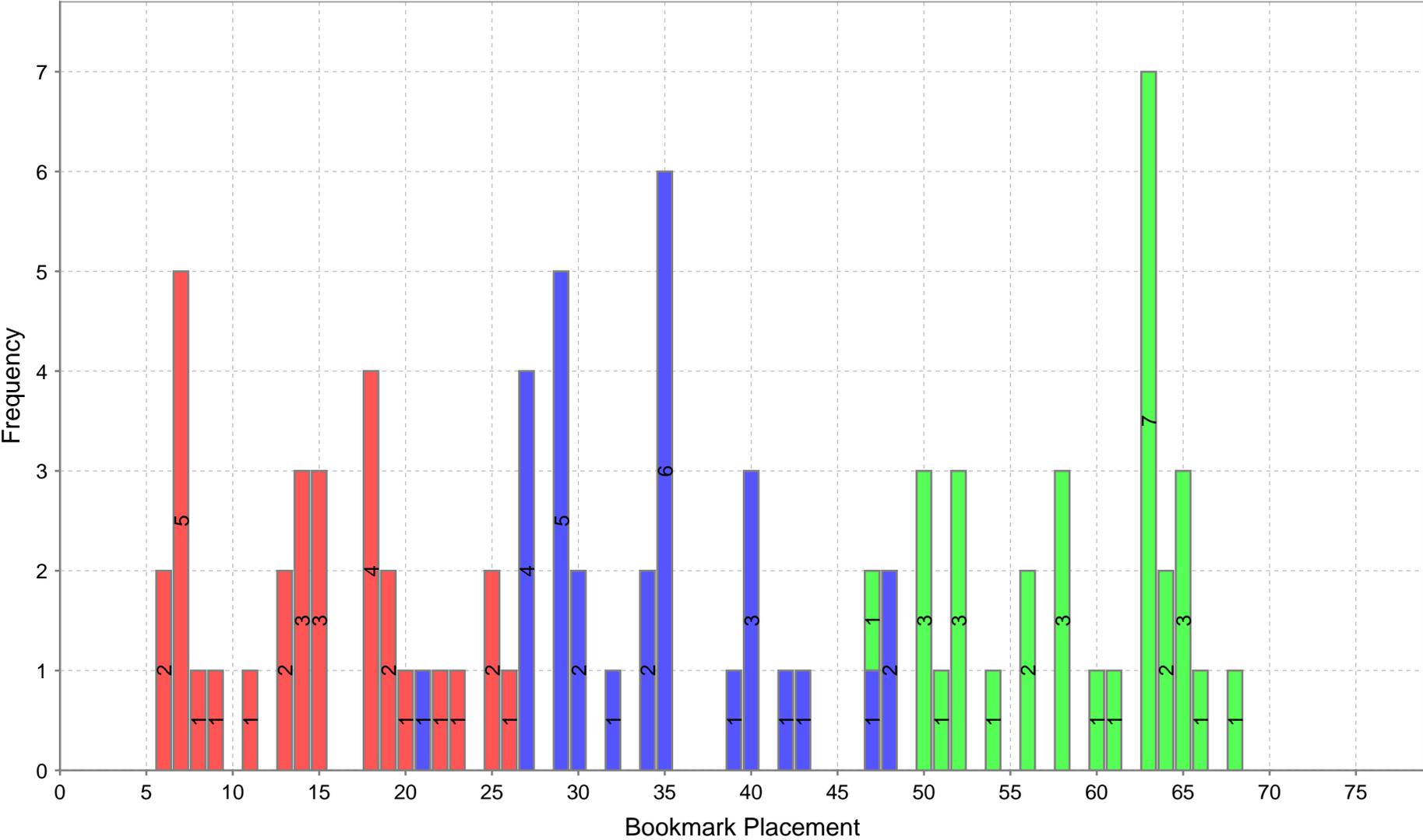
■ Level 2 (10.75-15-21.75)
 ■ Level 3 (27-35-46.25)
 ■ Level 4 (56.5-63-66.5)

Smarter Balanced In-Person ALS Grade 6 English Language Arts/Literacy Frequency of Bookmark Placements Round 2



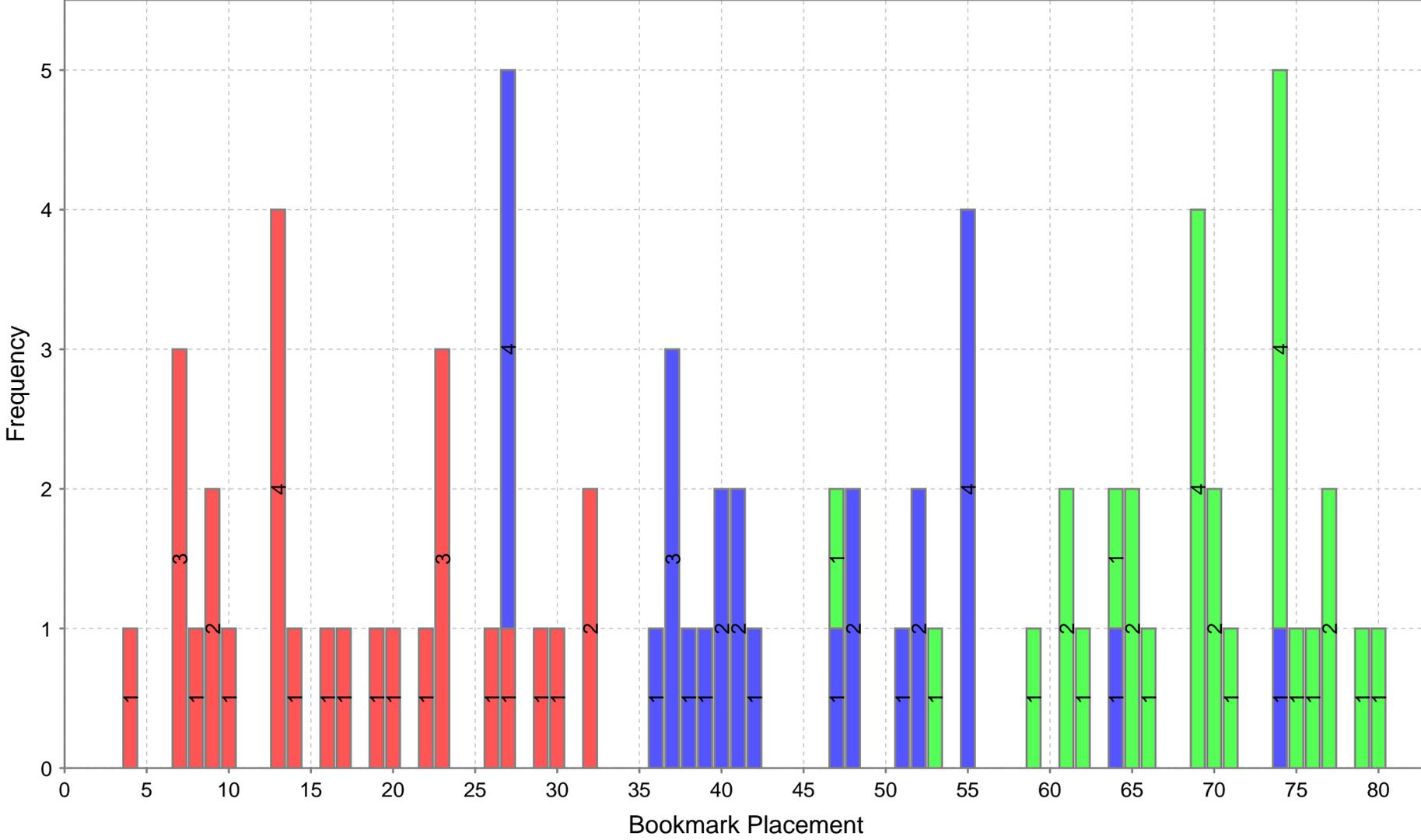
■ Level 2 (11-15-24)
 ■ Level 3 (29-36.5-43)
 ■ Level 4 (54.75-63-65)

Smarter Balanced In-Person ALS Grade 6 English Language Arts/Literacy Frequency of Bookmark Placements Round 3



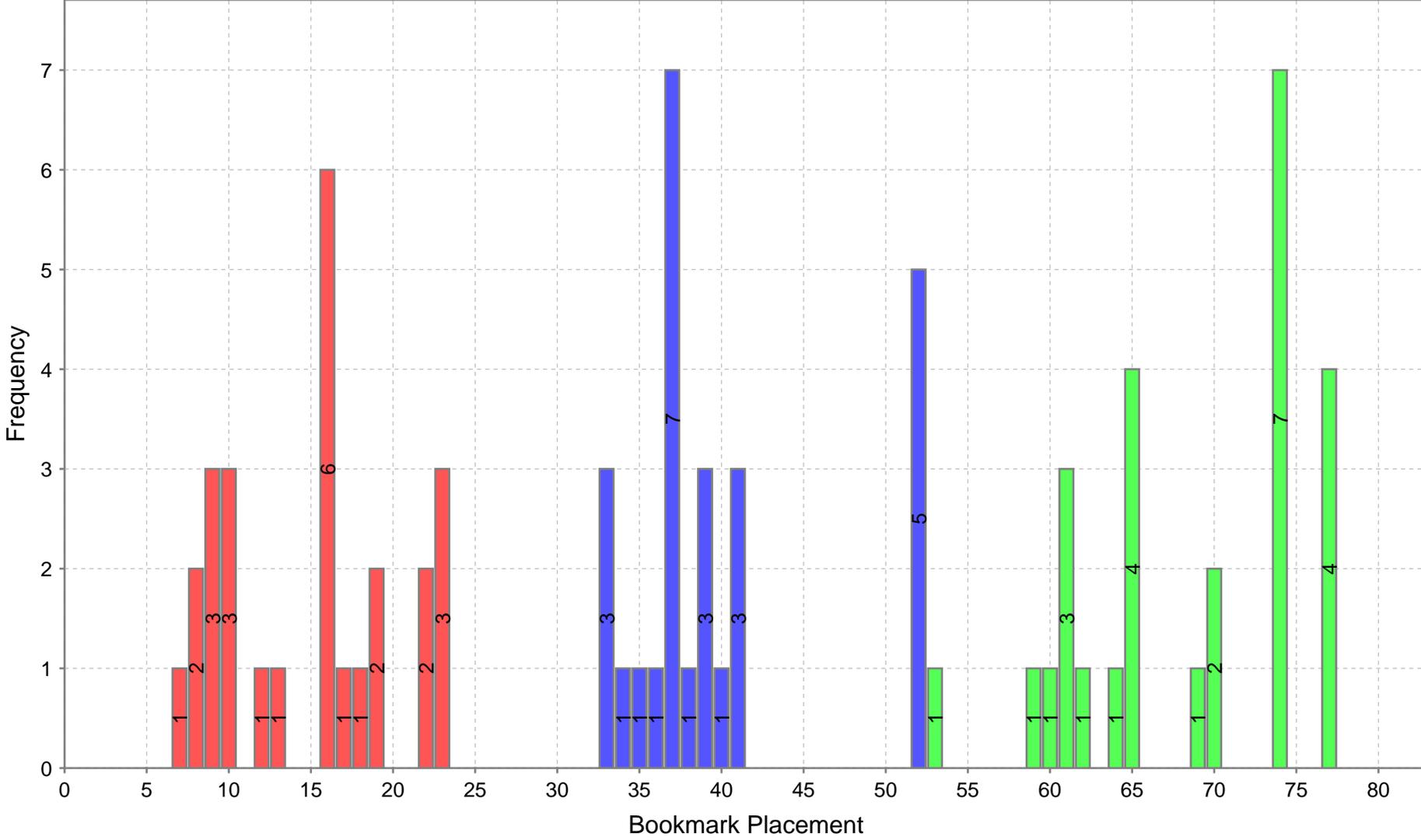
■ Level 2 (7.75-14.5-19)
 ■ Level 3 (29-34.5-40)
 ■ Level 4 (52-60.5-63.25)

Smarter Balanced In-Person ALS Grade 7 English Language Arts/Literacy Frequency of Bookmark Placements Round 1



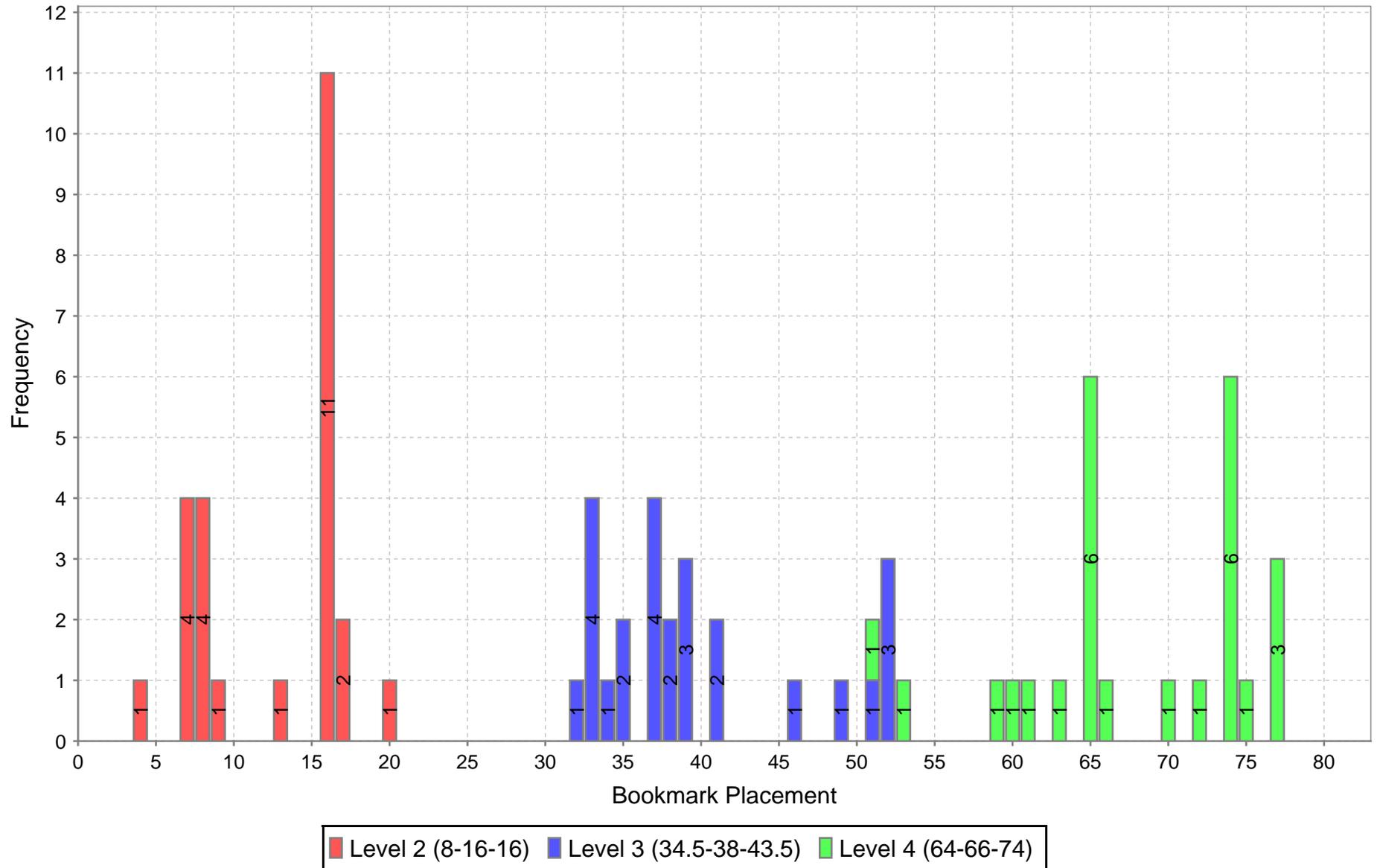
■ Level 2 (9-16-23)
 ■ Level 3 (37-41-52)
 ■ Level 4 (64-69-74)

Smarter Balanced In-Person ALS Grade 7 English Language Arts/Literacy Frequency of Bookmark Placements Round 2

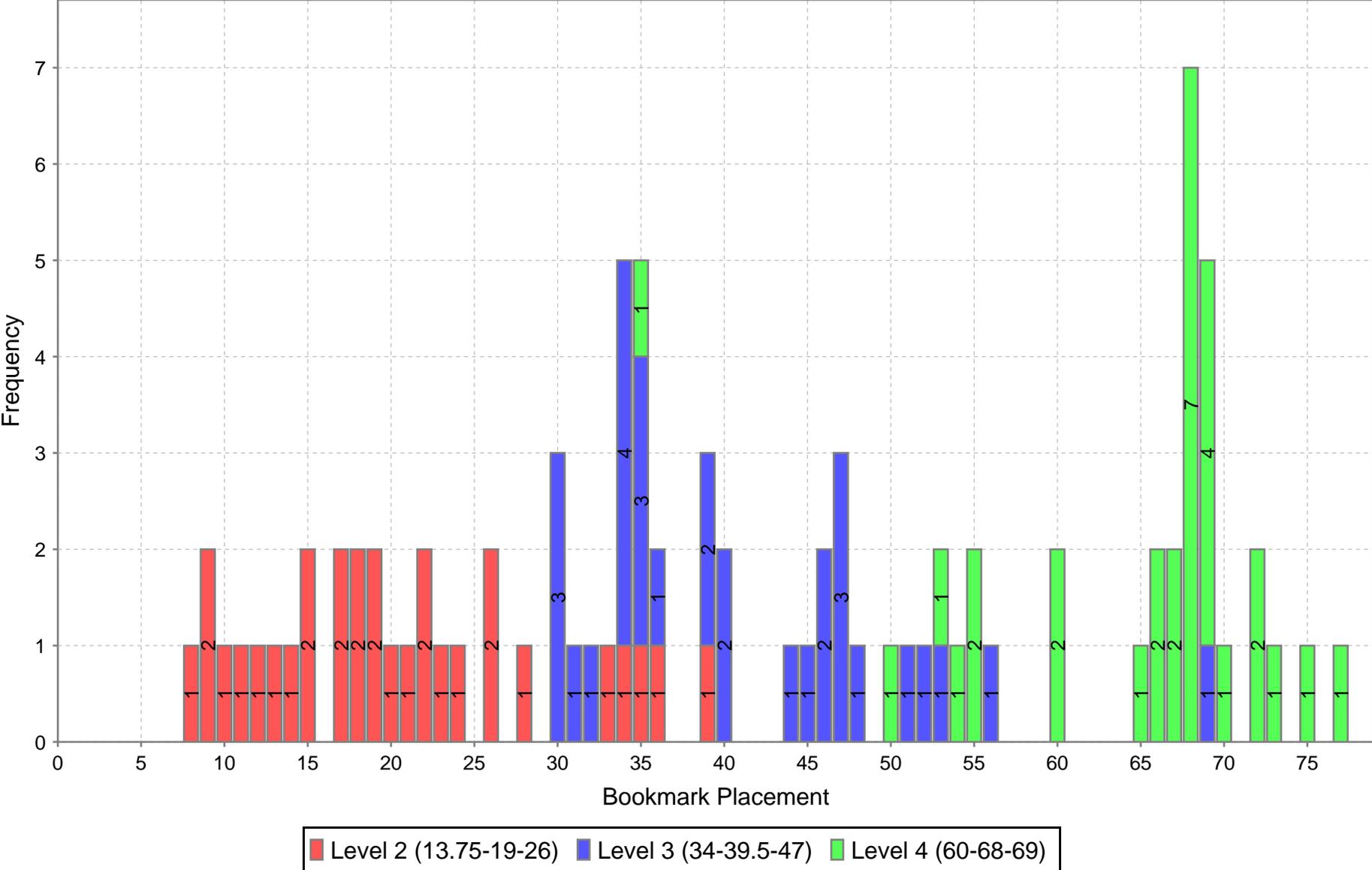


■ Level 2 (9.75-16-19)
 ■ Level 3 (36.75-37.5-41)
 ■ Level 4 (61.75-69.5-74)

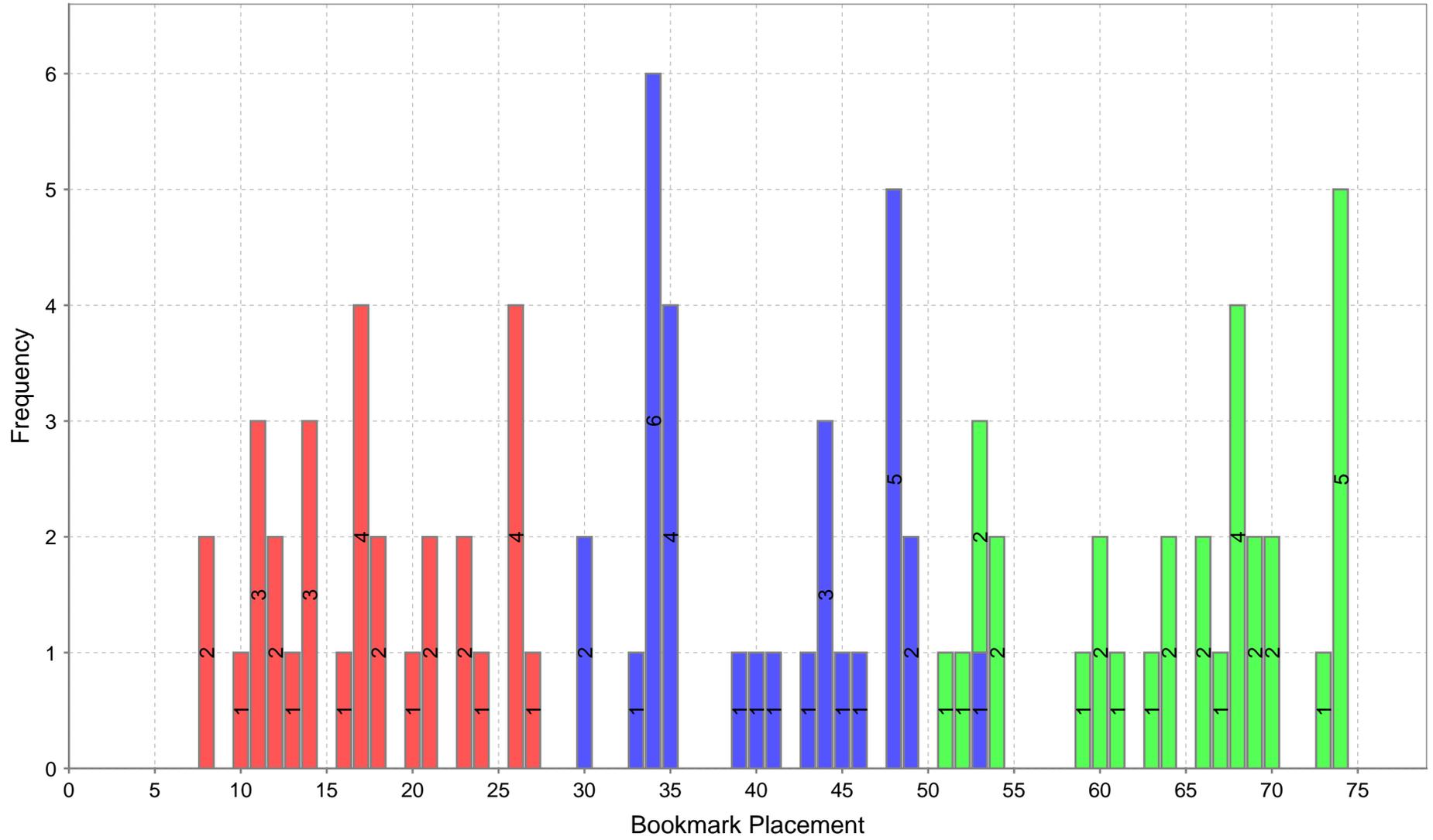
Smarter Balanced In-Person ALS Grade 7 English Language Arts/Literacy Frequency of Bookmark Placements Round 3



Smarter Balanced In-Person ALS Grade 8 English Language Arts/Literacy Frequency of Bookmark Placements Round 1

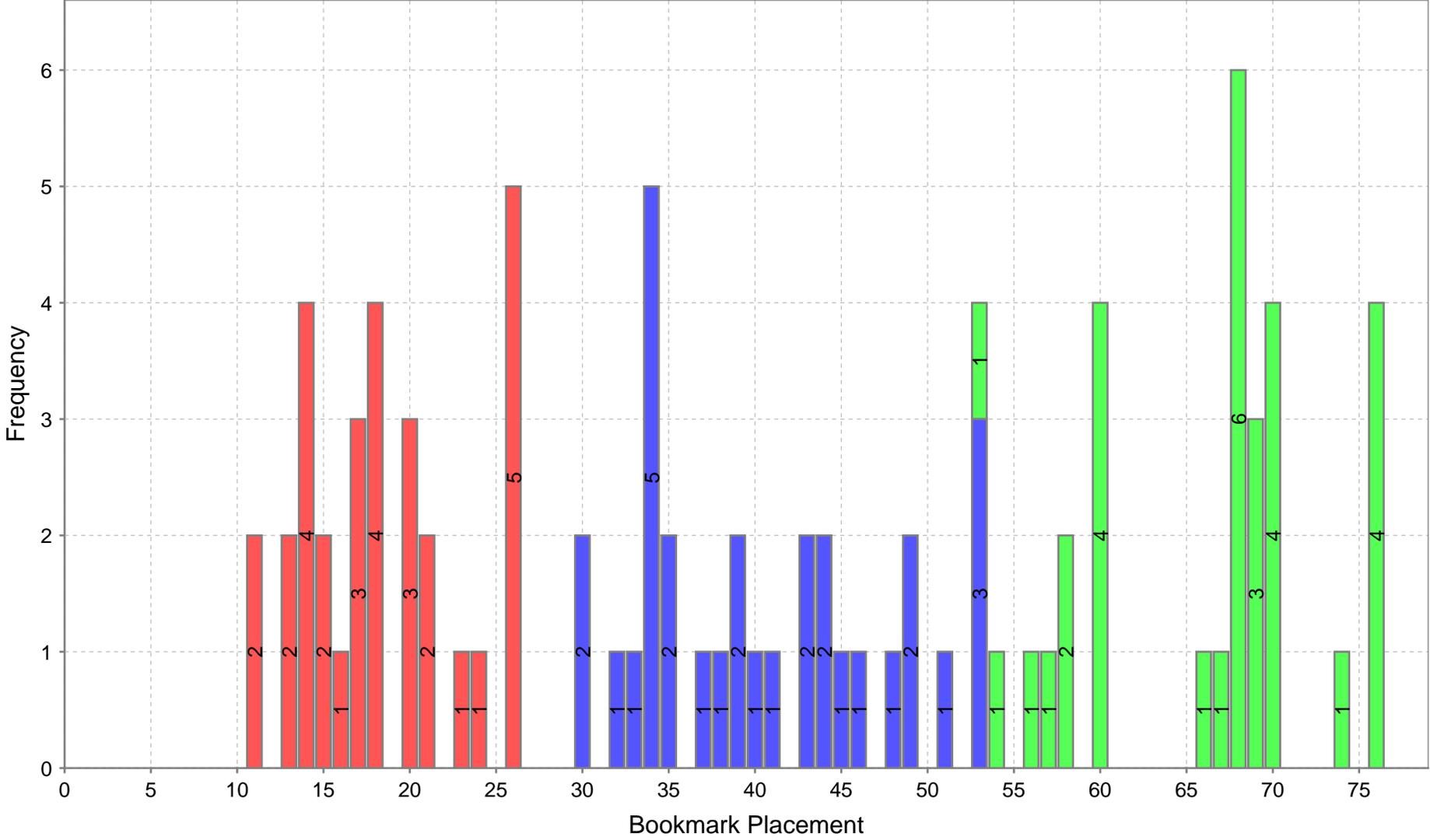


Smarter Balanced In-Person ALS Grade 8 English Language Arts/Literacy Frequency of Bookmark Placements Round 2



■ Level 2 (12-17-23)
 ■ Level 3 (34-40.5-48)
 ■ Level 4 (59.75-66.5-70)

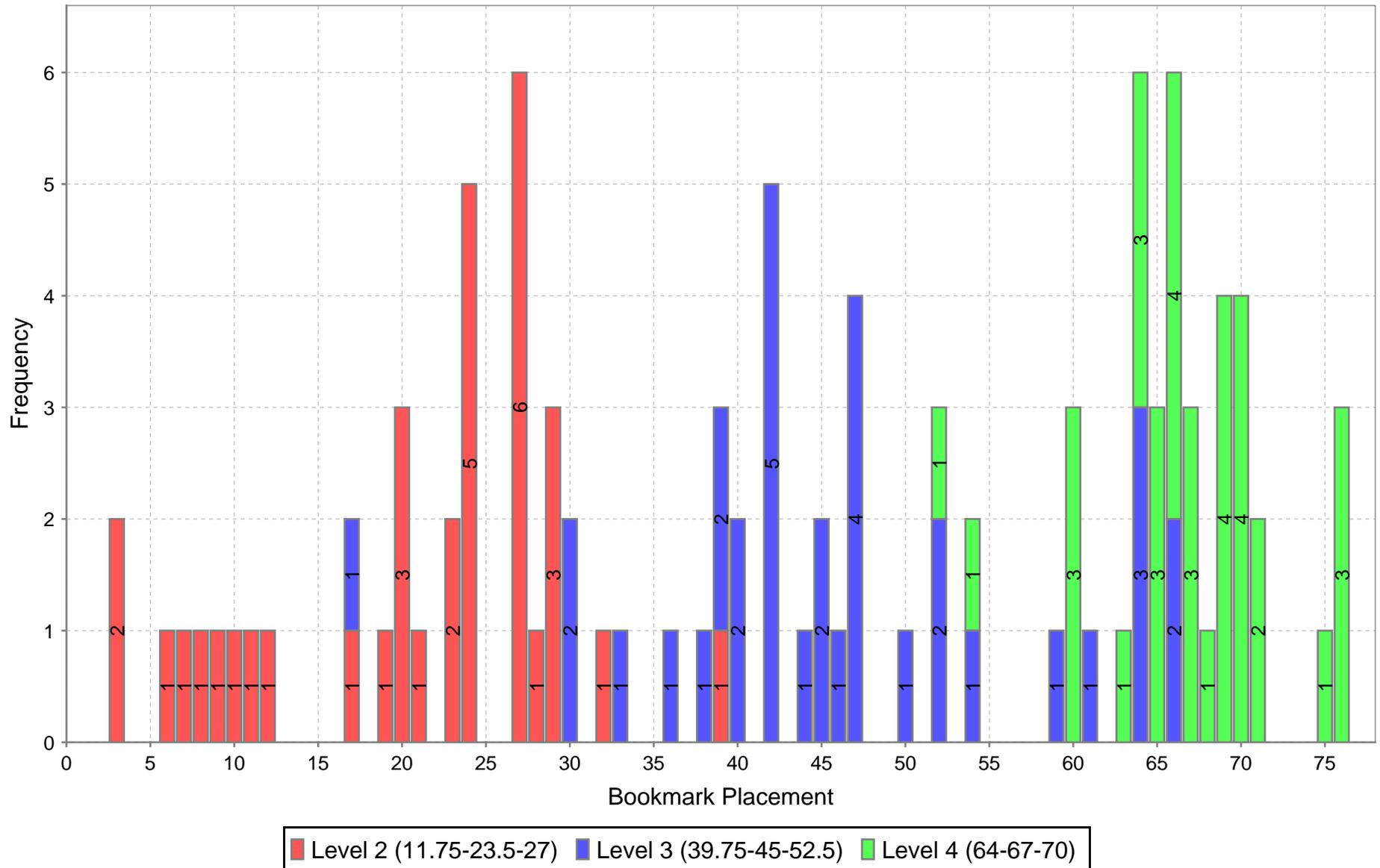
Smarter Balanced In-Person ALS Grade 8 English Language Arts/Literacy Frequency of Bookmark Placements Round 3



■ Level 2 (14-18-21.5)
 ■ Level 3 (34-39.5-46.5)
 ■ Level 4 (60-68-70)

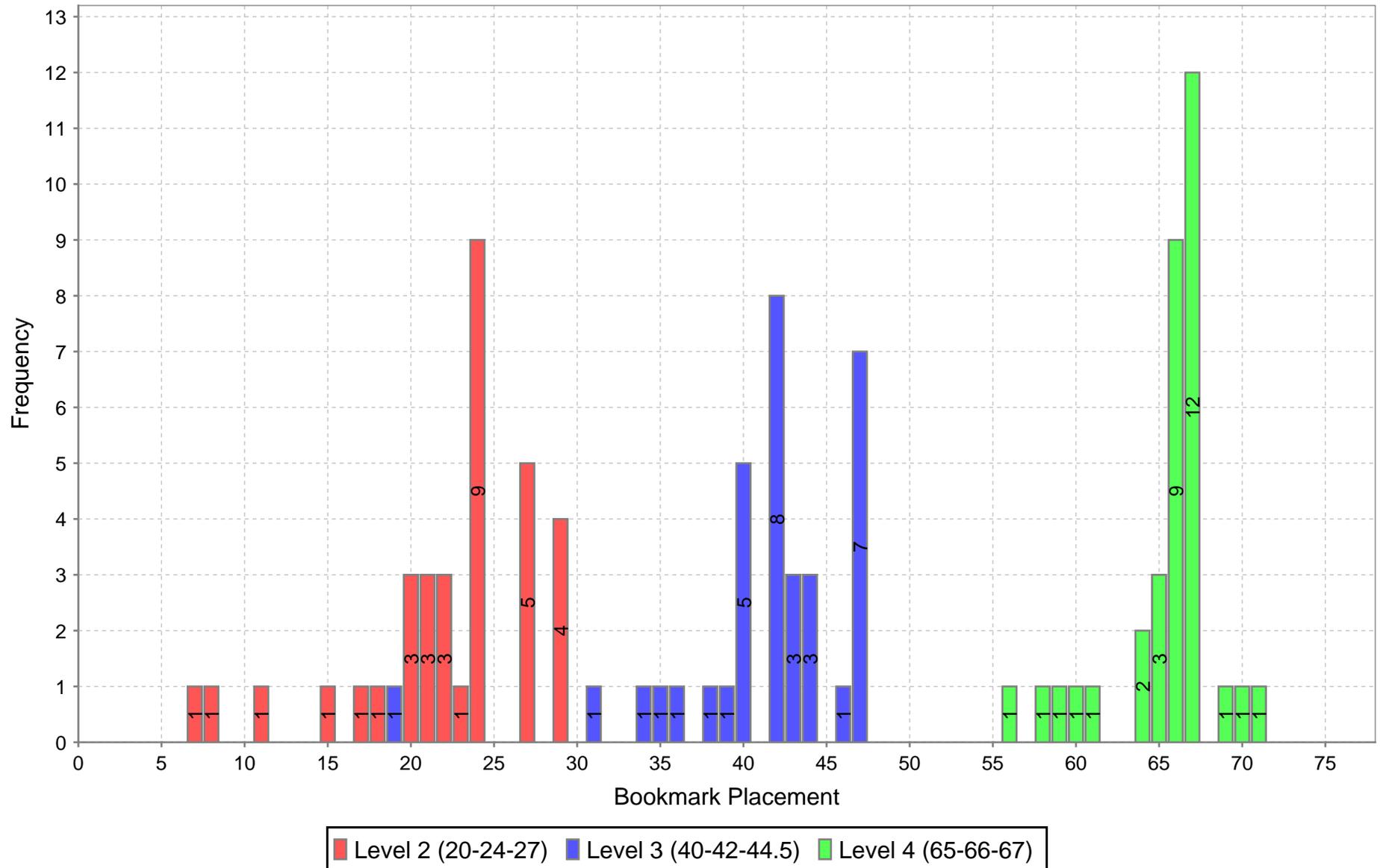
Smarter Balanced In-Person ALS Grade 11 English Language Arts/Literacy Panel A

Frequency of Bookmark Placements Round 1



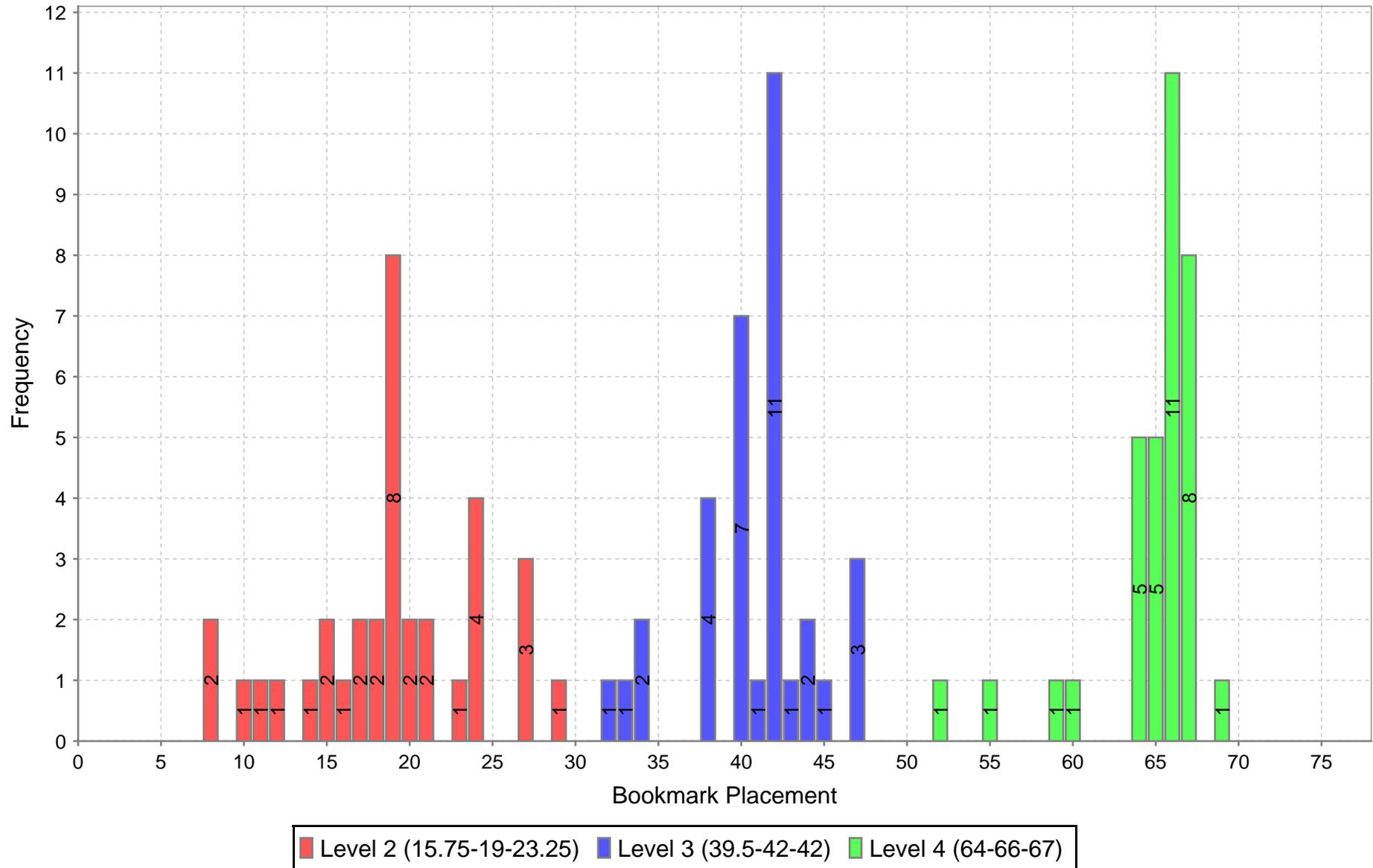
Smarter Balanced In-Person ALS Grade 11 English Language Arts/Literacy Panel A

Frequency of Bookmark Placements Round 2

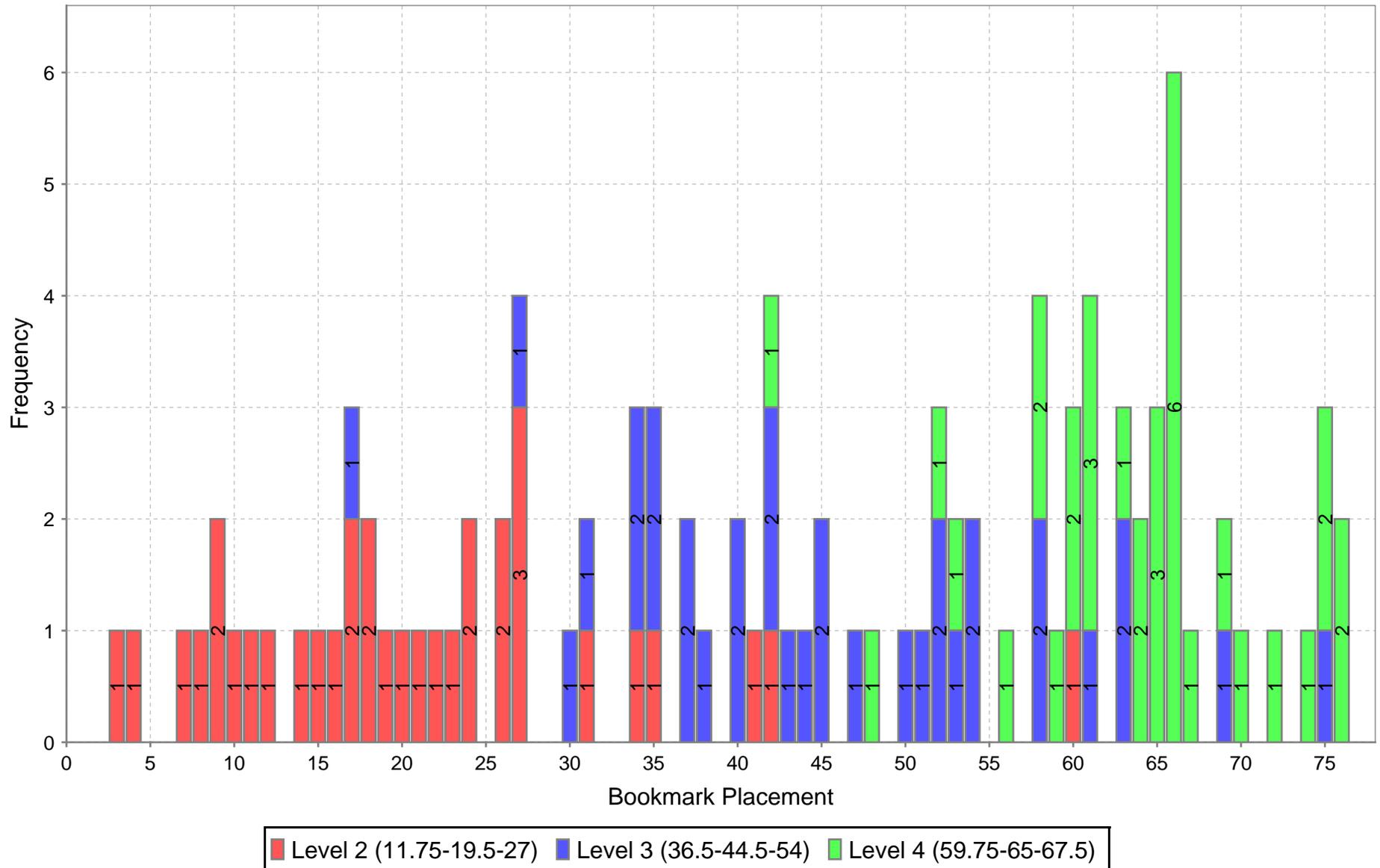


Smarter Balanced In-Person ALS Grade 11 English Language Arts/Literacy Panel A

Frequency of Bookmark Placements Round 3

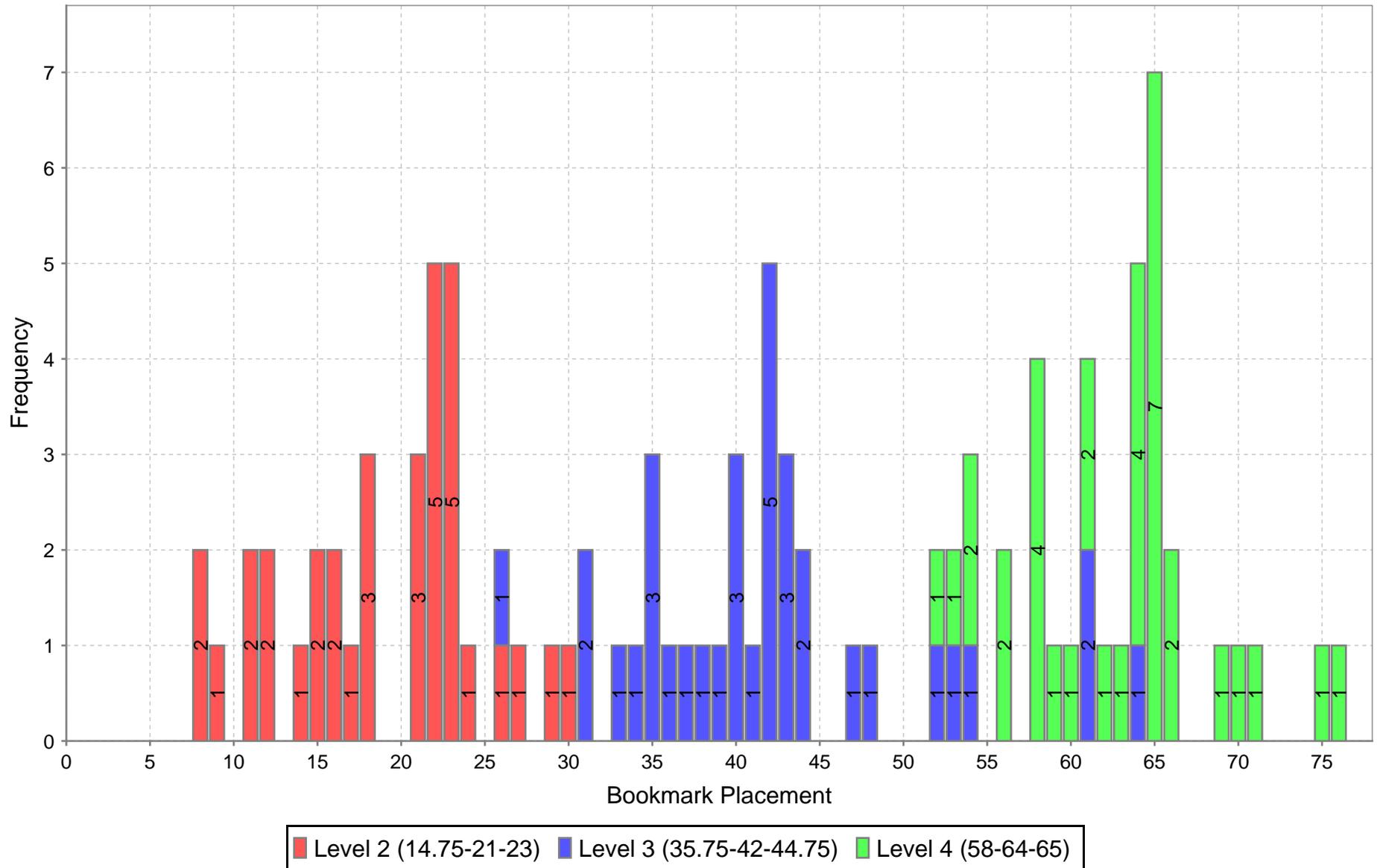


Smarter Balanced In-Person ALS Grade 11 English Language Arts/Literacy Panel B Frequency of Bookmark Placements Round 1

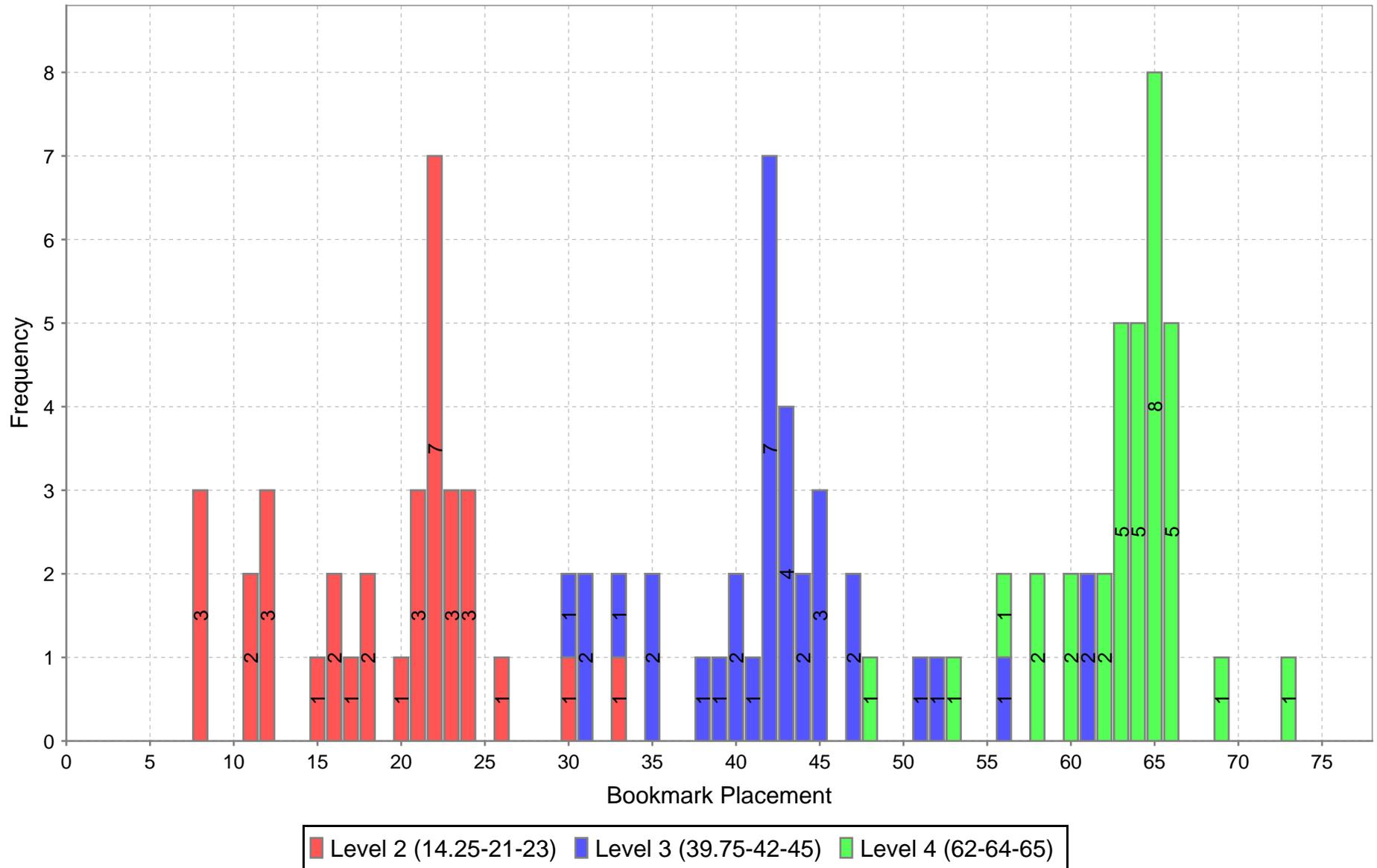


Smarter Balanced In-Person ALS Grade 11 English Language Arts/Literacy Panel B

Frequency of Bookmark Placements Round 2

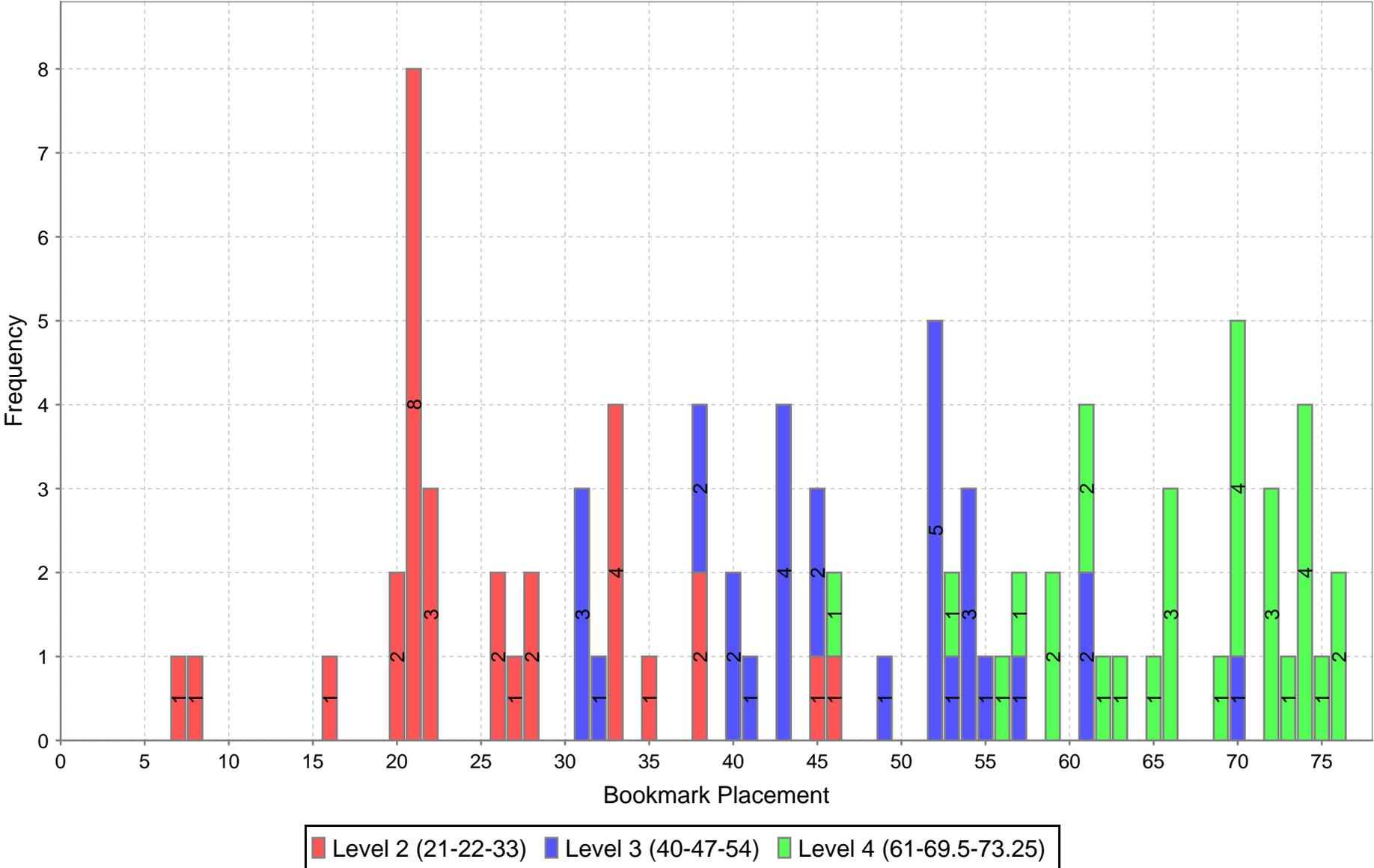


Smarter Balanced In-Person ALS Grade 11 English Language Arts/Literacy Panel B Frequency of Bookmark Placements Round 3



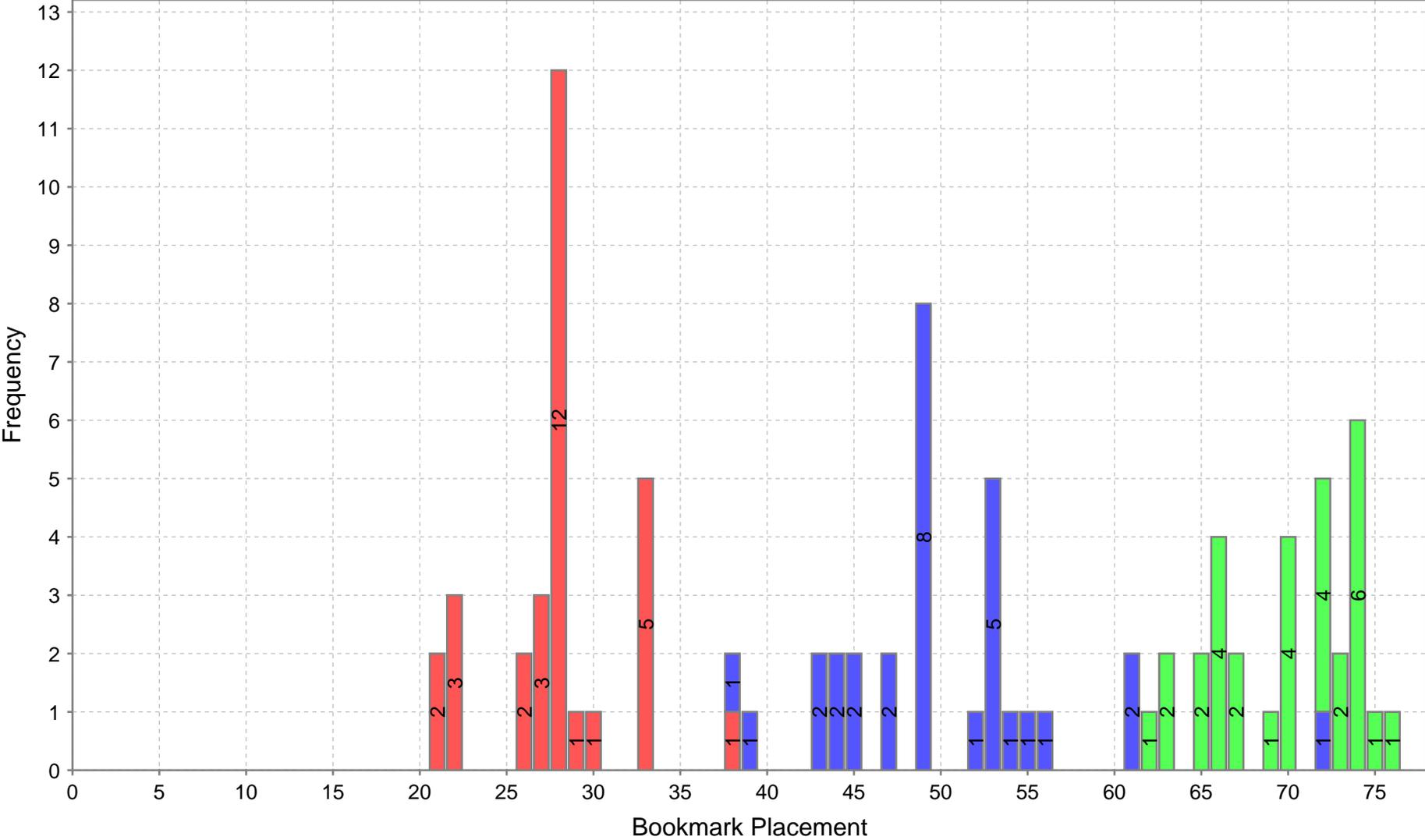
Smarter Balanced In-Person ALS Grade 3 Mathematics

Frequency of Bookmark Placements Round 1



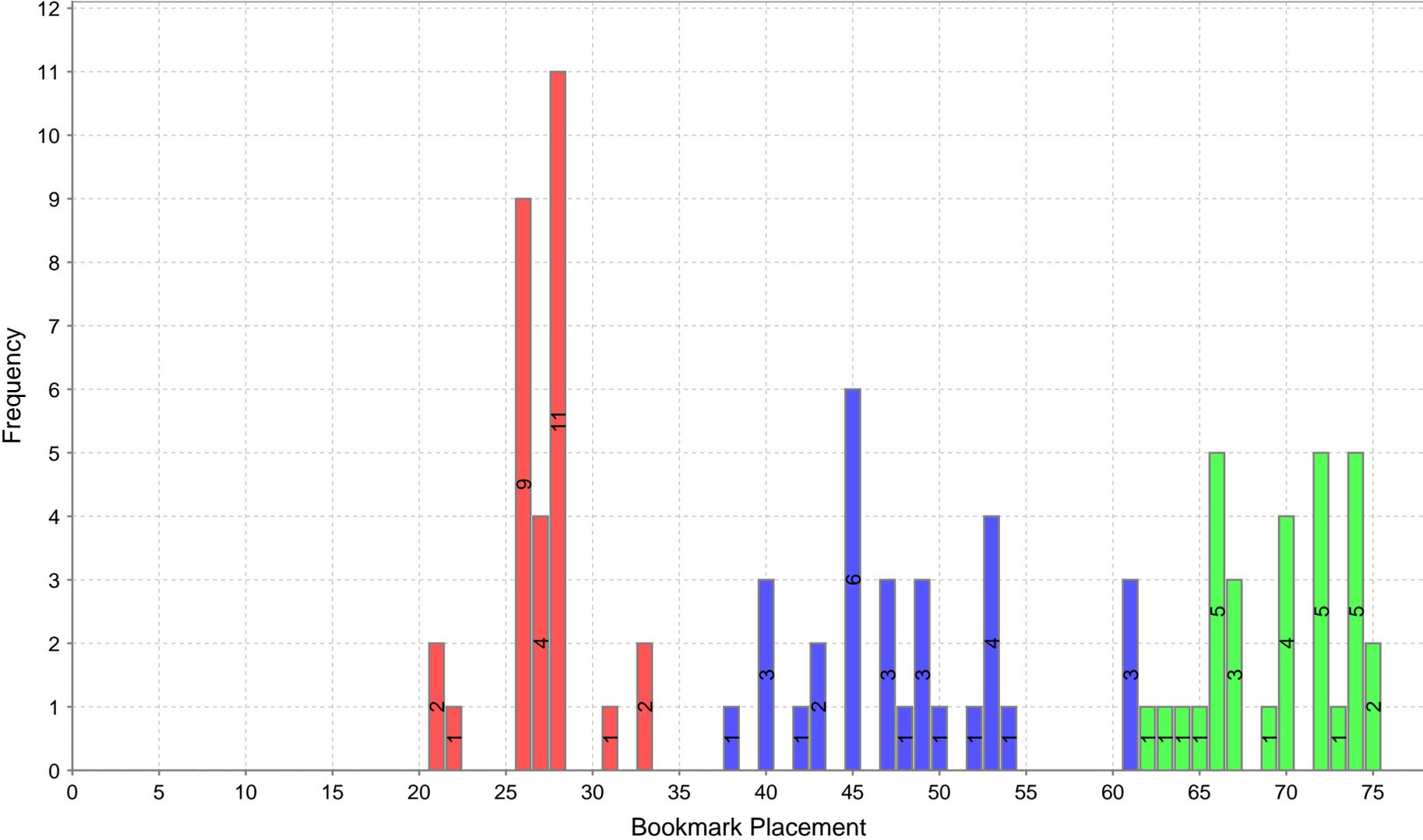
Smarter Balanced In-Person ALS Grade 3 Mathematics

Frequency of Bookmark Placements Round 2



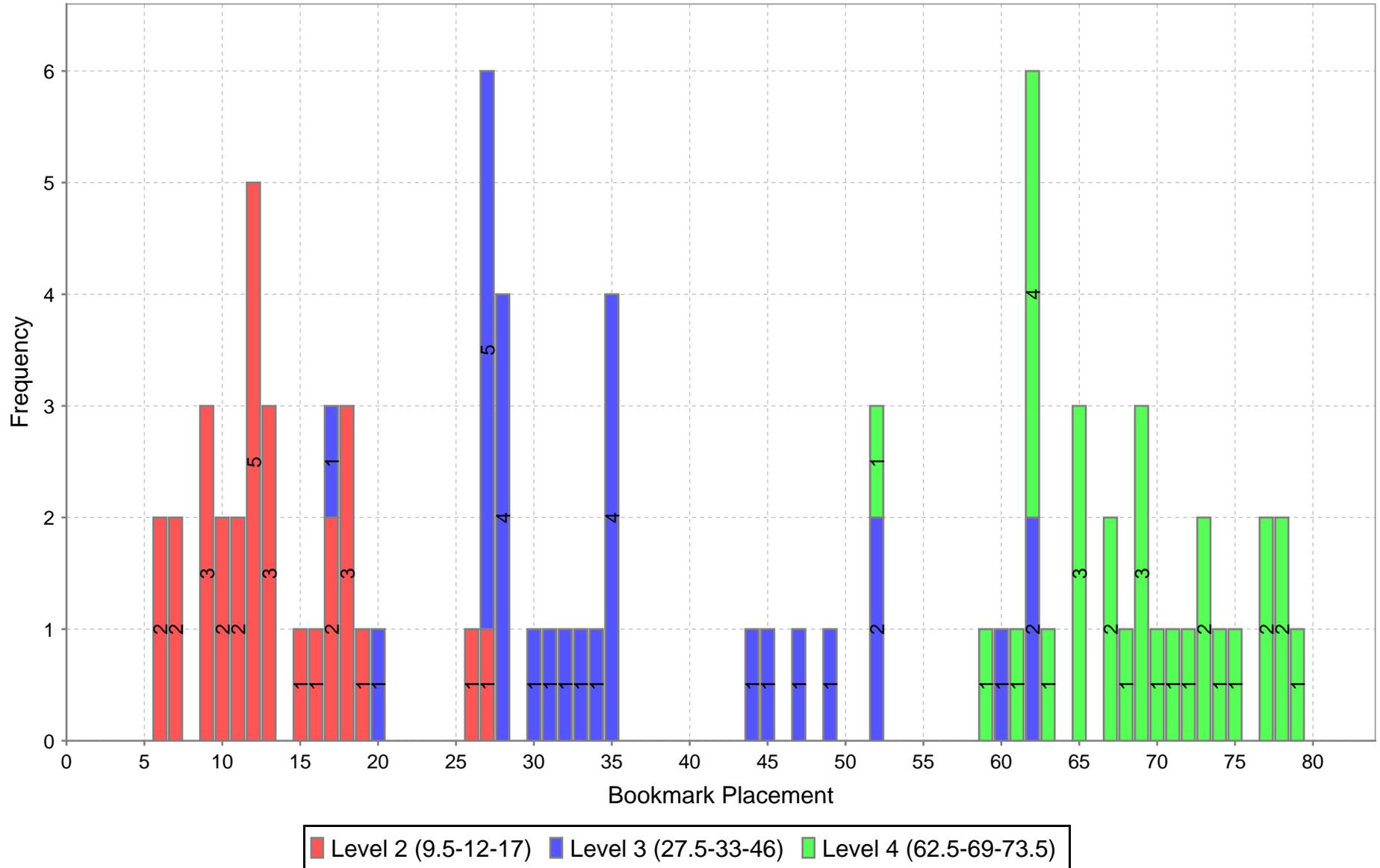
■ Level 2 (26.75-28-29.25)
 ■ Level 3 (45-49-53)
 ■ Level 4 (66-70-74)

Smarter Balanced In-Person ALS Grade 3 Mathematics Frequency of Bookmark Placements Round 3

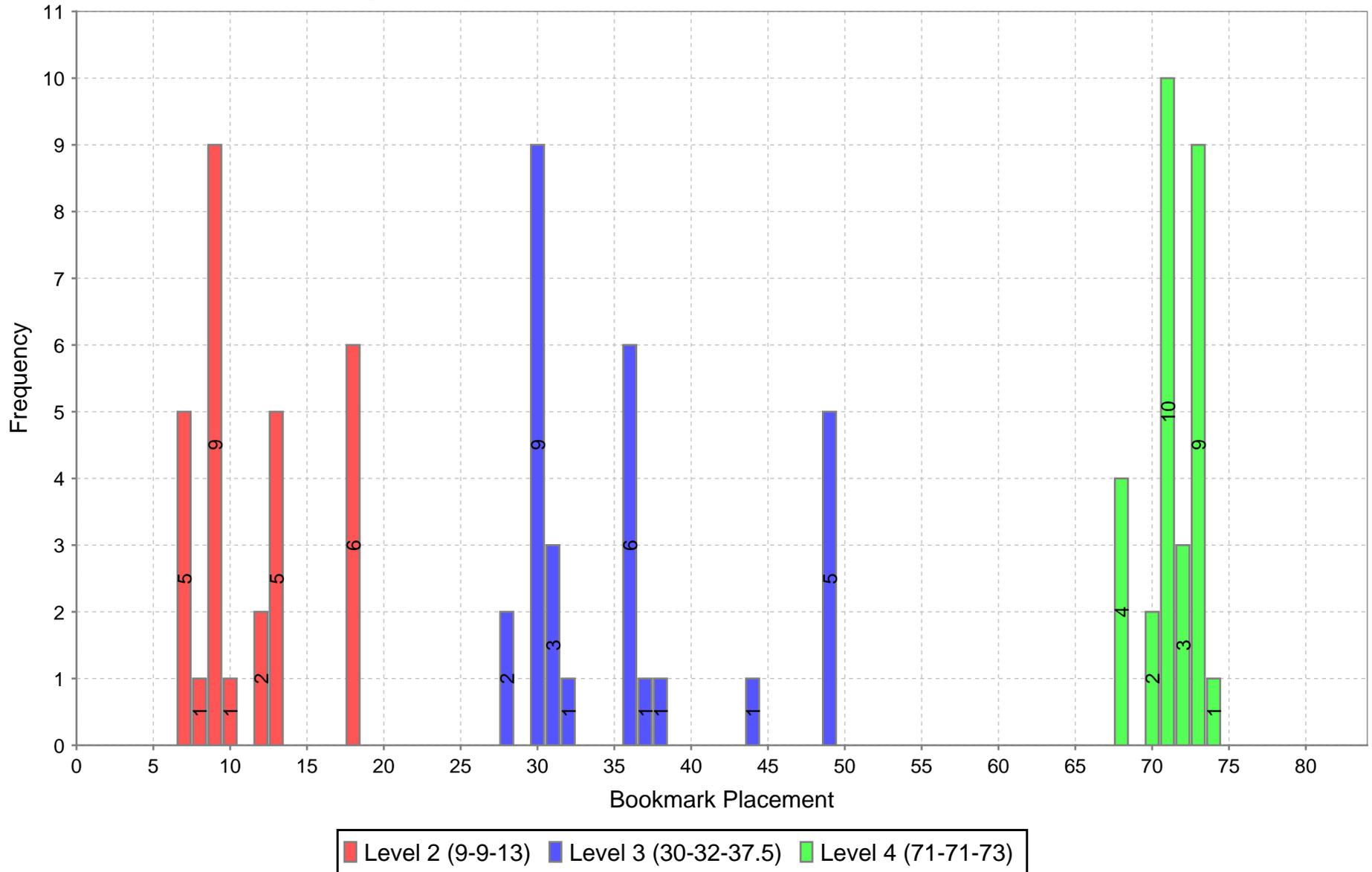


■ Level 2 (26-27-28)
 ■ Level 3 (44.5-47-53)
 ■ Level 4 (66-70-73.25)

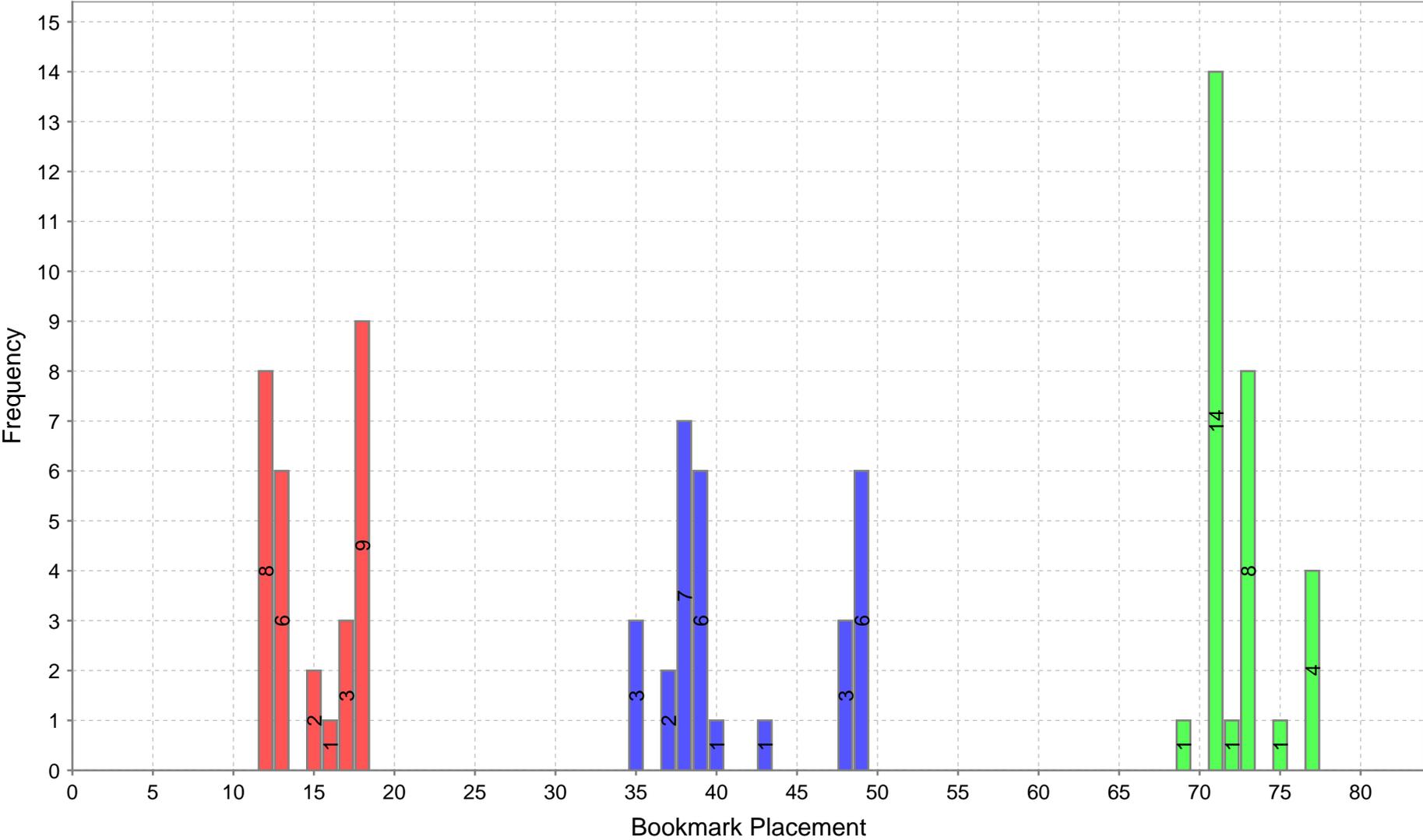
Smarter Balanced In-Person ALS Grade 4 Mathematics Frequency of Bookmark Placements Round 1



Smarter Balanced In-Person ALS Grade 4 Mathematics Frequency of Bookmark Placements Round 2



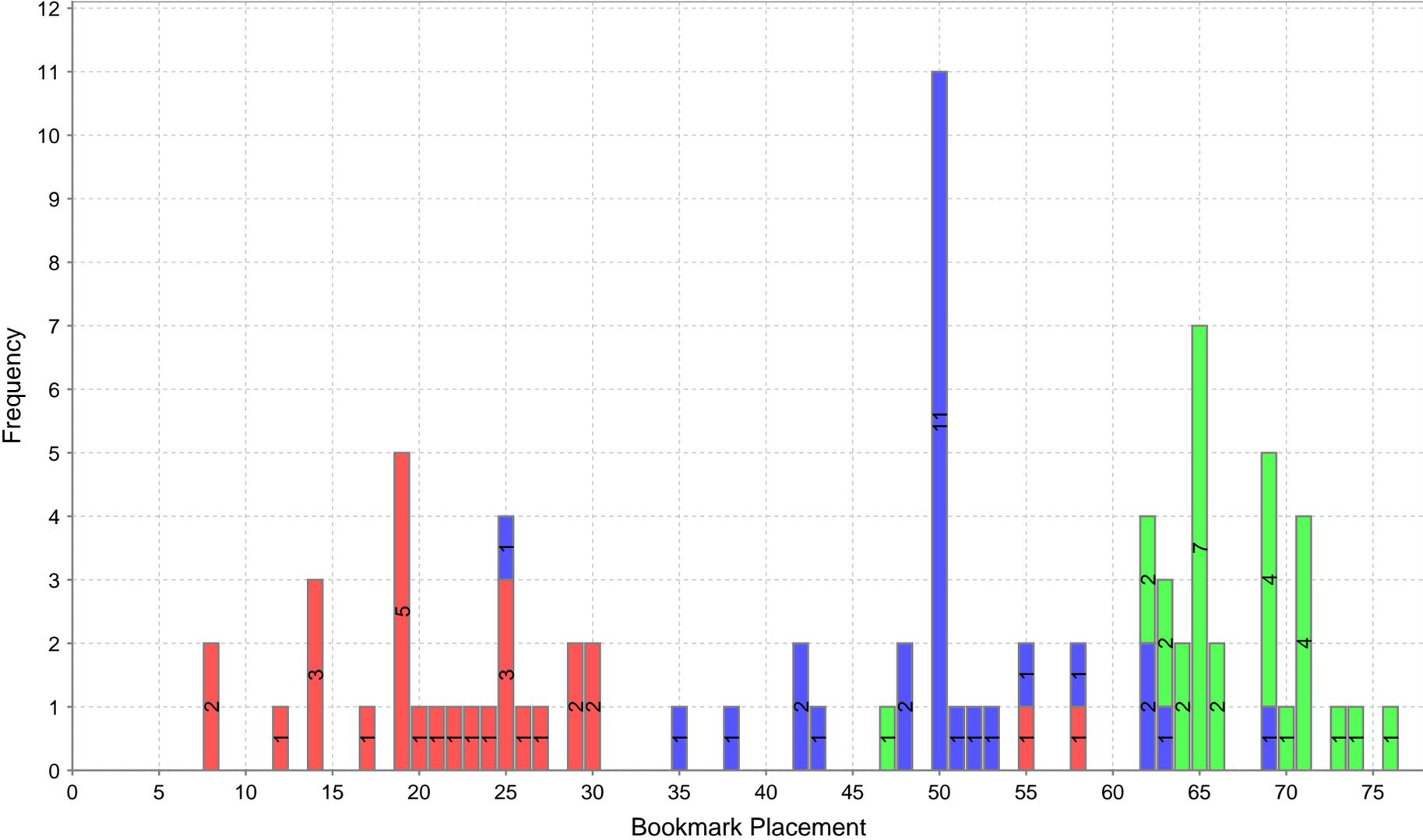
Smarter Balanced In-Person ALS Grade 4 Mathematics Frequency of Bookmark Placements Round 3



■ Level 2 (12-15-18)
 ■ Level 3 (38-39-48)
 ■ Level 4 (71-71-73)

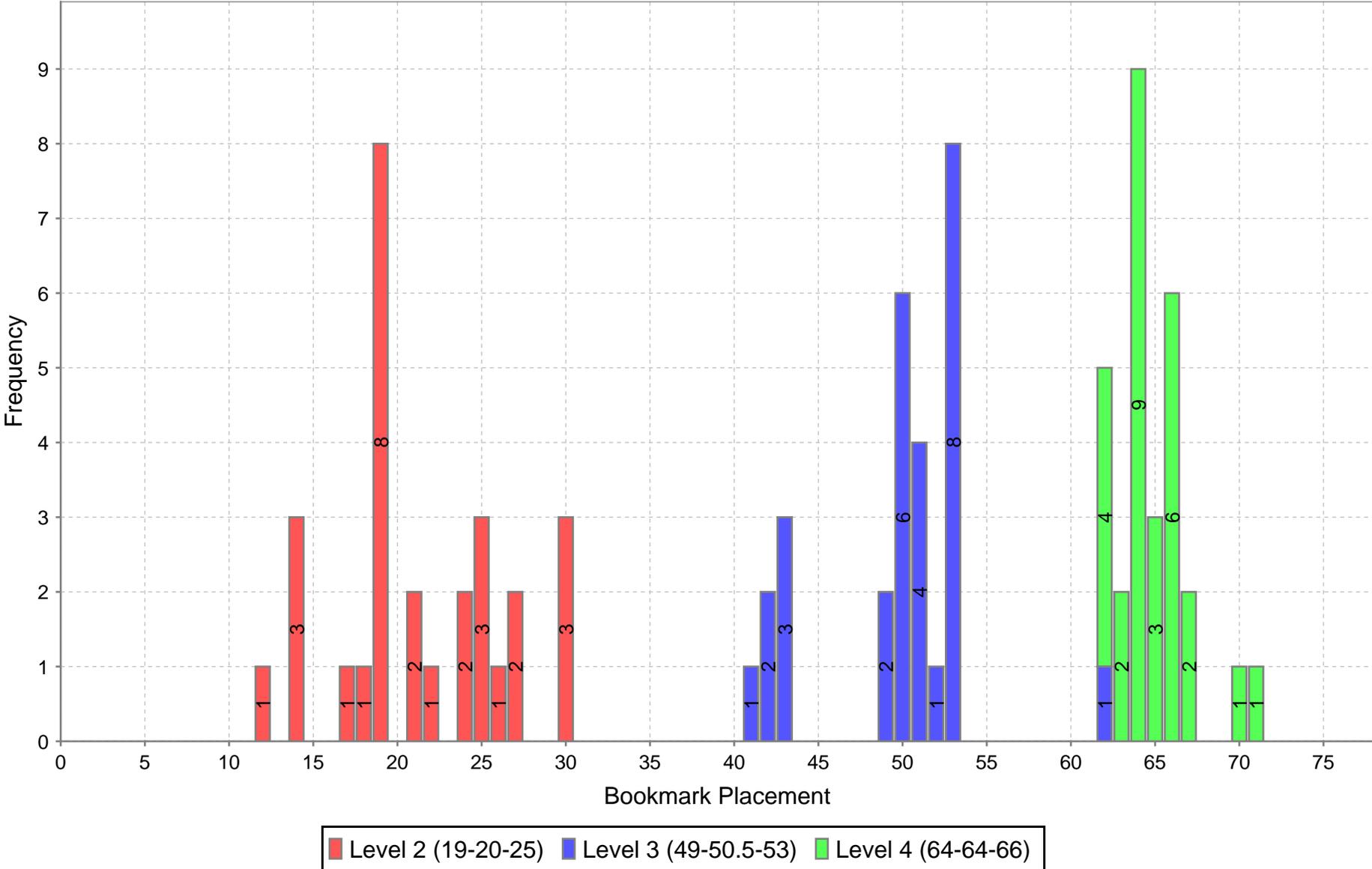
Smarter Balanced In-Person ALS Grade 5 Mathematics

Frequency of Bookmark Placements Round 1



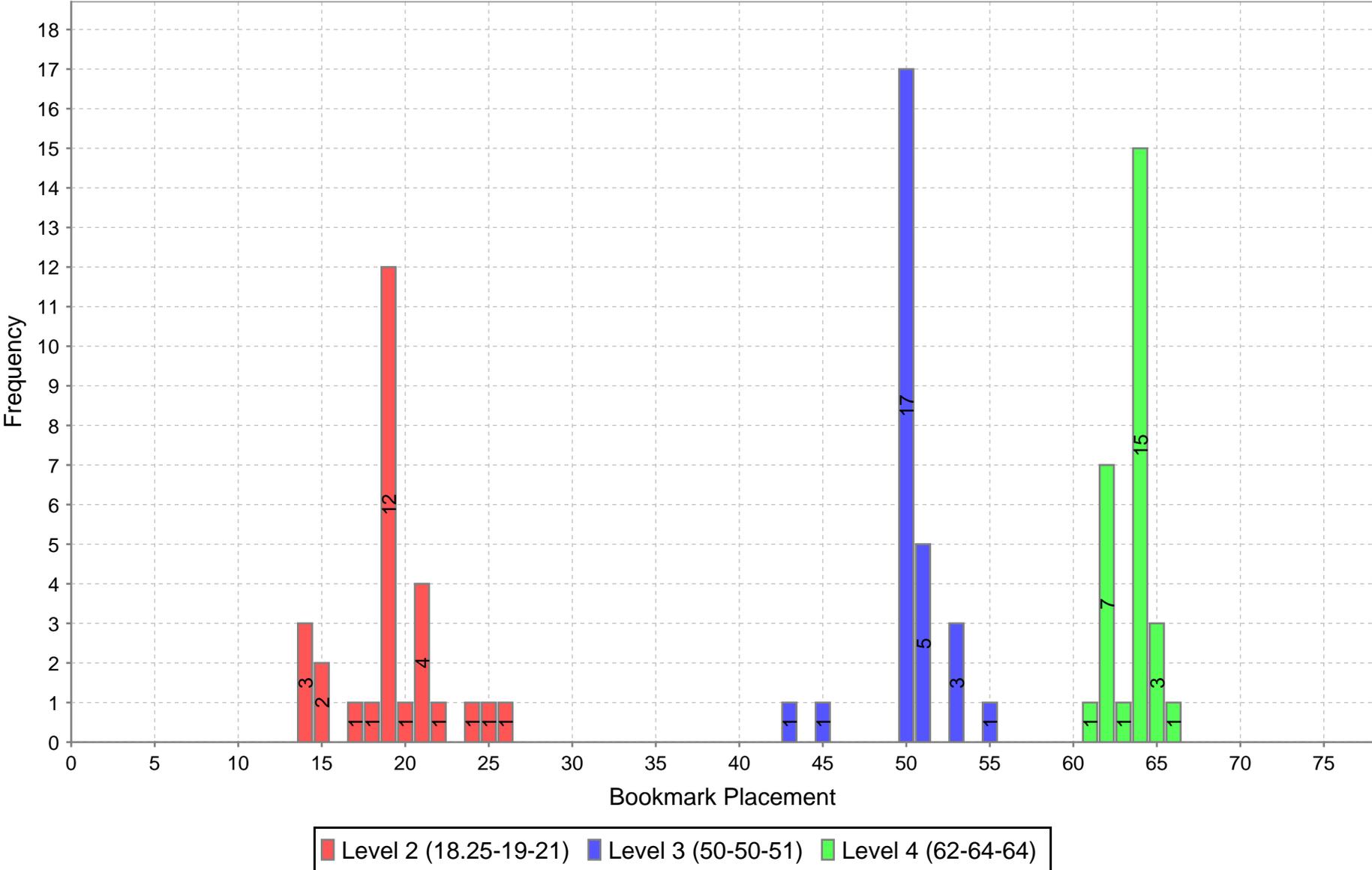
■ Level 2 (17.5-21.5-26.75)
 ■ Level 3 (48-50-52.75)
 ■ Level 4 (64.25-65.5-70.75)

Smarter Balanced In-Person ALS Grade 5 Mathematics Frequency of Bookmark Placements Round 2



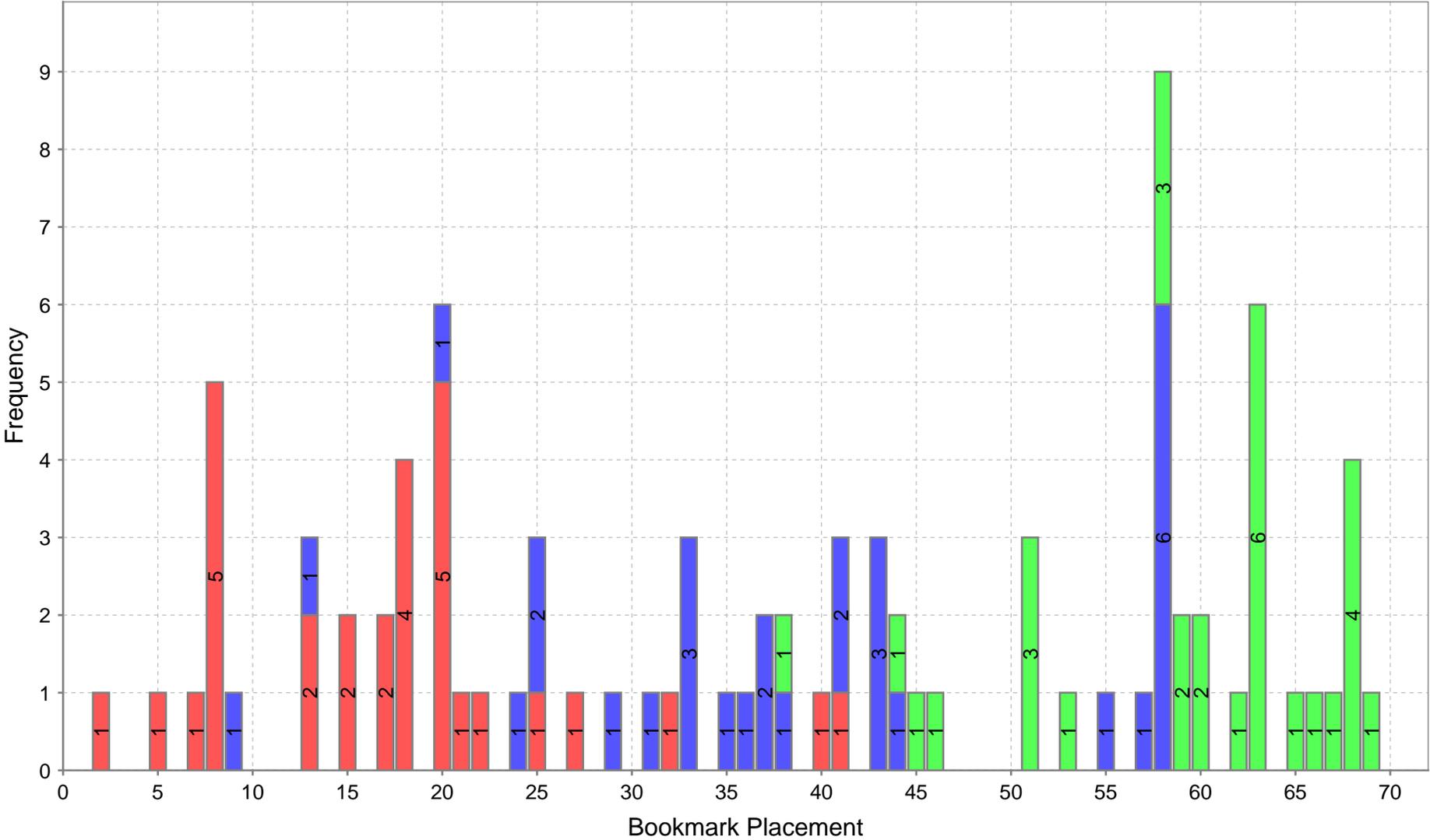
Smarter Balanced In-Person ALS Grade 5 Mathematics

Frequency of Bookmark Placements Round 3



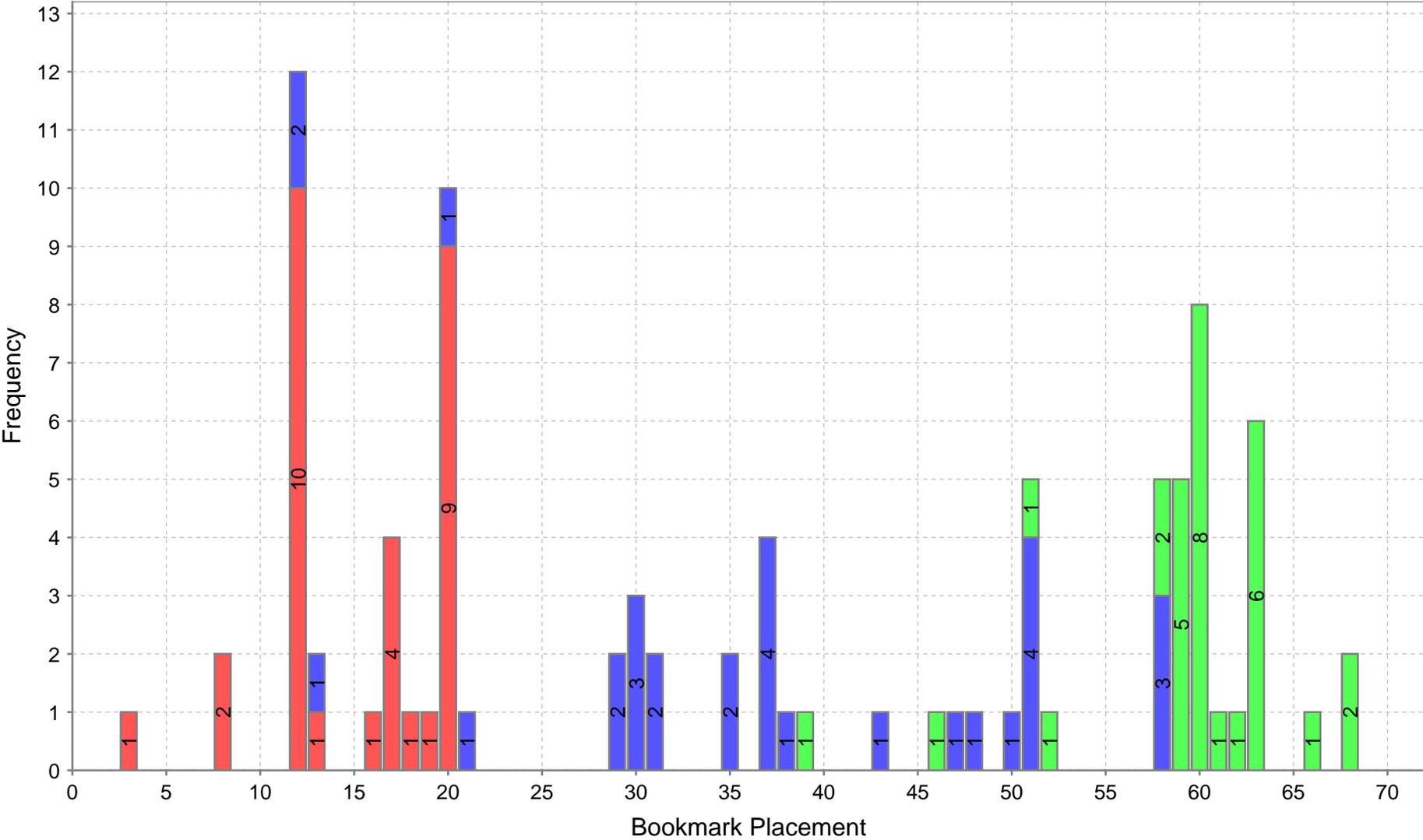
Smarter Balanced In-Person ALS Grade 6 Mathematics

Frequency of Bookmark Placements Round 1



■ Level 2 (8-18-20.25)
 ■ Level 3 (30.5-37.5-55.5)
 ■ Level 4 (52.5-61-65.25)

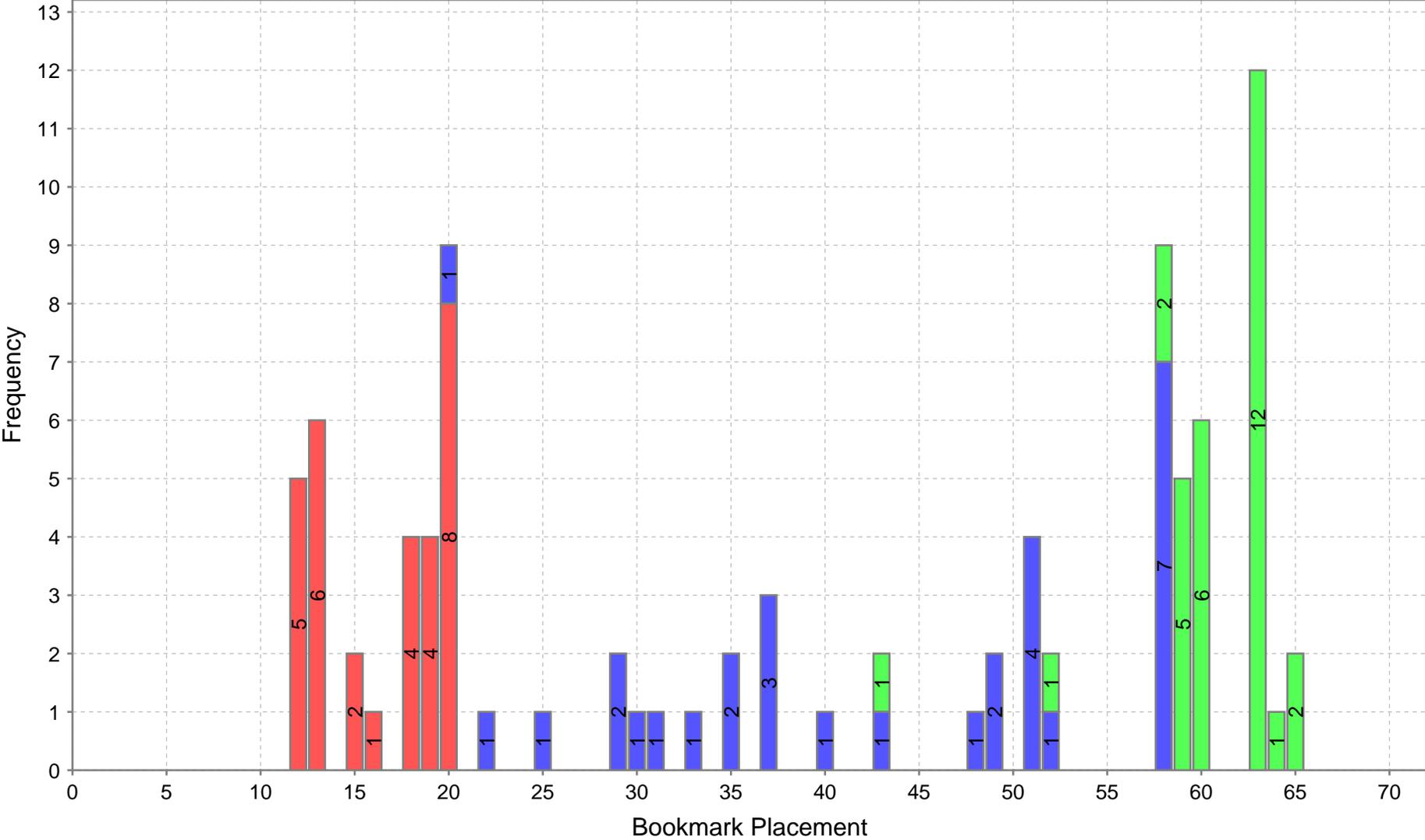
Smarter Balanced In-Person ALS Grade 6 Mathematics Frequency of Bookmark Placements Round 2



■ Level 2 (12-16.5-20)
 ■ Level 3 (29.75-37-50.25)
 ■ Level 4 (59-60-63)

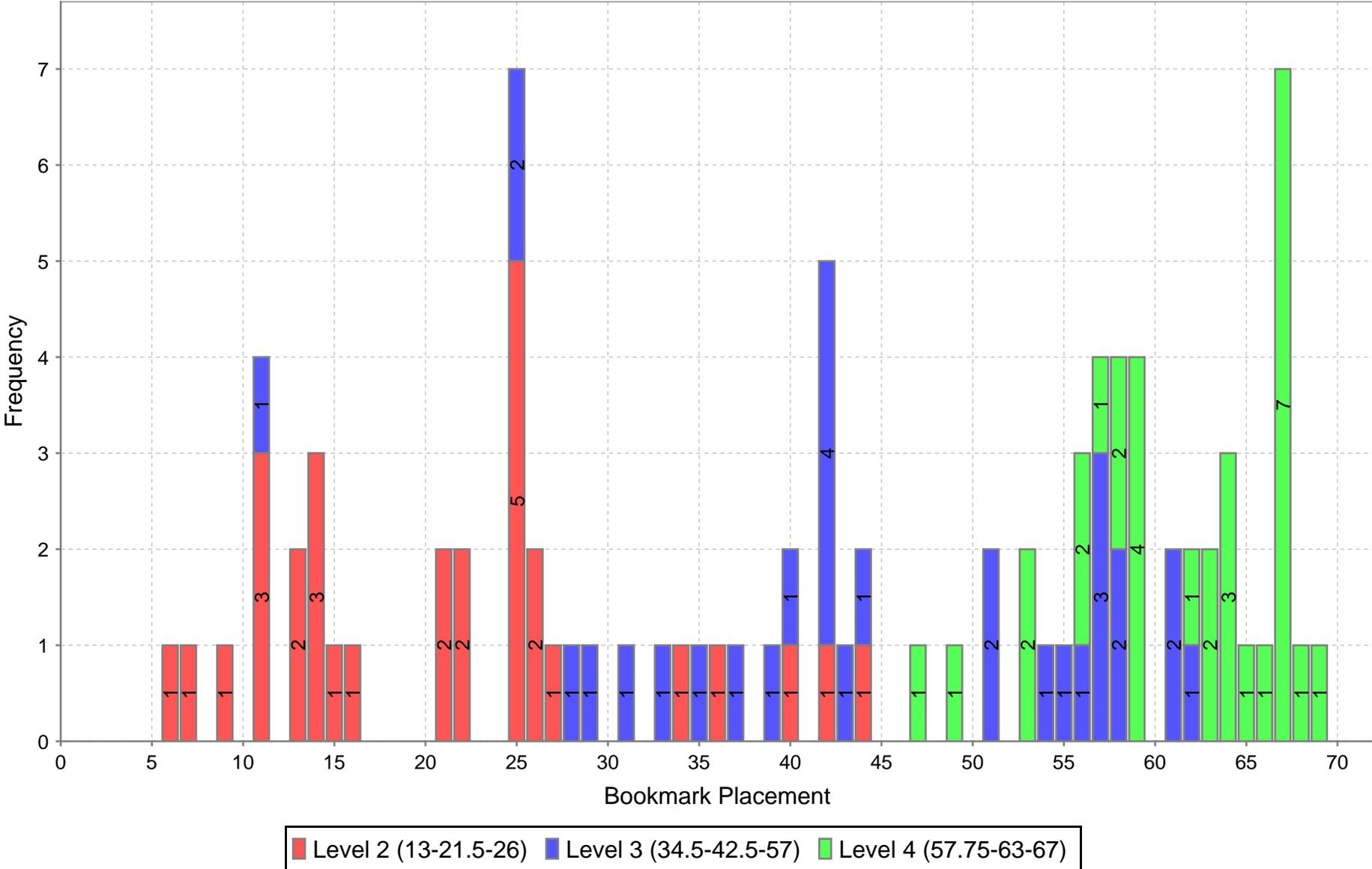
Smarter Balanced In-Person ALS Grade 6 Mathematics

Frequency of Bookmark Placements Round 3



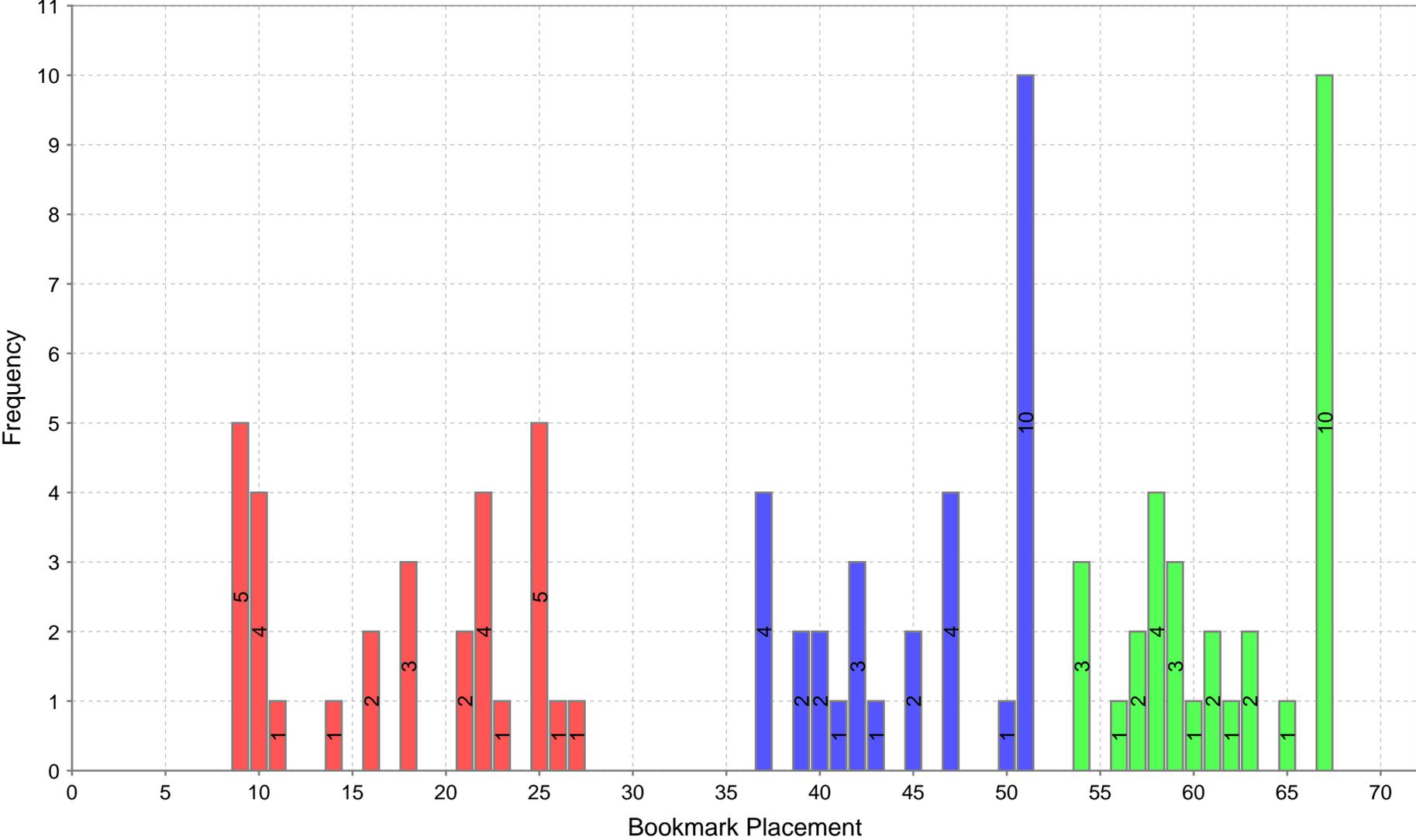
■ Level 2 (13-18-20)
 ■ Level 3 (32.5-45.5-53.5)
 ■ Level 4 (59-61.5-63)

Smarter Balanced In-Person ALS Grade 7 Mathematics Frequency of Bookmark Placements Round 1



Smarter Balanced In-Person ALS Grade 7 Mathematics

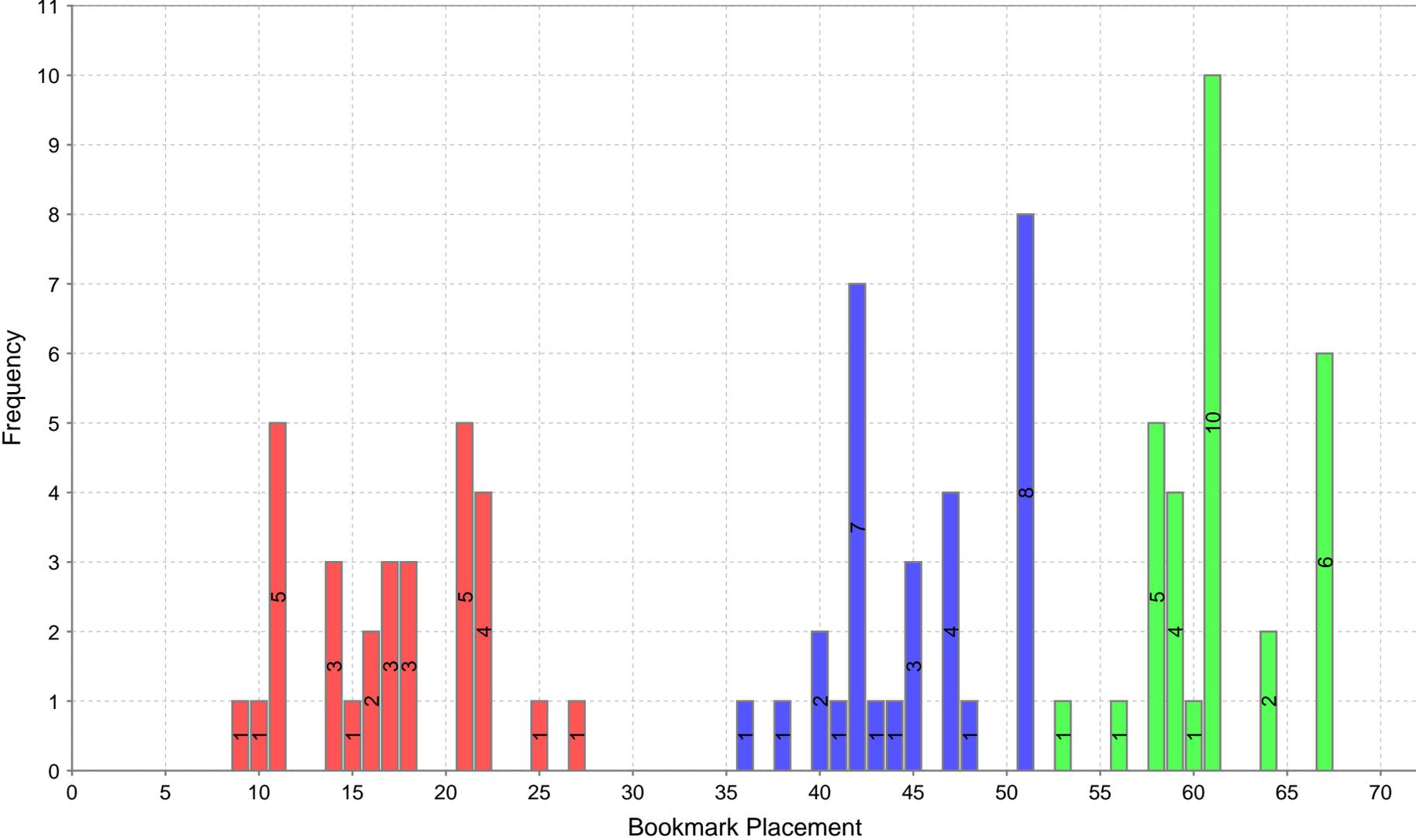
Frequency of Bookmark Placements Round 2



■ Level 2 (10-18-23.5)
 ■ Level 3 (40-46-51)
 ■ Level 4 (58-61-67)

Smarter Balanced In-Person ALS Grade 7 Mathematics

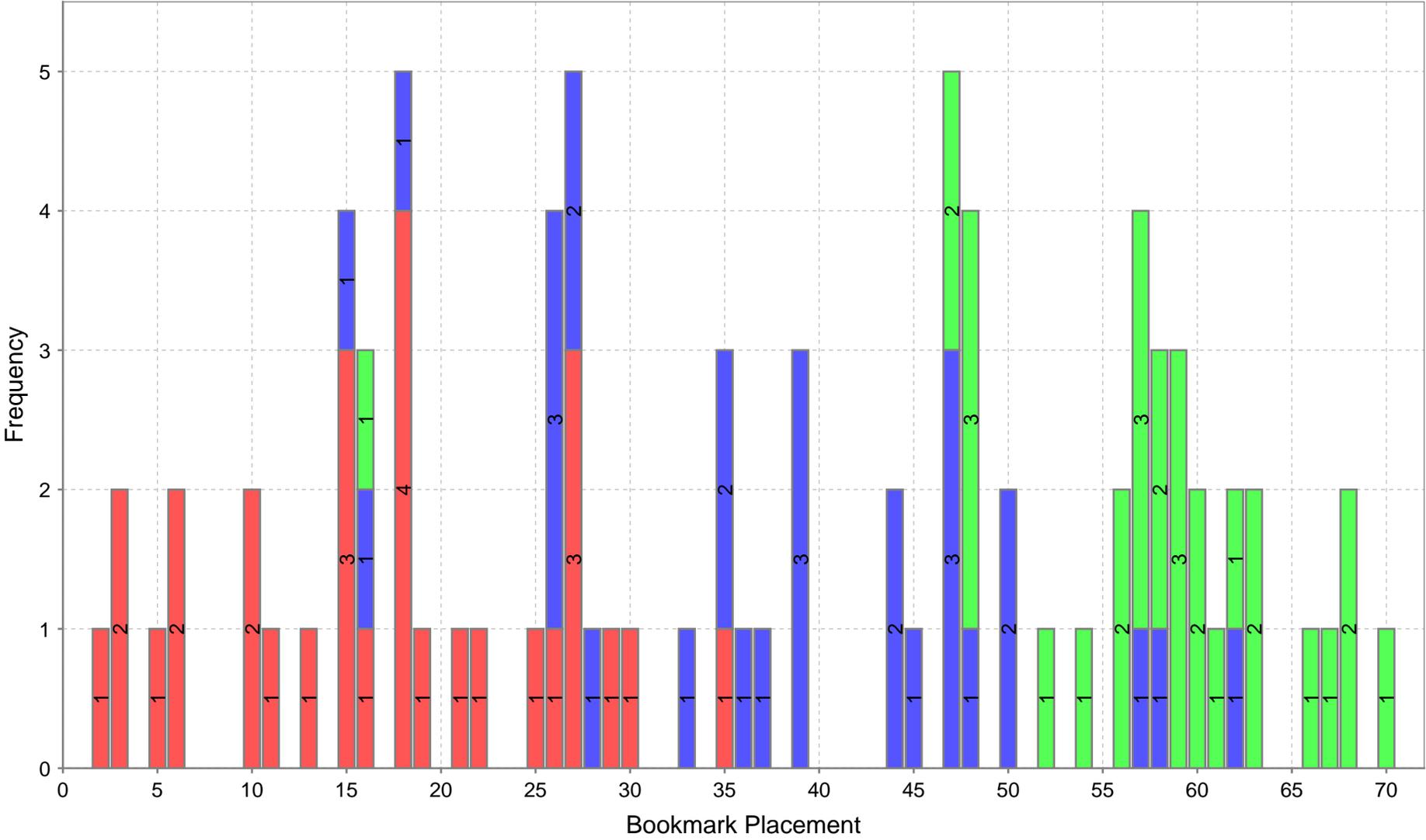
Frequency of Bookmark Placements Round 3



■ Level 2 (13.25-17-21)
 ■ Level 3 (42-45-51)
 ■ Level 4 (58.75-61-64)

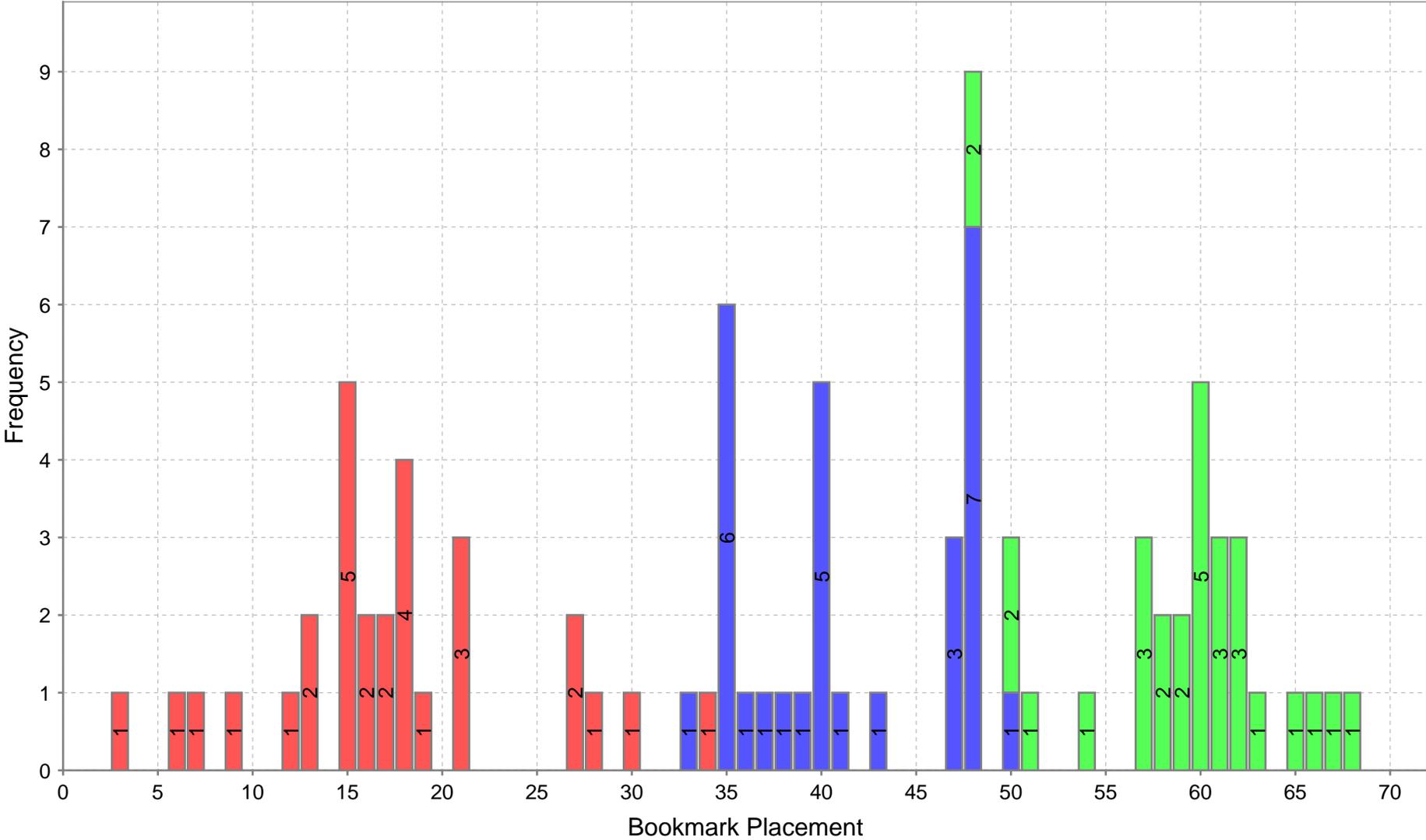
Smarter Balanced In-Person ALS Grade 8 Mathematics

Frequency of Bookmark Placements Round 1



■ Level 2 (10-18-25.5)
 ■ Level 3 (27-39-47)
 ■ Level 4 (53-58-62.5)

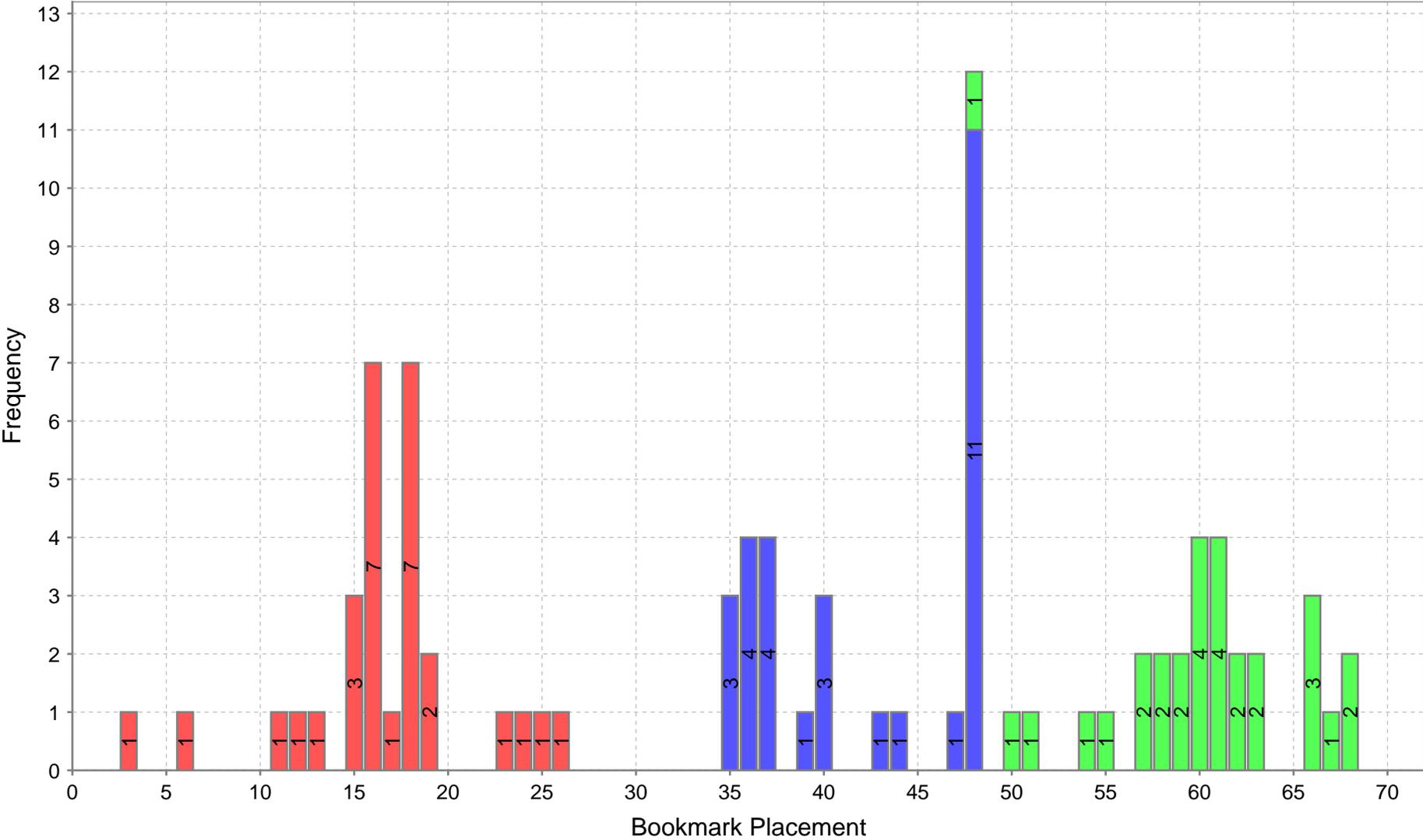
Smarter Balanced In-Person ALS Grade 8 Mathematics Frequency of Bookmark Placements Round 2



■ Level 2 (14-17-21)
 ■ Level 3 (35.5-40-48)
 ■ Level 4 (57-60-62)

Smarter Balanced In-Person ALS Grade 8 Mathematics

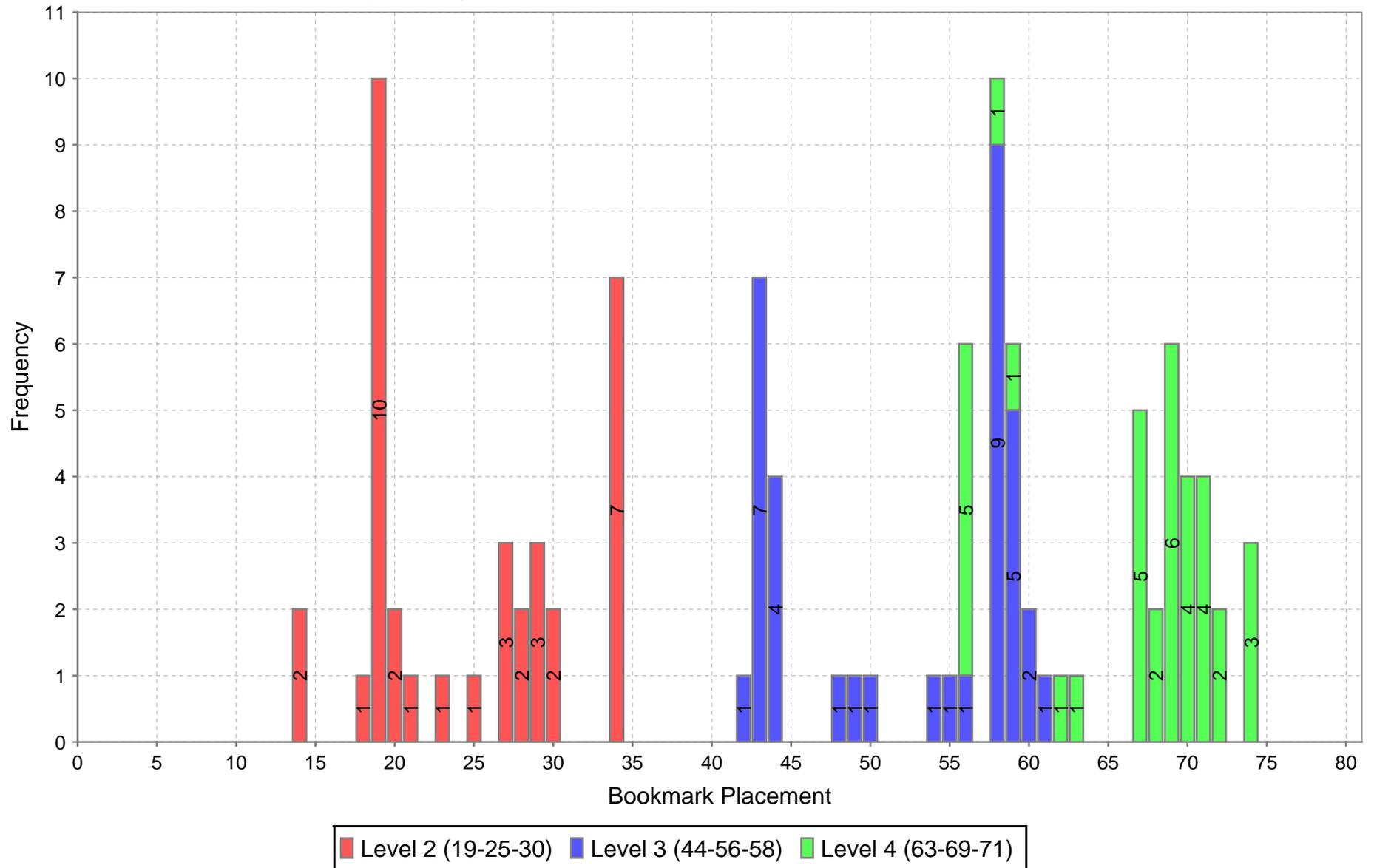
Frequency of Bookmark Placements Round 3



■ Level 2 (15-16-18)
 ■ Level 3 (36.5-40-48)
 ■ Level 4 (57.5-60-63)

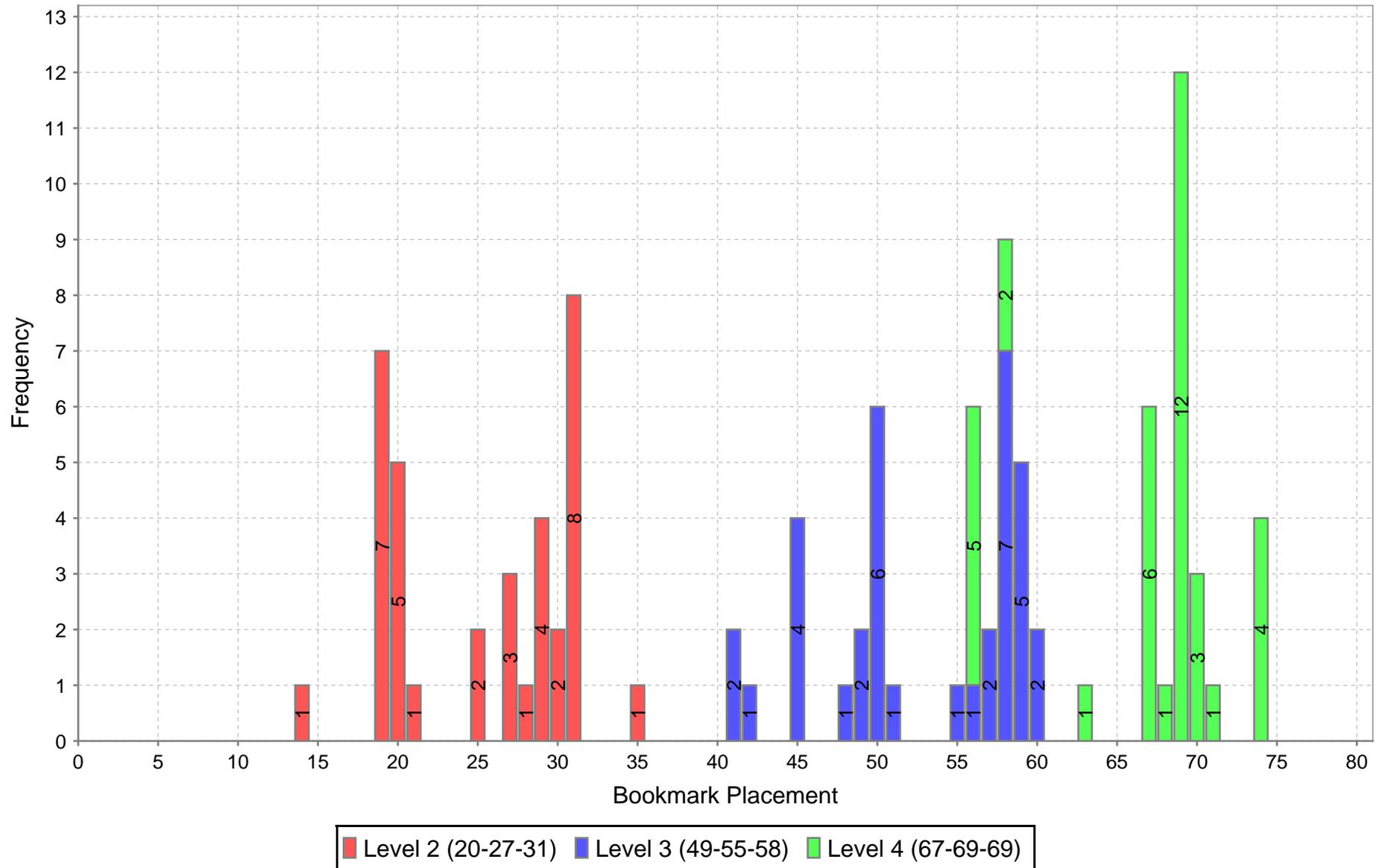
Smarter Balanced In-Person ALS Grade 11 Mathematics Panel A

Frequency of Bookmark Placements Round 1



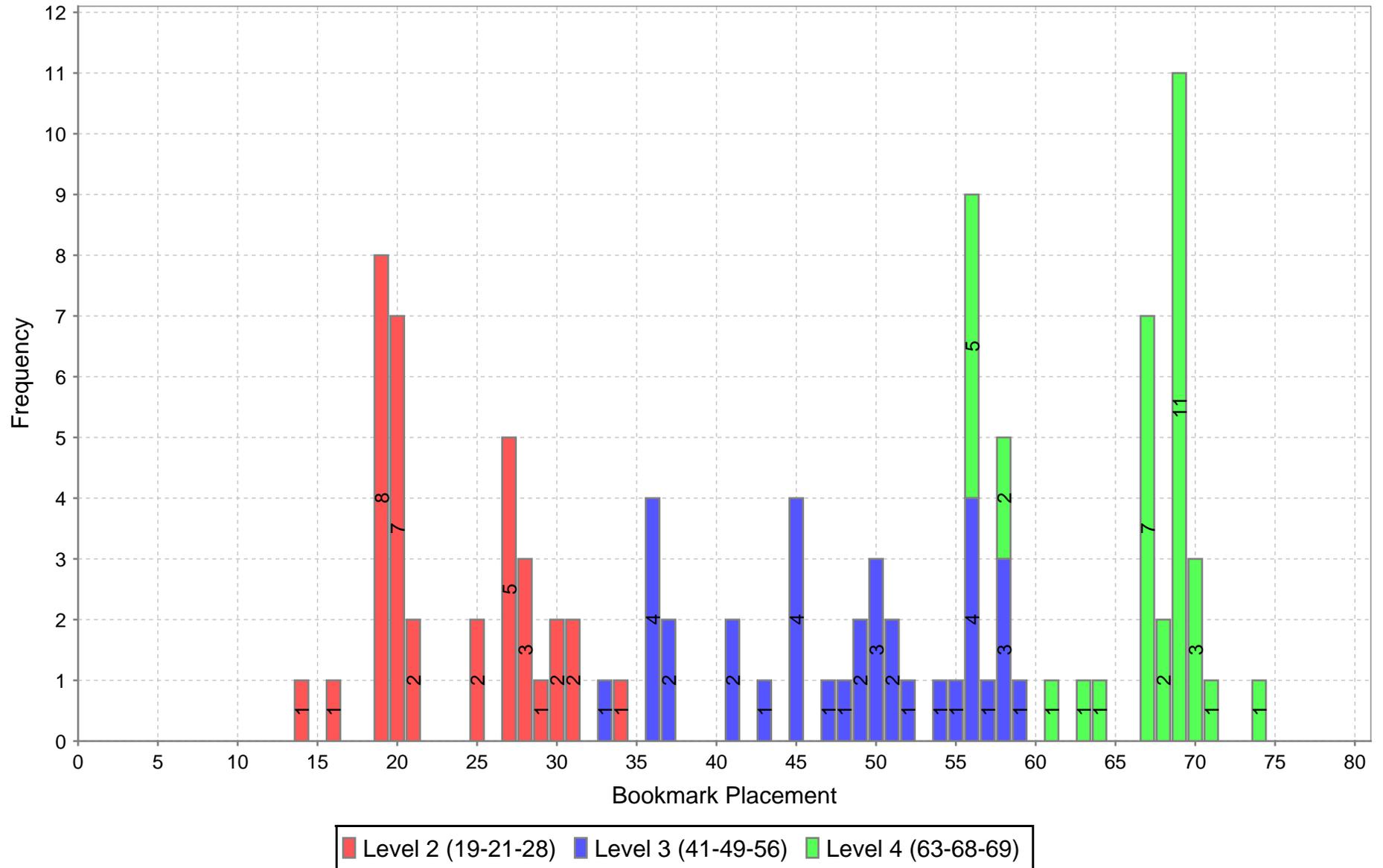
Smarter Balanced In-Person ALS Grade 11 Mathematics Panel A

Frequency of Bookmark Placements Round 2



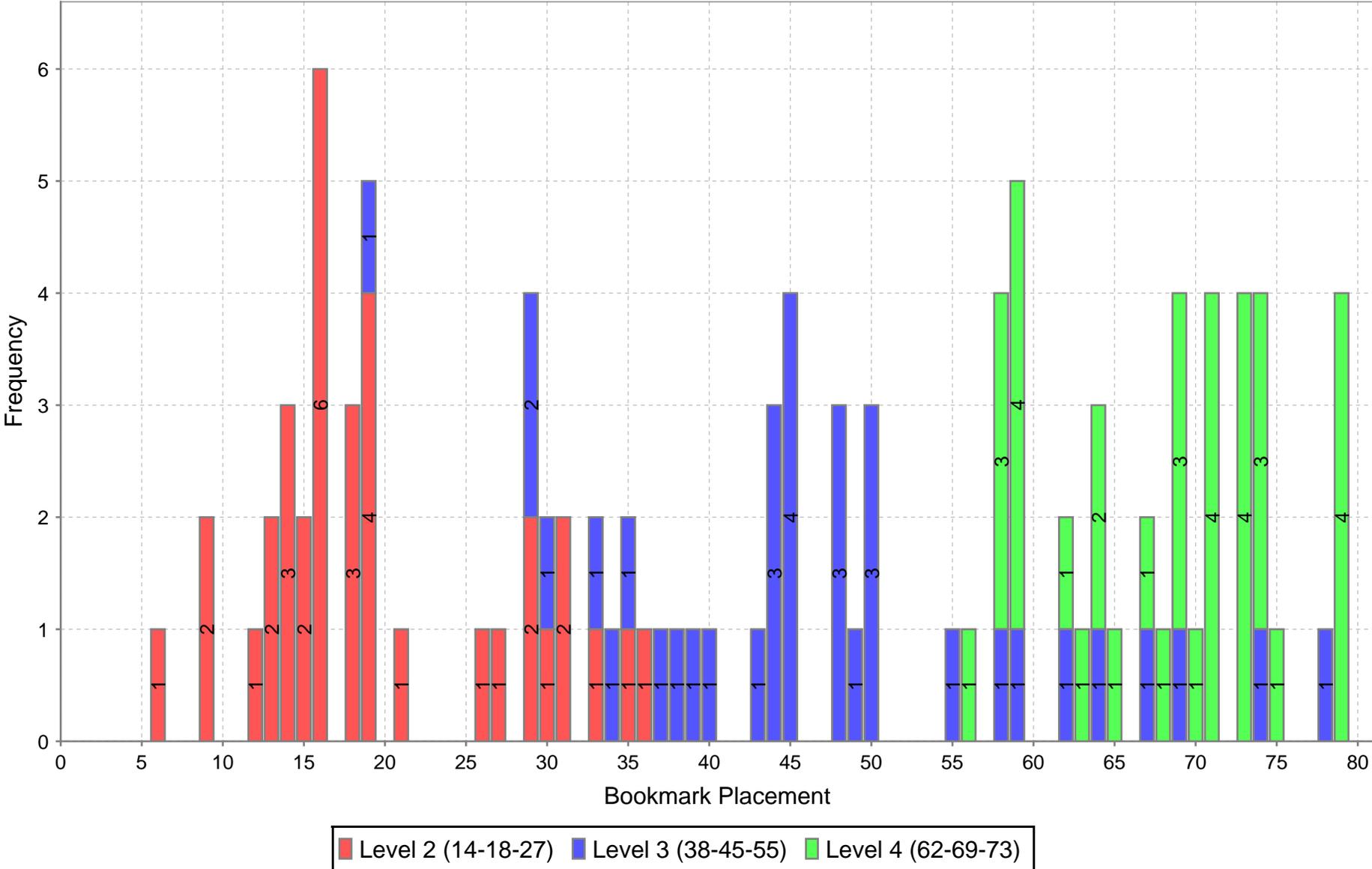
Smarter Balanced In-Person ALS Grade 11 Mathematics Panel A

Frequency of Bookmark Placements Round 3



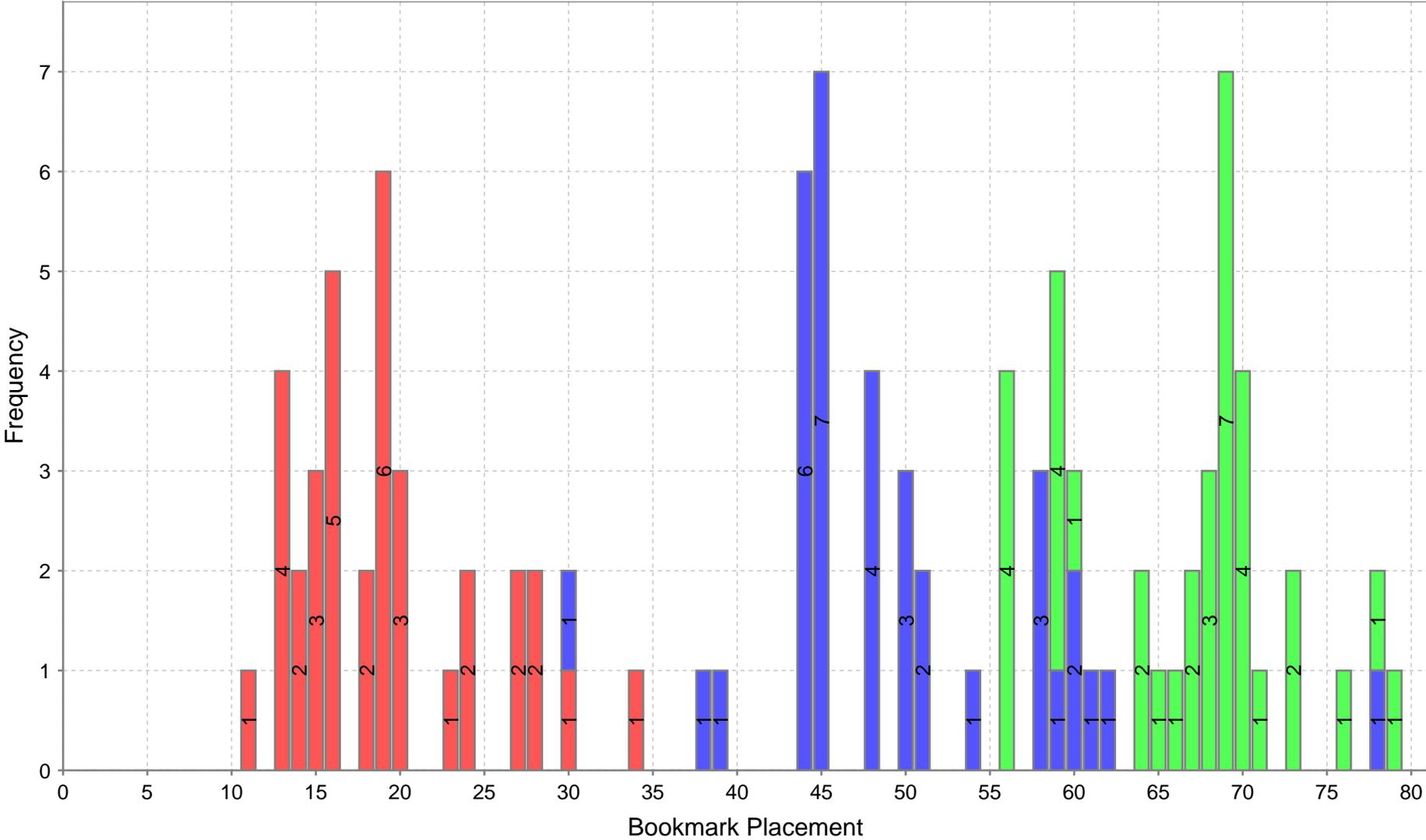
Smarter Balanced In-Person ALS Grade 11 Mathematics Panel B

Frequency of Bookmark Placements Round 1



Smarter Balanced In-Person ALS Grade 11 Mathematics Panel B

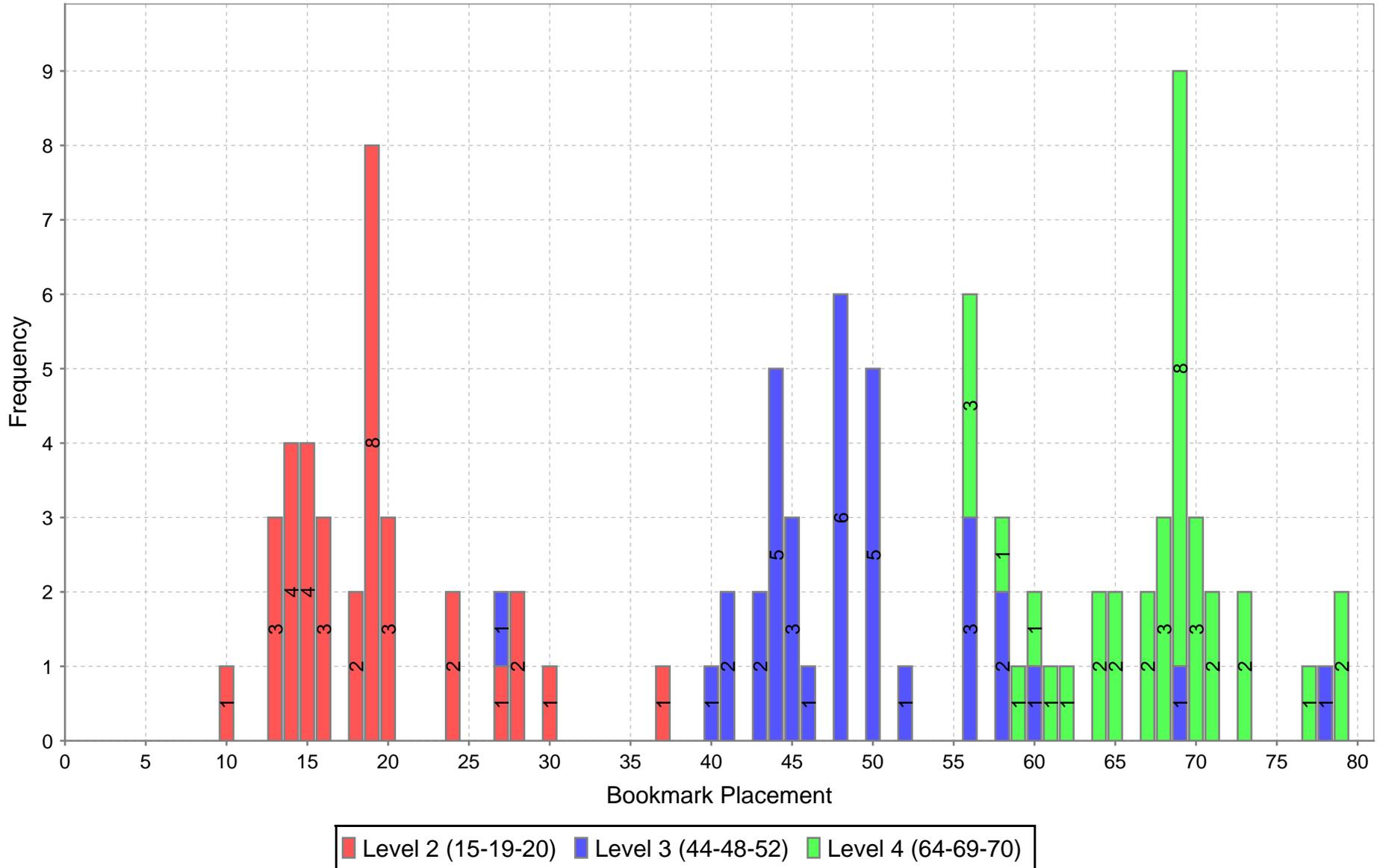
Frequency of Bookmark Placements Round 2



■ Level 2 (15-19-23)
 ■ Level 3 (44-48-58)
 ■ Level 4 (60-68-70)

Smarter Balanced In-Person ALS Grade 11 Mathematics Panel B

Frequency of Bookmark Placements Round 3



Appendix G – Cross-Grade Review Panel Documentation

- G01_Cross-Grade Review Agenda 926
- G02_Cross-Grade Review (PowerPoint Presentation) 927
- G03_Cross-Grade Review Results 952

Vertical Articulation Agenda

Day	Time	Event
Oct. 20	7:30 a.m.	Breakfast (Atrium I)
	8:30 a.m.	Reconvene in Breakout Rooms (Batik A/B)
	10:30 a.m.	Break
	12:00 noon	Lunch (Atrium I)
	1:00 p.m.	Reconvene in Breakout Rooms
	5:00 p.m.	Adjourn



Cross-Grade Review



October 20, 2014



Steps in the Process

- Online Panel
- In-Person Workshop
- Cross-Grade Review
- TAC Review
- Smarter Balanced Reviews
- Chiefs Review and Approval
- Technical Report



Steps in the Process

- Online Panel
- In-Person Workshop
- **Cross-Grade Review**
- TAC Review
- Smarter Balanced Reviews
- Chiefs Review and Approval
- Technical Report



The Cross Grade Review is your Final Round

- Round 1: Individual
- Round 2: Table level discussion
- Round 3: Room level discussion
- Round 4: System-wide discussion



Why Are You Here?

- You were selected from In-Person achievement level setting to provide
 - First-hand knowledge of all recommended cut scores
 - Balanced representation in grades 3-8
 - Larger representation in college- and career-ready grade
- We built the panel to be inclusive of all states

Your Task

- Examine system-wide bookmarks
- Establish coherence and reasonableness of cut scores across grades
- We will carefully review impact data and scale scores

Guiding Principles

- Group medians were based on different bookmarks—moving bookmarks modestly up or down is reasonable
- Scale score cuts should increase across grades
- New recommendations should align with the ALDs

Why across grade articulation?

- “VMSS (*Vertically Moderated Standard Setting*) is a procedure or set of procedures, typically carried out after individual standards have been set, that seeks to smooth out the bumps that inevitably occur across grades. Reasonable expectations are typically stated in terms of percentages of students at or above a consequential performance level, such as *Proficient...*”

(Cizek, *Setting Performance Standards*, 2012)

Why would we expect well-articulated performance standards?

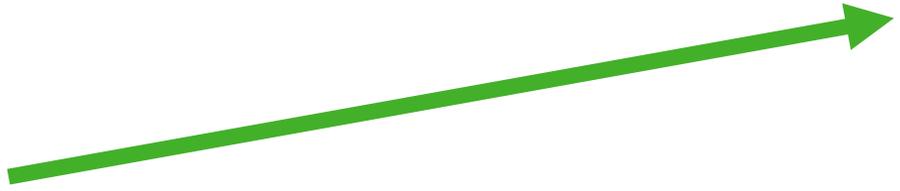
- The content standards at each of the contiguous grade levels were developed to reflect an ordered, increasing level of complexity...
- The tests at each of the grade levels were equally well-aligned to their respective content standards in terms of content coverage, depth of knowledge tapped by the assessments, and other alignment criteria (see e.g., Martone & Sireci, 2009; Webb, 1999, 2007).
- The PLDs (*sic*) across the covered grades were developed to represent comparable levels of achievement for each of the levels (e.g., Basic, Proficient, and Advanced).

(Setting Performance Standards; Cizek, 2012)



Expectations

- Scenario 1: Increasing



- Scenario 2: Equivalent

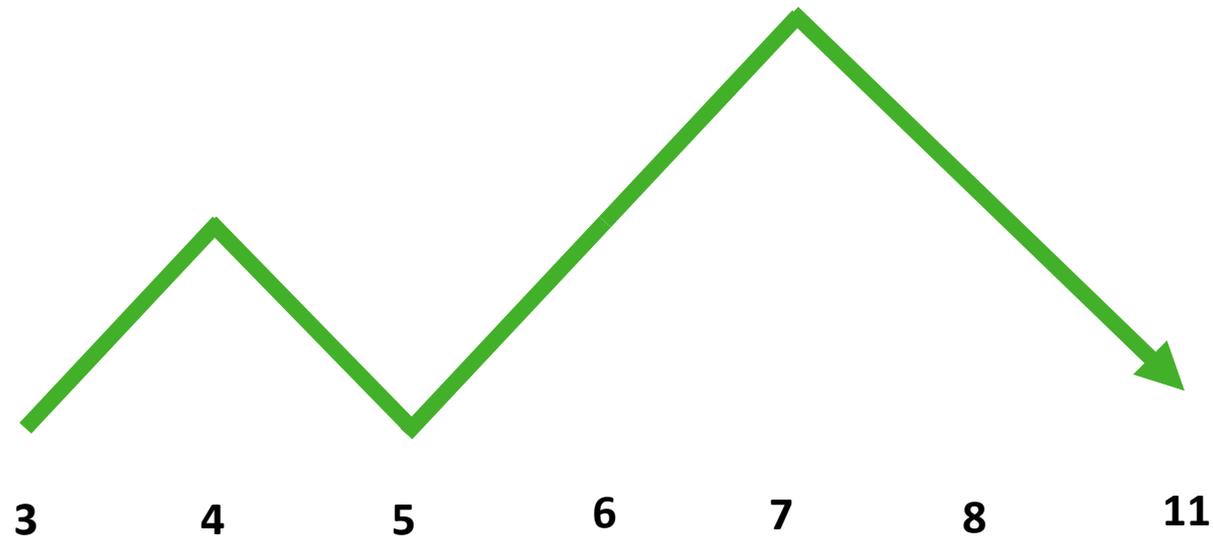


- Scenario 3: Declining





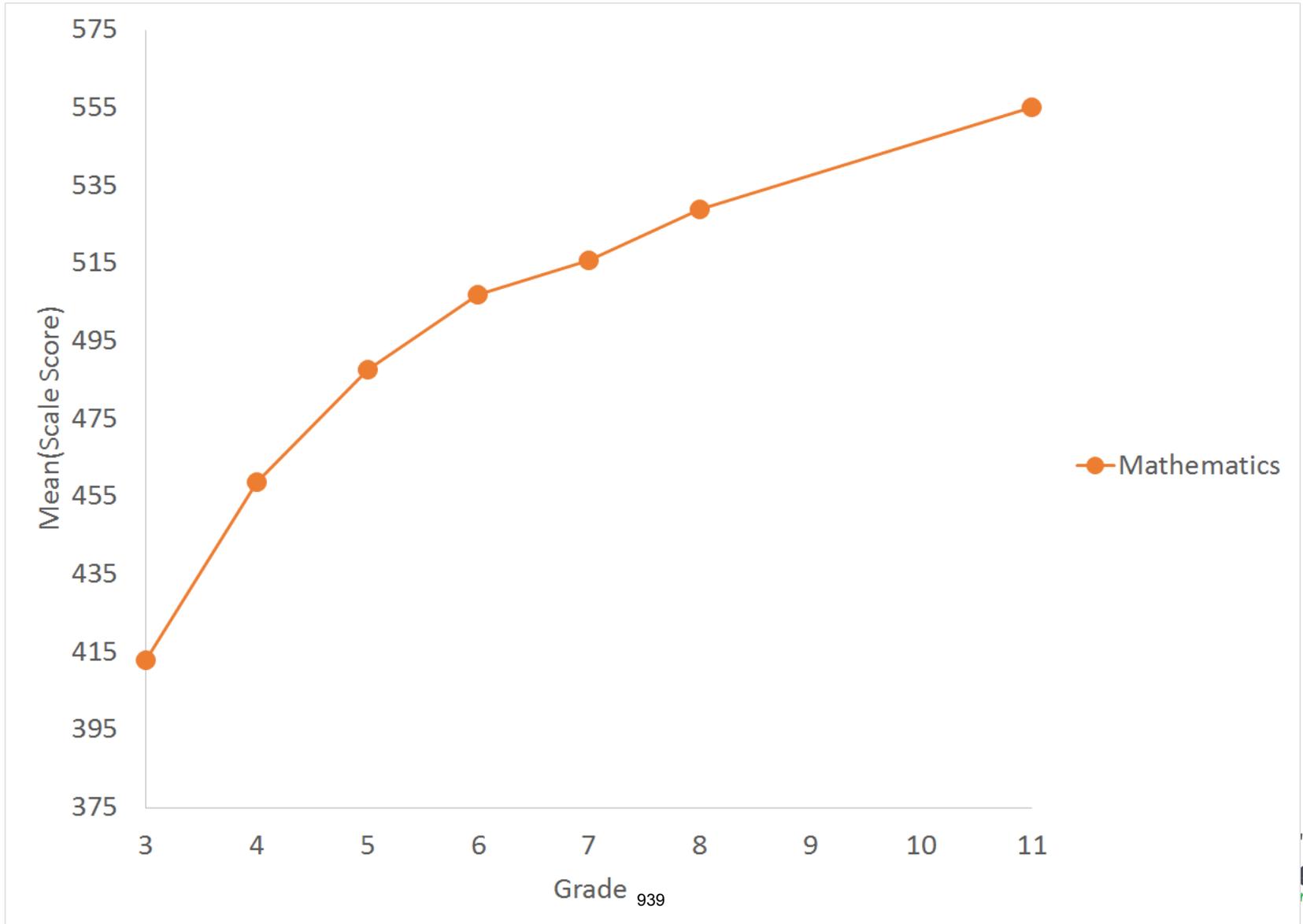
Not Expected



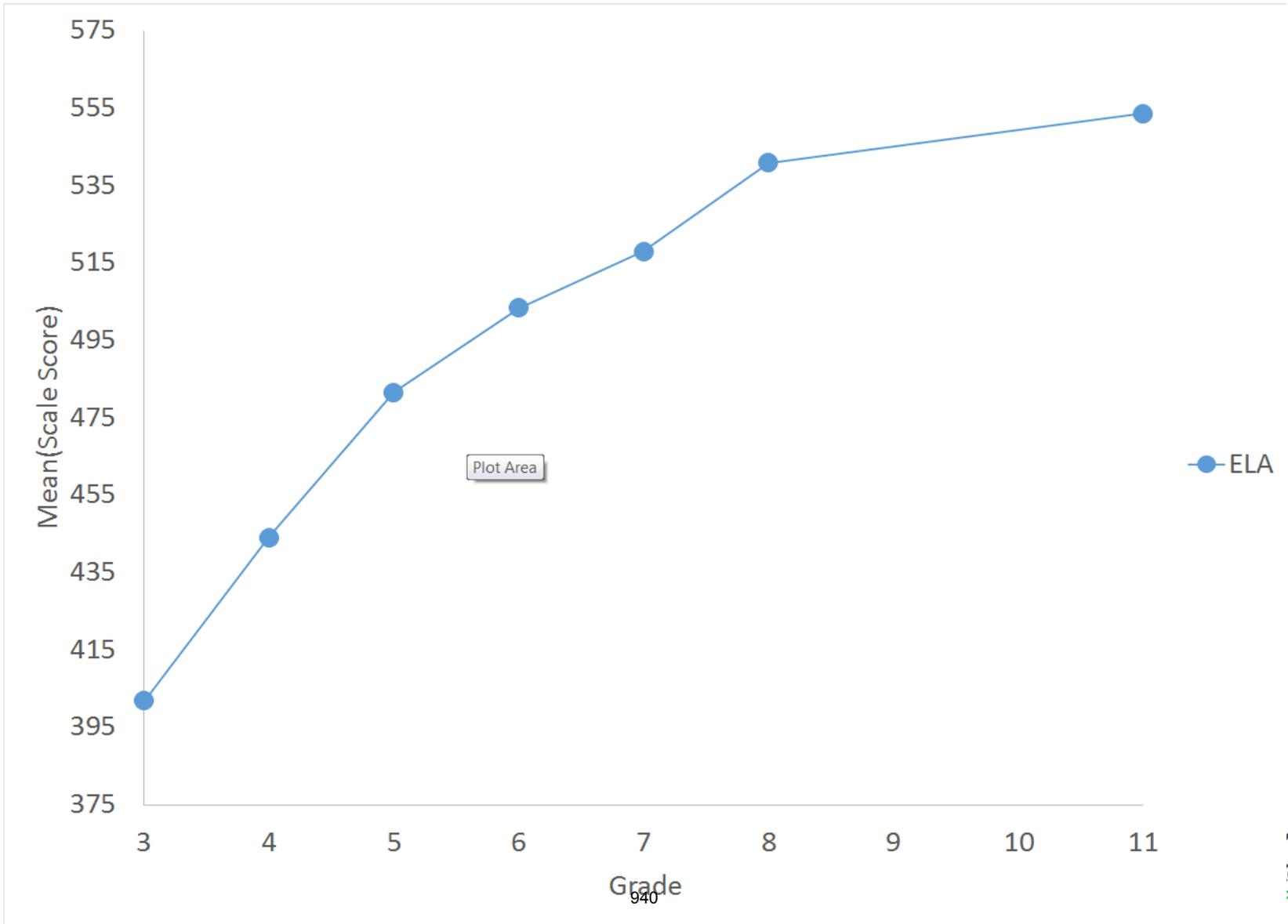
What should be articulated across grades?

- Bookmarks?
 - Not necessarily
- Percent of students at and above each performance level
- Scale score cut scores when scores are on a vertical scale

Math Mean Scale Scores (field test data)

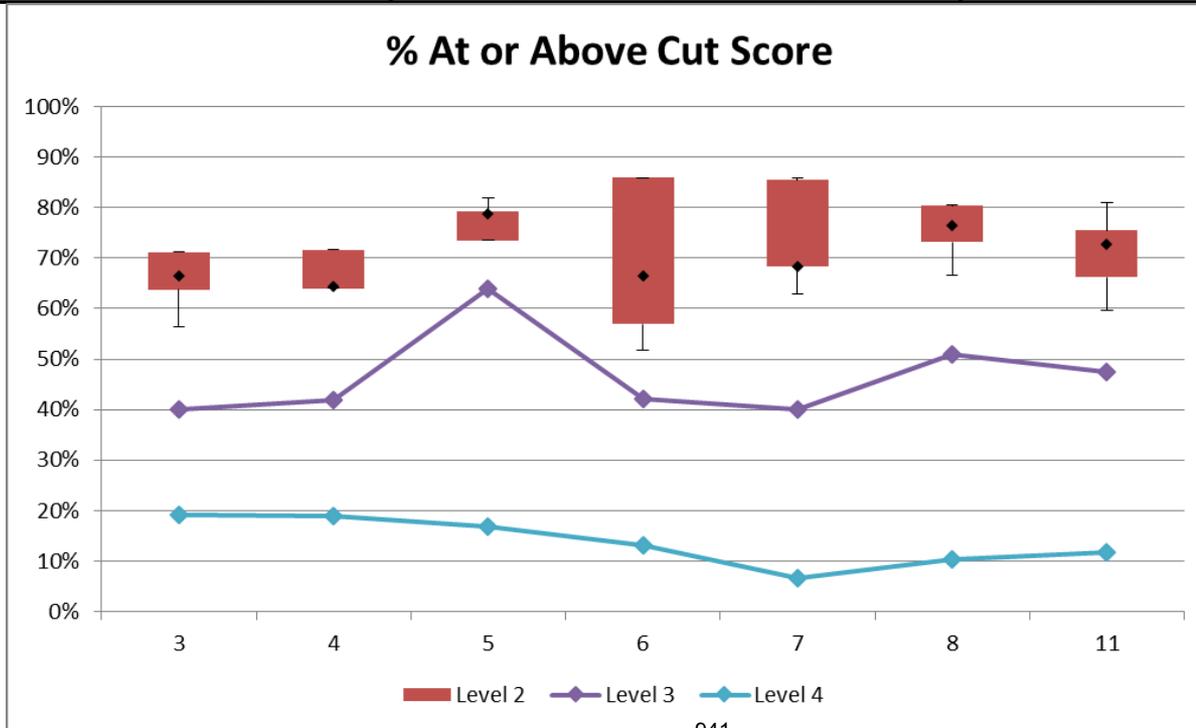


ELA Mean Scale Scores (field test data)



Orientation to Process Tools

Round 3 Bookmarks				% At or Above				% In Group				
Grade	Level 2	Level 3	Level 4	Grade	Level 2	Level 3	Level 4	Grade	Level 1	Level 2	Level 3	Level 4
3	13	33	54	3	67%	40%	19%	3	33%	26%	21%	19%
4	19	43	62	4	64%	42%	19%	4	36%	22%	23%	19%
5	11	27	63	5	79%	64%	17%	5	21%	15%	47%	17%
6	14.5	34.5	60.5	6	67%	42%	13%	6	33%	24%	29%	13%
7	16	38	66	7	68%	40%	7%	7	32%	28%	34%	7%
8	18	39.5	68	8	76%	51%	10%	8	24%	25%	41%	10%
11	19	42	65	11	73%	47%	12%	11	27%	25%	36%	12%



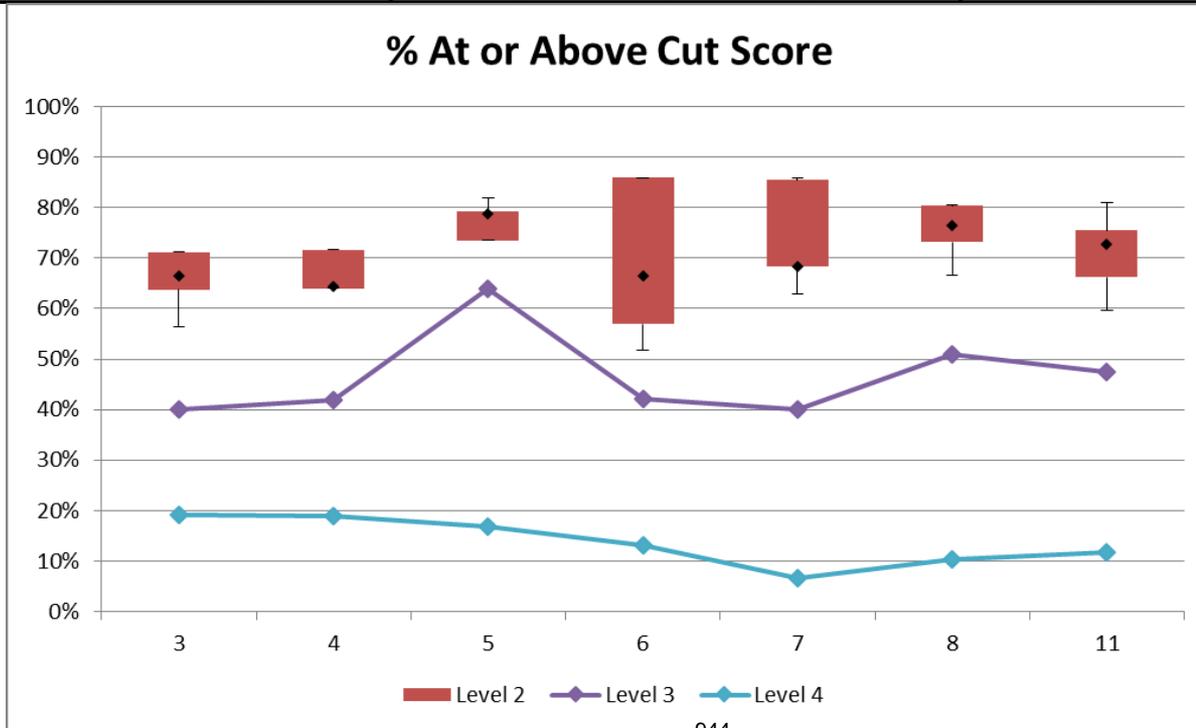
Round 3 Bookmarks

Grade	Level 2	Level 3	Level 4
3	28	49	70
4	9	32	71
5	20	50.5	64
6	18	45.5	61.5
7	17	45	61
8	16	40	60
11	19.5	48	68

% At or Above Cut Score				% In Group				
Grade	Level 2	Level 3	Level 4	Grade	Level 1	Level 2	Level 3	Level 4
3	65%	34%	11%	3	35%	31%	23%	11%
4	85%	52%	16%	4	15%	34%	36%	16%
5	61%	31%	14%	5	39%	30%	18%	14%
6	58%	29%	16%	6	42%	29%	14%	16%
7	53%	19%	6%	7	47%	34%	14%	6%
8	51%	26%	7%	8	49%	26%	18%	7%
11	59%	24%	6%	11	41%	35%	18%	6%

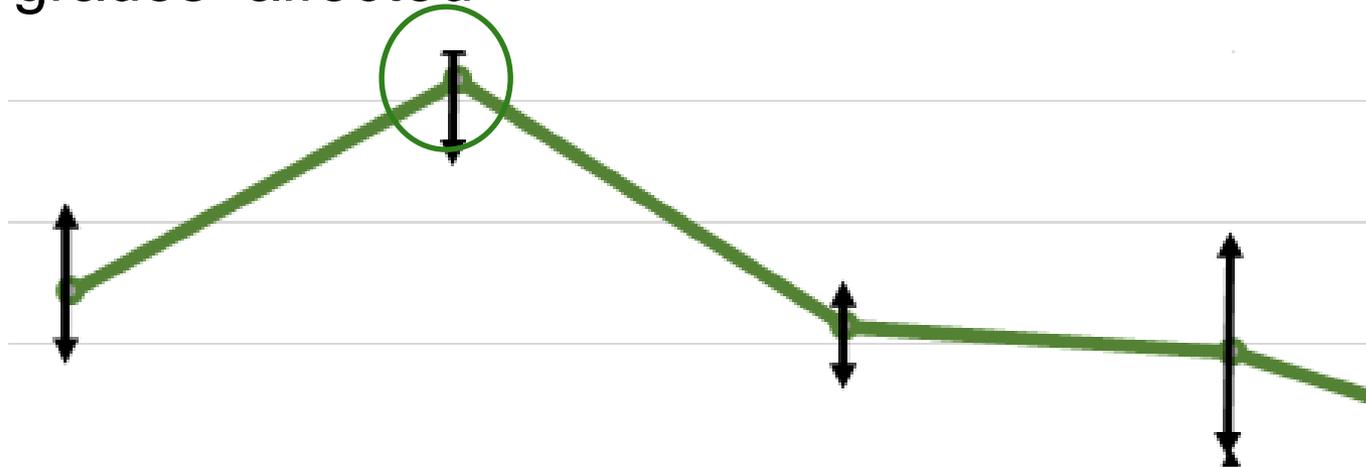
Orientation to Process Tools

Round 3 Bookmarks				% At or Above				% In Group				
Grade	Level 2	Level 3	Level 4	Grade	Level 2	Level 3	Level 4	Grade	Level 1	Level 2	Level 3	Level 4
3	13	33	54	3	67%	40%	19%	3	33%	26%	21%	19%
4	19	43	62	4	64%	42%	19%	4	36%	22%	23%	19%
5	11	27	63	5	79%	64%	17%	5	21%	15%	47%	17%
6	14.5	34.5	60.5	6	67%	42%	13%	6	33%	24%	29%	13%
7	16	38	66	7	68%	40%	7%	7	32%	28%	34%	7%
8	18	39.5	68	8	76%	51%	10%	8	24%	25%	41%	10%
11	19	42	65	11	73%	47%	12%	11	27%	25%	36%	12%



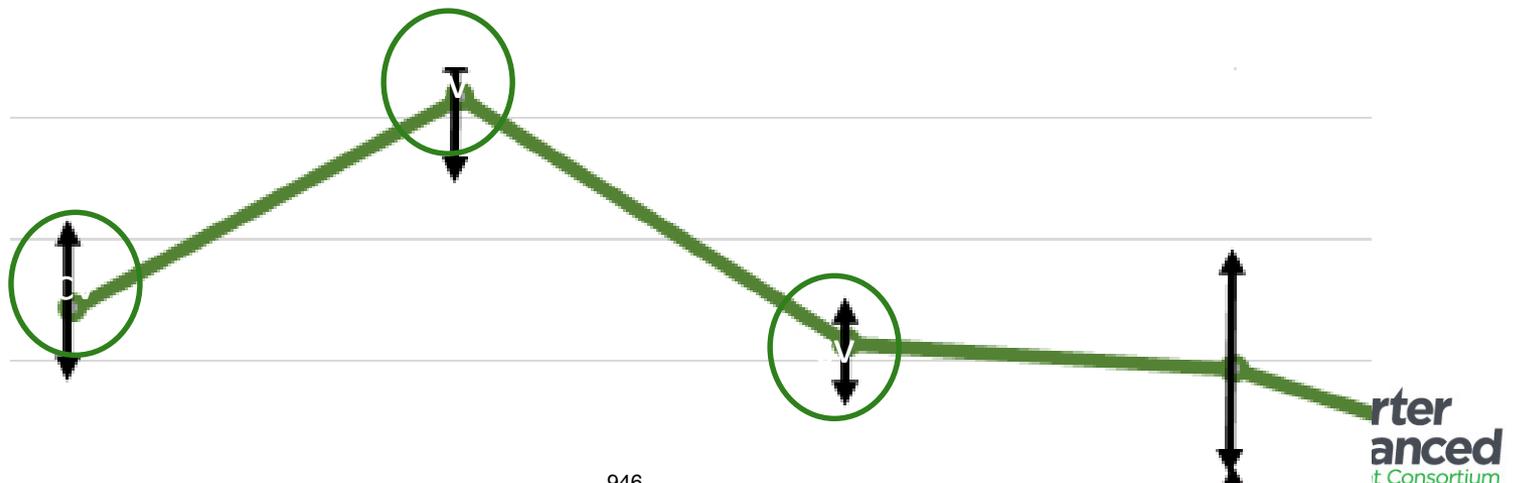
Guiding Principles

- Modest changes across three groups are more preferable than a large change in one grade
 - Assuming content based changes are acceptable in all grades affected



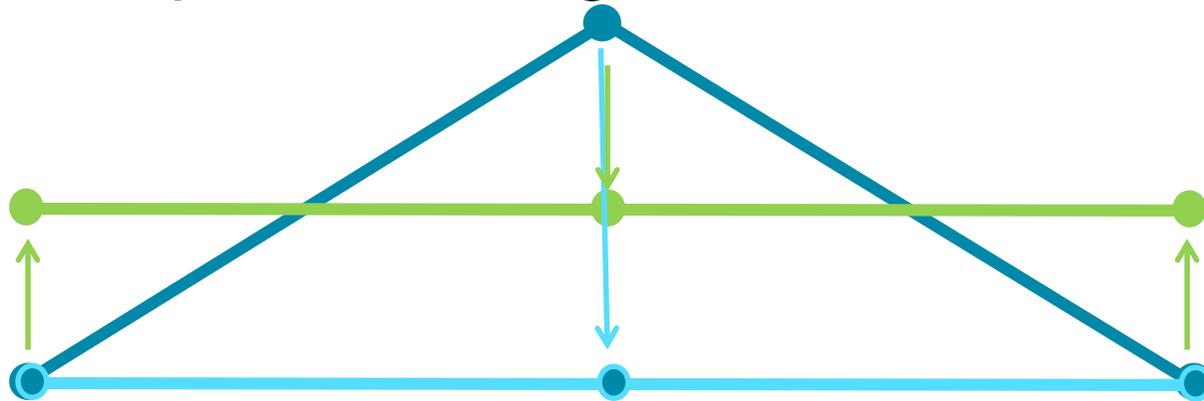
Orientation to Process Tools

- Lowering a point on line means:
 - a decrease in percent at or above (more rigor), and
 - a higher bookmark placement, further in the OIB (more rigor)
- Raising a point on the line means:
 - An increase in percent at or above (less rigor), and
 - a lower bookmark placement, earlier in the OIB (less rigor)



Guiding Principles

- Modest changes across three grades are more preferable than a large change in one grade
 - Assuming content based changes are acceptable in all grades affected



Grade 3

4

5

Panelist Materials

- 7 OIBs
 - Grades 3,4,5,6,7,8,11
- ALDs from your grade

Process

- Consider changes across grade-groups that may support articulation
 - Example: Grades 3 and 5 consider raising bookmarks and grade 4 consider lowering bookmarks
- Give time for each grade to review OIBs and ALDs to consider changes
- Each grade reports to whole group on their findings
- Recommend actions based on findings and resulting articulation
- A motion for recommended actions
- Vote and note (2/3 majority)
- Secretary takes notes of each of the above steps, carefully records motions, votes, and results

Process

Record of Actions of the Cross-Grade Review Committee
October 20, 2014
Mathematics

Motion	Time (CDT)	Second	Vote	Result/Action



Follow-Up Activities

- TAC Review
- Smarter Balanced Reviews
- Chiefs Review and Approval
- Technical Report

- **Questions?**

Cross-Grade Review Results

Table D.5.1 Cut Scores (Bookmarks) and Percentage of Students At or Above Cut Scores Before and After Cross-Grade Review, ELA

Grade	Before Cross-Grade Review						After Cross-Grade Review					
	Level 2		Level 3		Level 4		Level 2		Level 3		Level 4	
	Cut	%	Cut	%	Cut	%	Cut	%	Cut	%	Cut	%
Grade 3	13	66.5%	33	40.1%	54	19.1%	13	66.5%	33	40.1%	54	19.1%
Grade 4	19	64.4%	43	42.0%	62	18.9%	19	64.4%	43	42.0%	62	18.9%
Grade 5	11	78.7%	27	64.0%	63	16.9%	18	69.7%	37	47.1%	63	16.9%
Grade 6	14.5	66.6%	34.5	42.2%	60.5	13.1%	11	71.3%	34.5	42.2%	60.5	13.1%
Grade 7	16	68.2%	38	40.1%	66	6.6%	16	68.2%	38	40.1%	65	9.5%
Grade 8	18	76.4%	39.5	50.9%	68	10.2%	21	73.1%	44	43.3%	68	10.2%
Grade 11	19	72.7%	42	47.4%	65	11.6%	17	73.3%	45	42.6%	65	11.6%

Figure D.5.1 Percent of Students At or Above Cut Scores Before and After Cross-Grade Review, ELA

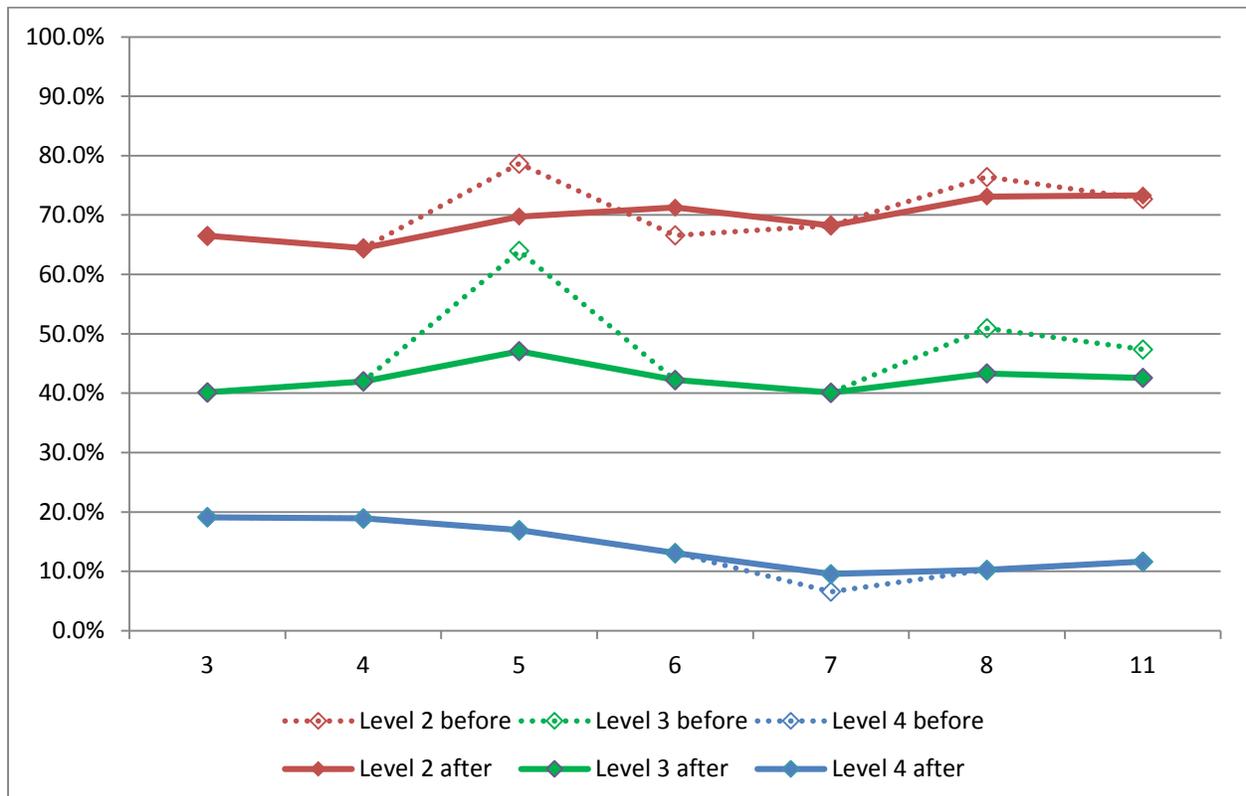
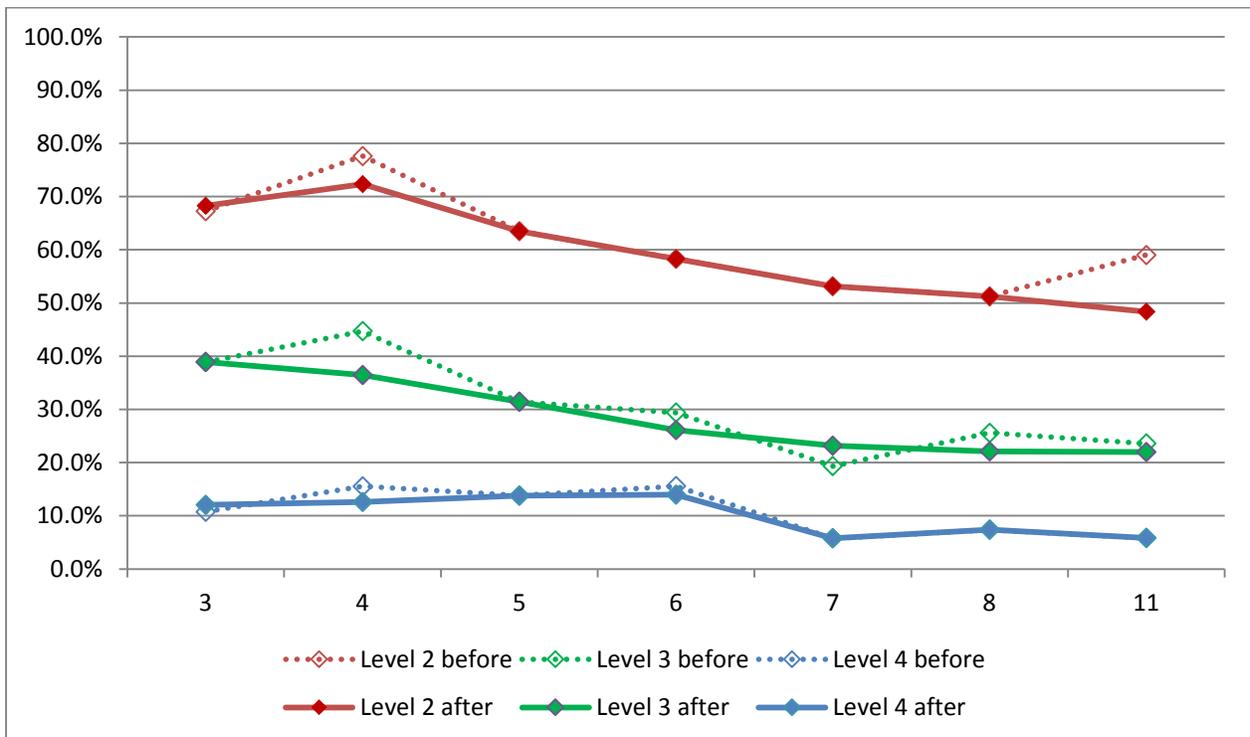


Table D.5.2 Cut Scores (Bookmarks) and Percentage of Students At or Above Cut Scores Before and After Cross-Grade Review, Math

Grade	Before Cross-Grade Review						After Cross-Grade Review					
	Level 2		Level 3		Level 4		Level 2		Level 3		Level 4	
	Cut	%	Cut	%	Cut	%	Cut	%	Cut	%	Cut	%
Grade 3	27	67.3%	47	38.9%	70	10.8%	26	68.3%	47	38.9%	69	12.1%
Grade 4	15	77.6%	39	44.7%	71	15.6%	18	72.3%	49	36.5%	73	12.6%
Grade 5	19	63.5%	50	31.4%	64	13.8%	19	63.5%	50	31.4%	64	13.8%
Grade 6	18	58.3%	45.5	29.4%	61.5	15.6%	18	58.3%	48	26.1%	63	14.0%
Grade 7	17	53.1%	45	19.3%	61	5.8%	17	53.1%	40	23.2%	61	5.8%
Grade 8	16	51.3%	40	25.6%	60	7.4%	16	51.3%	41	22.1%	60	7.4%
Grade 11	19.5	59.0%	48	23.6%	68	5.8%	23	48.3%	50	22.0%	68	5.8%

Figure D.5.2 Percent of Students At or Above Cut Scores Before and After Cross-Grade Review, Math



Appendix H – Follow-Up Activities

- H01_Disaggregated Field Test Data Summary..... 955
- H02_Disaggregated Field Test Data – ELA 956
- H03_Disaggregated Field Test Data – Math..... 960
- H04_November 6 Overview 964
- H05_November 14 Overview 989

Disaggregated Field Test Data

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels

Subject	Grade	Total	Level 1		Level 2		Level 3		Level 4		% At or Above		
			N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3	
ELA	3	23,223	8,241	35.5	6,127	26.4	4,764	20.5	4,091	17.6	64.5	38.1	
	4	35,689	13,112	36.7	8,082	22.7	8,093	22.7	6,402	17.9	63.3	40.6	
	5	31,594	10,456	33.1	7,281	23.0	9,237	29.3	4,620	14.6	66.9	43.9	
	6	31,535	9,507	30.1	9,245	29.4	9,239	29.3	3,544	11.2	69.9	40.5	
	7	30,913	10,414	33.7	8,704	28.1	9,169	29.7	2,626	8.5	66.3	38.2	
	8	35,913	10,189	28.4	10,861	30.2	11,538	32.1	3,325	9.3	71.6	41.4	
	9	7,714	2,638	34.2	2,428	31.5	2,096	27.1	552	7.2	65.8	34.3	
	10	11,924	2,956	24.8	3,686	30.9	3,914	32.8	1,368	11.5	75.2	44.3	
	11	31,019	8,663	27.9	9,667	31.2	9,303	30.0	3,386	10.9	72.1	40.9	
	MATH	3	24,799	7,867	31.7	7,283	29.4	6,659	26.8	2,990	12.1	68.3	38.9
		4	38,925	10,414	26.8	13,920	35.7	9,440	24.3	5,151	13.2	73.2	37.5
5		42,380	14,734	34.8	13,651	32.2	7,689	18.1	6,306	14.9	65.2	33.0	
6		29,946	10,363	34.6	9,721	32.5	5,764	19.2	4,098	13.7	65.4	32.9	
7		28,271	10,241	36.2	8,736	30.9	5,699	20.2	3,595	12.7	63.8	32.9	
8		34,880	13,230	37.9	10,489	30.1	6,752	19.4	4,409	12.6	62.1	32.0	
9		12,016	6,385	53.1	3,298	27.5	1,814	15.1	519	4.3	46.9	19.4	
10		14,342	6,093	42.5	4,157	29.0	2,937	20.4	1,155	8.1	57.5	28.5	
11	21,250	8,581	40.4	5,695	26.8	4,544	21.4	2,430	11.4	59.6	32.8		

Nov. 17, 2014

Disaggregated Field Test Data: ELA

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels:
ELA Grade 3

ELA Grade 3		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
ALL	23,223	8,241	35.5	6,127	26.4	4,764	20.5	4,091	17.6	64.5	38.1
Male	11,851	4,586	38.7	3,127	26.4	2,328	19.6	1,810	15.3	61.3	34.9
Female	11,372	3,655	32.1	3,000	26.4	2,436	21.4	2,281	20.1	67.9	41.5
Hispanic	7,411	3,591	48.5	2,058	27.7	1,157	15.6	605	8.2	51.5	23.8
Native	404	223	55.2	96	23.8	56	13.8	29	7.2	44.8	21.0
Asian	1,860	417	22.4	400	21.5	458	24.6	585	31.5	77.6	56.1
Black	1,595	828	51.9	420	26.3	224	14.1	123	7.7	48.1	21.8
White	10,914	2,839	26.0	2,876	26.4	2,657	24.3	2,542	23.3	74.0	47.6
Pacific Islander	703	323	45.9	192	27.4	116	16.5	72	10.2	54.1	26.7
Multi-Ethnic	1,055	362	34.3	269	25.5	216	20.5	208	19.7	65.7	40.2
Individualized Education Program	2,176	1,402	64.4	431	19.8	213	9.8	130	6.0	35.6	15.8
English Language Learner	4,356	2,621	60.2	1,144	26.2	438	10.1	153	3.5	39.8	13.6
Section 504	176	78	44.3	50	28.4	33	18.8	15	8.5	55.7	27.3
Economically Disadvantaged	12,554	6,005	47.8	3,490	27.8	2,010	16.0	1,049	8.4	52.2	24.4

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels:
ELA Grade 4

ELA Grade 4		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
ALL	35,689	13,112	36.7	8,082	22.6	8,093	22.7	6,402	17.9	63.2	40.6
Male	18,372	7,644	41.6	4,162	22.7	3,868	21.0	2,698	14.7	58.4	35.7
Female	17,317	5,468	31.6	3,920	22.6	4,225	24.4	3,704	21.4	68.4	45.8
Hispanic	9,579	4,898	51.1	2,268	23.7	1,590	16.6	823	8.6	48.9	25.2
Native	1,158	699	60.4	244	21.0	160	13.9	55	4.7	39.6	18.6
Asian	2,653	617	23.3	512	19.3	719	27.1	805	30.3	76.7	57.4
Black	2,493	1,359	54.5	531	21.3	429	17.2	174	7.0	45.5	24.2
White	19,905	5,803	29.2	4,540	22.8	5,127	25.7	4,435	22.3	70.8	48.0
Pacific Islander	542	279	51.5	124	22.9	88	16.2	51	9.4	48.5	25.6
Multi-Ethnic	1,548	531	34.3	362	23.4	374	24.1	281	18.2	65.7	42.3
Individualized Education Program	3,736	2,506	67.1	618	16.5	389	10.4	223	6.0	32.9	16.4
English Language Learner	4,509	3,155	70.0	914	20.2	344	7.7	96	2.1	30.0	9.8
Section 504	312	120	38.5	82	26.2	72	23.1	38	12.2	61.5	35.3
Economically Disadvantaged	18,432	9,185	49.8	4,357	23.7	3,255	17.6	1,635	8.9	50.2	26.5

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels:
ELA Grade 5

ELA Grade 5		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
ALL	31,594	10,456	33.1	7,281	23.0	9,237	29.2	4,620	14.6	66.8	43.8
Male	16,138	6,129	38.0	3,785	23.4	4,361	27.1	1,863	11.5	62.0	38.6
Female	15,456	4,327	28.0	3,496	22.6	4,876	31.6	2,757	17.8	72.0	49.4
Hispanic	8,909	4,187	47.0	2,231	25.0	1,885	21.2	606	6.8	53.0	28.0
Native	932	505	54.2	217	23.3	154	16.5	56	6.0	45.8	22.5
Asian	2,289	508	22.2	416	18.2	780	34.0	585	25.6	77.8	59.6
Black	2,483	1,224	49.3	576	23.2	515	20.7	168	6.8	50.7	27.5
White	19,438	5,414	27.9	4,430	22.7	6,276	32.3	3,318	17.1	72.1	49.4
Pacific Islander	517	230	44.5	132	25.5	112	21.7	43	8.3	55.5	30.0
Multi-Ethnic	1,289	430	33.4	285	22.1	389	30.1	185	14.4	66.6	44.5
Individualized Education Program	3,405	2,371	69.6	592	17.4	346	10.2	96	2.8	30.4	13.0
English Language Learner	3,463	2,469	71.3	712	20.6	253	7.3	29	0.8	28.7	8.1
Section 504	395	129	32.7	99	25.0	122	30.9	45	11.4	67.3	42.3
Economically Disadvantaged	15,836	7,351	46.4	3,928	24.8	3,541	22.4	1,016	6.4	53.6	28.8

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels:
ELA Grade 6

ELA Grade 6		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
ALL	31,535	9,507	30.1	9,245	29.3	9,239	29.3	3,544	11.2	69.8	40.5
Male	16,101	5,598	34.8	4,806	29.8	4,334	26.9	1,363	8.5	65.2	35.4
Female	15,434	3,909	25.3	4,439	28.8	4,905	31.8	2,181	14.1	74.7	45.9
Hispanic	8,845	3,741	42.3	2,824	31.9	1,897	21.5	383	4.3	57.7	25.8
Native	906	420	46.4	281	31.0	167	18.4	38	4.2	53.6	22.6
Asian	2,300	384	16.7	543	23.6	836	36.4	537	23.3	83.3	59.7
Black	2,158	982	45.5	642	29.8	414	19.1	120	5.6	54.5	24.7
White	17,335	4,050	23.4	5,009	28.9	5,878	33.9	2,398	13.8	76.6	47.7
Pacific Islander	371	161	43.4	109	29.4	86	23.2	15	4.0	56.6	27.2
Multi-Ethnic	1,274	382	30.0	367	28.8	369	29.0	156	12.2	70.0	41.2
Individualized Education Program	3,278	2,272	69.3	681	20.8	274	8.3	51	1.6	30.7	9.9
English Language Learner	3,058	2,173	71.1	744	24.3	127	4.1	14	0.5	28.9	4.6
Section 504	362	128	35.4	117	32.3	96	26.5	21	5.8	64.6	32.3
Economically Disadvantaged	15,977	6,665	41.7	5,006	31.3	3,493	21.9	813	5.1	58.3	27.0

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels:
ELA Grade 7

ELA Grade 7		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
ALL	30,913	10,414	33.7	8,704	28.2	9,169	29.7	2,626	8.5	66.4	38.2
Male	15,812	6,345	40.1	4,489	28.4	3,981	25.2	997	6.3	59.9	31.5
Female	15,101	4,069	26.9	4,215	28.0	5,188	34.3	1,629	10.8	73.1	45.1
Hispanic	13,040	5,826	44.7	3,900	29.9	2,833	21.7	481	3.7	55.3	25.4
Native	526	226	43.0	182	34.6	106	20.1	12	2.3	57.0	22.4
Asian	2,904	508	17.5	613	21.1	1,185	40.8	598	20.6	82.5	61.4
Black	1,582	787	49.7	439	27.8	307	19.4	49	3.1	50.3	22.5
White	12,537	3,056	24.4	3,511	28.0	4,559	36.3	1,411	11.3	75.6	47.6
Pacific Islander	254	126	49.6	62	24.4	55	21.7	11	4.3	50.4	26.0
Multi-Ethnic	867	222	25.6	233	26.9	310	35.7	102	11.8	74.4	47.5
Individualized Education Program	2,870	2,197	76.6	463	16.1	185	6.4	25	0.9	23.4	7.3
English Language Learner	3,818	2,967	77.7	703	18.4	139	3.7	9	0.2	22.3	3.9
Section 504	299	100	33.4	81	27.1	93	31.1	25	8.4	66.6	39.5
Economically Disadvantaged	17,348	7,696	44.4	5,119	29.5	3,841	22.1	692	4.0	55.6	26.1

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels:
ELA Grade 8

ELA Grade 8		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
ALL	35,913	10,189	28.4	10,861	30.2	11,538	32.1	3,325	9.3	71.6	41.4
Male	18,313	6,377	34.8	5,668	31.0	5,141	28.0	1,127	6.2	65.2	34.2
Female	17,600	3,812	21.7	5,193	29.5	6,397	36.3	2,198	12.5	78.3	48.8
Hispanic	9,997	3,809	38.1	3,310	33.1	2,457	24.6	421	4.2	61.9	28.8
Native	969	461	47.6	281	29.0	189	19.5	38	3.9	52.4	23.4
Asian	2,557	451	17.6	636	24.9	938	36.7	532	20.8	82.4	57.5
Black	2,869	1,282	44.7	926	32.3	541	18.8	120	4.2	55.3	23.0
White	19,419	4,299	22.1	5,709	29.4	7,295	37.6	2,116	10.9	77.9	48.5
Pacific Islander	330	136	41.2	110	33.3	65	19.7	19	5.8	58.8	25.5
Multi-Ethnic	1,355	340	25.1	384	28.3	469	34.6	162	12.0	74.9	46.6
Individualized Education Program	3,338	2,259	67.7	794	23.8	255	7.6	30	0.9	32.3	8.5
English Language Learner	2,750	1,979	72.0	638	23.2	128	4.6	5	0.2	28.0	4.8
Section 504	444	126	28.4	128	28.8	155	34.9	35	7.9	71.6	42.8
Economically Disadvantaged	17,525	6,777	38.7	5,711	32.6	4,237	24.1	800	4.6	61.3	28.7

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels:
 ELA Grade 11

ELA Grade 11		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
ALL	31,019	8,663	27.9	9,667	31.2	9,303	30.0	3,386	10.9	72.1	40.9
Male	15,445	5,300	34.3	4,774	30.9	4,081	26.4	1,290	8.4	65.7	34.8
Female	15,574	3,363	21.6	4,893	31.4	5,222	33.5	2,096	13.5	78.4	47.0
Hispanic	10,041	3,381	33.7	3,409	33.9	2,612	26.0	639	6.4	66.3	32.4
Native	777	284	36.6	286	36.8	175	22.5	32	4.1	63.4	26.6
Asian	2,344	477	20.3	599	25.6	800	34.1	468	20.0	79.7	54.1
Black	2,552	1,198	46.9	814	31.9	447	17.6	93	3.6	53.1	21.2
White	16,020	3,807	23.8	4,809	30.0	5,281	32.9	2,123	13.3	76.2	46.2
Pacific Islander	195	75	38.5	56	28.7	51	26.1	13	6.7	61.5	32.8
Multi-Ethnic	889	219	24.6	269	30.3	294	33.1	107	12.0	75.4	45.1
Individualized Education Program	2,084	1,386	66.5	511	24.5	158	7.6	29	1.4	33.5	9.0
English Language Learner	1,767	1,254	71.0	412	23.3	90	5.1	11	0.6	29.0	5.7
Section 504	366	112	30.6	122	33.3	98	26.8	34	9.3	69.4	36.1
Economically Disadvantaged	13,962	4,834	34.6	4,579	32.8	3,605	25.8	944	6.8	65.4	32.6

Disaggregated Field Test Data: Mathematics

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels:
Mathematics Grade 3

MATH Grade 3		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
ALL	24,799	7,867	31.7	7,283	29.4	6,659	26.9	2,990	12.1	68.4	39.0
Male	12,530	3,955	31.6	3,651	29.1	3,318	26.5	1,606	12.8	68.4	39.3
Female	12,269	3,912	31.9	3,632	29.6	3,341	27.2	1,384	11.3	68.1	38.5
Hispanic	7,589	3,355	44.2	2,407	31.7	1,480	19.5	347	4.6	55.8	24.1
Native	792	400	50.5	224	28.3	140	17.7	28	3.5	49.5	21.2
Asian	1,874	281	15.0	455	24.3	630	33.6	508	27.1	85.0	60.7
Black	2,158	1,066	49.4	665	30.8	341	15.8	86	4.0	50.6	19.8
White	14,503	3,937	27.1	4,159	28.7	4,350	30.0	2,057	14.2	72.9	44.2
Pacific Islander	292	132	45.2	90	30.8	57	19.5	13	4.5	54.8	24.0
Multi-Ethnic	903	248	27.5	279	30.9	237	26.2	139	15.4	72.5	41.6
Individualized Education Program	2,409	1,447	60.1	519	21.5	318	13.2	125	5.2	39.9	18.4
English Language Learner	4,000	2,072	51.8	1,270	31.7	555	13.9	103	2.6	48.2	16.5
Section 504	188	62	33.0	54	28.7	48	25.5	24	12.8	67.0	38.3
Economically Disadvantaged	12,973	5,554	42.8	4,043	31.2	2,672	20.6	704	5.4	57.2	26.0

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels:
Mathematics Grade 4

MATH Grade 4		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
ALL	38,925	10,414	26.8	13,920	35.8	9,440	24.3	5,151	13.2	73.3	37.5
Male	19,954	5,351	26.8	6,814	34.2	4,882	24.4	2,907	14.6	73.2	39.0
Female	18,971	5,063	26.7	7,106	37.4	4,558	24.1	2,244	11.8	73.3	35.9
Hispanic	10,467	4,242	40.5	4,062	38.8	1,677	16.1	486	4.6	59.5	20.7
Native	1,563	666	42.6	597	38.2	221	14.1	79	5.1	57.4	19.2
Asian	2,444	320	13.1	679	27.8	688	28.1	757	31.0	86.9	59.1
Black	4,137	2,074	50.1	1,463	35.4	489	11.8	111	2.7	49.9	14.5
White	25,697	5,535	21.5	9,218	35.9	7,097	27.6	3,847	15.0	78.5	42.6
Pacific Islander	331	114	34.4	127	38.4	66	19.9	24	7.3	65.6	27.2
Multi-Ethnic	1,857	441	23.7	682	36.8	456	24.5	278	15.0	76.3	39.5
Individualized Education Program	4,219	2,569	60.9	1,104	26.2	410	9.7	136	3.2	39.1	12.9
English Language Learner	4,374	2,349	53.7	1,577	36.1	368	8.4	80	1.8	46.3	10.2
Section 504	406	88	21.7	150	36.9	103	25.4	65	16.0	78.3	41.4
Economically Disadvantaged	19,774	7,805	39.5	7,605	38.4	3,303	16.7	1,061	5.4	60.5	22.1

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels:
Mathematics Grade 5

MATH Grade 5		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
ALL	42,380	14,734	34.8	13,651	32.2	7,689	18.1	6,306	14.9	65.2	33.0
Male	21,600	7,618	35.3	6,591	30.5	3,924	18.1	3,467	16.1	64.7	34.2
Female	20,780	7,116	34.2	7,060	34.0	3,765	18.1	2,839	13.7	65.8	31.8
Hispanic	10,065	5,235	52.0	3,138	31.2	1,135	11.3	557	5.5	48.0	16.8
Native	1,395	776	55.6	407	29.2	146	10.5	66	4.7	44.4	15.2
Asian	2,779	500	18.0	707	25.4	641	23.1	931	33.5	82.0	56.6
Black	3,602	2,152	59.7	990	27.5	331	9.2	129	3.6	40.3	12.8
White	29,256	8,808	30.1	9,801	33.5	5,914	20.2	4,733	16.2	69.9	36.4
Pacific Islander	395	192	48.6	118	29.9	55	13.9	30	7.6	51.4	21.5
Multi-Ethnic	2,127	693	32.6	705	33.1	394	18.6	335	15.7	67.4	34.3
Individualized Education Program	4,671	3,377	72.3	849	18.2	267	5.7	178	3.8	27.7	9.5
English Language Learner	3,666	2,604	71.0	836	22.8	166	4.6	60	1.6	29.0	6.2
Section 504	564	198	35.1	198	35.1	91	16.1	77	13.7	64.9	29.8
Economically Disadvantaged	20,393	10,035	49.2	6,548	32.1	2,535	12.4	1,275	6.3	50.8	18.7

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels:
Mathematics Grade 6

MATH Grade 6		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
ALL	29,946	10,363	34.6	9,721	32.5	5,764	19.2	4,098	13.7	65.4	32.9
Male	15,207	5,333	35.1	4,772	31.3	2,895	19.1	2,207	14.5	64.9	33.6
Female	14,739	5,030	34.1	4,949	33.6	2,869	19.5	1,891	12.8	65.9	32.3
Hispanic	7,013	3,405	48.6	2,271	32.3	941	13.5	396	5.6	51.4	19.1
Native	855	448	52.4	249	29.1	114	13.4	44	5.1	47.6	18.5
Asian	1,918	351	18.3	490	25.5	478	25.0	599	31.2	81.7	56.2
Black	2,061	1,116	54.1	630	30.6	233	11.3	82	4.0	45.9	15.3
White	19,032	5,669	29.8	6,397	33.6	4,045	21.3	2,921	15.3	70.2	36.6
Pacific Islander	264	110	41.7	88	33.3	46	17.4	20	7.6	58.3	25.0
Multi-Ethnic	1,383	467	33.8	429	31.0	273	19.7	214	15.5	66.2	35.2
Individualized Education Program	2,808	1,900	67.7	637	22.6	185	6.6	86	3.1	32.3	9.7
English Language Learner	2,010	1,445	71.9	461	22.9	74	3.7	30	1.5	28.1	5.2
Section 504	437	169	38.7	138	31.6	69	15.7	61	14.0	61.3	29.7
Economically Disadvantaged	13,562	6,461	47.6	4,442	32.8	1,857	13.7	802	5.9	52.4	19.6

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels:
Mathematics Grade 7

MATH Grade 7		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
ALL	28,271	10,241	36.2	8,736	30.9	5,699	20.2	3,595	12.7	63.8	32.9
Male	14,268	5,201	36.5	4,316	30.2	2,858	20.0	1,893	13.3	63.5	33.3
Female	14,003	5,040	36.0	4,420	31.6	2,841	20.2	1,702	12.2	64.0	32.4
Hispanic	10,095	5,065	50.2	3,259	32.3	1,317	13.0	454	4.5	49.8	17.5
Native	660	337	51.1	216	32.7	71	10.7	36	5.5	48.9	16.2
Asian	2,703	490	18.1	643	23.8	713	26.4	857	31.7	81.9	58.1
Black	1,602	885	55.2	478	29.9	180	11.2	59	3.7	44.8	14.9
White	12,776	3,388	26.5	4,000	31.3	3,304	25.9	2,084	16.3	73.5	42.2
Pacific Islander	391	164	41.9	138	35.3	66	16.9	23	5.9	58.1	22.8
Multi-Ethnic	922	315	34.2	282	30.6	173	18.7	152	16.5	65.8	35.2
Individualized Education Program	2,469	1,779	72.1	472	19.1	159	6.4	59	2.4	27.9	8.8
English Language Learner	2,842	2,119	74.6	555	19.5	124	4.4	44	1.5	25.4	5.9
Section 504	323	116	35.9	100	31.0	65	20.1	42	13.0	64.1	33.1
Economically Disadvantaged	14,592	7,084	48.5	4,641	31.9	2,076	14.2	791	5.4	51.5	19.6

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels:
Mathematics Grade 8

MATH Grade 8		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
ALL	34,880	13,230	37.9	10,489	30.1	6,752	19.4	4,409	12.6	62.1	32.0
Male	17,575	6,771	38.5	5,099	29.0	3,369	19.2	2,336	13.3	61.5	32.5
Female	17,305	6,459	37.3	5,390	31.2	3,383	19.5	2,073	12.0	62.7	31.5
Hispanic	13,554	6,860	50.6	4,179	30.8	1,856	13.7	659	4.9	49.4	18.6
Native	595	332	55.8	164	27.6	64	10.7	35	5.9	44.2	16.6
Asian	3,548	669	18.9	814	22.9	887	25.0	1,178	33.2	81.1	58.2
Black	1,934	1,098	56.8	525	27.1	216	11.2	95	4.9	43.2	16.1
White	14,846	4,212	28.4	4,703	31.6	3,606	24.3	2,325	15.7	71.6	40.0
Pacific Islander	223	119	53.4	58	26.0	35	15.7	11	4.9	46.6	20.6
Multi-Ethnic	1,177	403	34.2	352	29.9	242	20.6	180	15.3	65.8	35.9
Individualized Education Program	2,856	2,094	73.3	539	18.9	151	5.3	72	2.5	26.7	7.8
English Language Learner	3,217	2,397	74.5	650	20.2	126	3.9	44	1.4	25.5	5.3
Section 504	401	172	42.9	111	27.7	73	18.2	45	11.2	57.1	29.4
Economically Disadvantaged	17,722	8,784	49.6	5,416	30.5	2,519	14.2	1,003	5.7	50.4	19.9

Number and Percent of Students in Field Test Demographic Sample at Different Achievement Levels:
Mathematics Grade 11

MATH Grade 11		Level 1		Level 2		Level 3		Level 4		% At or Above	
Sub Group	Total	N	Pct	N	Pct	N	Pct	N	Pct	Level 2	Level 3
ALL	21,250	8,581	40.4	5,695	26.8	4,544	21.4	2,430	11.4	59.6	32.8
Male	10,518	4,204	40.0	2,731	25.9	2,249	21.4	1,334	12.7	60.0	34.1
Female	10,732	4,377	40.8	2,964	27.6	2,295	21.4	1,096	10.2	59.2	31.6
Hispanic	9,637	5,059	52.5	2,665	27.6	1,480	15.4	433	4.5	47.5	19.9
Native	187	103	55.1	43	23.0	30	16.0	11	5.9	44.9	21.9
Asian	3,184	599	18.8	715	22.5	981	30.8	889	27.9	81.2	58.7
Black	1,076	619	57.5	269	25.0	150	14.0	38	3.5	42.5	17.5
White	6,772	2,099	31.0	1,892	27.9	1,783	26.4	998	14.7	69.0	41.1
Pacific Islander	141	79	56.0	36	25.6	19	13.4	7	5.0	44.0	18.4
Multi-Ethnic	472	164	34.7	126	26.7	122	25.9	60	12.7	65.3	38.6
Individualized Education Program	1,158	876	75.6	195	16.9	60	5.2	27	2.3	24.4	7.5
English Language Learner	1,592	1,257	79.0	245	15.3	71	4.5	19	1.2	21.0	5.7
Section 504	261	94	36.0	79	30.3	59	22.6	29	11.1	64.0	33.7
Economically Disadvantaged	11,064	5,524	49.9	2,975	26.9	1,872	16.9	693	6.3	50.1	23.2

Overview of Achievement Level Setting Process

Joe Willhoft, Ph.D., Smarter Balanced
Michael B. Bunch, Ph.D., Measurement Incorporated



Smarter Balanced Chiefs' Meeting
November 6, 2014

Motions



Motion 1: Suspension of Rules

Smarter Balanced temporarily suspends its rules to allow Affiliate states to be included as voting members in matters related to Achievement Level Setting, in accordance with the voting rules described in the Smarter Balanced Governance Document as amended on August 26, 2014.

Motion 2: Endorse ALS Process

After review of the auditors' report, In-Person panelist feedback and endorsements by both the Achievement Level Setting Advisory Panel and the Smarter Balanced Technical Advisory Committee, Smarter Balanced states affirm that the Achievement Level Setting events and activities thoroughly and faithfully adhered to the plans for Achievement Level Setting that had previously been approved in April 2014.

Motion 3: Cut Scores for High School

Smarter Balanced states approve the achievement levels for Grade 11 English Language Arts/Literacy and for Grade 11 Mathematics as shown in the table below, displayed in vertically scaled logits.

Motion 4: Cut Scores for Grades 3-8

Smarter Balanced states approve the achievement levels for Grades 3-8 English Language Arts/Literacy and for Grades 3-8 Mathematics as shown in the table below, displayed in vertically scaled logits.

Overview



Completion of Recruiting Activities

- Online Panel
 - Over 10,000 registered
- In-Person Panel
 - 504 panelists confirmed
 - 10+ alternates confirmed
- VAC
 - 64 panelists confirmed

Completion of the Online Panel

- Opening of the Window
 - October 6-17
 - Individual windows and total window extended
- Level of Participation
 - 10,099 registered
 - 5,840 logged in
 - 2,660 submitted
- Support Provided
- Results Shared with In-Person Panel

Completion of the In-Person Panel

- Training Activities
 - Software
 - Common Core and ALDs
 - Ordered Item Booklet
- Panel Activities
 - Bookmark placement
 - Discussion

Agenda

Day 1

- Morning: Orientation
 - Common Core
 - Achievement Level Descriptors
 - Smarter Balanced Tests
 - Software
- Afternoon: Review of Ordered Item Booklet
 - Discuss items with others at table
 - Study additional resource materials

Agenda

Day 2

- Morning: Round 1
 - Practice in breakout rooms
 - Make bookmark recommendations individually
- Afternoon: Round 2
 - Discuss Round 1 bookmark placements at each table
 - Place Round 2 bookmarks individually

Day 3

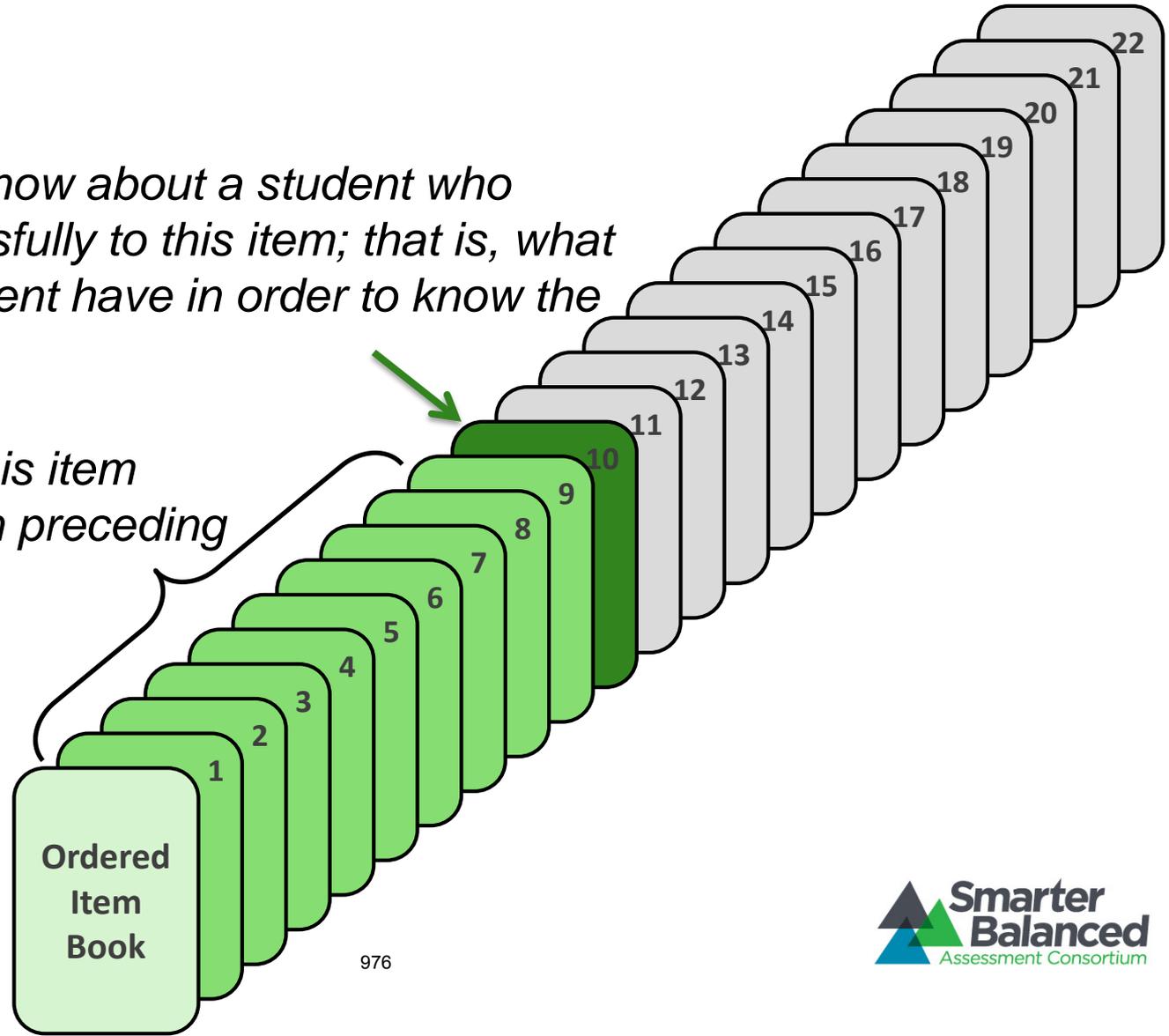
- Morning: Round 3
 - Discuss Round 2 bookmark placements for the entire room
 - View supporting data based on Round 2 bookmarks
 - Place Round 3 bookmarks individually
 - Review final recommendations
 - Evaluate the process

Training

As you study each item in the OIB, discuss two questions with your fellow panelists:

1. *What do you know about a student who responds successfully to this item; that is, what skills must a student have in order to know the correct answer?*

2. *What makes this item more difficult than preceding items?*



2. *What makes this item more difficult than preceding items?*

35758

Read the sentence and the question that follows.

My aunt was hungry and wanted to eat before the movie but the rest of the family
wanted to see it before dinner.

What is the **best** way to correct the error in the sentence?

- (A) Add a hyphen (-) after *movie*.
- (B) Add a comma (,) after *movie*.
- (C) Add a semicolon (;) after *movie*.
- (D) Add an exclamation mark (!) after *movie*.

Ordered Item Booklet: Achievement Level Setting

English Language Arts/Literacy, 4

Page: 60

[Back to Item Map](#)

◀ Previous

Next ▶

Item

Question Information

Passages and Other Materials **3**

Achievement Level Descriptors **1**

57202

In Source #1, the author says that "friend behavior can be important to an animal's survival." Explain what the author means and give **two** details, using **two** sources, that support your explanation. For each detail include the source title or number.

Achievement Level Descriptors

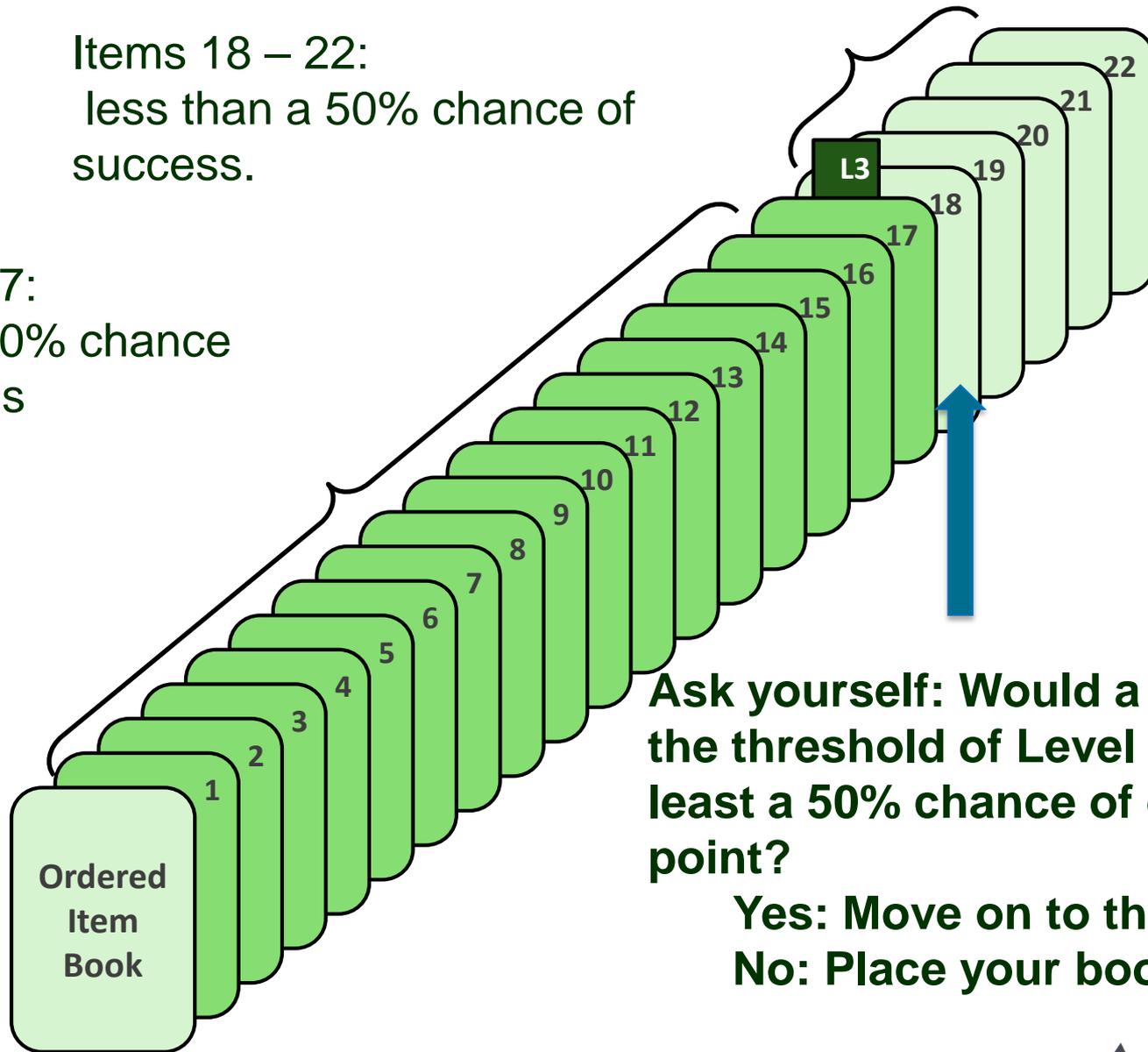
The student who just enters Level 3 should be able to:

**Targets
1–7:
Reading
Literary
Text**

- Use details and information from texts of moderate complexity to support answers and inferences.
- Identify or summarize central ideas/key events in texts of moderate complexity.
- Begin to determine the intended meanings of words, including words with multiple meanings, based on context, word relationships, word structure, and use of resources in texts of moderate complexity.
- Use supporting evidence to justify/explain own inferences in texts of moderate complexity.
- Interpret, specify, or compare how information is presented across texts of moderate complexity.
- Begin to relate knowledge of text structures, genre-specific features, or formats to obtain, interpret, explain, or connect information within texts of moderate complexity.
- Determine or interpret figurative language, literary devices, or connotative meanings of words and phrases used in context and partially explain the impact of those word choices on meaning and tone in texts of moderate complexity.

Items 18 – 22:
less than a 50% chance of
success.

Items 1-17:
At least 50% chance
of success



Ask yourself: Would a student at the threshold of Level 3 have at least a 50% chance of earning this point?

**Yes: Move on to the next item.
No: Place your bookmark here.**

Evaluations

How confident are you about the three bookmarks you just entered?

Bookmark	Very Confident	Confident	Uncertain	Very Uncertain	Total
Level 2	222 (47%)	237 (51%)	10 (2%)	0 (0%)	469
Level 3	234 (50%)	220 (47%)	15 (3%)	0 (0%)	469
Level 4	245 (52%)	217 (46%)	7 (1%)	0 (0%)	469

Overwhelming endorsement of process (92-99% positive) on 14 separate measures.

Completion of Vertical Articulation

- Training
 - Orientation
 - Groundrules
- Procedures
 - Motion
 - Second
 - Discussion
 - Vote: 2/3 majority required
- Results
 - Aligned cut scores across grades
 - Eliminated scaled score reversals

Audit



Focus of Audit

- Pilot Test
- Software
- Online Panel Process
- In-Person Workshop
- Vertical Articulation
- Adherence to Plan

Audit Report

“It is my conclusion that the standard setting activities described in this report were designed and conducted appropriately so as to yield defensible performance standards grounded in the knowledge, skills, and expectations represented by the ALDs.”

(Auditors’ Report, p. 52)

Advisory Panel Statement

After reviewing the information about the Achievement Level Setting activities and the auditors' report, the Advisory Panel confirms the design and procedures for the Achievement Level Setting and the Vertical Articulation were implemented as planned, represent a valid process that is consistent with best practices in standard setting, and support the defensibility of the content-based performance standards.

Unanimously endorsed October 28, 2014

TAC Statement

The Technical Advisory Committee concludes that the Smarter Balanced achievement level setting design and implementation reflect contemporary professional practice and represent a valid process that supports the defensibility of the content-based performance standards.

Unanimously endorsed October 30, 2014

Recommendations from ALS Panels



Overview of Achievement Level Setting Process

Joe Willhoft, Ph.D., Smarter Balanced
Michael B. Bunch, Ph.D., Measurement Incorporated



Smarter Balanced Chiefs' Meeting
San Diego, CA Nov. 14, 2014

Agenda

- 6:30 – 7:30 Dinner Buffet
- 7:00 – 7:15 Introductions/Overview
- 7:15 – 7:45 Achievement Level
Recommendations
- 7:45 – 9:00 Discussion
- 9:00 Next Steps / Adjourn

Motions



Motion 1: Adoption of Position Paper

Smarter Balanced member states endorse the position paper “Interpretation and Use of Scores and Achievement Levels” dated November 14, 2014, as amended, and recommend that member states attend to it when considering the use of scores from Smarter Balanced assessments.

Motion 2: Cut Scores for High School

Smarter Balanced states approve the achievement levels for Grade 11 in Mathematics and English language arts/Literacy as shown in the accompanying table, displayed in Smarter Balanced scale scores.
(staff recommendations)

Motion 3: Cut Scores for Grades 3-8

Smarter Balanced states approve the achievement levels for Grades 3-8 in Mathematics and English language arts/Literacy as shown in the accompanying table, displayed in Smarter Balanced scale scores.
(staff recommendations)

What Staff Has Done Since 11/6



1. Corrected Gr. 11 Impact Data

- A file transfer error was identified with Gr. 11 impact data. (Gr. 9 & 10 responses had been included with Gr. 11 files) (No effect on Ordered Item Booklet or on In-Person recommendations)
- Gr. 11 ELA impact data had been 42.6% “At or above level 3”; is now 42.8%
- Gr. 11 Math impact data had been 22.0% “At or above level 3”; is now 26.4%

2. Created Smarter Balanced Scale Scores

- Scale scores on Smarter Balanced assessments
 - Use a **four digit number**
 - Range roughly from **2000-to-3000**
- Avoids confusion with other, commonly-used scales such as: percentile ranks, percent correct, ACT, SAT, Lexiles, etc.
- Provides adequate “space” on the scale to see growth in whole-number units

3. Position Statement About Use of Ach. Levels & Scale Scores

- Collaboration among staff and advisors for purpose of articulating need for and pitfalls of using achievement levels
- Will be proposed for adoption by Smarter Balanced states on Nov 12

4. Developed Three Principles to Guide Staff Recommendations

1. Honor the work of the panelists who contributed their content expertise to the process
1. Give consideration to external evidence of student readiness for credit-bearing college work: Notably NAGB Grade 12 college preparedness data
1. Maintain a “system perspective” by using information from all grades (and both content areas) to make recommendations that support a coherent system

5. Developed Options for Consideration

- 2 options in Math and ELA respectively.
 - Original recommendation from VAC
 - Option 2 is the staff recommendation.
- The staff recommendation:
 - Adheres to the “Guiding Principles”
 - Closes the gap between Panel recommendations and NAGB estimates of “College Preparedness” – giving equal weight to Panel recommendations and NAGB findings
 - Leaves other Panel recommendations (L2 and L4) as is, with minor exceptions to separate L4 from L3

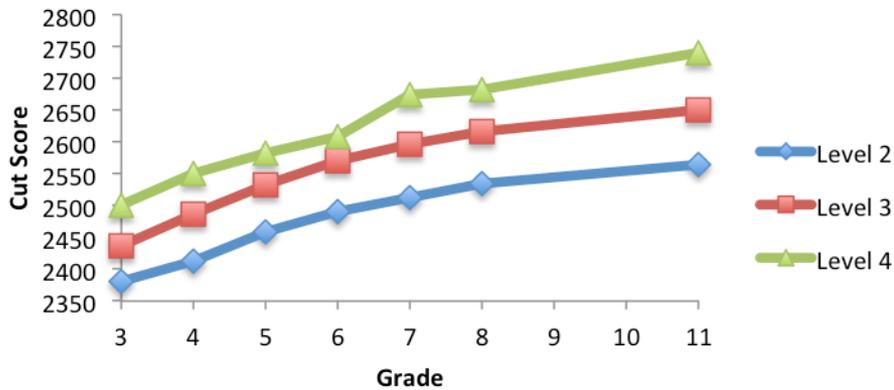
Recommendations -- Math



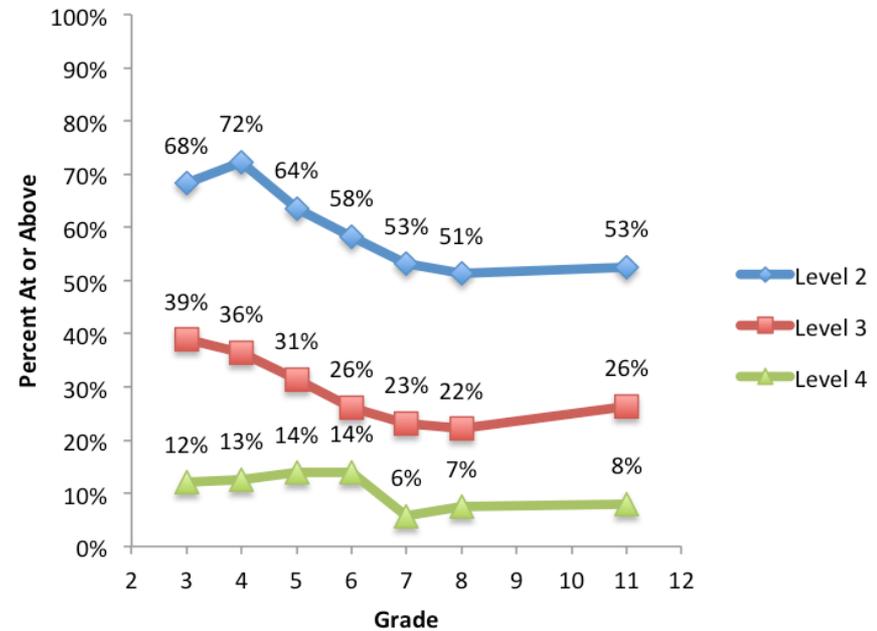
Math Option 1

VAC Recommendations

**Cut Scores - Math
(VAC Recommended)**



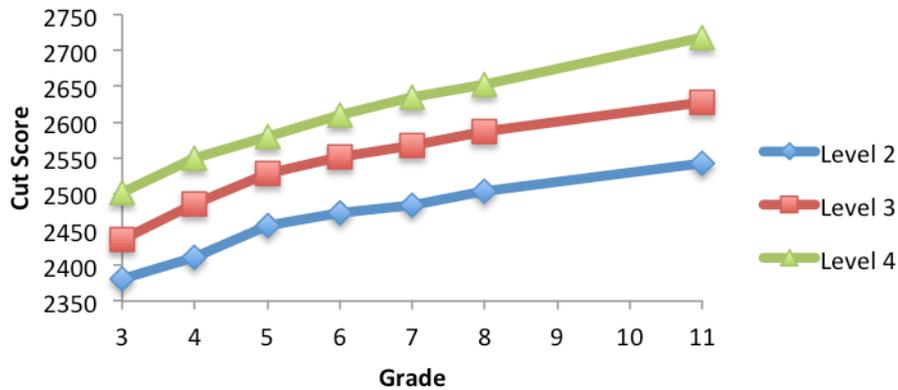
**Pct. At or Above - Math
(VAC Recommended)**



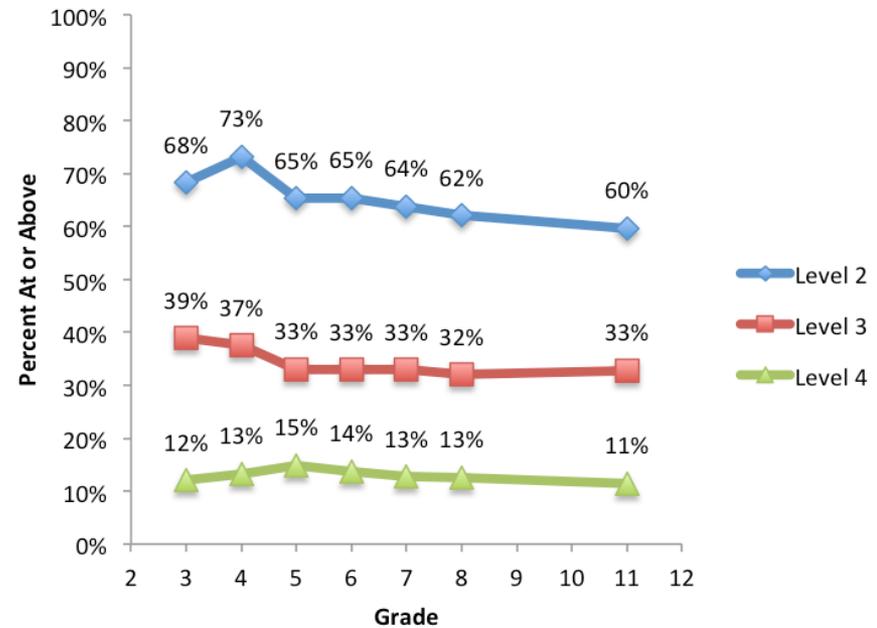
Math Option 2

Staff Recommendations

**Cut Scores - Math
(Staff Recommended)**



**Pct. At or Above - Math
(Staff Recommended)**



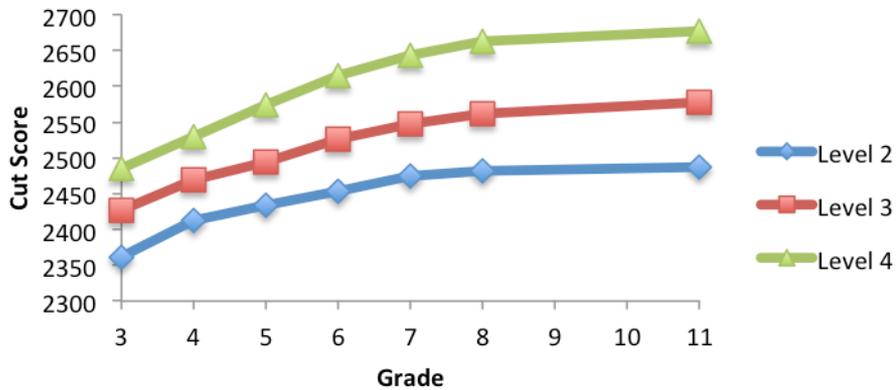
Recommendations -- ELA



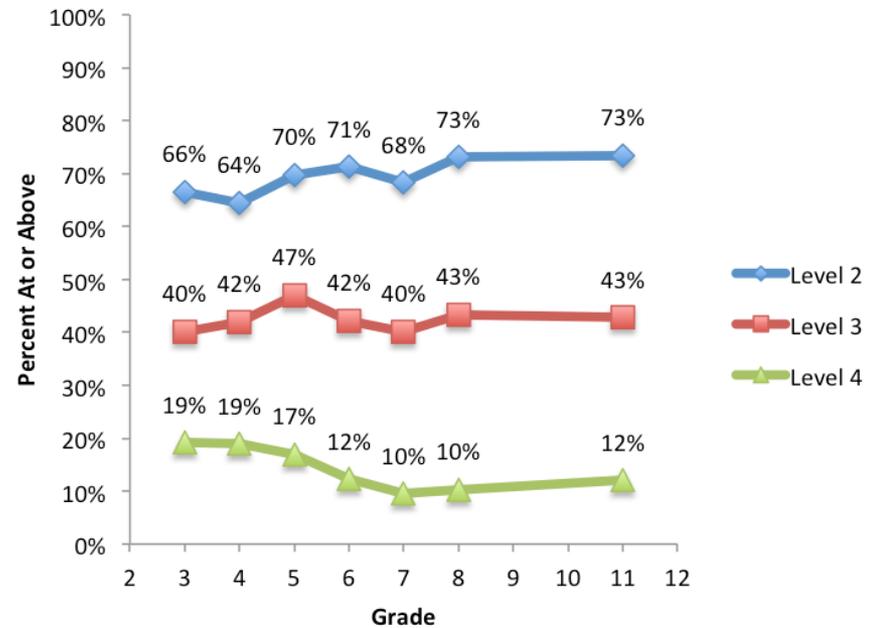
ELA Option 1

VAC Recommendations

**Cut Scores - ELA
(VAC Recommended)**



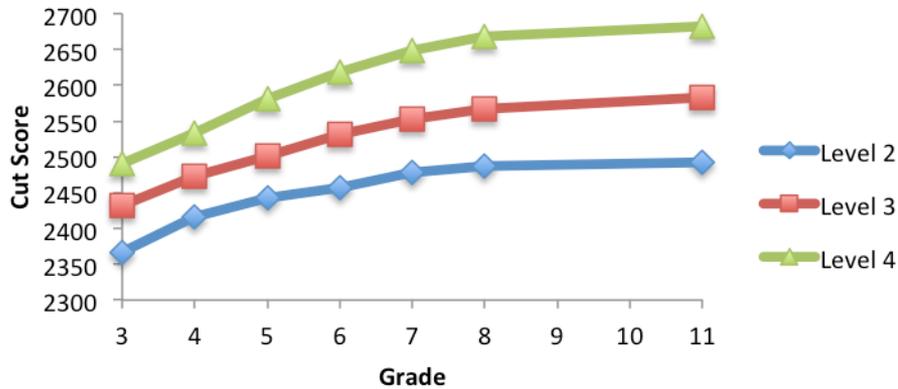
**Pct. At or Above - ELA
(VAC Recommended)**



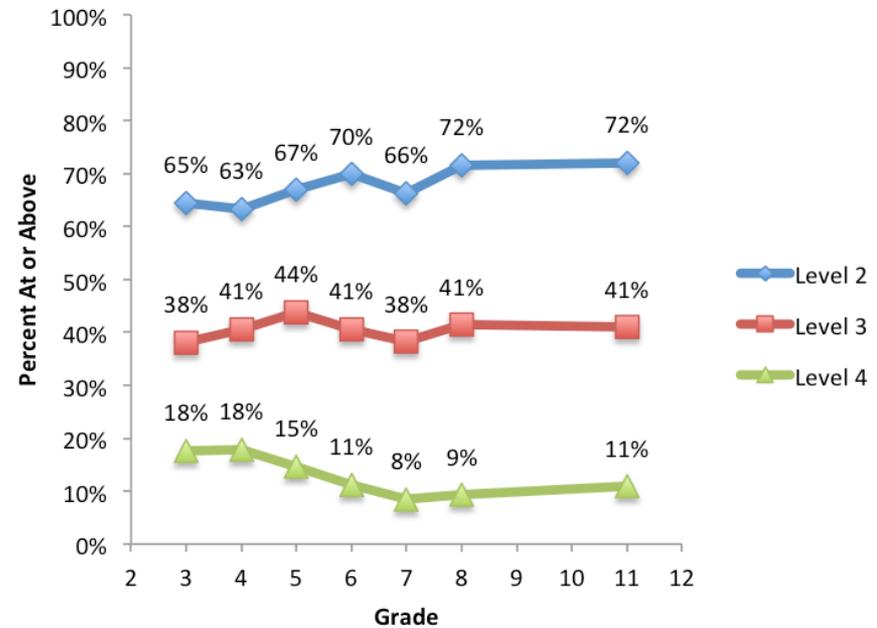
ELA Option 2

Staff Recommendations

**Cut Scores - ELA
(Staff Recommended)**



**Pct. At or Above - ELA
(Staff Recommended)**



Discussion



Gr. 11 Recommended Cut Scores (VAC and Staff)

Grade 11 (VAC)

Math Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
11	2565	2650	2740

ELA Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
11	2488	2578	2677

Grade 11 (Staff)

Math Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
11	2543	2628	2718

ELA Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
11	2493	2583	2682

Gr. 3-8 Recommended Cut Scores (VAC and Staff)

Grades 3-8

(VAC)

Math Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
3	2381	2436	2501
4	2413	2487	2551
5	2459	2532	2583
6	2491	2570	2609
7	2513	2596	2674
8	2534	2616	2683

ELA Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
3	2362	2427	2485
4	2413	2470	2530
5	2434	2494	2574
6	2453	2527	2614
7	2474	2547	2644
8	2482	2562	2663

Grades 3-8

(Staff)

Math Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
3	2381	2436	2501
4	2411	2485	2549
5	2455	2528	2579
6	2473	2552	2610
7	2484	2567	2635
8	2504	2586	2653

ELA Scale Score Cuts			
Grade	Level 2	Level 3	Level 4
3	2367	2432	2490
4	2416	2473	2533
5	2442	2502	2582
6	2457	2531	2618
7	2479	2552	2649
8	2487	2567	2668

Appendix I – Field Test Documentation

Online Panel

- I01_Bookmark Training OP Field Test1011
- I02_Bookmark OP Field Test Instructions1039
- I03_Online Field Test Survey Results1046

In-Person Workshop

- I04_List of Field Test Materials1059
- I05_In-Person Field Test Intro1060
- I06_In-Person Field Test Survey1087

Vertical Articulation

- I07_Vertical Articulation Intro1089
- I08_VAC Readiness form1102
- I09_VAC Script1103

Evaluation

- I10_Field Test Debrief1107



The Bookmark Procedure



Online Panel Field Test
August 14-15, 2014



Bookmark Overview

- The Bookmark Procedure is a process used to set cut scores that define achievement levels by comparing and aligning items and content to Achievement Level Descriptors
- It is so named because you place bookmarks in an ordered item booklet of test items to make your cut score recommendations.
- Item maps guide you through the ordered item booklet and include additional information to provide context.



Ordered Item Books (OIBs)

- One item per page
- Easiest item first, hardest item last
- Items ascend by difficulty
- Multi-point items appear on multiple pages in the OIB— one page per positive score point.



Item Map

Online Panel for Achievement Level Setting



Instructions Item Map

Hello dmcdaniel@measinc.com!

Log off

Item Map for Booklet: Demo booklet

Language Arts, 7

Round: 1

display standard comparisons:

ACT SAT

Enter Comment

Page	Ability	Standard Code	DOK	Bookmark	Comments	Reviewed
01	4.1	S.B.123	1			<input type="checkbox"/>
02	4.11	S.B.123	1			<input type="checkbox"/>
03	4.12	S.B.123	1			<input type="checkbox"/>
04	4.13	S.B.123	1			<input type="checkbox"/>
05	4.14	S.B.123	1			<input type="checkbox"/>
06	4.15	S.B.123	1			<input type="checkbox"/>
07	4.16	S.B.123	1			<input type="checkbox"/>
08	4.17	S.B.123	1			<input type="checkbox"/>
09	4.18	S.B.123	1			<input type="checkbox"/>
10	4.19	S.B.123	1			<input type="checkbox"/>
11	4.2	S.B.123	1			<input type="checkbox"/>
12	4.21	S.B.123	1			<input type="checkbox"/>
13	4.22	S.B.123	1			<input type="checkbox"/>

Set Bookmark

Submit Bookmarks





Study the Items in the OIB

- When we examine each item in the OIB we will discuss two questions with our fellow participants:
 - *What do you know about a student who responds successfully to this item or score point? Think about what a student needs to know and be able to do in order to earn this point?*
 - *What makes this item more difficult than preceding items?*



Ordered Item 1

1

27. Kitty is taking a trip on which she plans to drive 300 miles each day. Her trip is 1,723 miles long. She has already driven 849 miles. How much farther must she drive?
- A. 574 miles
 - B. 874 miles
 - C. 1,423 miles
 - D. 2,872 miles



Ordered Item 2

2

6. Compare how Billy feels when he first realizes there is a mountain lion nearby to the way he feels by the end of the story. Use information from the story to support your response.

Sample response that received a score of 1 out of 4



Ordered Item 3

3

CARTONS OF EGGS SOLD LAST MONTH

Farm A ○ ○ ○ ○
Farm B ○ ○ ○ ○ ○ ○
Farm C ○ ○ ○
Each ○ = 100 Cartons

4. According to the graph how many cartons of eggs were sold by farms A, B, and C month?
- A. 13
 - B. 130
 - C. 1,300
 - D. 13,000



Ordered Item 4

17

- 6 Compare how Billy feels when he first realizes there is a mountain lion nearby to the way he feels by the end of the story. Use information from the story to support your response.

Sample response that received a score of 2 out of 4



Item Map for Training

Online Panel for Achievement Level Setting



Instructions Item Map

Hello dmcdaniel@measinc.com!

Log off

Item Map for Booklet: Demo booklet

Language Arts, 7

Round: 1

display standard comparisons:

ACT SAT

Enter Comment

Page	Ability	Standard Code	DOK	Bookmark	Comments	Reviewed	ACT	SAT
01	4.1	S.B.123	1			<input type="checkbox"/>		
02	4.11	S.B.123	1			<input type="checkbox"/>		
03	4.12	S.B.123	1			<input type="checkbox"/>		
04	4.13	S.B.123	1			<input type="checkbox"/>	[Green Block]	
05	4.14	S.B.123	1			<input type="checkbox"/>		[Green Block]
06	4.15	S.B.123	1			<input type="checkbox"/>		[Green Block]
07	4.16	S.B.123	1			<input type="checkbox"/>		[Green Block]
08	4.17	S.B.123	1			<input type="checkbox"/>		[Green Block]
09	4.18	S.B.123	1			<input type="checkbox"/>		[Green Block]
10	4.19	S.B.123	1			<input type="checkbox"/>		[Green Block]
11	4.2	S.B.123	1			<input type="checkbox"/>		[Green Block]
12	4.21	S.B.123	1			<input type="checkbox"/>		[Green Block]
13	4.22	S.B.123	1			<input type="checkbox"/>	[Green Block]	[Green Block]

Set Bookmark

Submit Bookmarks



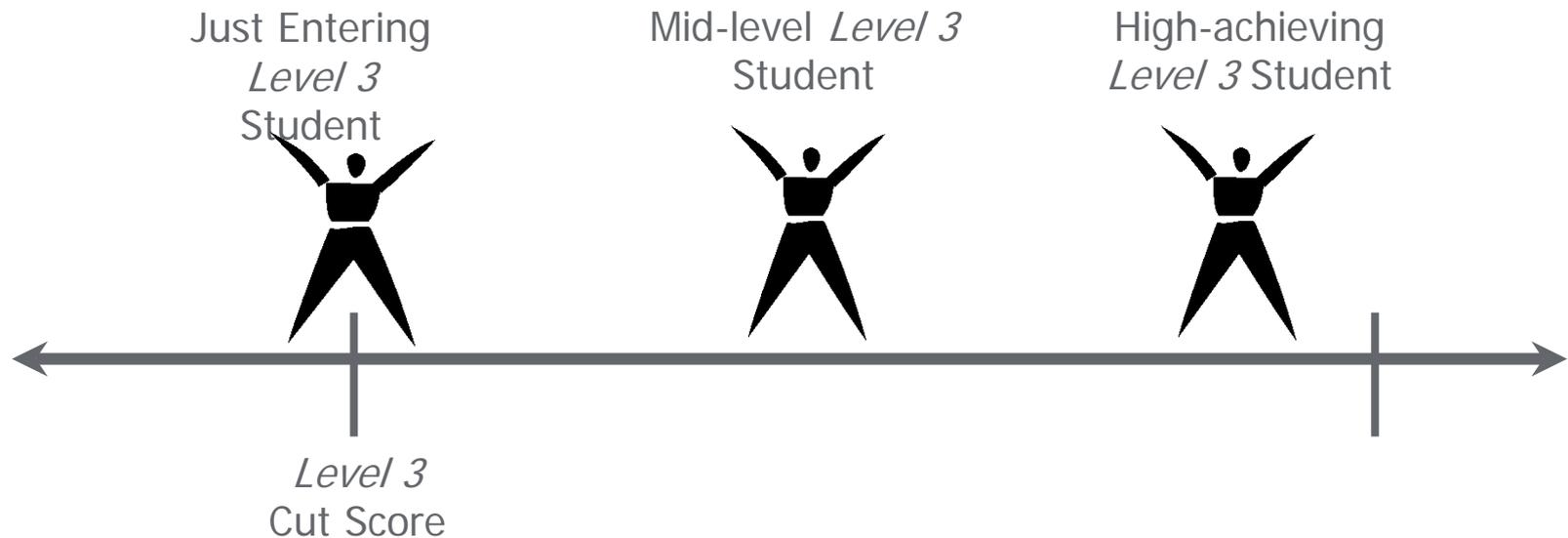


Our First Task is to Understand What the Test Measures

- We will study the ordered item book, item by item, answering the two questions, until we complete the OIB
 - *What do you know about a student who responds successfully to this item or score point? Think about what a student needs to know and be able to do in order to earn this point?*
 - *What makes this item more difficult than preceding items?*
- Only then have we acquired the knowledge to make content-based recommendations for cut scores

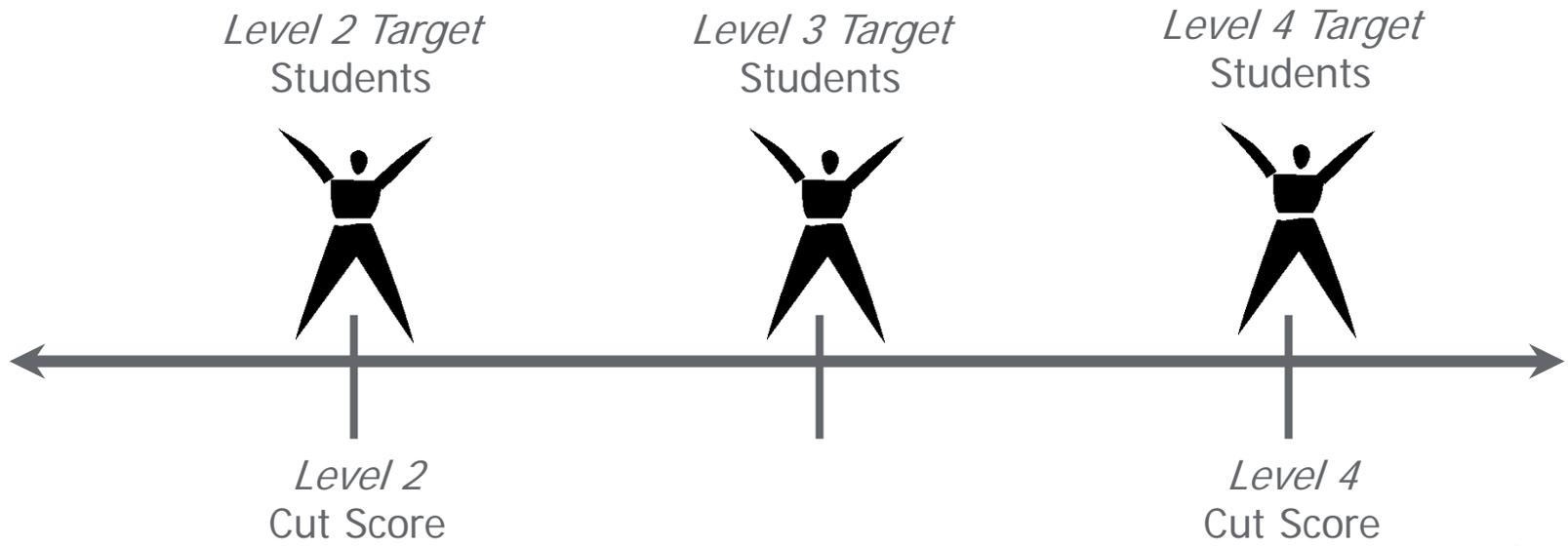
Target Student

- We recommend cut scores for Target Students—the “just-entering” student
 - These are the skills of the student just entering the performance level.

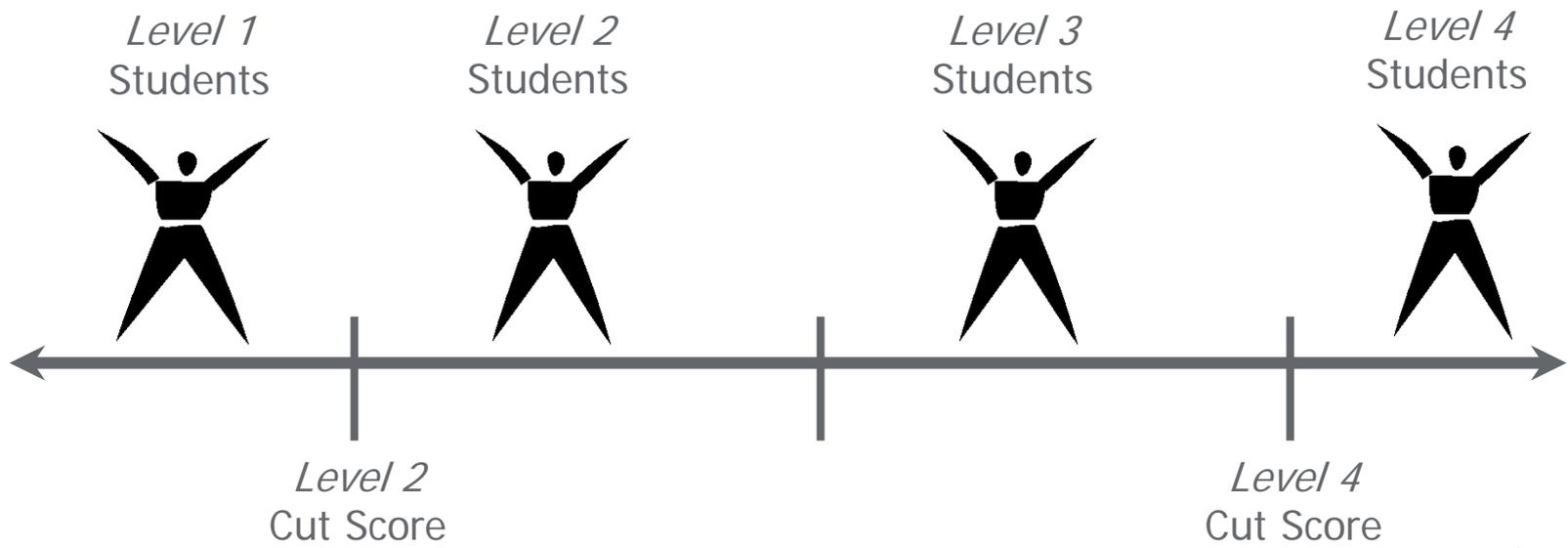




Three Target Students



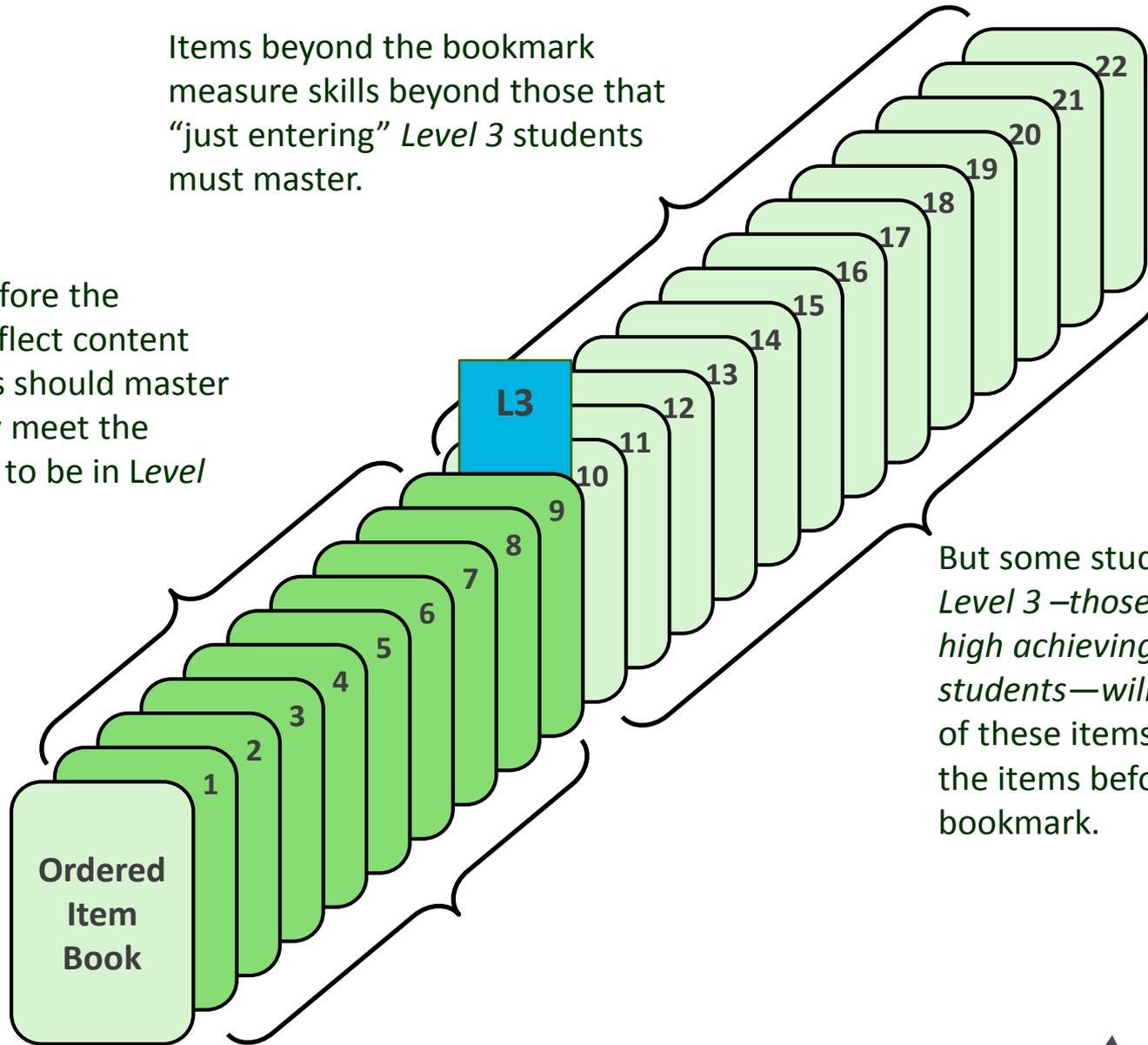
Three Cut Scores Define Four Achievement Levels





Items beyond the bookmark measure skills beyond those that “just entering” *Level 3* students must master.

The items before the bookmark reflect content that students should master to just barely meet the expectations to be in *Level 3*.

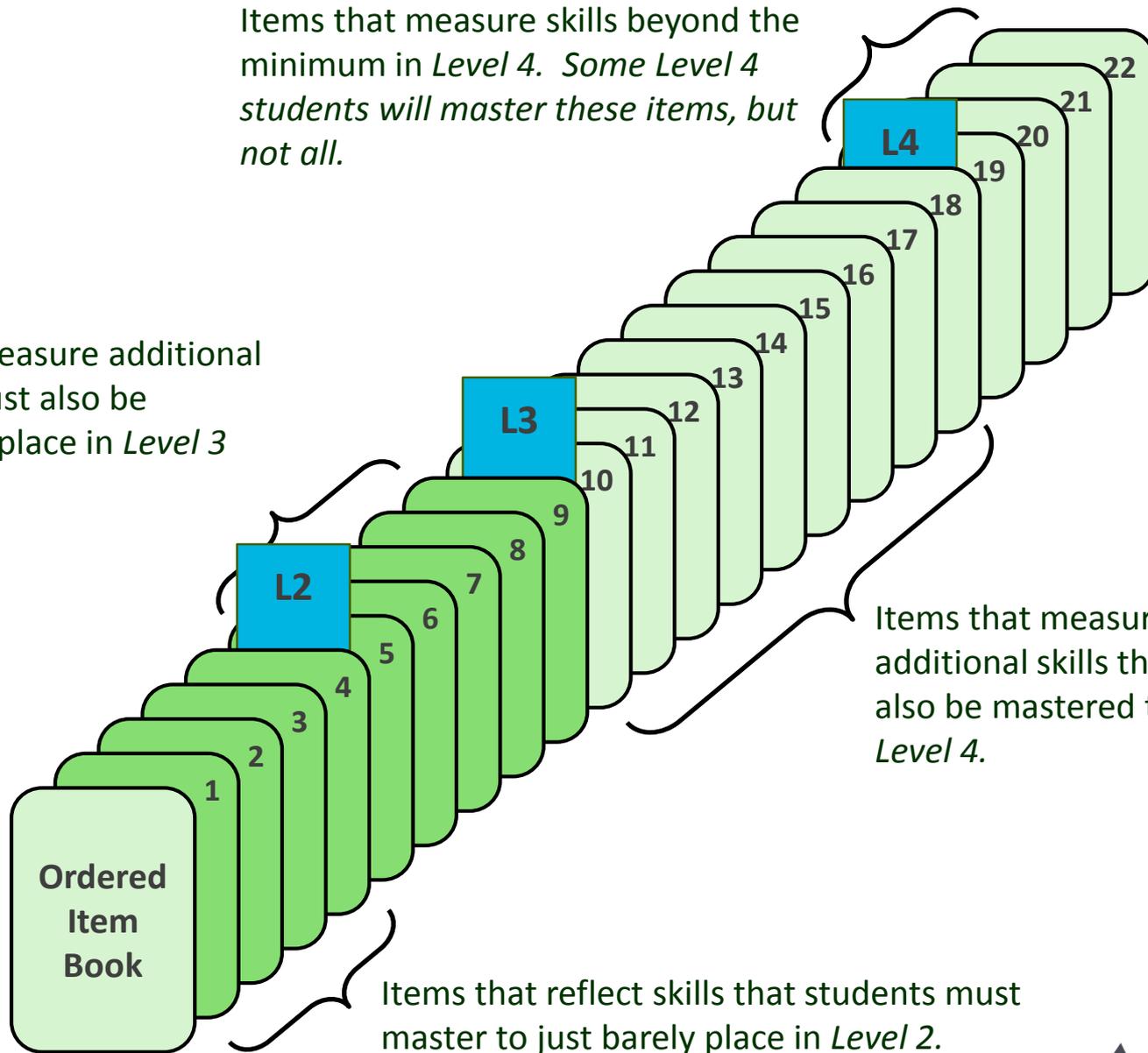


But some students in *Level 3*—those *mid-level* or *high achieving Level 3* students—will master some of these items in addition to the items before the bookmark.



Items that measure skills beyond the minimum in *Level 4*. Some *Level 4* students will master these items, but not all.

Items that measure additional skills that must also be mastered to place in *Level 3*

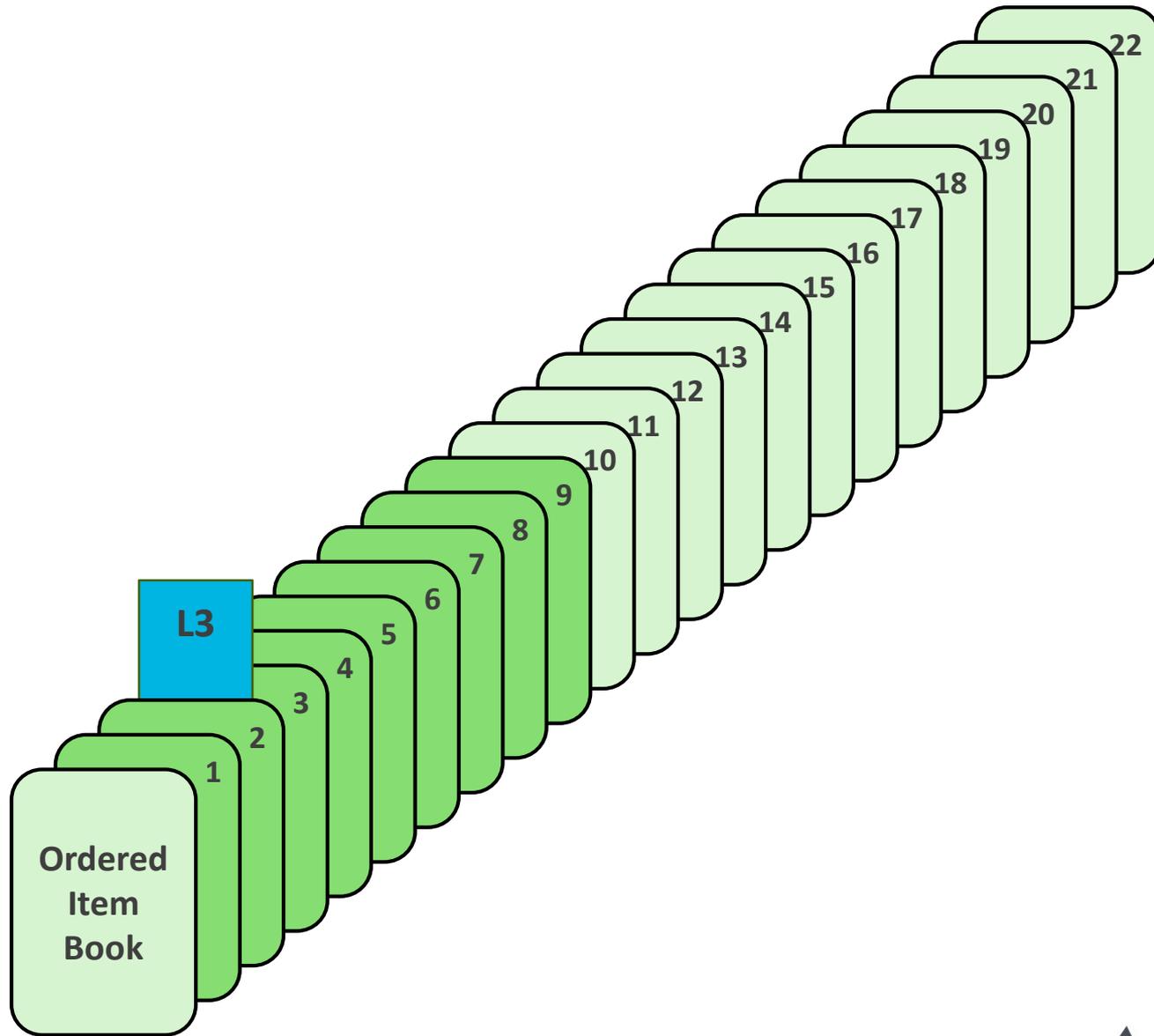


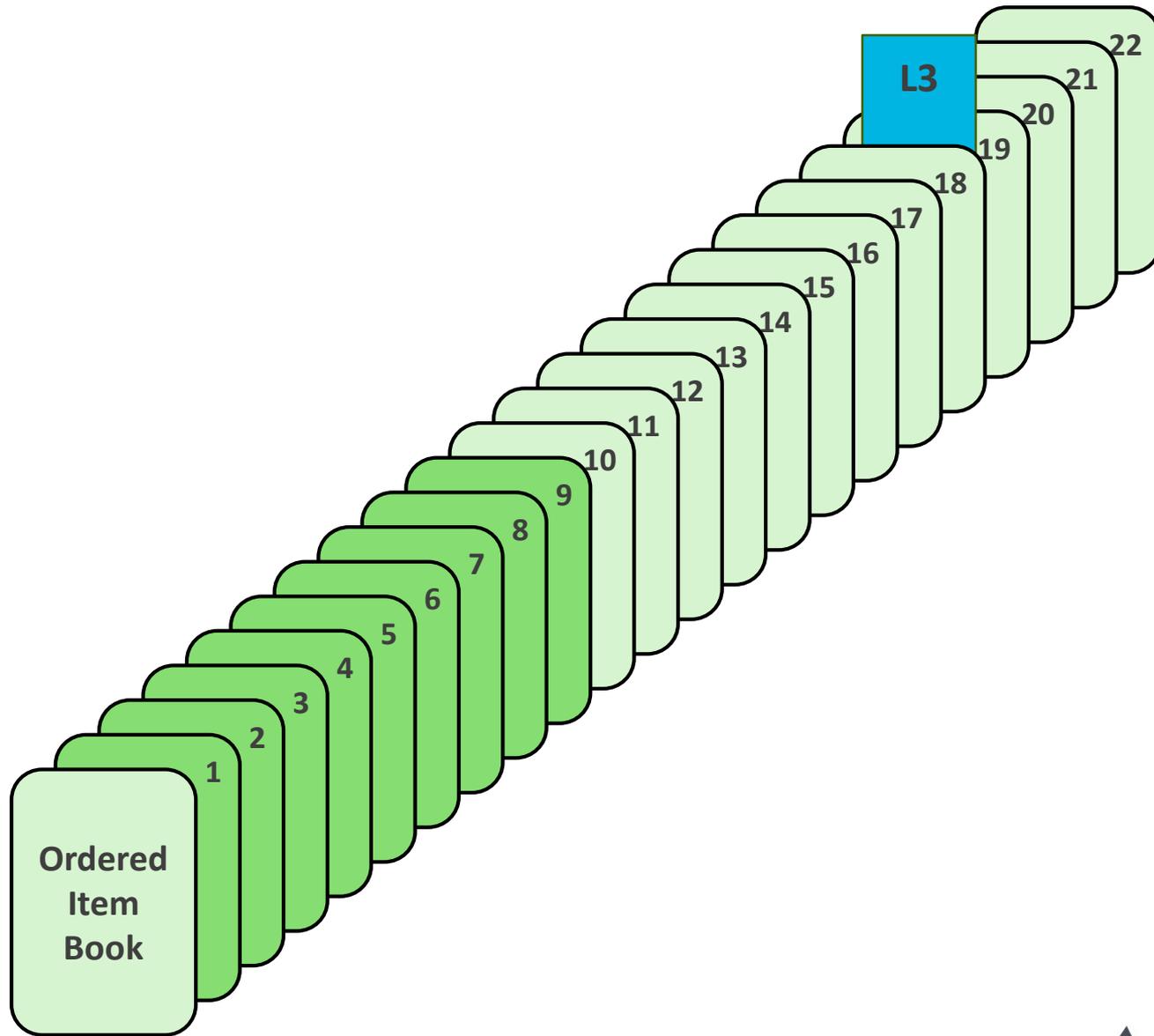
Items that measure additional skills that must also be mastered to place in *Level 4*.

Items that reflect skills that students must master to just barely place in *Level 2*. Students that can't demonstrate this level of mastery will be in *Level 1*.

Level 3 Bookmark Placement

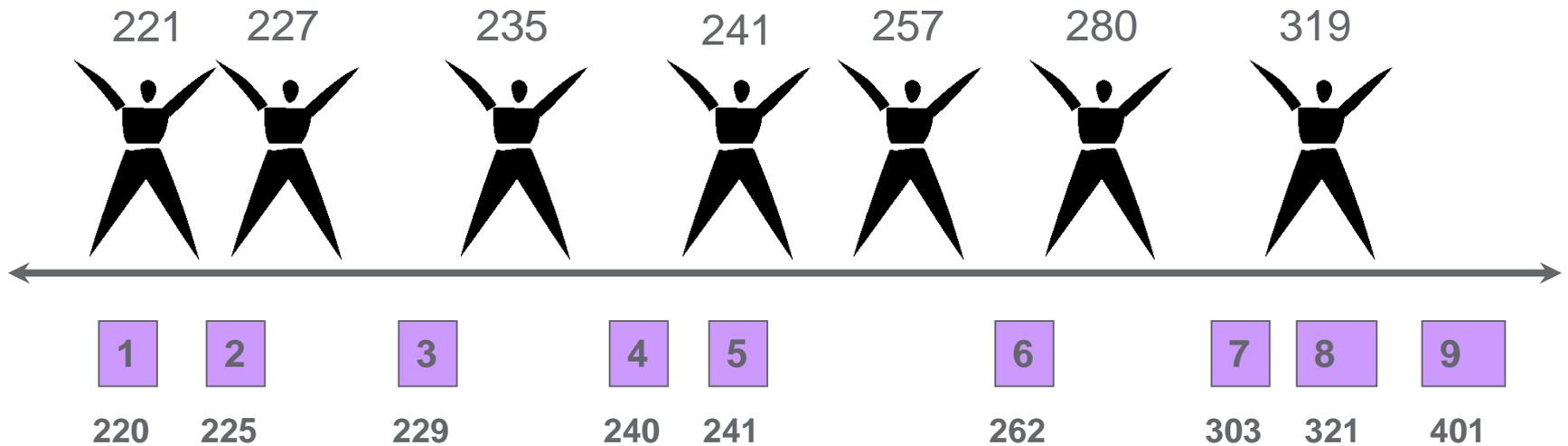
- When you place your *Level 3* bookmark, think about a student who has mastery of the content measured by the items before the bookmark.
- Examine the knowledge, skills, and abilities held by that student, and compare them to the *Level 3* target student descriptor.
 - Ask yourself: Would a student who has mastered the content prior to the bookmark demonstrate sufficient skills to infer that the student is in *Level 3*? Do those skills align with the target student descriptors?
 - Place your *Level 3* bookmark at the first point where you could make this inference.





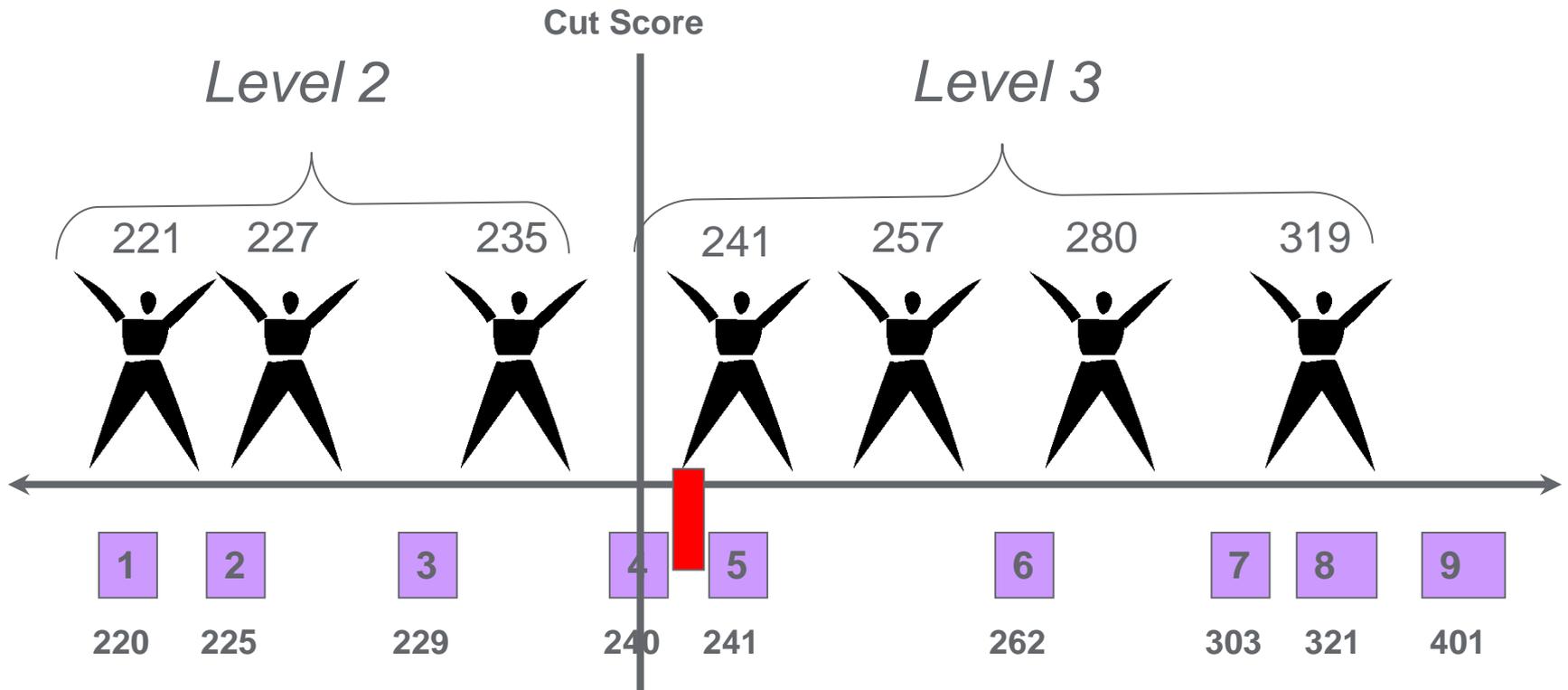


Test Location



Items are ordered by difficulty.
Students are ordered by ability.

The Bookmark and the Cut Score



The bookmark separates items and content.

The cut score separates students.

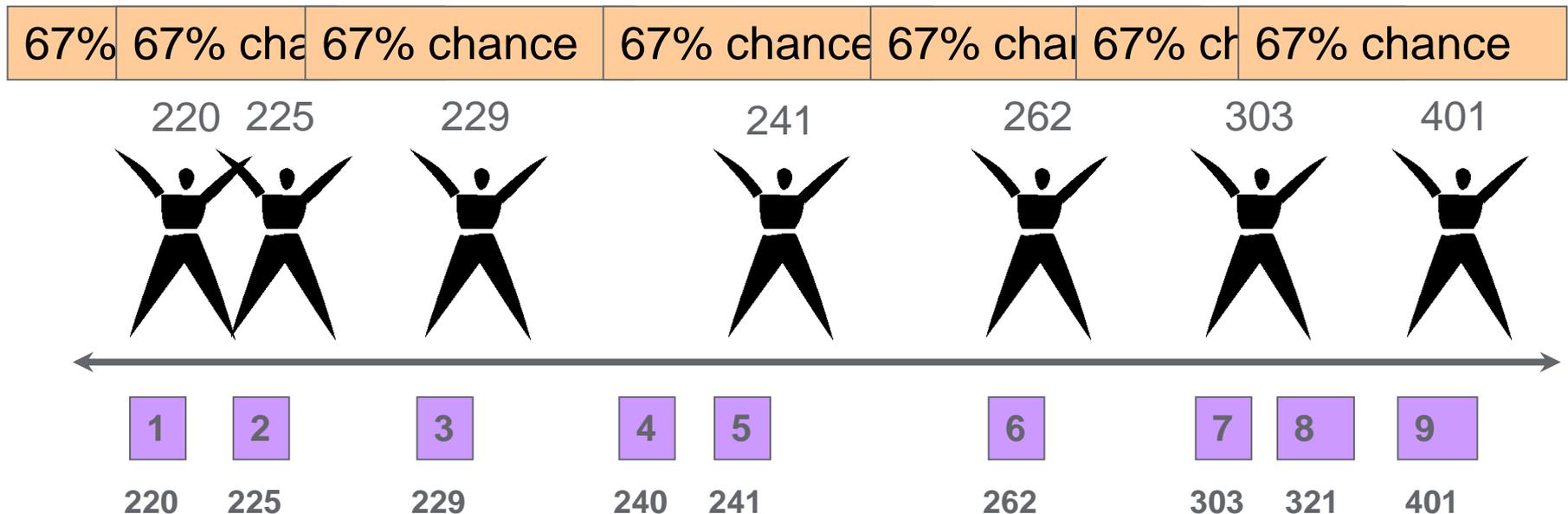


Mastery

- Students demonstrate mastery when they have at least a [XX]% chance of answering an item correctly.
 - The decision to use [XX]% was based on research.



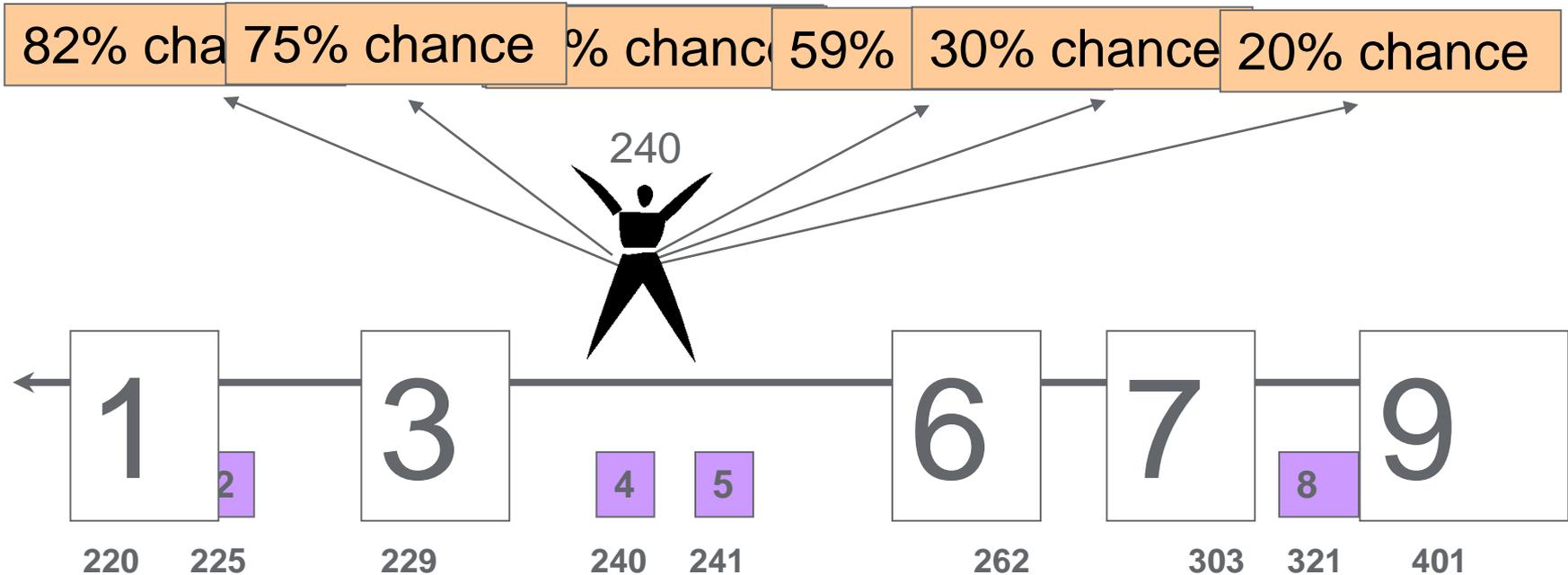
Item Location



Location is an indication of difficulty.

Location represents the ability level necessary to have a [XX]% chance of answering the item correctly.

Target Student

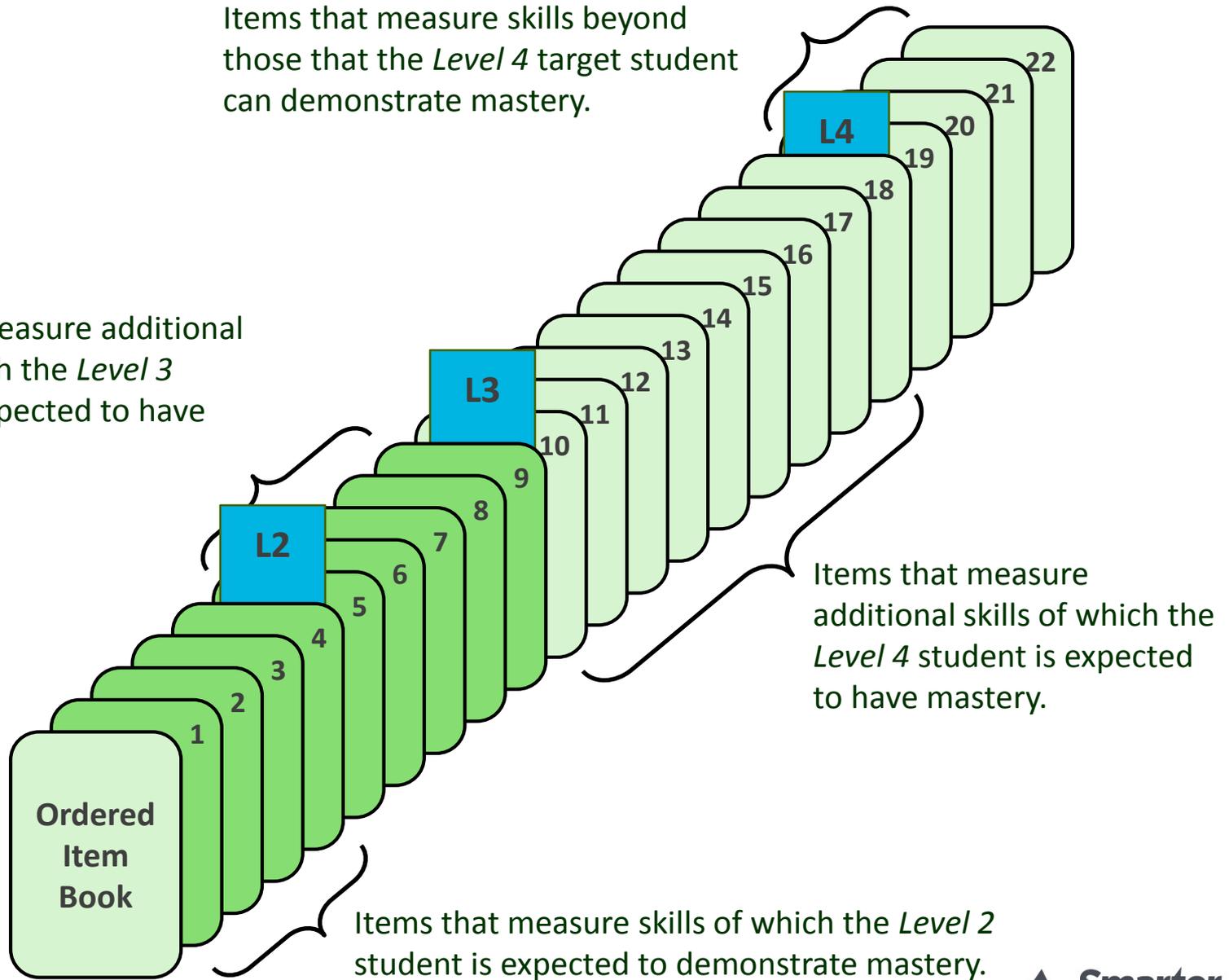


A student right at the cut score will have at least a **[XX]**% chance of answering the items correctly at and below the cut score.



Items that measure skills beyond those that the *Level 4* target student can demonstrate mastery.

Items that measure additional skills of which the *Level 3* student is expected to have mastery.



Setting Bookmarks

Online Panel for Achievement

Instructions Item Map

Smarter Balanced Assessment Consortium

dmcdaniel@measinc.com! Log off

Item Map for Booklet: D

Language Arts, 7

Round: 2

display standard comparisons:
 ACT SAT

Enter Comment

Page	Ability	Standard Code	DOK	Bookmark	Comments	Reviewed	ACT	SAT	At or Above
01	4.1	S.B.123	1			<input checked="" type="checkbox"/>	0%	0%	98%
02	4.11	S.B.123	1			<input checked="" type="checkbox"/>	0%	0%	90%
03	4.12	S.B.123	1			<input checked="" type="checkbox"/>	0%	0%	84%
04	4.13	S.B.123	1			<input checked="" type="checkbox"/>		0%	82%
05	4.14	S.B.123	1			<input checked="" type="checkbox"/>			74%
06	4.15	S.B.123	1			<input checked="" type="checkbox"/>			70%
07	4.16	S.B.123	1			<input checked="" type="checkbox"/>			65%
08	4.17	S.B.123	1			<input checked="" type="checkbox"/>			63%
09	4.18	S.B.123	1			<input checked="" type="checkbox"/>			62%
10	4.19	S.B.123	1			<input checked="" type="checkbox"/>		0%	60%
11	4.2	S.B.123	1			<input checked="" type="checkbox"/>	0%		58%
12	4.21	S.B.123	1			<input checked="" type="checkbox"/>	0%		57%
13	4.22	S.B.123	1			<input checked="" type="checkbox"/>			55%

Submit Bookmark Cancel

Set Bookmark

Submit Bookmarks

Setting Bookmarks

Online Panel for Achievement

Instructions Item Map

Smarter Balanced Assessment Consortium

dmcdaniel@measinc.com! Log off

Item Map for Booklet: D

Language Arts, 7

Round: 2

display standard comparisons:
 ACT SAT

Enter Comment

Page	Ability	Standard Code	DOK	Bookmark	Comments	Reviewed	ACT	SAT	At or Above
01	4.1	S.B.123	1			<input checked="" type="checkbox"/>	0%	0%	98%
02	4.11	S.B.123	1			<input checked="" type="checkbox"/>	0%	0%	90%
03	4.12	S.B.123	1			<input checked="" type="checkbox"/>	0%	0%	84%
04	4.13	S.B.123	1			<input checked="" type="checkbox"/>		0%	82%
05	4.14	S.B.123	1			<input checked="" type="checkbox"/>			74%
06	4.15	S.B.123	1			<input checked="" type="checkbox"/>			70%
07	4.16	S.B.123	1			<input checked="" type="checkbox"/>			65%
08	4.17	S.B.123	1			<input checked="" type="checkbox"/>			63%
09	4.18	S.B.123	1			<input checked="" type="checkbox"/>			62%
10	4.19	S.B.123	1			<input checked="" type="checkbox"/>		0%	60%
11	4.2	S.B.123	1			<input checked="" type="checkbox"/>	0%		58%
12	4.21	S.B.123	1			<input checked="" type="checkbox"/>	0%		57%
13	4.22	S.B.123	1			<input checked="" type="checkbox"/>			55%

Set Bookmark

Submit Bookmarks

Attention

Are you sure you want to set this cut score?

Clicking "Yes" will set "Level 2" as the selected Ability value. If you've already set "Level 2" at a different Ability value, clicking "Yes" will change it to the new selected value. Clicking "No" will take you back to the item map.

Yes No



Questions?

- Thank you for your participation!

Setting Achievement Levels Using the Bookmark Procedure

The purpose of this document is to acquaint you with the Bookmark procedure for setting achievement levels. It provides the basics and will give you a preview of some of the things you will be seeing when you log in to set a cut score.

The Bookmark procedure is one of the most widely used methods for setting cut scores on tests. A cut score is a score that divides one category, or Level in this case, from another. Rather than directly identifying a cut score, you will place a bookmark in a specially constructed test booklet between the most difficult question you think a student just entering Level 3 can answer correctly and the first question you think that student could not answer correctly. The details of this procedure are explained below.

By now, you have studied the **Common Core State Standards (CCSS)** and the **Achievement Level Descriptors (ALDs)**. Your task now will be to translate those Achievement Level Descriptors into cut scores by identifying the most difficult test question you believe students at the threshold of Achievement Level 3 should be expected to master.

The content in the ordered item booklets (OIBs) is aligned to the content in the Achievement Level Descriptors. Questions early in the booklet can probably be answered correctly by nearly all students. Later questions are likely to be too difficult for students at lower ability levels but still not too difficult for students at Levels 2, 3, and 4. Further into the booklet, questions become even more difficult, but not too difficult for students at Levels 3 and 4. At some point, the questions become too difficult for the lowest-performing Level 3 student. It is at this point that you will find the one question that differentiates between the “Just Entering Level 3 Student” and the rest of the Level 3 students.

You will examine real test questions in a test booklet whose questions have been ordered from easiest to hardest, and as you go through, you will place one bookmark, delineating where you believe the questions become too difficult to expect students just entering Level 3 to master. You will place a bookmark after the most difficult question you believe the just entering Level 3 student will be able to answer correctly.

In addition to the online test booklet, you will receive an online map of the questions. These maps provide information about the questions in the ordered item booklets, such as standards alignment and depth of knowledge (DOK).

The OIB has many pages, each of which has a single test question displayed. If you examine these questions in order, you will see that the easiest question on the test appears on Page 1 of the OIB. The hardest question appears on the final page. Questions ascend in difficulty all the way through the OIB. Each question is more difficult than the one before it and easier than the one after it.

The questions are ordered based on the actual performance of students on the test. Sometimes you may be surprised at the ordering of the questions. For example, you may

see a question that you perceive as easy near the end of the OIB. Just remember that the ordering of the questions is based on actual student performance, not our opinion.

Because some of the tests are based on reading passages, and because the questions are placed in difficulty order, rather than passage order, you won't see the reading passage next to the questions. Instead, each passage will be available through a hyperlink. As you go through the ordered item booklet, you can click on the passage name to link directly to that passage. For example, in the screenshot below, you can see that there is a passage associated with the question on this page (highlighted in red). By clicking the passage number (here, [1659.JPG](#)), you could see the entire passage. This hyperlink will appear on every page in the booklet that contains a question based on it.

In-Person Panel for Achievement Level Setting



[Instructions](#) [Reference Materials](#) [Item Map](#)

Hello dmcdaniel@measinc.com [Log off](#)

Ordered Item Booklet: English Language Arts/Literacy, Grade 4

English Language Arts/Literacy, 4

Page: 07

Round: 1

[Back to Item Map](#)

[Set Bookmark](#) [Enter Comment](#) [◀ Previous](#) [Next ▶](#)

Item	Item Info	Resource(s)	ALD
Passage 1659.JPG			

[Back to Item Map](#)

Since there is a lot of information in the OIB, you also have a guide called the item map.

In-Person Panel for Achievement Level Setting



Instructions Reference Materials Item Map

Hello dmcDaniel@measinc.com! Log off

Item Map for Booklet: English Language Arts/Literacy, Grade 4

English Language Arts/Literacy, 4

Round: 1

[Enter Comment](#)

Page	Standard Code	DOK	Bookmark	Comments	Reviewed
01	4.L.6	1			<input type="checkbox"/>
02	4.L.2.c	1			<input type="checkbox"/>
03	4.RL.4	2			<input type="checkbox"/>
04	4.SL.3	1			<input type="checkbox"/>
05	4.W.3.d	1			<input type="checkbox"/>
06	4.L.1.a	1			<input type="checkbox"/>
07	4.RI.4	2			<input type="checkbox"/>
08	4.RI.5	2			<input type="checkbox"/>
09	4.L.1.d	1			<input type="checkbox"/>
10	4.RL.2	2			<input type="checkbox"/>
11	4.SL.3	2			<input type="checkbox"/>
12	4.SL.3	1			<input type="checkbox"/>
13	4.W.2.c	2			<input type="checkbox"/>

[Set Bookmark](#)

[Submit Bookmarks](#)

This is an example of an item map from grade 4 English language arts/literacy. Take a look at each column of the item map, just to see what kind of information is available to you during the achievement level setting.

First, you will see the **Page** number. Remember that easier questions appear toward the beginning of the OIB, so the question on Page 3 was easier for students than the question on Page 4, for example. You can click on any page number in the item map to go directly to that page in the ordered item booklet.

Standard Code refers to the CCSS on which the question is based. The question on page 1 is based on the 4th grade standard 4.L.6 (Grade 4, Language, Standard #6). In fact, if you hovered over the standard with the cursor, that particular content standard would appear on the page.

DOK refers to depth of knowledge, which ranges from 1 to 4, depending on the type of task the student is asked to perform. DOK 1 questions, for example, ask students to recall information, while DOK 3 questions ask students to solve complex problems, draw conclusions, or provide detailed arguments in support of a position. As with Standard Code, you can hover over the DOK with the cursor to call up the complete description of that DOK.

The Bookmark column is where you will be able to place your bookmark recommendation.

Comments provides a place for you to enter comments and your thoughts regarding each of the questions. Click on Comments, and a box will appear. You can type your comments in the box, and they will be saved.

Reviewed is a box you check to indicate when you have reviewed a question so you can keep track of your place in the process. As you work through the OIB, the system will automatically mark the questions as Reviewed; you do not have to return to the item map each time.

As you review the questions in the OIB, consider two questions about each one.

- *What do you know about a student who responds successfully to this question? Think about what a student needs to know and be able to do in order to answer this question correctly.*
- *What makes this question more difficult than preceding questions?*

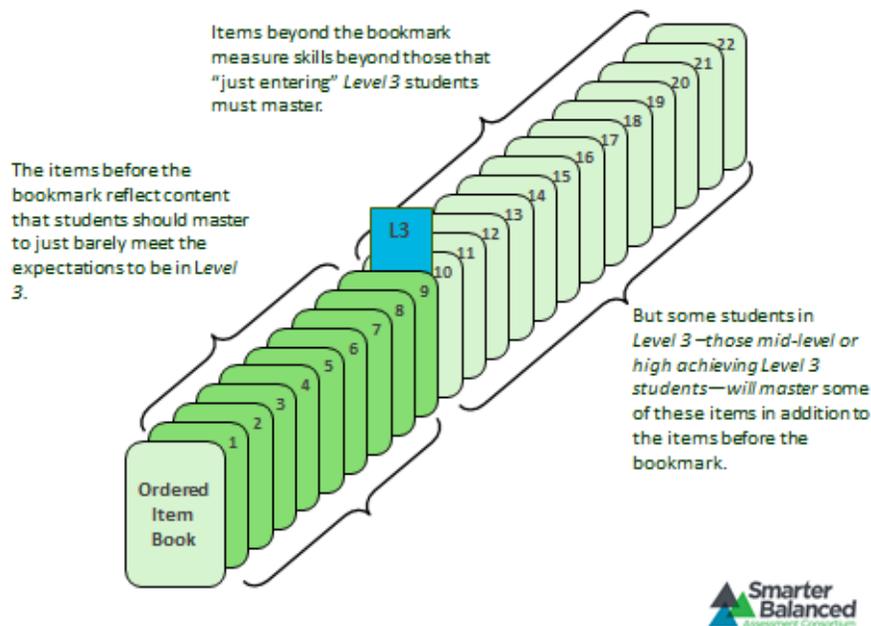
When you set your bookmark for Level 3, keep the Target Student in mind. Here, we have a **Just Entering Level 3 student** – a student who just barely makes it into the Level 3 category, not the Middle of Level 3 student or the High Performing Level 3 student.



When you set your bookmark, you are identifying a point that separates students in one level from the level below. So you want to think about the expectations for the “just entering” student at the threshold of the level (as defined in the Level 3 Achievement Level Descriptor). We consider the knowledge, skills, and abilities held by the Just Entering Level 3 student. When you set your bookmark, think about the minimum level of demonstrated mastery that you will accept for a student to just make it into Level 3.

After you study the questions in the ordered item booklet, you will place one bookmark to indicate the beginning of Level 3. When you place your *Level 3* bookmark, you are saying that the Level 3 target student—the student just barely meeting your expectations to be in Level 3—is expected to master the content measured by the questions before the bookmark. In the example below, that includes questions 1 through 9. When you think about the content measured by these questions, you should be able to compare it with the content described in the Level 3 achievement level descriptor (ALD).

Note that mid-level and high performing Level 3 students will master more content than just the questions before the bookmark. The bookmark denotes the minimum level of achievement to be in Level 3.



15

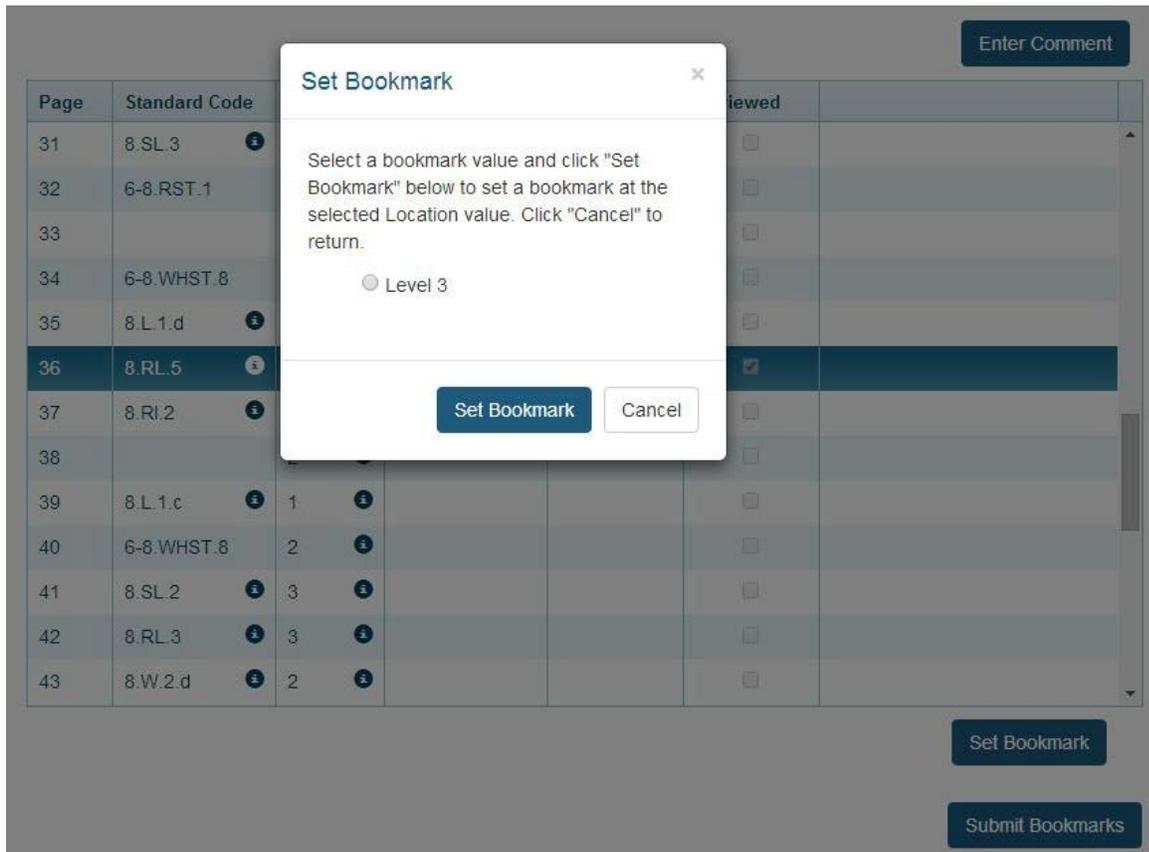
When you place your Level 3 bookmark, think about a student who has mastery of the content measured by the questions before the bookmark. Examine the knowledge, skills, and abilities held by that student, and compare them to the Level 3 Achievement Level Descriptor.

Ask yourself: Would a student who has mastered the content of the questions prior to the bookmark just meet the requirements of Level 3? Place your Level 3 bookmark right after the most difficult question you believe the student just entering Level 3 could answer correctly, as shown above.

A critical part of the placement of your bookmark is deciding what kind of chance you think a student should have to answer the question on this page correctly. For this test, we are using 50 percent. If you think the student will have a 50 percent chance of answering the question correctly, go on to the next page and ask yourself the same question. Keep going through the booklet until you reach a question you do not believe a student just entering Level 3 would have a 50 percent chance of answering correctly. Look at the next couple of pages, just to make sure you are not stopping too soon. Then come back and place your bookmark after the last page you believe a student just entering Level 3 would have a 50 percent chance of answering correctly. So, for example, if the last "Yes" question was on page 20, place your bookmark on page 21.

There are a couple of ways to consider 50 percent. Imagine 100 students just entering Level 3. Would about 50 of them answer this question correctly? Or consider a student just entering Level 3 encountering 100 questions like this one. Would he or she answer about 50 of them correctly? Whichever way works best for you, think in terms of the student just entering Level 3 having a 50 percent chance of answering the question correctly.

Now let's turn our attention to how you enter your bookmark. Here we see where someone has entered a bookmark on page 36 of the ordered item booklet to indicate that the question on page 35 was the last one a student just entering Level 3 would have a reasonable (50 percent) chance of answering correctly.



Once you have examined enough questions to be convinced that you have found the location in the ordered item booklet where the student just entering Level 3 would have less than a 50 percent chance of answering the next item (and presumably any remaining ones), place your bookmark there, as shown above. Before you do, however, check the ALD for Level 3 to make sure the content of the question you have bookmarked is consistent with the ALD. By placing the bookmark on page 36, this panelist is saying that the question on page 35 is the last question a student just entering Level 3 would have a 50 percent chance of answering correctly and that the questions on pages 36 through the end of the booklet are beyond the ability of a student just entering Level 3.

This concludes the introduction to the Bookmark procedure.

When you close this document, return to the home page and click “Booklets.” Select the Orientation Round Booklet. The Orientation Round Booklet consists of a 6-question ordered

The Bookmark Procedure

item booklet and item map. Your task will be to review the questions and place a bookmark after the most difficult question that you believe a student just entering Level 3 would have a 50 percent chance of answering correctly. When you have finished the Orientation Booklet and entered a bookmark for Level 3, proceed to the Readiness Form and complete Part 3. After you do that, return to the home page, click “Booklets,” and select the Ordered Item Booklet.

Online Field Test Survey Monkey

8/15/2014 – 8/20/2014

N=43

Did you have any difficulty logging on to the Measurement Incorporated website?

Yes – 2 (5%)

No – 41 (95%)

Comments:

- Once I received the correct one.
- Only difficulty was logging into the TES
- Failure to connect-Domain Name R
- This has always been fine; Getting into Turbo meeting has always been fine, too

How much time did it take you to review the Common Core State Standards for your test?

Less than 30 minutes – 14 (33%)

30 minutes to an hour – 23 (53%)

More than an hour – 6 (14%)

[If Survey Monkey permits, make this open ended: _____hours and _____minutes]

Mean = 41.63 minutes

Minimum = 10 minutes

Maximum = 90 minutes

How much time did it take you to review the Achievement Level Descriptors?

Less than 30 minutes – 16 (37%)

30 minutes to an hour – 26 (61%)

More than an hour – 1 (2%)

[If Survey Monkey permits, make this open ended: _____hours and _____minutes]

Mean = 34.42 minutes

Minimum = 10 minutes

Maximum = 90 minutes

Did you have any difficulty logging on the Smarter Balanced portal to take the practice test?

Yes – 10 (23%)

No – 33 (77%)

Comments:

- Confusing. No instructions given that were useful. A quick run through with info left out, was no good.
- Couldn't access after finished second questionnaire. Couldn't get into log map.
- Couldn't access after finished second questionnaire. Couldn't get into log map.
- Tried multiple times to log onto the link before experiencing success
- navigating within the Smarter Balanced site was confusing.

- Couldn't get past question #16 on the practice test for a while. rebooted to firefox and was able to continue
- Once I found the web address, no problem
- I had stop and load Firefox after starting the practice set and was unable to save my work. Lost time having to start over.
- DNR error requests
- I didn't have my assigned grade level. There was a lot of confusion navigating the entire site for the practice test. We did not get step-by-step directions until late afternoon.
- I was confused about which test(s) to take.
- Needed corrected access code (wasn't the one given in mass email the day before); can't recall the other problems - I had started Practice Exam by 2:30 ET, though (Turbo meeting began at 1 p.m. ET)
- did not have event code
- did not have event code

How long did it take you to complete the practice test?

Less than 30 minutes – 2 (5%)

30 minutes to an hour – 20 (47%)

More than an hour – 20 (47%)

Omit – 1 (2%)

[If Survey Monkey permits, make this open ended: _____ hours and _____ minutes]

Mean = 73.26 minutes

Minimum = 20 minutes

Maximum = 180 minutes

How long did it take you to read the information about the Bookmark procedure?

Less than 30 minutes – 14 (33%)

30 minutes to an hour – 21 (49%)

More than an hour – 2 (5%)

Omit – 6 (14%)

[If Survey Monkey permits, make this open ended: _____ hours and _____ minutes]

Mean = 32.09 minutes

Minimum = 10 minutes

Maximum = 120 minutes

How long did it take you to complete the 6-item Bookmark Practice Round?

Less than 30 minutes – 23 (53%)

30 minutes to an hour – 15 (35%)

More than an hour – 0 (0%)

Omit – 5 (12%)

[If Survey Monkey permits, make this open ended: _____ hours and _____ minutes]

Mean = 22.09 minutes

Minimum = 10 minutes

Maximum = 60 minutes

How long did it take you to complete the main Bookmark exercise?

Less than 30 minutes – 1 (2%)

30 minutes to an hour – 18 (42%)

More than an hour – 16 (37%)

Omit – 8 (19%)

[If Survey Monkey permits, make this open ended: _____ hours and _____ minutes]

Mean = 68.14 minutes

Minimum = 10 minutes

Maximum = 180 minutes

How helpful was the document Setting Achievement Levels Using the Bookmark Procedure?

Not At All Helpful – 5 (12%)

Not Very Helpful – 10 (23%)

Somewhat Helpful – 17 (40%)

Very Helpful – 7 (16%)

Extremely Helpful – 1 (2%)

Omit – 3 (7%)

Comments:

- some of the items in the level 2 category seemed more complex
- not enough time to digest info in document for assigned 4 hour slot
- Did not get to do the bookmark procedure
- didn't get there... I think what happened was: On the bottom of the page where I was supposed to click on the "practice test" highlight, it's right next to a BIG tab marked "Continue" so I just clicked on that thinking it was the right thing to do. Then I didn't know if I was in a "practice test" or not... Why should I know?... So I just got all involved in that then kinda realized that it wasn't a practice test but felt under time constraints so just continued thinking well maybe I can go back and take the practice test later... or something... Seems to me now that it should be set up differently so that one HAS to take the practice test before being able to move on. I'm guessing that there's some kind of trigger involved in the prac test section which enables one to connect with the whole bookmark process... ???
- Couldn't download Benchmark.
- Difficulty understanding it. Like a lot of this project I thought there was a lot of information overload.
- I have to admit my frustration with this. I never really understood how to complete this pilot.
- I was only able to place a "3" level bookmark
- It did not come up immediately when I logged in, so unfortunately I read it after I completed some of the other tasks.
- Very confusing. Needed to be more comprehensive.

- Like the other resources, it was a mass of verbiage which I found daunting to get through. What I needed to know was buried in there but I did not see a Highlighter feature and felt under pressure to meet the time frame.
- I thought I was supposed to read all of it.
- Especially helpful was the example - it clarified whether you were selecting the last 'doable' one for student or first 'not doable' (so to speak) for student
- could not access for a long time
- no communication w/ Turbo for help
- I had trouble finding the information and following the steps. During the turbo meeting, there were many different people asking and answering questions at the same time. This made it a little tough to find answers to questions regarding the next steps.
- Needed more direction on how this applied the actual process
- took a long time to download after finally realizing how to find it. Felt unsure of procedures.

Did you have any difficulty navigating through the item map?

Yes, a great deal – 11 (26%)

Yes, some – 19 (44%)

No – 10 (23%)

Omit – 3 (7%)

Comments:

- Not once I figured out where it was.
- It took a while to figure out how to navigate through this, but once I figured it out it wasn't that bad.
- directions could be clearer if assigned time to complete task that important
- Did not get to this part
- not sure what you mean by "item map"
- Couldn't download to write on item map.
- No download once I got there.
- Definitely a learning curve which I didn't seem to grasp.
- There was no instructions to tell you how to get into the item
- Difficulty with the order of thins, and finding the item map.
- The first task in my list was the bookmark task, so I first went to that instead of the practice. It might be better for others to have the practice at the top.
- better and specific directions
- I felt we didn't receive enough training and I wasn't sure what to do.
- I didn't know what to click on, it seemed like I was supposed to be answering on the item map itself. To bring up yet another verbose/technical/repetitive/unfamiliar document only confused me. I found it difficult to keep it all straight.
- After I finally got through it, I didn't know how I did it or even if I was finished.
- Proportionately speaking - didn't figure out how to get to it till 7 p.m. ET (took off for dinner) - I think somehow I missed going to it when I completed the second part of Questionnaire; I had to re-do the Questionnaire Parts 1 and 2 to get back to it (and I made notes on one page saying I'd already submitted those answers and didn't want to type

them again - hopefully you got my first submission - I did indeed know what was technologically-enhanced, what accommodations there were for special needs students, and what the standards code meant)

- when I finally got to the screen, it went ok, after I figured out what you all expected.
- Used TI list from Heide to navigate...w/o that I would still be trying
- The item map was somewhat confusing. For me, it was the way in which the information was grouped. Sometimes it would log me out, and I would have to start from the beginning.
- Needed further instruction from on steps to take
- Initial directions not clear about this being the main event
- Did not realize level 3 was the only choice. Kept going back to chat to get further direction.
- Kept closing out the test window, but logged back on with no difficulty

Did you have any difficulty navigating through the ordered item booklet?

Yes, a great deal – 7 (16%)

Yes, some – 19 (44%)

No – 14 (33%)

Omit – 3 (7%)

Comments:

- just not familiar enough with format for SBAC benchmark analysis
- Did not get to this part
- not sure what you mean by "ordered item booklet"... the 30 items?... no problems
- yes, never got that far.
- Never got there. couldn't download.
- The questions came up sequentially and automatically. The navigation buttons were obvious.
- I could not move past item number 16 on the practice test. I was not able to use the function required for that task.
- clearer directions
- More training would have helped.
- What a strange name.
- Was not able to open something resembling a booklet, just questions.
- Some of the item bullets couldn't be read-scroll was inactive
- It was not clear how to access the booklets until we were given step-by-step instructions much later.
- What's the ordered item booklet? Are you talking about the test
- When looking at 'Resources' from the instructions page (in black bar menu) the core values document did not come up - error page message - so I assumed, when in the main event that would be the case, too - I opened another page on the Internet and got them that way, so didn't need them to bookmark the test; however, thus I missed the full text for a few of the items; luckily, I figured this out the next day and went back to the items and read the texts as needed (iceberg, Wilma, and beach closed - didn't get to the breakfast/beans/boxes one; Otherwise, no problems at all - very clear and helpful
- Thanks to TI list, it got easier when I got started.

- poor if any directions, felt like I was flying blind
- Unclear procedure, spent time waiting for further instruction on chat or email
- No support for questions about procedure. Spent time waiting for answers on chat
- Was not certain of the "main event" then felt rushed to complete
- See above - I also made comments about some of the resource materials not being readable.

How confident are you that the bookmark you placed represents the ability of a student just entering Level 3?

Very Confident – 0 (0%)

Confident – 9 (21%)

Unsure – 24 (56%)

Not At All Confident – 7 (16%)

Omit – 3 (7%)

Comments:

- Did not get to do the bookmark
- didn't get there...
- Couldn't download.
- Only because I never got that far in project.
- I did not feel confident so I did not complete the pilot
- I found the bookmarking procedure to be very confusing. I wasn't sure if the level represented students just entering or exiting level 3.
- it's been a minute since I have had to do any of this kind of testing so took a minute for me to adjust
- the item tasks weren't consistent, so I felt like there were items after it that easily could have been completed and a few items beforehand they would have had trouble with.
- It took me awhile to understand how I was to place the bookmark and process each level so I could relate this to a student.
- I completed this task before completing the practice because I was in a hurry to get it all done in 4 hours. I did not have much time to review my choices and consult the guidelines. Instructions via email came very later in the process. I tried to just complete everything between 1pm and 5pm, since I had plans for the evening and Friday morning. I was stressed out because of that and missed some of the important instructions on first glance.
- I had no opportunity to place a bookmark.
- Hard to say where all students are.
- I hardly knew what I was doing!
- I entered the project as a Reader Evaluator, which usually means student responses not evaluating the actual questions. There were no student responses to compare to the rubric as there are in on-site reader-evaluator projects. I found it hard to transport myself to a 4th grade mentality while studying curriculum-design oriented technical papers with unfamiliar/repetitive terminology.
- I understood the bookmark placement should be placed at mid-point (placed where 50% of students can grasp material) That wouldn't seem to me to mean those just entering level 3

- I didn't find the necessary explanation as to what I was supposed to do regarding bookmarks. I guessed at what was expected, but even then, every time I bookmarked something I would get a message saying that it would cancel out any other. Very confusing.
- I am not a middle school teacher so can only make a guess (having just graded thousands of gr. 6-8 SBAC census tests probably helped a bit) - the standards were very helpful in making the call - I might have picked an earlier cut score than I did without their guidance, honestly.
- Ranking of items were questionable....some easier items after bookmark,
- felt the items were in poor order (some less difficult items were after book mark)
- I was unable to place a bookmark due to confusion with the procedures to do so. I have little doubt that I would have been able to complete this task with a little more clarity through Turbo Meeting. This is not to place blame on any person. It was just tough to get clarification when everyone in our group was posting questions at the same time.
- Felt rushed to finish by the time I was able to access the booklet
- Did not immediately realize that only level 3 was available
- Again, felt rushed to complete and not enough instruction
- As I continued through the pack, there were some that appeared to be easier, but out of place in the sequence.

In the space below, describe what you consider to be the most helpful part of the training for this activity. Please be specific with regard to how it was helpful.

- When Heide got us a step by step guide to navigate the website.
- All my questions were answered quickly via chat.
- The training gave us good general details about the activity and its purpose.
- Turbo meeting was helpful - it helps and would help to guide the team to complete the process in a timely and efficient way.
- The practice test gave me a good idea of what the students had to do in the test, and aided in placement of the bookmark
- the practice test was ok....because I am a retired teacher (last 22 years I was a 3rd grade self-contained teacher) I understood what the standards and the achievement indicators meant...I do believe we should have taken it a little slower as the VP described those documents. Giving us time to read over them was time consuming, but not having the actual test item until the main event meant I couldn't look at the item and consider how well it correlated with the standard and the descriptor. I have taught in inner-city and neighborhood (suburban) settings. I have had students at all levels of achievement in those settings. I understood what looking for an item where a just entering level 3 student would have a 50% chance of understanding and correctly attempting the item meant, but I think I felt pressure to analyze each item in a timely manner using what I know a 4th grader should understand.
- info received from scoring site supervisor from Durham. Heide replied with correct info thru emails as fast as she could.

- The training pertaining to the Common Core State Standards and the Achievement Level Descriptors was very helpful. It helped to fully understand what was expected of the students when reaching a specific level.
- nothing comes to mind...
- Helpful practice tests questions. Good use of multimedia.
- The practice test questions were very well written and interesting. Liked the multimedia and shuffling of types of questions approach.
- Since I never really understood the process I can't say any of it helped me
- The most helpful part of this experience was being able to email Heide Kober with questions about logging on and accessing the different portions of the test. She was in constant contact with participants with answers to our most pressing questions. I realize she was not directly involved with the training aspect, but her assistance made the difference between me completing the task and just throwing up my hands and logging out!
- I spent 30 minutes reviewing the standards but your survey will not accept those answers...had to enter 1 hour or it kicked back with red message, "answer is required"
- The site was user friendly.
- The information that was provided later on from IT was very helpful, and definitely the initial training material was helpful
- The most helpful part of training was reading the materials. I do better with an assignment when I can read the instructions myself, and refer to them during the exercise.
- I switched back and forth between the Common Core State Standards and the Achievement level standards worksheets while I was trying to set the bookmark. They were very helpful. It would have been much more helpful to have someone explain them rather than attempt to figure them out on our own. Personally, I have never done a project like this before and it was very difficult for me to figure out what I was supposed to be doing without much instruction.
- I'm not sure the training was helpful at all. We were given no instructions ahead of time on how to navigate through the system, and then when we began asking questions, no one was there to assist. It was extremely frustrating
- Providing a webinar overview was helpful so we knew the purpose of what we were to do. Each section had a visual that also helped us understand what we would be seeing in each task.
- The training was a bit brief, but I liked the overview that gave me some idea of what I would be doing.
- Looking at the actual test was the most important and only useful training. The formal and generalized training materials were really no help at all.
- Turbomeeting was helpful to an extent.
- The most helpful part was when Heide Kober sent us an email with the flow process - step by step. I was able to check off each item as it told me exactly what to do.
- The rubric and the Descriptors. At last there were key words to guide my thinking (at least I was able to extract a general idea). And Heidi's email with the sequence /flow list,

otherwise I'd have been really at sea - (In fact that was the first note I made to give as feedback when the opportunity came).

- There was no advance explanation of what the expectation or outcome was for this project. As a participant it would have been helpful to know what we were trying to accomplish before we embarked on this project. We should have been given the steps before we were turned loose, not after we started. At the very least, we should have known what grade we were expected to be scoring. There was no mention when training ended to look for an email giving us our grade level. Time was wasted because of this. Four hours does not seem like a sufficient amount of time for this project.
- Although unsure of how those entering level 3 means 50% of students, the bookmark instructions were of some help.
- I'm sorry I did not find anything helpful about this training. It seemed very rushed and not well organized. I did appreciate Heide Kober trying to answer our questions via email.
- Explaining what a bookmark meant was helpful in setting them.
- Dr. Bunch's presentation/overview was very helpful - I was grateful I had sound and video in the Turbo meeting (apparently some people did not). I was alerted to most steps and aspects of this pilot project. [All that was missing, honestly, were a few URL and item map directions.] His comments gave me a good understanding of the project's scope and purpose.
- The most help was the TI list that Heide sent after I got knocked off from Thursday session. It was simply a list with short directions on how to proceed in order.
- The most helpful part of the training was the explanation of what the bookmarks represent.
- Instructions at beginning stressing the need to take notes, as I referred to them throughout. This shortened the need to open the tabs for common core & descriptors.
- Nothing really stood out.

In the space below, describe what you consider to be one aspect of the training that needs to be improved. Please be specific with regard to what should be improved and how it might be improved.

- A thorough guide, step by step, for us to follow once the training booklets were read and understood.
- Clearer navigation guidelines through the various steps of the project and a walk-through of the various locations of available instructional/procedural texts.
- At first it was confusing and difficult to figure out what to do. The list that was emailed to show the exact steps to go through was helpful.
- The bookmarking was very unclear - how to do it made sense eventually - but there needs to be a guided - step by step - to know where to go and how to proceed.
- The clarity and accuracy of the material presented during the TurboMeeting prior to the exercise
- The practice test was ok, but I would have liked for a sample of the real main event to have been explained and had someone discuss why and why not a random item may be suitable for a bookmark. Just reading how to set a bookmark with no real item and discussion would

be very hard for a person who had no experience in the classroom. I do appreciate being assigned grade 4, because I felt I could be helpful as I judged the content.

- Opening presentation was too fast. Main speaker seemed unfamiliar with technology and confused readers. After brief opening remarks readers faced confusion and frustration as many questions were unanswered or perhaps ignored. Instructions need to be clear and allow readers time to digest and familiarize with process before beginning their task.
- I do not know of anything that needed to be improved. Because of computer problems on my end I did not get a chance to finish.
- Too twisty/turny ... instructions muddled... not straightforward. Especially that page I referred to previously where there's an unclear choice to "take the prac test" AND continue right next to each other. I just naturally clicked on "continue". Seems like it oughtta be set up so that one HAS to take the prac test before being able to move on. i.e. I didn't really know whether I was taking the prac test or not until I was WAY into the main 30-item test... then it seemed impossible to go back.
- Better communication between leaders. Info could have been transferred earlier to scorers to make it easier. A lot of guess work at times on our part.
- Communication between Turbo meeting team and scoring director, Many of the things we were told were ex post facto. I believe they could have been given before Turbo Meeting.
- Step by step through whole pilot. Maybe a run through. I never understood why I even took the test. I didn't even receive the grade I was assigned to until after I started the practice. I thought the whole training did not prepare me. It was very frustrating. I felt like I wasn't shown the whole picture so didn't understand the concept. Probably just me!!
- SBAC 21 was my first experience with a pilot test. I have successfully scored field tests both as onsite and online projects but this was an entirely new experience for me. I personally needed more training than the minimal session that we received. The chat room was random but the comments posted indicated that others were experiencing the same frustration as myself. I would not have been able to finish the testing session without Heide Kober's answering questions about how to proceed.
- clicked on links to support documents and got "file not found 404" while I was (desperately) trying to do the job.
- The bookmarking procedure was unclear. I was only able to place one bookmark. The directions for this portion could be revised.
- Would like to see more of that information we got later in the process available earlier...also if we got our grade assignments earlier and if there were greater details given in the presentation as to what is expected and how to do it, that definitely would have been helpful.
- An aspect of training that needs to be improved is the TurboMeeting that prepares the readers for the activity. It was somewhat rushed, and I get the idea that most of us were still not sure what to do once we started the activity.
- The instructions. I felt like I was left on my own without really knowing what was going on. I also still have no idea what "technology enhanced items" are, so I had no idea how to answer that specific question. I believe it would also be very helpful to have someone read and explain, even briefly, the achievement level standards and Common Core State

Standards. Going into this project with no knowledge of the achievement level standards and no instruction was very difficult and frustrating.

- Take out all the stuff we don't need. There were pages and pages with tabs that weren't pertinent to our task, and you could really get lost in all that, and it was difficult to navigate back and forth from one place to the other. There were no clear instructions the first day on how to complete the assigned tasks, and we wasted a lot of time waiting on answers that never came during the allotted time. We were emailed that evening on how to complete the tasks, but that should have been done ahead of time, and there should have been a contact person online at all times to answer questions. Chat rooms aren't the best way to communicate in that circumstance because everyone is in different places in the process and asking different questions and getting different responses, and we were trying to assist each other as much as possible, because there was no leader in turbo meeting with us. It was again, extremely frustrating.
- There is a big difference between providing an explanation of the content of the test (and task) and actually providing a step by step process that people can refer to, on how to do it! Much of the time I was confused as to how to get to the next screen, so I could not complete the task. There is a lot of information to try to absorb in a short period of time. I tried to bring up the levels of the core curriculum when I got to the part where I was to place the bookmark, and the links that were provided did not work. Even after contacting the helpline I was not able to access these, which would have been helpful in placing the bookmark. After receiving the steps in print (from Heidi) "from the IT guy" I was able to walk through each step, and complete the process. Please provide a step by step process to future field testers that they can refer to when they have questions (and their "host" on chat does not answer).
- While the power-point type presentation conveyed some idea of the task we were asked to do, a tutorial of what we would be seeing in each of the pages we'd be logging into and reviewing would have been very helpful. Some of this info could have been sent in an email beforehand. I really had no idea what I'd be doing that afternoon. When I was trying to complete it, I was overwhelmed for a while.
- Specific examples that have been scored which could be used as a reference would be helpful.
- Bookmark section was very difficult to understand. The whole project seemed rushed. I felt as if I had failed. Not a lot of fun.
- I feel that the brief time on TurboMeeting for the webinar was not enough. As was obvious by all the chats, no one seemed to know exactly what to do and how to do it. I think the training needs to be more specific and detailed.
- The oral presentation at the beginning seemed far-ranging and open-ended, I felt unsure what exactly I was to do...In the Descriptors, I noticed that level 2 used concrete statements of things student could do, implying they could, but level 3 used the term "BEGIN TO" (and then the same concrete statement as level 2 - wondered if that was a typo). I tried to keep track of the time I spent on each task but frankly they blended into one another so much that now I'm not sure. The Rubric reversed the order of the levels introduced in the Descriptors (ok, that's minor). The chat was pretty distracting and random and the Turbo

icon blinked distractingly in the Tray at the bottom of my screen where I put it since it kept disappearing anyway every time I clicked while doing a task. Navigating from task to task and resource to resource was destabilizing to say the least. There was too much information that could have been identified as "background" as opposed to "You'll need to keep this accessible". I understand that you are immersed in this language and the project, and thought more information would round out the experience for us and make it clearer, but I didn't find it that way. I recall hearing "study this" in reference to the Resources, but a better way to say it may be "familiarize yourself with this background information, but keep the rubric handy".

- Better training. Access to someone online to answer questions while readers are working. Training materials available ahead of time so they can be properly digested and then questioned during training.
- The amount of achievement & common core standards information for all potential items isn't helpful if you aren't aware of what type of item(s) you will encounter. The jargon alone requires deciphering and would need to be easily assessable for reference.
- I felt that this training really needed a lot more time and explanation. The bookmark procedure needed to be explained very specifically during the turbo meeting. Additionally, we needed the step-by-step instructions also explained prior to doing the actual practice test, etc. I know from the turbo chat that continued after the meeting had ended that many people had difficulty navigating the site, as well as needing A LOT more time to complete the practice test. Overall, I think this training needed to be practiced a few times before doing the actual pilot test in order to work out the kinks.
- Up front, specific expectations. I thought I would be given my target grade on the website, so I read all of the information on grades 4 - 12, then found out I was to do grade 8 in an email after I had finished. I did not understand what kind of time keeping I was supposed to do. A copy of the survey ahead of time would have been most helpful. It took me some time to figure out exactly which test to take and how to get to it. For those of us who felt completely lost, directions that begin with "log on" and don't say where are a little scary. I know we're all educated adults, but next time you might explain it like we're very bright 6-year-olds. Heidi, thank you for your infinite patience on the other end of the line.
- Pre-information' - The mass emails sent prior to project start were 95% there, in terms of completeness of information needed. Had the correct access code been there (i.e, 08...._FT instead of 3587191) and maybe also included the step-by-step instructions an IT person provided us, via Heide by email, Thur. afternoon sometime, I would have no problems at all, I think.
- The computer presentation for training lacked specific directions which matched the opening presentation which matched the lack of help from Turbo. The complaint chats on Turbo over rode the CRIES for help. Make whole presentation more user friendly w/ more specific directions,
- One aspect that I think needs to be improved, at least for this particular project, is more time spent during the initial Turbo Meeting. Perhaps a walk-through of steps, so that we can see what we are supposed to do in sequential order.

- Needed more support through the process. Many questions were unanswered and time was wasted waiting for direction to the next step. Downloading the booklet was slow and caused rushed decisions at the end of the process. Initial training should include slides of what to do after the practice test...that is where most of the confusion was.
- Communication - relying on email when people are freaking out in the turbomeeting was not optimal. Turbo should have been monitored for questions based on what I was seeing. Also, once I got started, my earlier assumption that I understood the standards became shakier as I proceeded. Need more time explaining what the State is looking for.

Materials and Equipment Needed for In-Person ALS - Field Test

Materials	Count
Name Tents (Table Tents)	55
Name Badges	55
Common Core State Standards (ELA Hardcopy)	8
ID labels	55
Flip charts (PostIt)	5
Notepads	40
Pens	40
ELA Common Core State Standards (hard copy)	8

Equipment	Count
Computers with sound cards (1 per facilitator)	4
LCD projectors (1 per room)	4
Speakers (sets)	4
USB drives for storing data	4
Cables and connectors	40
Computers with sound cards (1 per panelist)	40
Extension cords for presenters/facilitators	4
Power strips for presenters/facilitators	4
Spare bulbs	1
Batteries (AA)	8
Earphones	40

Other	Count
Internet connections for facilitators and panelists	44
Secure storage site	1
Work room	1

To be Sent in Advance
Security Agreement
Room/Panel assignment

Achievement Level Setting



In-Person Workshop Field Test
August 18-20, 2014



Welcome!

- Goal of Smarter Balanced Assessment Consortium
 - Ensure that all students leave high school prepared for postsecondary success in college or a career
- Achievement Level Setting in October
- Your Part in Meeting That Goal
 - Help us fine tune the procedures and programs we will use



Your Job

- Review Support Materials
 - Smarter Balanced tests
 - Common Core State Standards
 - Achievement Level Descriptors
- Learn an Achievement Level Setting Procedure
 - Receive instruction
 - Practice
 - Apply

Overview





Goals for This Workshop

- Understand Test Contents
- Understand Achievement Level Descriptors
- Learn Bookmark Procedure
- Recommend Cut Scores



Activities

- Review Common Core State Standards
- Review Achievement Level Descriptors
- Take Practice Tests
- Receive Bookmark Instruction
- Practice Using Bookmark Procedure
- Apply Bookmarks
 - 3 Rounds
 - Discussion Between Rounds



Agenda

Day - Time	Event(s)
Monday A.M.	Overview, training on CCSS, ALDs, tests
Monday P.M.	Training on Bookmark procedure, training round; begin Round 1
Tuesday A.M.	Review Day 1; complete Round 1
Tuesday P.M.	Review Round 1; complete Round 2
Wednesday A.M.	Review Round 2; complete Round 3; evaluate process
Wednesday P.M.	Vertical Articulation/Debrief
Thursday A.M.	Vertical Articulation

Common Core State Standards





Key Shifts

English Language Arts	Mathematics
Complex texts	Greater focus on fewer topics
Use of evidence from texts	Coherence – linking across grades
Building knowledge – nonfiction texts	Rigor <ul style="list-style-type: none">• Conceptual understanding• Procedural skills and fluency• Application

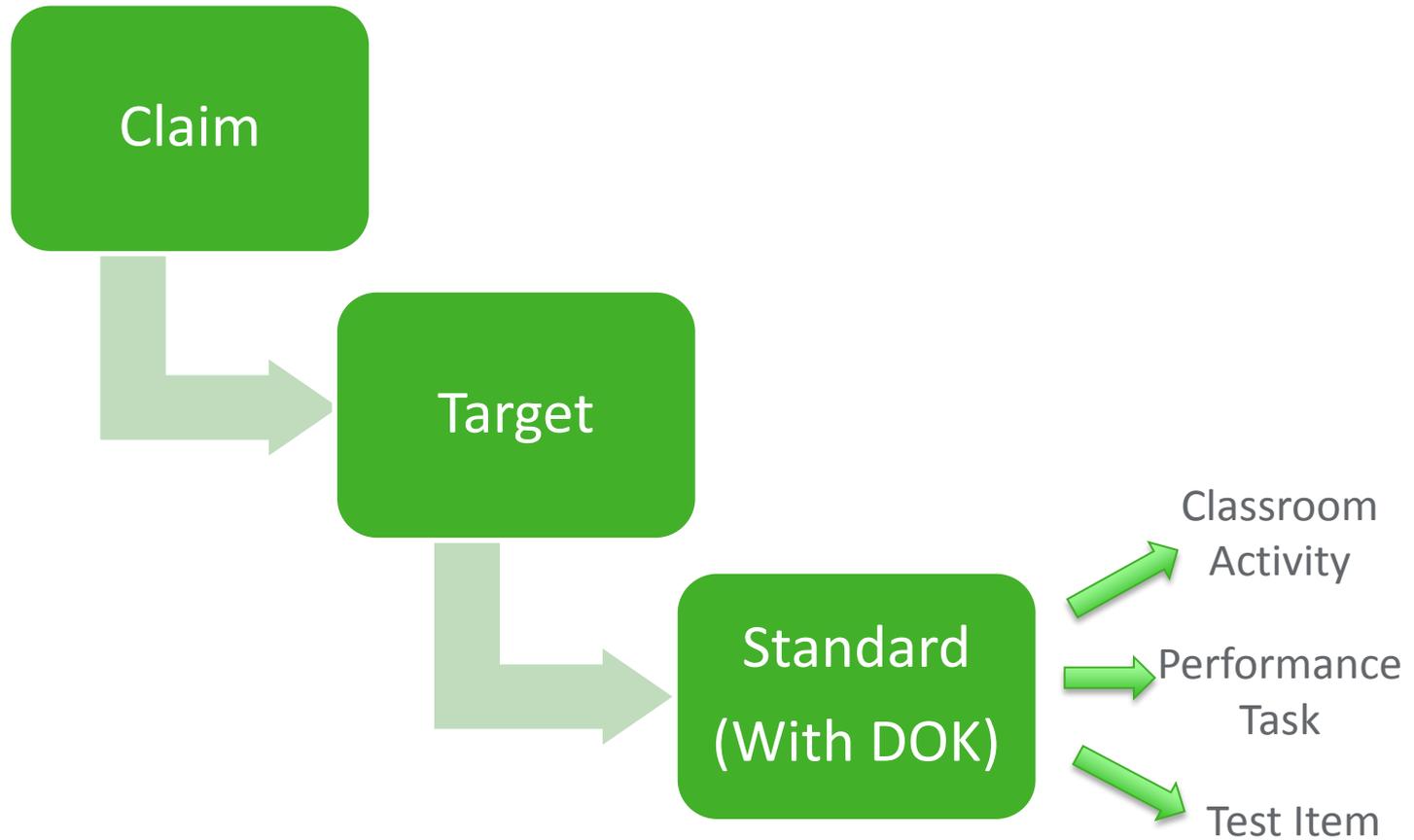


Claims

English Language Arts	Mathematics
Reading	Concepts & Procedures
Writing	Problem Solving
Speaking/Listening	Communicating/Reasoning
Research/Inquiry	Modeling & Data Analysis



Organization





Depth of Knowledge (DOK)

- 1 – Recall & Reproduction
- 2 – Basic Skills & Concepts
- 3 – Strategic Thinking & Reasoning
- 4 – Extended Thinking

Achievement Level Descriptors





Achievement Levels

- Level 1 – Minimal
- Level 2 – Partial
- Level 3 – Adequate
- Level 4 – Thorough



Types of ALDs

- Policy
- Range
- Threshold
- Reporting



Types of ALDs

- Policy
- Range
- **Threshold – students just entering Level 2, 3, 4**
- Reporting

Smarter Balanced Assessments





Review Tests

- Examine online test supports
- Take practice tests
- Discuss contents



Item and Test Development

- Common Core State Standards
- Development of Blueprints
- Development of Templates
- Item Development
- Item Review
- Field Testing
- Statistical Review
- Item Retention and Banking



Item Types

- Traditional Multiple-Choice Items
- Traditional Constructed-Response Items
- Technology Enhanced Items
- Viewing/Listening Items



Tools and Supports

Universal Tools	Designated Supports
Calculator	Color Contrast
Digital Notepad	Highlighter
English Dictionary	Magnification
English Glossary	Masking
Expandable Passages	Text to Speech
Math Tools	Translations (Glossary)
Spell Check	
Writing Tools	
Zoom	



Accommodations

- Braille
- Abacus
- Alternate Response Options
- Calculator
- Multiplication Table
- Print on Demand
- Read Aloud

Practice Test Preview

Welcome to the Smarter Balanced Practice and Training Tests

The Smarter Balanced Practice and Training Tests are available to schools and districts for practice and training purposes, professional development activities, and for discussions with parents, policymakers, and other interested stakeholders.

- Updated classroom activities and scoring guides for the Smarter Balanced Practice Test are available under [Resources and Documentation](#).

Note that some braille tests have a separate scoring guide, but the general scoring guide is applicable to all other accommodations and any braille test for which an independent scoring guide is not posted.

Added May 18, 2014

- [Calculators](#) are available for students to preview and practice with outside of the testing environment.

Added September 24, 2013

- Practice Test performance task writing rubrics are available under [Resources and Documentation](#).

Added September 24, 2013

- The Practice and Training Tests can be taken on any Internet-connected computer using a current Web browser including:

- Mozilla Firefox
- Google Chrome
- Microsoft Internet Explorer 10
- Apple Safari

Added August 20, 2012



[Diagnostic Tools](#) | [FAQs](#) | [Resources and Documentation](#) | [Provide Test Feedback](#) | [Contact Us](#)

Getting Started





Putting It All Together

- Relate ALDs to test items
- Consider students who are just barely at Level 2, 3, or 4
- Consider which questions those students can answer correctly
- Place 3 bookmarks



Groundrules

- Security/Confidentiality
- Group Process
- All Voices Equal
- Recommend – Not Set



Room Assignments

Panel	Room	Facilitator
High School ELA	Shoofly	Craig Deville
Grade 8 ELA	Orange Grove	Jennie Bowen
Grade 6 ELA	Orange Grove	Gretchen Schultz
Grade 4 ELA	Hurdle Mills	Kelly Connelly

Survey Monkey for In-Person Field Test Panelists

Part 1: Opening Session (Monday Morning)

The purpose of the activity we were about to undertake was clearly explained by the presenters.

Definitely Agree Agree Disagree Definitely Disagree

The opening session provided useful information.

Definitely Agree Agree Disagree Definitely Disagree

The amount of time devoted to the opening session was appropriate.

Definitely Agree Agree Disagree Definitely Disagree

Part 2: General Training (Monday Morning)

My group facilitator made me feel welcome.

Definitely Agree Agree Disagree Definitely Disagree

My group facilitator clearly explained the importance of the Common Core State Standards to the activity we were about to undertake.

Definitely Agree Agree Disagree Definitely Disagree

My group facilitator clearly explained the importance of the Achievement Level Descriptors to the activity we were about to undertake.

Definitely Agree Agree Disagree Definitely Disagree

The Smarter Balanced Practice Test was a useful exercise.

Definitely Agree Agree Disagree Definitely Disagree

Part 3: Bookmark Training (Monday Afternoon)

The purpose of the activity we were about to undertake was clearly explained by the presenter.

Definitely Agree Agree Disagree Definitely Disagree

The Bookmark practice round helped me prepare for the task I was about to undertake.

Definitely Agree Agree Disagree Definitely Disagree

Part 4: Bookmark Placement (Rounds 1-3)

My group facilitator helped me understand how to apply the bookmarks.

Definitely Agree Agree Disagree Definitely Disagree

I was confident of my bookmark placements in Round 1.

Definitely Agree Agree Disagree Definitely Disagree

The discussion after Round 1 helped me understand the process better.

Definitely Agree Agree Disagree Definitely Disagree

I was confident of my bookmark placements in Round 2.

Definitely Agree Agree Disagree Definitely Disagree

The discussion after Round 2 helped me understand the process better.

Definitely Agree Agree Disagree Definitely Disagree

I was confident of my bookmark placements in Round 3.

Definitely Agree Agree Disagree Definitely Disagree

The process was fair and open.

Definitely Agree Agree Disagree Definitely Disagree

A few people dominated the discussion in my room.

Definitely Agree Agree Disagree Definitely Disagree

The facilitator managed the process well.

Definitely Agree Agree Disagree Definitely Disagree

I was able to use the software without much difficulty.

Definitely Agree Agree Disagree Definitely Disagree

When there were problems with the software or procedures, they were quickly corrected.

Definitely Agree Agree Disagree Definitely Disagree

Part 5: Comments

In the space below, describe what you think went particularly well during the activity. Be as specific as possible.

In the space below, describe what you think should be improved, offering any specific suggestions you care to add.



Vertical Articulation

Field Test



August 20, 2014



Purpose

- Align cut scores across grades
- Link cut scores to expectations



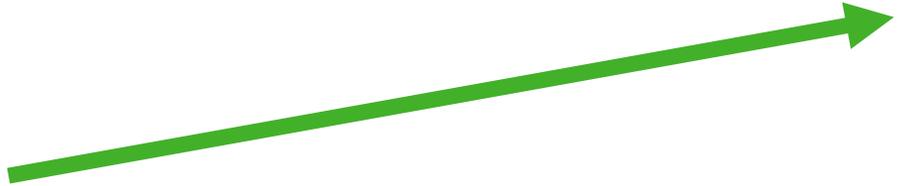
Why Are You Here?

- You were selected from In-Person achievement level setting to provide
 - First-hand knowledge of all recommended cut scores
 - Balanced representation
- You came highly recommended



Expectations

- Scenario 1: Increasing



- Scenario 2: Equivalent

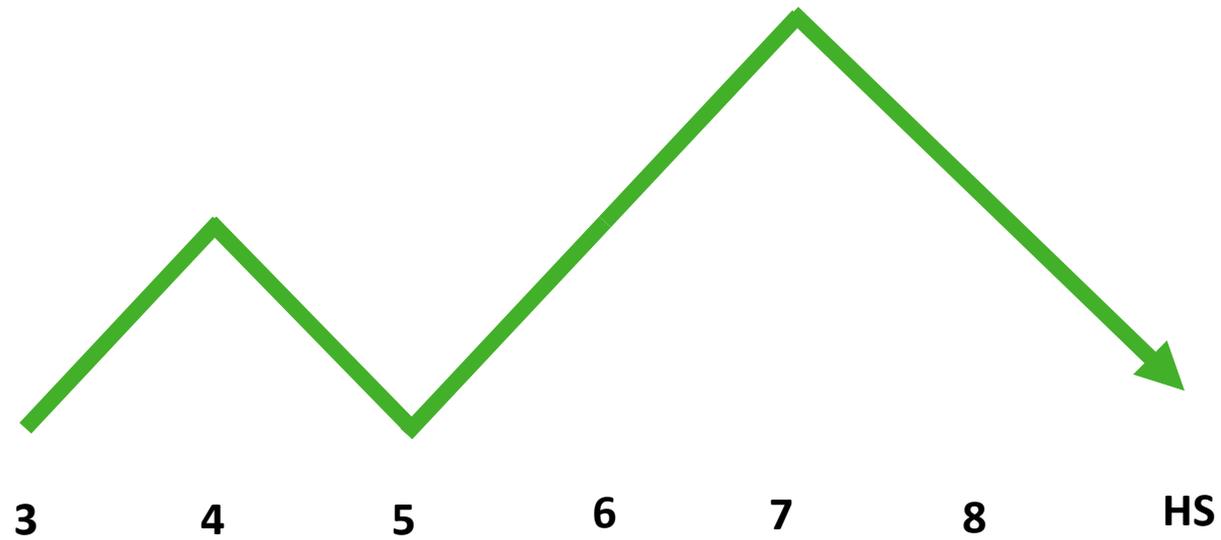


- Scenario 3: Declining





Not Expected





Basis for Expectations

- History
- Knowledge of student population
- Knowledge of program conditions
- Related programs



Questions to Ask

- Does this seem reasonable?
- If not, what can I do about it?
- What would happen if I changed this cut score?
- How will I explain what I've done?



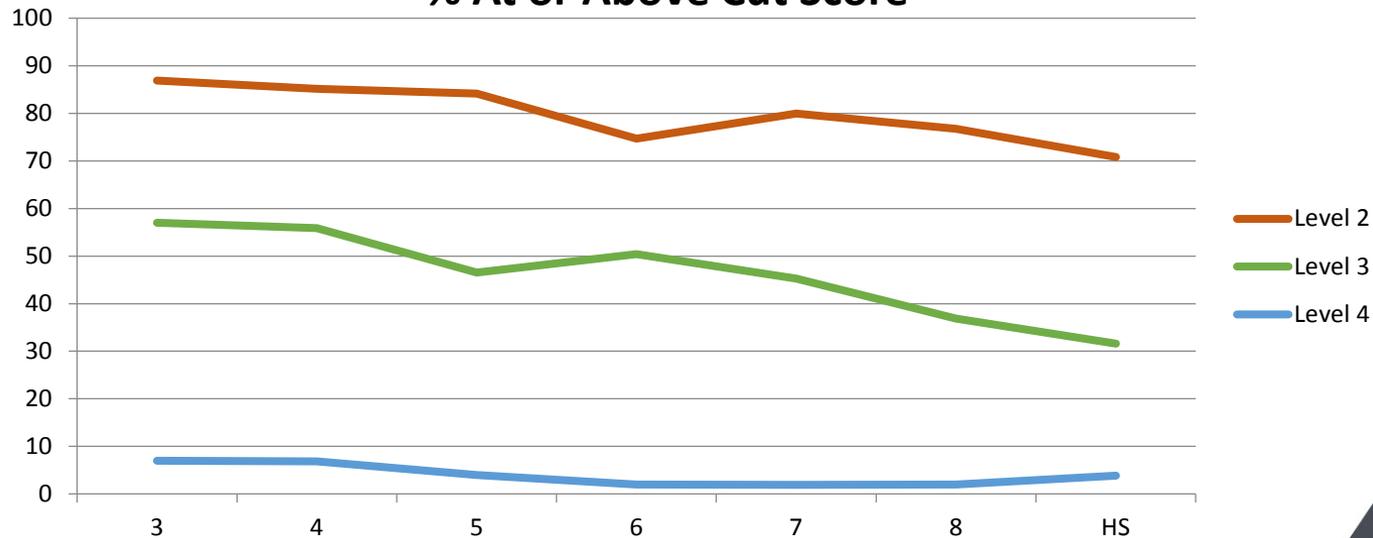
Procedures

- Review data summaries
- Discuss what your group did
- Consider “bumps” and “dips”
- Consider options
- Decide
 - One Level at a time across grades
 - Start with Level 3
 - Vote (2/3 majority rules)

Vertical Articulation

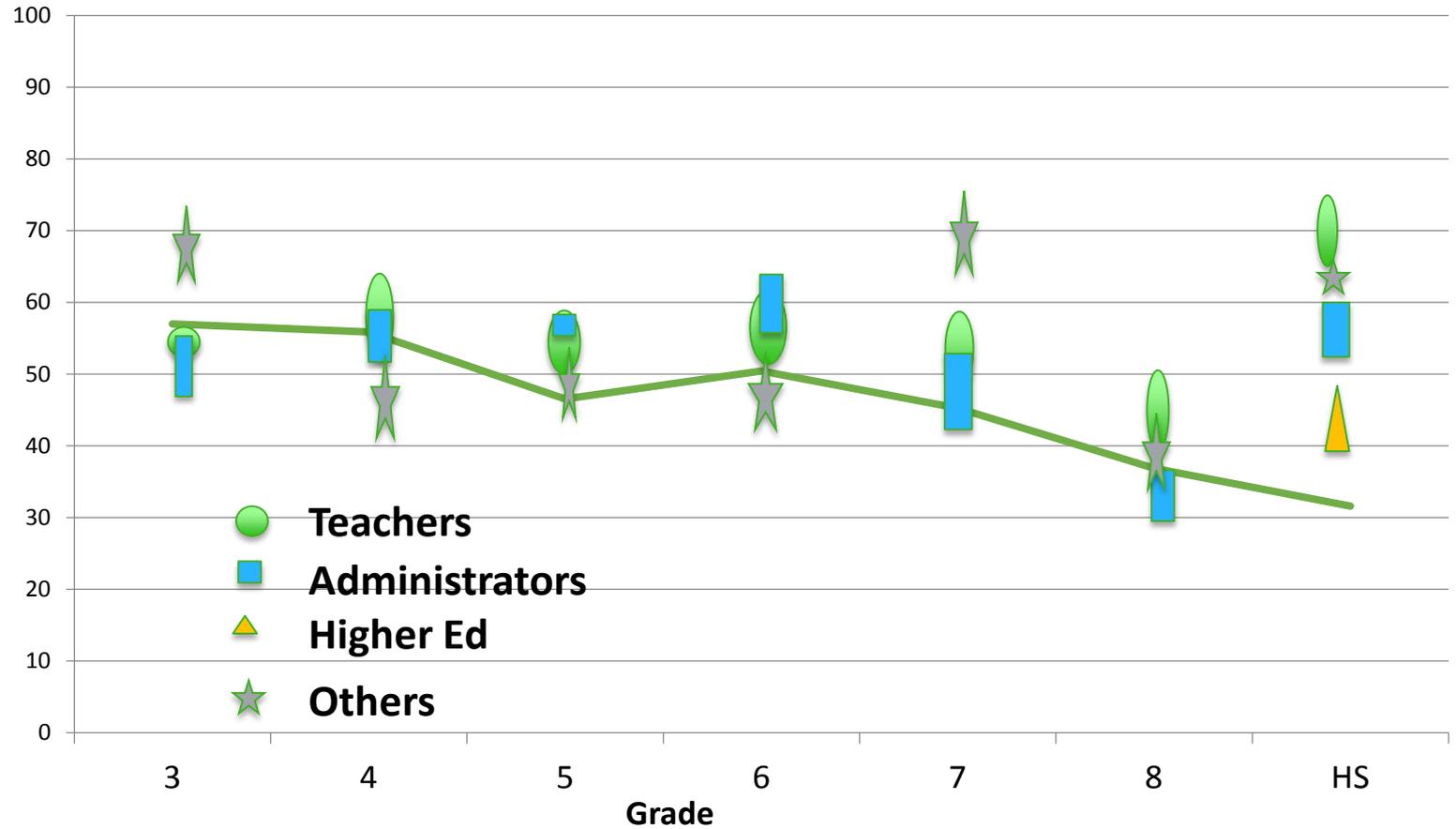
Cut Score				% At or Above				% in Group				
Grade	Level 2	Level 3	Level 4	Grade	Level 2	Level 3	Level 4	Grade	Level 1	Level 2	Level 3	Level 4
3	2.15	2.53	4.11	3	87	57	7	3	13	30	50	7
4	3.25	3.85	7.39	4	85	56	7	4	15	29	49	7
5	4.22	5.22	9.25	5	84	47	4	5	16	38	43	4
6	5.49	6.04	11.01	6	75	50	2	6	25	24	49	2
7	5.60	6.87	12.11	7	80	45	2	7	20	35	43	2
8	6.77	8.47	13.88	8	77	37	2	8	23	40	35	2
HS	8.45	10.94	15.99	HS	71	32	4	HS	29	39	28	4

% At or Above Cut Score





Using Online Panel Data

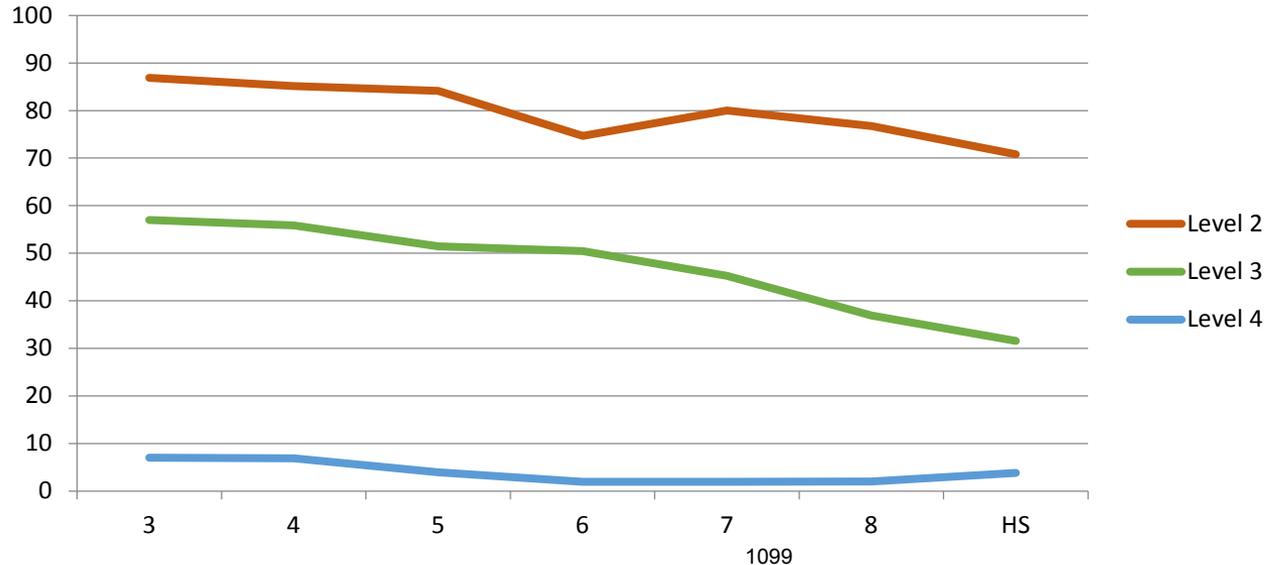




Vertical Articulation

Grade	Cut Score			Grade	% At or Above			Grade	% in Group			
	Level 2	Level 3	Level 4		Level 2	Level 3	Level 4		Level 1	Level 2	Level 3	Level 4
3	2.15	2.53	4.11	3	87	57	7	3	13	30	50	7
4	3.25	3.85	7.39	4	85	56	7	4	15	29	49	7
5	4.22	5.06	9.25	5	84	51	4	5	16	33	48	4
6	5.49	6.04	11.01	6	75	50	2	6	25	24	49	2
7	5.60	6.87	12.11	7	80	45	2	7	20	35	43	2
8	6.77	8.47	13.88	8	77	37	2	8	23	40	35	2
HS	8.45	10.94	15.99	HS	71	32	4	HS	29	39	28	4

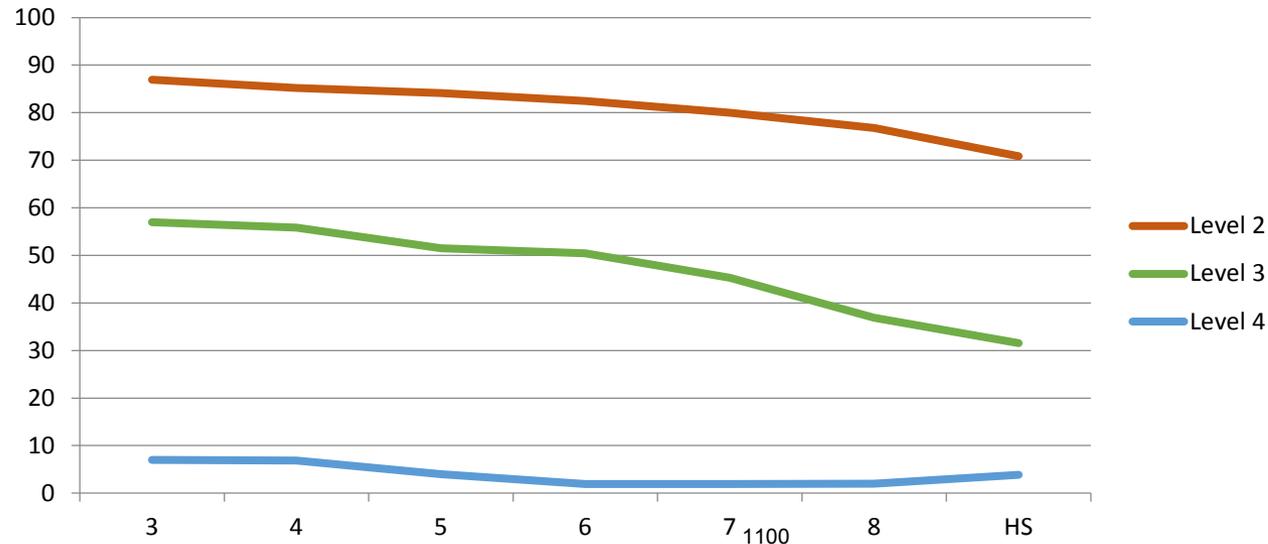
% At or Above Cut Score



Vertical Articulation

Grade	Cut Score			Grade	% At or Above			Grade	% in Group			
	Level 2	Level 3	Level 4		Level 2	Level 3	Level 4		Level 1	Level 2	Level 3	Level 4
3	2.15	2.53	4.11	3	87	57	7	3	13	30	50	7
4	3.25	3.85	7.39	4	85	56	7	4	15	29	49	7
5	4.22	5.06	9.25	5	84	51	4	5	16	33	48	4
6	5.09	6.04	11.01	6	82	50	2	6	18	32	49	2
7	5.60	6.87	12.11	7	80	45	2	7	20	35	43	2
8	6.77	8.47	13.88	8	77	37	2	8	23	40	35	2
HS	8.45	10.94	15.99	HS	71	32	4	HS	29	39	28	4

% At or Above Cut Score





Follow-Up Activities

- TAC Review
- Smarter Balanced Reviews
- Chiefs Review and Approval
- Technical Report

- **Questions?**

Smarter Balanced Achievement Level Setting Readiness Form

Vertical Articulation

I have seen the presentation about vertical articulation of cut scores.

Yes

No

I understand the purpose of vertical articulation.

Yes

No

I understand that recommendations to change a cut score will be done in accordance with a recognized democratic process.

Yes

No

I understand that there will be other reviews of the cut scores after the Vertical Articulation Committee completes its work.

Yes

No

Which of these patterns of student performance is/are expected (check all that apply)?

Generally increasing over grades or time

Generally decreasing over grades or time

Generally stable over grades or time

Up and down over grades or time

Which type of majority will be required to change a cut score?

Simple Majority

2/3 Majority

I am ready to begin vertical articulation.

Yes

No

Vertical Articulation Script

The key to successful vertical articulation is careful study of the cut scores and impact data for all grades (from In-Person Workshop and Online Panel) well in advance of the meeting. We will have several days to study trends from grade 6 to high school and at least a day to study trends across all grades after grades 3-5 complete the IPW on October 19. These trends should be the focus of intense discussion among facilitators and Smarter Balanced staff over the course of the seven days of the In-Person Workshop. By the morning of October 20, 2014, we should have a clear idea not only of what anomalies may exist but why they exist and what may be done to eliminate them. Therefore, this script will evolve considerably between now and October 20.

Points to consider in the data include the following:

- Reversals – higher theta cuts in lower grades
- Differential pacing – with a vertical scale, uneven size of increase from one grade to the next (e.g., a small difference between theta cuts for grades 6 and 7 but a large difference between grades 7 and 8)
- Differential opportunity to learn – better implementation of the CCSS at certain grades
- Differential test difficulty – potentially interacting with Pacing
- Idiosyncrasies in one or more panels – systematically stringent or lenient interpretation of the ALDs; differences in interpretation of the charge

Keep all of these possibilities in mind, but be prepared for other suggestions that may come from the Vertical Articulation Committee.

Once VAC training is over, display the actual Round 3 impact graphic. Review the process by which these impact data were obtained, and give panelists a few minutes to size it up. Then ask:

Do these results look more like the generally increasing, stable, or generally decreasing examples we saw earlier?

Give them time to form an opinion. Depending on the overall shape of the graphics, they could be any of the above or no particular shape at all. Say:

Keep in mind that each point on this graph represents the performance of a whole grade of students in a particular grade. For example, the point for Grade 6, Level 3, represents the percentage of grade 6 students scoring at or above Level 2. Looking just at the graph for Level 3, do you see any points that seem out of place; that is, a point that seems higher or lower than you would expect, given performance in the adjacent grades?

Give panelists time to identify any anomalous points on the Level 3 graph only. If anyone identifies a point on the Level 2 or Level 4 line, note that we will get to Levels 2 and 4 later; we are focusing on Level 3 now. It is possible that two or more points will be identified. Ask each panelist who identifies a point to tell how it relates to adjacent grades' points and how it seems out of place. After each panelist explains his or her data point, select one for the group to discuss. The following is a hypothetical example in which one panelist identified a dip at grade 6 and another identified an unusual rise at grade 4.

Before exploring any one cut score, introduce the Online Panel results. Superimpose the cut scores and impact for each of the OP groups (teachers, administrators, etc.) on the Round 3 results for Level 3 and say:

Last week, each of you saw Online Panel results for one grade. Now, we're going to take a look at those results for all grades simultaneously. Here you see the percentages of students at Level 3 or

above across grades, based on your Round 3 cut scores, as well as the corresponding percentages based on the recommendations of each of the various groups who logged in and completed this exercise online. You will notice that we have represented their recommendations not as exact cut points but as ranges, as, for example, here for third grade, where the Level 3 cut score recommended by teachers ranged from [X to Y], and the percentage of third graders at or above Level 3 would be [A to B]%. In most cases, the points on the In-Person Workshop chart fall within the ranges of the Online Panels' cut scores.

If the representation of all subgroups plus the IPW is too busy, show the IPW results for Level 3 with superimposed results for just one or two OP subgroups at a time.

Give the panelists a few minutes to consider the OP data and ask questions about the process and outcomes of that activity. Where any of the ranges for any OP subgroup are remarkably different from that of the IPW panel, note the discrepancy in advance and be prepared to show items in the OIB within that range but outside the interquartile range of the IPW cut score recommendations. Then reiterate the ground rules:

We can change any one of the 42 cut scores set last week or none of them. To make any change, we will need a formal motion to change a specific cut score for a specific level for a specific grade to a specific value, associated with a specific page or set of pages in the ordered item booklet. For example, if you want to raise the cut score for grade 7, Level 2, you would say, "I move that the cut score for grade 7, level 2 be moved from 1.47, which is between pages 16 and 17 to 1.63, which is on page 19." Then someone will need to second the motion. Once we have a motion and a second, we will discuss the pros and cons of the motion. I will call the question, and we will vote on the change. Given the level of effort that has already gone into each of these cut scores, Smarter Balanced has decided that a 2/3 majority will be required for passage of any motion. If a motion fails to get a 2/3 majority, it fails; if it receives a 2/3 majority, it passes, and the cut score is changed. Are there any questions about the procedure?

Let's take a look at grade 6. From grade 5 to grade 8, there is a general increase in the percentage of students performing at Level 3 or higher, except for grade 6. Does it seem logical that there should be such a "dip" at grade 6?

Wait for responses, and in each instance, ask the panelist to tell why it does or does not seem logical for such a "dip" to occur, without any further information about the tests or students who took them. Use panelists' comments to guide the next portion of the discussion.

It looks like most of you found this "dip" unusual. Let's take a look at the actual Level 3 cut score for grade 6, relative to those for grades 5 and 7, for example. We can see from the first table at the upper left of the screen that

[Scenario 1: Differential pacing] *The difference between the cut scores for grades 5 and 6 is much larger than the difference between the grade 6 cut score and the grade 7 cut score or between the grade 7 cut score and the grade 8 cut score. In other words, we have asked sixth graders to make more progress from grade 5 to grade 6 than we have asked seventh graders or eighth graders to make from the previous grade. Is it reasonable to have such an expectation?*

Allow panelists to respond to this question. If such an expectation is reasonable, it may not be necessary to consider this part of the graph at all. However, if most panelists believe it is not reasonable, move forward:

All right; it seems most of you believe that it is unreasonable to expect students to make larger gains grade 5 to grade 6 than from grade 6 to grade 7 or from grade 7 to grade 8. We could conclude that the grade 5 cut score is on target, but the others are off target; or we could conclude that grade 5 is off target, and the others are on. For the moment let's assume the others are on target, and grade 5

is off. Let's see how we arrived at this level 3 cut score for grade 5. Can someone from the grade 5 panel give the rest of us a brief summary of the last round of discussion you had surrounding the Level 3 cut score?

[Scenario 2: Reversal] *When we look at the Level 3 cut score for grade 5, we see that it is actually higher than the Level 3 cut score for grade 6. We have placed these tests and items on a common scale, so what this seems to be telling us is that we expect the fifth grader just entering Level 3 to know more and be able to do more than the sixth grader just entering Level 3. Is it reasonable to have such an expectation?*

Allow panelists to respond to this question. If such an expectation is reasonable, it may not be necessary to consider this part of the graph at all. However, if most panelists believe it is not reasonable, move forward:

All right; it seems most of you believe that it is unreasonable to expect fifth graders just entering Level 3 to know and be able to do more than sixth graders just entering Level 3. Can someone from the grade 5 panel give the rest of us a brief summary of the last round of discussion you had surrounding the Level 3 cut score?

For both scenarios (plus others that might be imagined), if no one responds immediately, call on a couple of people from the fifth grade panel, using your roster and making sure not to put any one person on the spot. After they have given a review of the panel's rationale, ask all panelists to call up the grade 5 OIB and turn to the region in which the Level 3 cut score was set. This should include 2-4 items on consecutive pages. Direct the panel to consider these items for a moment. Then say,

Given the Level 3 cut score set by the fifth grade panel, we are saying that the performance of students just entering Level 3 would be best characterized by these items and tasks or others similar to them. In fact, it comes closest to the item on page [X, the page with the RP value closest to the cut score but not below it]. I would also point out that half the panel placed their Level 3 bookmarks between pages [XX and YY, (the interquartile range for this panel for this cut)]. Now, considering the items in the vicinity of the cut, would anyone care to suggest that one of these other items on a previous page would be a better indicator of what students just entering Level 3 can do?

Wait for someone to suggest a page number. If no one suggests a page number, ask again if the items on the two pages surrounding the cut seem to be reasonable representations of what students just entering Level 3 know and can do. If there are still no recommendations, ask,

Do I hear a motion that we alter the cut score for grade 5, Level 3?

If someone responds, make sure the motion is specific, as described above. Record the motion and call for a second. If there is no second, the motion fails for lack of a second. If there is a second, record the second, and proceed to the discussion.

It has been moved and seconded that the Level 3 cut score for grade 5 be changed from [XX] to [YY]. Is there any discussion? Raise your hand if you would like to comment. We will have [AA] minutes to discuss this cut score, after which I will call the question.

Call on those who raise their hands, in order. The assistant will help with this task. The key feature of this discussion will be to urge all respondents to ground their comments in the ALDs and the contents of the items at or near the cut scores they are discussing. During the discussion, panelists will have access to all the OIBs and item maps, all the content standards, and all the ALDs. Facilitators will have the same access and the ability to project their own screens on the large screen in the front of the room. For every comment not based in the ALDs or content standards, politely ask the respondent to check the appropriate online document in defense of the point he or she is

making. At key points, it may be necessary to show on the big screen a particular page of the CCSS, ALDs, or some other document to make sure everyone is focusing on the same information.

If no one mentions any of the pages corresponding to cuts suggested by the Online Panel, mention some examples and ask the VAC what they think of those items as representative of what students should know and be able to do. Give ample opportunity to consider these items and debate their strengths and weaknesses.

Gauge in advance the likely number of cut scores that may need to be considered and pace the discussion accordingly. Level 3 will generate the most discussion, so if it seems to be running a bit long, keep this in mind. Levels 2 and 4 will go considerably faster. Call the question at the appointed time. Ask for a show of hands of those who favor the motion. Have the assistant and one other person count hands. Then ask for a show of hands of those who oppose the motion. Have the assistant and one other person count hands. Tally the votes and announce the result. If the motion received a 2/3 majority, it passes; otherwise, the motion fails. If the motion fails, ask for another recommendation, and repeat the process. If no one recommends a new cut score, discussion of that cut score for that grade ceases. At that point, call the panel's attention back to the chart and ask if there are other cut scores that should be considered. For each such cut score, follow this same procedure.

When it is apparent that there will be no more changes for Level 3, ask for a motion to cease making changes for Level 3. On obtaining a motion and second, call for discussion. Call the question, and vote as before. Again, a 2/3 majority is required for passage of this motion. It is important to include this vote simply to preclude later charges that someone wanted to consider a particular cut score but was denied the opportunity. A formal vote by the full panel will show that the panel itself, not the facilitator or Smarter Balanced, foreclosed further discussion. If the motion to cease discussion of any level fails, open the floor for additional motions for any cut score for that level. Give a full minute for the panel to think about it. If no one offers a motion in that time, it is entirely appropriate to call that level closed. This action on the part of the facilitator may later be reversed if changes in another level make it clear that discussion at this level should be reopened. For example, if a Level 2 cut score is raised so much that it is very near the previously approved (or unchallenged) Level 3 cut score, it is entirely appropriate to reopen discussion of Level 3 for that grade. However, this could reopen the entire set of Level 3 cut scores. Exercise caution and point out such consequences when the recommendation to raise the Level 2 cut score that high is made.

Take a 15-minute break after Level 3 and prepare for Level 2. While panelists are on break, check the voting records for Level 3, and check your notes for Level 2. Upon resumption of the vertical articulation, follow the same procedures as above, with the exception of OP results (there won't be any).

Take a 15-minute break after Level 2 and prepare for Level 4. Follow the same procedures as in Level 2.

At the close of voting for Level 4, display the final cut scores and impact data for all grades and levels. On the cut score chart, highlight those that have been changed. Thank panelists for their assistance, and ask them to turn to the final evaluation form on their screens.

Joe, Tony and Christyan were highly complementary of the field test process, plan and execution. Mike's abilities managing staff and keeping the conversation going was excellent. Mike showed superlative skill in debriefing and they have a lot of confidence in plan and execution for Dallas meeting. The fact that validity was close to real was impressive, they can say it worked. The software looks good, the facilities impressive and there was good cross-contractor interaction. Review of feedback each night is important.

In Dallas, a command center is needed to review ourselves. Tony commented they did not get to practice challenges they may face, we need to identify the escalation process when an emergency arises. MI will have Joe McClintock in Dallas, a psychologist to help "council" panelists with an issue who has experience and has been involved in the teacher evaluation process. Mike explained the facilitator will discuss with panelist and escalate to Mike or Dan Lewis who will decide to escalate to Joe or Tony. Feedback survey will be done via Survey Monkey, Nelson working on incorporating a Pause button place holder to come back to.

- We plan to act on feedback daily in Dallas with a 5:15 PM debrief
- Lunch hour debrief-opportunity for panelists to "park" questions that can be answered in the last 20 minutes of the lunch hour.
- Coordinate Joe and Tony on who answers what? Joe/ELA; Tony/Math
- Mike envisions he and Dan handling any issues from facilitators and bring policy issue to Joe and Tony and take response back to facilitators to dissemble to panelists as a single message to whole group, it's more efficient.
- We will delegate if it's Joe or Tony "on call" and Kathleen needs to know your schedule and how to reach you so Mike has that info.

Social media- Tony raised concern on managing Tweets. Mike said no cell phones are allowed in breakouts, no paper notes leave the room. Hager Sharp will be monitoring Tweet and has a plan to get the message out. After the 9/22 Advisory Panel in Chicago Ted will hold a Contingency meeting 9/23 until 11:30 AM, he will have materials to lay out plan.

Facilitators-Mike assured there will be 2 in each room which works out well; if 1 has to leave, the room has an alternate monitor. All did a good job at FT.

Round 1 debrief- confusion on bookmark procedure/instructions-this needs works. Joe expressed concern over Dan's analogy of people v. people not people v. item. He suggested to try a skiing analogy instead i.e., person v. mountain. A sense of moving from one class to another based on skills learned, onto a new set of challenges.

- Christyan suggested a 2 sentence "script" posted in each room the reiterate the message from opening meeting, he spoke to Juan about tightening this up so each facilitator says the same thing to guide the training.
- Suggested to move the modeling of setting the bookmark to Day 2 morning session for all in ballroom so they are more familiar with the process and have Dan demonstrate making it a group activity. This reinforces what's heard in general meeting with statement in breakouts.

- MI to check on ballroom availability- Mandy is on this.

Joe needs a script and to be hounded to read, revise and practice with Mike on Saturday or Sunday Oct. 11 or 12th for approximately two hours.

Training: There will be webinar training available to all MI and CTB facilitators prior to the Dallas meeting to become familiar with the software. On Sunday, Tuesday and Thursday evenings in Dallas:

- 4-6pm Facilitators led by Rick Mercado
- 6-8pm Table Leaders led by Jennifer Lord-Bessen

Emotional drain-off-discussion of when and how long? Handle in the larger group forum.

- We need to let panelists know ahead of time if we field questions during the last 20 minutes of lunch.
- Note cards can be collected and the most common question answered in this way by Joe or Tony. Joe also suggested an “open mike” for anyone to ask.
- Mike and Joe to script for large meeting which Mike will address group and Joe will suggest topics that are out of bounds. Joe should not address as Executive Director.

Food/beverage/dinner plan: arrange a conference call to include Joe, Tony, Jaci, Nancy, Christyan, Barbara, Mike, Mandy and Kathleen to discuss meal plan as it is and available options, this could be the last 20 minutes of a catchall meeting lasting 2 hours that gets into meeting details, as well as desegregating data and bookmark procedure.

- How many food stations? Where is location? What is being served and when?
- Layout breaks, lunch, coffee service, if dinner on their own; where are restaurants located and how to get there, explain shuttle service

Hager-Sharp-Joe questioned HS activities. Mike will have Barb address to Joe and Tony what they have done so far and their future plan of action. Joe’s concerned as his only interaction concerns the video concept. Tony is concerned that HS get out in front of the “SBAC letting just anyone except teachers be involved in this process”. The state leads meeting is in September and we need to be concerned with what happens after. What is the communication to go out to professional organizations? Their 2 key responsibilities are:

- get excitement brewing before the event
- get the state buy in after the results on 11/7
- create high level document on strategies before and after the event

Mike said they will be providing materials at the Contingency meeting.

Disaggregated data- Mike will send models along with panel demographics to Nancy, Joe and Tony. Just Level 3 and above and on the VAC. Tony asked for sub-groups across grades; ethnicity; ELL; SWD and Male/Female. Mike said we are tracking panelist demographics and will check with David Breen on this.

Security - Joe concerned on casual security here at FT. We need to tighten up in Dallas. He mentioned a coded note paper (bar code or IP#) handed out on the way in and collected on the way out at the end of the day.

TAC- 10/30 in Portland, OR, all day meeting. 10/31 for USED, 2 hour meeting. Greg to be at both. At TAC show software screenshots, raw footage from Hager Sharp video to get an impression of scope. Mike needs to attend. Susan Phillips should attend USED only.

- TAC – 1 representative per state
- CCSSO invited to both meetings
- HE Leads invited to TAC

Christyan asked Mike for tiered escalation process written up. Mike has as part of original work plan to send Christyan. He also suggested Mike expand bookmark introduction to 20 minutes.

External data – Mike explained we have data from Michigan only. Other states, no response. We will use NAEP and PISA data. OK that PISA is 14 years old and pre NCLB, as long as we have 1 near proficient per item, PISA would not allow the use of their operational data.

11/6 meeting with Chiefs – what are we asking? Joe “ we are asking to approve not for themselves but for SBAC (states) Mike will craft sentences around this statement on ALS and send to Joe and Tony to review and revise.

Mike and Joe have been invited to speak at NCME April 2015 in San Diego.