

Setting Cut Scores on Smarter Balanced Assessments: A Ground-Level View

Michael B. Bunch
Measurement Incorporated

Citation: Bunch, M. B. (2015). Setting Cut Scores on Smarter Balanced Assessments: A Ground-Level View. In L. Keng (Moderator), *Standard Setting in the Common Core World: PARCC and SBAC Experiences*. Invited symposium presented at the annual meeting of the National Council on Measurement in Education, Chicago, IL.

Setting Cut Scores on Smarter Balanced Assessments: A Ground-Level View

Michael B. Bunch

Abstract

This paper focuses on the preparations for and execution of a Bookmark procedure for an on-site/online achievement level setting involving over 3,000 panelists in the setting of cut scores for 14 Smarter Balanced tests in the fall of 2014. Recruiting, communications, and logistics are also discussed.

The outcomes of the fall 2014 achievement level setting are by now quite well known. The online and in-person components of that project were a massive undertaking, involving over 3,000 individuals. Detailed documentation of these activities is available at the Smarter Balanced website: www.smarterbalanced.org. Therefore, this paper focuses primarily on the behind-the-scenes activities necessary to achieve that massive undertaking and arrive at those outcomes. The view is decidedly ground-level.

Preparations

In 2013, I presented a presession at NCME with the title “Setting Cut Scores on 21st Century Tests.” That presession was a follow-up to an NCME presession I had done in 2012, titled “Standard Setting in the Real World.” Both presessions laid out preliminary plans for setting cut scores for large-scale, high-stakes tests such as those associated with PARCC and Smarter Balanced. Module 2 of the 2013 presession addressed five aspects of preparation: People, meeting sites, materials, lodging, and meals. We took the basic elements of that module and applied them to the Smarter Balanced achievement level setting project.

People. Smarter Balanced achievement level setting required the services of five organizations working together: Smarter Balanced, WestEd, Measurement Incorporated (MI), McGraw-Hill Education (CTB), Hager Sharp, and an independent audit team. Within each organization, there were public-facing as well as behind-the-scenes personnel. Roles and responsibilities are provided below.

- **Smarter Balanced/OSPI** – Overall supervision of contractors and advisors, overseen by Executive Director Joe Willhoft (now retired) with assistance from the Washington Office of the Superintendent of Public Instruction (OSPI) and Nancy Arnold, contract manager. Additional public-facing Smarter Balanced staff included Tony Alpert (now Executive Director), Jacqueline King, Dacia Hopfensperger, Marty McCall, Magda Chia, Shelbi Cole, Nikki Elliott-Schuman, and Nicole Siegel. Additional OSPI staff monitored the contract’s financial aspects. State Leads for K-12 and higher education in each Smarter Balanced state were also actively involved.
- **WestEd** – General project monitoring, led by Christyan Mitchell, assisted by Kate Darling, Mira Weinstein and Ann Appert (meetings, calls, and webinars).

- **MI** – Prime contractor responsible for all contract deliverables. Barbara Scherich served as program manager, and Michael Bunch served as lead facilitator. Additional public-facing staff included psychometricians and content specialists who served as panel facilitators and trainers, meeting specialists, and network operations (netops) specialists who assisted during the in-person sessions whenever any panelist had a hardware or software problem.
- **CTB** – Subcontractor responsible for panelist recruiting and auditing and assisting with panel facilitation. This group was led by subcontract manager David Breen and co-lead facilitator Daniel Lewis. Additional psychometricians and content specialists assisted in panel facilitation and data analysis as well as facilitator and table leader training.
- **Hager Sharp** – Subcontractor for communications. This team was led by Debra Silimeo and Ted Eismeier, with assistance from writers, copy editors, graphic artists, and other behind-the-scenes staff. In addition to drafting communications for Smarter Balanced, Hager Sharp worked with media outlets throughout the country and conducted dozens of interviews of panelists in Dallas to produce a video recording of the proceedings to share with Chiefs and others in Smarter Balanced states.
- **Audit Team** – Greg Cizek and Heather Koons served as the independent auditors, paid by CTB but reporting directly to Smarter Balanced. Susan Phillips performed a separate legal audit.

Meetings, webinars, and conference calls. The project began with a kickoff meeting on February 24-25, 2014, in Minneapolis, the purpose of which was to define roles for all parties and to sketch out the timeline and deliverables for the next nine months. Over the life of the project, there were several more such meetings, to review progress toward those deliverables and more clearly define upcoming deliverables. There were additional meetings specifically for the Technical Advisory Committee (TAC) as well as an Advisory Panel. Meeting dates, purposes, and participants are summarized in Table 1. All meetings took place in 2014. This table does not include the October achievement level setting meetings or the November meetings with Chiefs to approve the cut scores.

Table 1
Planning and Related Meetings

Date(s)	Meeting/Location	Purpose(s)/Outcome(s)	Participants
February 24–25	Kickoff Meeting/ Minneapolis, MN	Review proposal, clarify roles and deliverables	Smarter Balanced, MI, CTB
March 12	Hager Sharp Planning Meeting/Washington, D.C.	Outline Communications Plan	Smarter Balanced, MI, Hager Sharp
March 19	Hager Sharp Planning Meeting/Washington, D.C.	Revise and approve Communications Plan	Smarter Balanced, MI, Hager Sharp
April 3–7	AERA/NCME Conference/ Philadelphia, PA	Present plan for Smarter Balanced achievement level setting to educational measurement community	Joe Willhoft

Date(s)	Meeting/Location	Purpose(s)/Outcome(s)	Participants
April 29– May 2	Collaboration Conference/Minneapolis, MN	Meet with other contractors; present plan to Chiefs for approval	Smarter Balanced, MI, Chiefs, WestEd, Test Development & Validation Committee
April 30	Technical Advisory Committee/Minneapolis	Present plan to TAC; respond to technical concerns	Smarter Balanced, MI, TAC, WestEd
June 25–27	National Conference on Student Assessment/ New Orleans, LA	Present plan to the educational measurement community	Joe Willhoft, Greg Cizek
July 16–17	Technical Advisory Committee/Portland, OR	Review and revise plans for online and in-person ALS	TAC members, Smarter Balanced, WestEd, MI
July 17–18	Achievement Level Setting Advisory Panel/ Portland	Present preliminary plans to Advisory Panel for policy recommendations	Smarter Balanced, MI, WestEd, Advisory Panel
August 18- 21	Field Test/Durham, NC	Test procedures and materials to be used in October	Smarter Balanced, WestEd, MI, CTB, auditors
September 22	Achievement Level Setting Advisory Panel/ Chicago, IL	Present revised plans for October and receive policy recommendations	Smarter Balanced, MI, Advisory Panel
September 23	Contingency Communications Training/Chicago	Plan for dissemination of information during and after achievement level setting	Smarter Balanced, MI, Hager Sharp
October 30	Technical Advisory Committee/Portland	Present results of achievement level setting; receive policy recommendations	Smarter Balanced, MI, TAC
October 31	United States Education Department Briefing/ Portland	Brief USED officials on conduct and outcomes of achievement level setting	Smarter Balanced, MI, USED

It is noteworthy that staff, both across and within organizations, were located in many states, from coast to coast. Face-to-face meetings were fairly rare; the most common form of meeting was the webinar. These were typically set up by MI staff, with agendas and review materials sent electronically to participants in advance. A typical webinar would last from one to two hours and include PowerPoint presentations, open discussion, private chat, and a question-and-answer session. Some webinars were conducted with WebEx, and others were conducted with TurboMeeting. Table 2 summarizes the various webinars conducted over the life of the project.

Table 2
Webinars

Date(s)	Meeting	Purpose(s)/Outcome(s)	Participants
March 3-4	Preparation for Technical Advisory Committee (TAC) meeting	Identify TAC topics and develop responses	Smarter Balanced, WestEd, MI, CTB
March 6	TAC	Present plan for achievement level setting	Smarter Balanced, WestEd, MI, CTB, TAC
March 11	Leads Conference	Present achievement level setting plan to State Leads and Chiefs	Smarter Balanced, WestEd, MI, State Leads, Chiefs
March 25	Leads Conference	Delineate roles of State Leads in recruiting	Smarter Balanced, WestEd, MI, CTB, State Leads
April 7	Recruiting Materials Review	Review recruiting materials developed by CTB and Hager Sharp	Smarter Balanced, MI, CTB, Hager Sharp
April 11	Recruiting Materials Review	Review recruiting materials developed by CTB and Hager Sharp	Smarter Balanced, WestEd, MI, CTB, Hager Sharp
April 15	State Leads	Review of recruiting schedule	Smarter Balanced, WestEd, MI, CTB, Hager Sharp, Leads
April 17	Management Meeting	Review of recruiting schedule	Smarter Balanced, WestEd, MI, CTB, Hager Sharp, WestEd
April 22	K-12 Leads	Review of recruiting schedule	Smarter Balanced, WestEd, MI, CTB, Hager Sharp, Leads
May 14	Advisory Panel Plan	Review of changes in make-up of Advisory Panel	Smarter Balanced, MI, WestEd
May 30	TAC	Present updated plan for achievement level setting	Smarter Balanced, WestEd, MI, CTB, TAC
June 13	Public Information Officers	Present updated plan for achievement level setting	Smarter Balanced, WestEd, MI, CTB, Public Information Officers
August 27	TAC	Present updated plan for achievement level setting	Smarter Balanced, WestEd, MI, CTB, TAC
September 10	Background Briefing for Trade Media	Trade media briefing	Smarter Balanced, WestEd, MI, Hager Sharp

Date(s)	Meeting	Purpose(s)/Outcome(s)	Participants
September 12	USED	Present updated plan for achievement level setting	Smarter Balanced, MI, USED representatives
October 28	Achievement Level Setting Advisory Panel	Present results of achievement level setting	Smarter Balanced, WestEd, MI, Advisory Panel

In addition to the meetings and webinars, there were weekly project management conference calls, scheduled for each Thursday at 11 A.M. Eastern Time. Starting in April, ongoing cross-contractor calls, facilitated by WestEd, included this project. These calls were scheduled for every Tuesday and Friday for the duration of the project and included opportunities for MI and its subcontractors to interact with contractors carrying out other Smarter Balanced projects on matters of mutual concern. Also starting in April were weekly recruiting calls and webinars. Beginning on May 16, 2014, weekly reminders were distributed to State Leads and teacher involvement coordinators (TICs) 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. Additional telephone calls and e-mail exchanges were a daily occurrence. It was not unusual for 30-40 e-mail messages to be exchanged on any given day.

Hardware and software. Hardware for this project included not just computers and peripherals but other equipment as well. Most of the equipment was shipped to the in-person workshop site in Dallas by MI, but some was purchased or rented on site:

- 400 Celeron Lenovo PCs running Windows 7 with wireless internet access (shipped from MI)
- 400 flat-screen 21" monitors (shipped from MI)
- 400 headsets (shipped from MI)
- 2 servers (shipped from MI)
- 12 LCD projectors (1 per break-out room; shipped from MI)
- 12 projector screens (1 per break-out room; shipped from MI)
- 3 portable hand sanitizer stands (shipped from MI)
- 1 flat-screen 32" monitor for a sight-impaired panelist (purchased on site)
- 400 internet connections (rented on site)
- 4 tripod speaker stands (rented on site)
- 4 12" speakers (rented on site)
- 2 4-channel mixers (rented on site)
- 2 headset microphones (rented on site)
- 4 handheld microphones (rented on site)
- 1 LCD projector (6500 lumen XGA) and screen (7'6" x 13'4") for large-group instruction (rented on site)
- Cabling for sound systems (rented on site)

The software used to drive the online as well as the in-person achievement level setting grew out of two NCME presessions (Bunch, 2012b, 2013). In those presessions, I outlined procedures for setting up and conducting large-scale standard setting sessions involving hundreds of panelists. In the 2013 presession, I introduced Virtual Bookmark Plus (VBM+), an online approach to standard setting. Requirements gathering and specifications for the software used in October 2014 thus began in the summer of 2012, with building and testing continuing through the summer of 2014.

In August 2014, MI and CTB conducted a field test of the online and in-person achievement level setting process, including the software and hardware to be used in October. That field test revealed that the software was fundamentally sound but required minor adjustments to accommodate the much larger number of anticipated online panelists who would not have the same access to MI netops staff that field test panelists had and that in-person panelists would have. With these changes made, the software was ready for operation in October. Figures 1-10 show some of the key features of the software as it was implemented for the online and in-person panels in October.

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.

The screenshot displays a three-step instructional guide. Step 1, 'Access Reference Materials', is in a blue bar with a 'More Info' link. Step 2, 'Practice Setting Your Achievement Level Bookmarks', is in a green bar with a 'Less Info' link. Step 3, 'Set and Submit Your Achievement Level Bookmarks', is in a blue bar with a 'More Info' link. The content for Step 2 is expanded, providing a list of prerequisites and a three-step process for setting and submitting achievement level bookmarks.

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.

Figure 1. Home Page with One Instruction Bar Expanded

The home page contained a list of all resource materials, accessible through hyperlinks, as shown in Figure 2.

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

Figure 2. List of Resources Accessible From Home Page

The online item map page (see Figure 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 panelists 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.

Item Map for Booklet: Achievement Level Setting Booklet

Mathematics, High School

Round: 3

Display additional information:

☒ ACT

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	ACT
25	576	Claim: 1, Target: P	2			<input type="checkbox"/>	
26	580	Claim: 1, Target: J	2			<input type="checkbox"/>	
27	583	Claim: 4, Target: F	3			<input type="checkbox"/>	
28	584	Claim: 1, Target: G	1			<input type="checkbox"/>	
29	586	Claim: 3, Target: C	3			<input type="checkbox"/>	
30	589	Claim: 1, Target: J	2			<input type="checkbox"/>	
31	597	Claim: 4, Target: A	3			<input type="checkbox"/>	
32	599	Claim: 1, Target: L	1			<input type="checkbox"/>	
33	600	Claim: 2, Target: A	2			<input type="checkbox"/>	
34	602	Claim: 1, Target: O	1			<input type="checkbox"/>	
35	605	Claim: 2, Target: C	2			<input type="checkbox"/>	
36	608	Claim: 1, Target: J	2			<input type="checkbox"/>	
37	608	Claim: 3, Target: F	2			<input type="checkbox"/>	

Set Bookmark

Figure 3. Sample Item Map

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 4 shows a sample selected-response item, while Figure 5 shows the associated item information page, and Figure 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 being used.

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 4. Sample OIB Page with Selected-Response Item

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 5. Item Information Page

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 effectively 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
I_x




Figure 6. OIB Page for a Constructed-Response Item

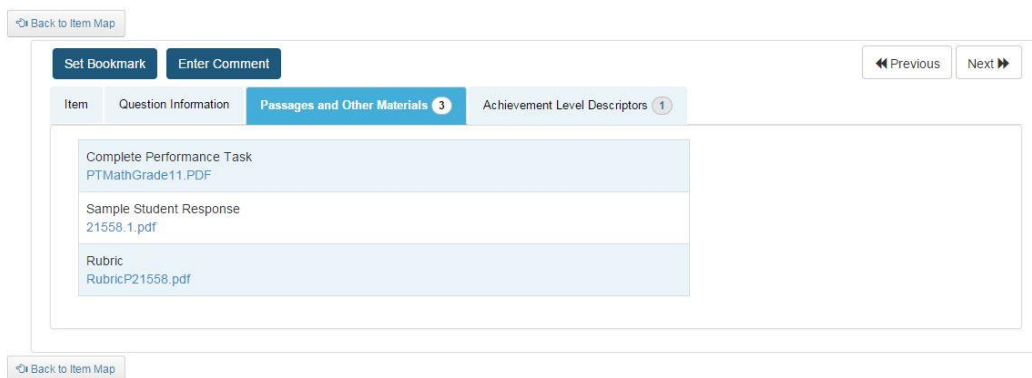
By clicking “Passages and Other Materials,” panelists could see resource materials such as reading or listening passages, sample student responses, and scoring rubrics, as shown in Figure 7.

Ordered Item Booklet: Achievement Level
Setting Booklet

Mathematics, High School

Page: 31

Round: 2



The screenshot shows the 'Ordered Item Booklet' interface. At the top, there are buttons for 'Set Bookmark' and 'Enter Comment'. Below these are navigation tabs: 'Item', 'Question Information', 'Passages and Other Materials' (which is active and has a '3' in a circle), and 'Achievement Level Descriptors' (which has a '1' in a circle). To the right of the tabs are 'Previous' and 'Next' buttons. The main content area displays three links: 'Complete Performance Task' with 'PTMathGrade11.PDF', 'Sample Student Response' with '21558.1.pdf', and 'Rubric' with 'RubricP21558.pdf'. There are 'Back to Item Map' buttons at the top left and bottom left of the content area.

Figure 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 8 illustrates the “Comment” function.

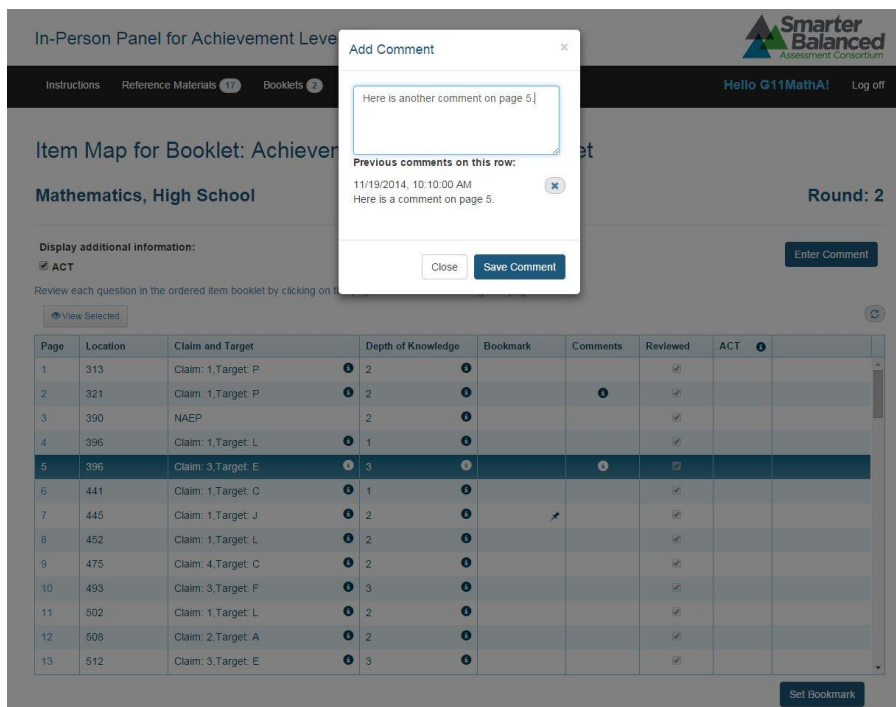


Figure 8. Comment

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 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 10.

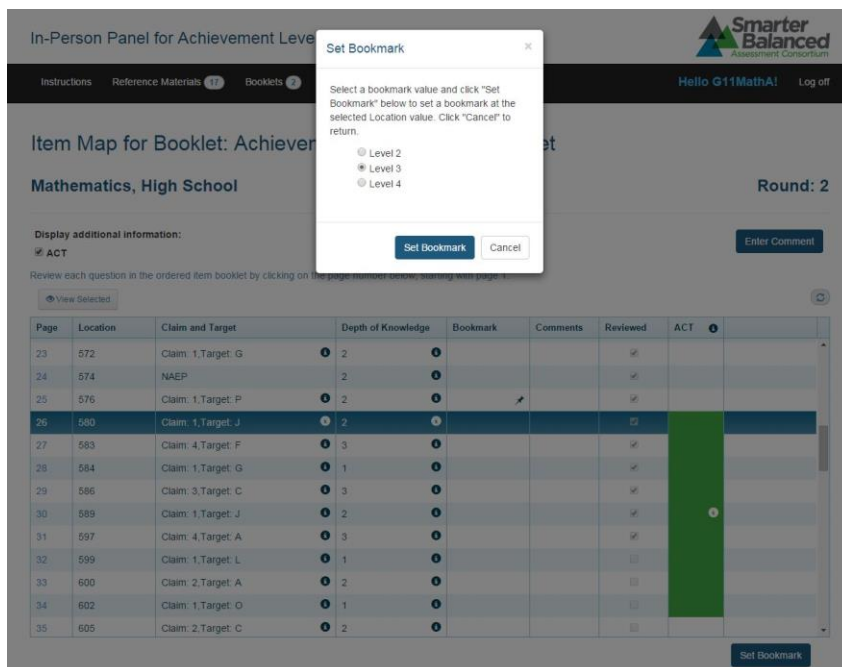


Figure 9. Set Bookmark Dropdown Box in the Item Map

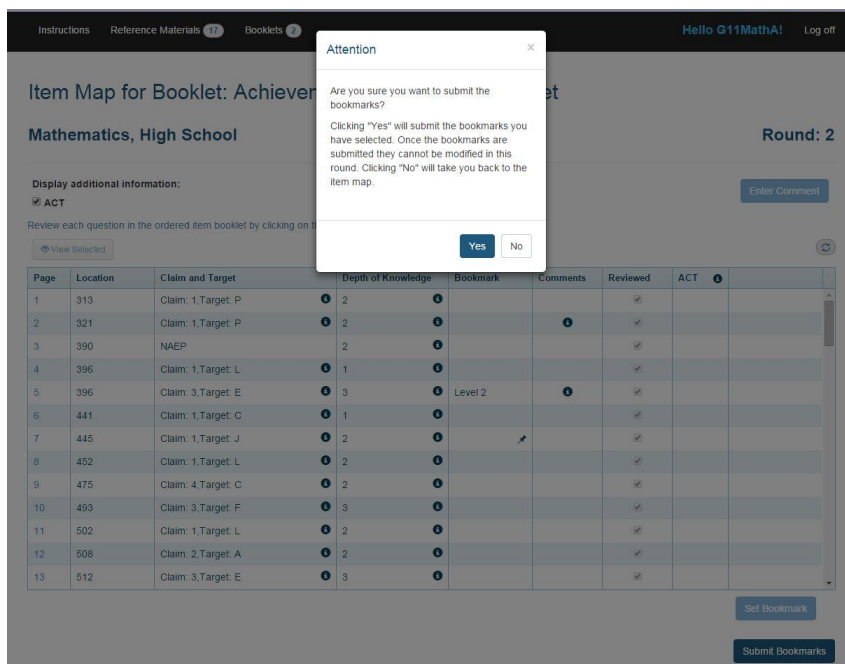


Figure 10. Submitting Bookmarks

Ordered item booklets. MI staff prepared ordered item booklets (OIBs) in September. The timing of OIB development was one of the most intricate parts of the project. Tests were developed under one contract and administered under another. Administration, data collection, analyses, and calibration all occurred outside the MI achievement level setting contract. One of the benefits of the twice-weekly cross-contractor calls was the opportunity to track the progress of the field test administration, data collection, scoring, and analysis tasks being conducted by other contractors.

American Institutes for Research (AIR) was responsible for the online item bank and the software to deliver the tests. MI and AIR held a planning session at the April-May Collaboration Conference in Minneapolis to make sure the handoff of calibrated items, scheduled for September 11, was successful. That meeting focused on the transfer of item text and graphics as well as calibrations and other metadata. Throughout the spring and summer, MI and Smarter Balanced staff worked with the TAC and other groups to clarify the requirements of the OIBs. On May 19, MI and AIR conducted a very successful dry run of the handoff. When the handoff finally occurred on September 11, it was flawless.

An important consideration when creating an ordered item booklet (OIB) is to ensure appropriate content coverage. Psychometricians and content specialists from MI worked together 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 no reading passages with fewer than three items presented.

Item statistics 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 item 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.

Some of the booklets also had links to external benchmarks. Specifically, OIBs for grades 4 and 8 had items from the National Assessment of Educational Progress (NAEP). High school tests were linked to ACT scores via a special regression analysis done in September. NAEP items were selected to have values as close to cut scores as possible and were so designated in the OIBs and item maps. For the ACT link, we estimated the Smarter Balanced scale score equivalent of the ACT college/career ready score and represented it on the item map as a score range rather than a single point (see Figure 3). Panelists received instruction as to how the external benchmark data were chosen and how to consider them in recommending cut scores.

Training materials. All training materials were based on materials MI and CTB had used in the execution of Bookmark activities for many years. These materials were adapted to the special circumstances of the joint online/in-person activities to be carried out in October 2014. We field tested all training materials in August, made adjustments as necessary, and obtained final approval from Smarter Balanced before implementing them in October. Table 3 summarizes the training materials developed and used in this project.

Table 3
Training Materials

Product	Purpose	Brief Description
Amplified Agenda	Keep facilitators and panelists on schedule	Complete task-by-task agenda of the workshop
Facilitator Script/Supplement	Provide consistent instruction across panels	Detailed bulleted points for facilitators to make, accompanied by detailed script and reference materials
Table Leader Training PowerPoint	Orient table leaders to their task	34 slides describing ALDs, the Bookmark procedure, and round-by-round tasks
Observer Script	Orient observers to the process and groundrules	6-page description of the schedule and process
In-Person Intro PowerPoint	Orient panelists to the tasks to be performed	35 slides providing background information as well as detailed instructions regarding schedule and logistics
Achievement Level Descriptors	Orient panelists to ALDs	Specially formatted ALDs for each test highlighting expectations of students just barely at each cut score

Product	Purpose	Brief Description
OIB Training/Introduction to the Bookmark Procedure PowerPoint	Orient panelists to the Bookmark procedure	39 slides showing how the Bookmark procedure translates panelists' bookmarks into cut scores
Links to Smarter Balanced Materials	Orient panelists to ALDs, and Practice Test	All panelists took one Smarter Balanced Practice Test and reviewed other materials on the website.
Cross-Grade Review PowerPoint	Orient panelists to the cross-grade review process	25 slides showing purpose and process of cross-grade review

Recruiting

Recruiting people to serve on the online, in-person, and cross-grade review panels required the efforts of a team of CTB, Hager Sharp, State Leads, TICs, and MI staff working for six months to identify, contact, recruit, and retain representative panels for each subject and grade. CTB, MI, and Smarter Balanced established networks in March to involve state-level higher education and public school leaders in each of the Smarter Balanced states. Hager Sharp crafted several recruitment messages and expanded the network to national educational media. State Leads and TICs operated at the grass-roots level.

The recruitment process was just the first step. Having found representative panelists, the next step was to keep, inform, and motivate them and get them to Dallas for the in-person activity or to the website for the online activity. Over 3,000 people were involved. Some district and school administrators changed their minds about allowing their teachers time off to participate. People got sick. One had a death in the family. Some cancelled for no stated reason. It was therefore necessary to have a list of fully qualified and representative back-up panelists. The back-up list proved its value when news of the Ebola case in Dallas broke in September.

Online panel. The purpose of the online panel was to allow as wide an audience as possible to see and comment on the Smarter Balanced assessments. From the outset, it was the goal of Smarter Balanced to use feedback from the online panel to inform, but not necessarily direct, the deliberations of the in-person panels.

Recruitment of the online panel began in April, with primary emphasis on educators. In June, recruitment opened up to the general public. Prior to any recruitment activities, however, the project team tested various recruitment messages. Specifically, Hager Sharp drafted and tested messages on focus groups of parents, educators, and other stakeholders. In addition to learning about the efficacy of various messages, Hager Sharp identified a number of concerns about Common Core, Smarter Balanced assessments, and other issues that might affect recruitment efforts. They brought those issues to the attention of the larger project team who, in turn, crafted messages to address them. Sample findings are highlighted below (Hager Sharp, 2014):

- Parents are confused about the plethora of assessments; they need easy-to-understand information about the objectives of the various assessments.
- They are also interested in specific examples of how information included in the standard is incorporated into the 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 the standard setting 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. They need informative and encouraging messaging about this.

Once the project team understood the issues involved, the recruitment process commenced along multiple lines. State Leads for higher education and K-12 education were identified and brought into the process. Hager Sharp and Smarter Balanced staff crafted specific messages for each potential audience and prepared toolkits for leads in each state to use to recruit teachers and higher education faculty and staff. They produced additional materials, deposited on the Smarter Balanced website, to answer frequently asked questions. From April to October, project staff sent thousands of e-mail messages and employed various social media to get the message out.

Online panelists signed up for one of six 48-hour windows, the first of which started on October 6, 2014. 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 4.

Table 4
Numbers of Online Panelists Entering a Bookmark, by Role, Grade, and Subject

	Teachers		Administrators		Higher Education		Other	
Grade	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
8	87	115	27	39	4	7	11	22
11	193	267	55	64	60	83	13	26
Total	1,730		481		205		244	

In-person panel. 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, Smarter Balanced implemented a nomination process in which Governing States nominated potential panelists based on grade level and subject panel targets specific for their state. Table 5 outlines the overall panel targets by role used to recruit and select panelists. The target was 252 panelists for ELA and 252 for mathematics.

Table 5
Target Counts by Role In-Person Panels

Role	Grade							Total
	3	4	5	6	7	8	11	
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

It is noteworthy that the grade 11 panels had 30 higher education participants out of a total of 72 (43%). Moreover, each panel for grades 3-8 had five above-grade educators. The rationale for the higher education personnel and above-grade teachers on the panels was to focus the panel on next steps. The focus of the high school panels was college and career readiness. The focus of the panels for grades 3-8 was preparation for the next grade. Who better to provide that focus than those who would receive those students a year or two after testing?

CTB staff 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.

The selection process began on May 27, 2014, after states submitted their nominees. CTB incorporated a multi-stage selection process to ensure state representation. All 504 panelists and 10 alternates were selected and confirmed by August 5; however a number of selected panelists dropped out. Throughout the selection process states were notified of their nominees' selection status and when applicable were requested to provide additional nominees or support to obtain a representative panel. Table 6 provides the criteria used and notifications provided for each stage of the selection process.

Table 6
Panelist 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 nominees' 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 to participate. Therefore, CTB continued to fill newly vacant panel positions until September 29, 2014.

A total of 601 nominees were invited to attend, and of those, 482 participated in the in-person panel for achievement level setting. Tables 7 and 8 present the numbers of panelists who actually served on each panel. Table 9 shows the actual numbers of panelists participating, by sex and race, along with a comparison to national and Consortium norms. Hispanic panelists were somewhat underrepresented, while Asian and multiracial panelists were somewhat overrepresented.

Table 7
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	1 panel of 30	1 panel of 30	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

Table 8
In-Person Panelist Representation by Role

Role	3	4	5	6	7	8	11	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
% of Target	93	93	93	100	95	98	96	96

Table 9
Demographic Representation for All In-Person Panels

Group	Panels	Nation	Consortium
Female	76.6%	75.9%	
Male	23.4%	24.1%	
Hispanic	3.9%	7.8%	6.7%
Multiple Race, not Hispanic	3.5%	1.0%	1.2%
American Indian, not Hispanic	0.2%	0.5%	0.0%
Asian, not Hispanic	4.6%	1.8%	2.1%
Black, not Hispanic	6.4%	6.8%	3.3%
Native Hawaiian, not Hispanic	0.2%	0.1%	0.0%
White, not Hispanic	81.1%	81.8%	85.0%

Cross-grade review panel. Cross-grade review panel members were nominated by State Leads after those panelists had confirmed their participation in the in-person panel for achievement level setting. Thus, each panelist, knowing he or she would be serving on this panel, would have opportunity to hear discussion during the single-grade achievement level setting activity and bring those insights to the cross-grade review panel.

During the selection process, nominees indicated their interest in and availability for participating in the panel. On August 15, 2014, CTB staff provided each state with a list of confirmed nominees and requested the State Lead or designee 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

CTB staff used the recommendations received from State Leads to create a cross-grade review panel as depicted in Table 10. Both panels met the specified representation goals.

Table 10
Desired Composition of Cross-Grade Review Panel for Each Content Area

Group	Number Desired	
	Grades 3-8	High School
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 Grade Span	16	16
Total per Content Area	32	

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 cross-grade review panel. This goal was met. 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 who had completed their on-grade panel work by October 15 or October 17 had the option to stay in Dallas to attend the meeting on October 20 or to leave and return to Dallas on October 19, if they wished.

Communications

Measurement Incorporated partnered with Hager Sharp, Inc. for the communications activities. Hager Sharp, in conjunction with MI, CTB, and Smarter Balanced, developed a communications plan encompassing two key phases. Phase 1 focused on building awareness about the achievement level setting process with each of our target audiences. Phase 2 focused on synchronizing messaging across the Consortium and explaining the process to key audiences.

The goals of Phase 1 were as follows:

- Gain buy-in and support for the achievement levels in advance of launch through a 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.

The goals of Phase 2 were as follows:

- 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 information on the new performance standards to state audiences.

Hager Sharp and the rest of the project team achieved these goals through a series of interrelated activities:

- **Environmental scan** – Starting with source materials 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”), Hager Sharp staff created and tested messages for stakeholder groups such as statewide organizations, districts, colleges and universities, education associations and PTA affiliates to inform them about the upcoming achievement level activities and encourage them to participate in them.
- **Focus group study** – To gather feedback from key audience segments on messages and materials for the Smarter Balanced standard setting activities, Hager Sharp conducted online focus groups among parents, teachers and school administrators from two Smarter Balanced governing states, California and Michigan. 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. Out of one of these studies came the name “achievement level setting” to replace “standard setting.”
- **Social media campaign/paid promotion** – Hager Sharp created a variety of social media posts 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. In addition, Hager Sharp worked closely with *Education Week*, an online magazine focusing on education.
- **Communications planning for key activities and state approval** – One of the most critical events in the process was the vote of the Governing States, originally scheduled for November 6, 2014, and ultimately taken on November 14. Hager planned and carried out a series of external briefings and conducted a run-of-show for a press conference and staffing plans for each day’s events.

- **Intake interviews: Development of contingency plan, messaging and protocols** – As part of a contingency planning exercise, Hager Sharp conducted a kick-off meeting, a series of telephone interviews with key MI and Smarter Balanced staff, and a contingency training session.
- **Ebola contingency response** – America’s first reported case of the Ebola virus was in Dallas, Texas, and occurred one week prior to the in-person achievement level setting activity in that city. Hager Sharp monitored media outlets throughout the event, including daily review of stories in the media and on social media sites, and maintained continuous contact with the CDC and local Dallas health officials. 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. Dr. Willhoft delivered these messages to panelists each morning during the In-Person Panel event.
- **Video production** – Hager Sharp developed a generic video as well as 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.
- **Achievement level naming research** – Hager Sharp interviewed nationally recognized testing experts and conducted online and in-person focus groups with parents, teachers, and students. They synthesized the results of their research and presented options to Smarter Balanced, primarily for the naming of Level 3 but secondarily for the other levels, which had previously been named Levels 1, 2, 3, and 4, with Level 3 corresponding to Proficient for accountability reporting purposes.

Logistics

In 2013, I presented a pre-session at NCME with the title “Setting Cut Scores on 21st Century Tests.” Module 2 of that pre-session addressed five logistical elements: People, meeting sites, materials, lodging, and meals. We took the basic elements of that module and applied them to the Smarter Balanced achievement level setting project.

People. In 2012, I contributed a chapter on practical issues in standard setting (Bunch, 2012a) to *Setting Performance Standards, Edition 2*. In that chapter, I identified several groups of people who require attention before, during, and after standard setting. Panelists constitute a key stakeholder group, and I noted four goals with respect to them:

- Gain and maintain panelist trust and respect
- Establish rapport with panelists
- Elicit feedback
- Guarantee equity and transparency

We focused on each of these goals from the project’s outset. From the first contact with potential panelists through State Leads, we made sure we shared only reliable, well-vetted information and responded quickly to all questions and concerns. Our travel specialists worked diligently with each of the in-person panelists to obtain airline, ground transportation, and lodging appropriate to their schedules and needs. During achievement level setting, we made sure that all facilitators elicited feedback in an encouraging and welcoming environment. Panelists who tended to dominate discussion were gently

redirected, and those who were reticent were encouraged to contribute. From the opening plenary session to the final cross-grade review vote, leaders made it clear that everyone had an equal voice in the establishment of cut scores. In the end, evaluations of the process reflected the fact that we had been successful in our pursuit of these goals.

Panelists for the in-person workshop constituted only a fraction of the people involved in this process. The online panelists represented some major logistical challenges, aside from the recruiting challenges. We were contractually obligated to allow for as many as 250,000 online panelists, so we had to make arrangements for that group to log on to our website and use our software *en masse*. We mitigated the risk by setting up six 2-day windows and dividing online panelists randomly across those windows. We also had to test access with a variety of computing devices and browsers. We discovered that some worked better than others and that some did not work at all. Thus, getting the word out to all potential online panelists about these limitations became part of the communications process. In the end, some of the logistical challenges proved to be too daunting. Many people who had signed up to participate never logged on. Many who logged on gave up after a brief time; others persisted for hours without entering a bookmark. There was no way to know how sincere their efforts were, whether the system or task was too complicated, or if they were just browsing.

Meeting sites. The in-person meeting site was the Hilton Anatole in Dallas, Texas. MI's meeting planner visited the site, armed with the agenda and meeting room requirements, and secured a contract that proved to be quite adequate, and Smarter Balanced approved the site. It featured large meeting rooms for plenary sessions and multiple break-out rooms for the 30-person and 36-person panels. The meetings were set up in waves: High school on October 13-15, grades 6-8 on October 15-17, and grades 3-5 on October 17-19. All cross-grade review was done on October 20. There was never a time when all panels were meeting simultaneously, but there was one time (the morning of October 17) when 12 of the 16 panels were meeting, and all had to be accommodated in separate meeting rooms. Figure 11 shows the layout of the meeting facilities, including Smarter Balanced's workroom on the same floor as the breakout rooms. The workroom for MI and CTB was on the lower floor near the Carpenter Ballroom, immediately below the breakout rooms. Figure 12 shows a diagram of a sample breakout room.

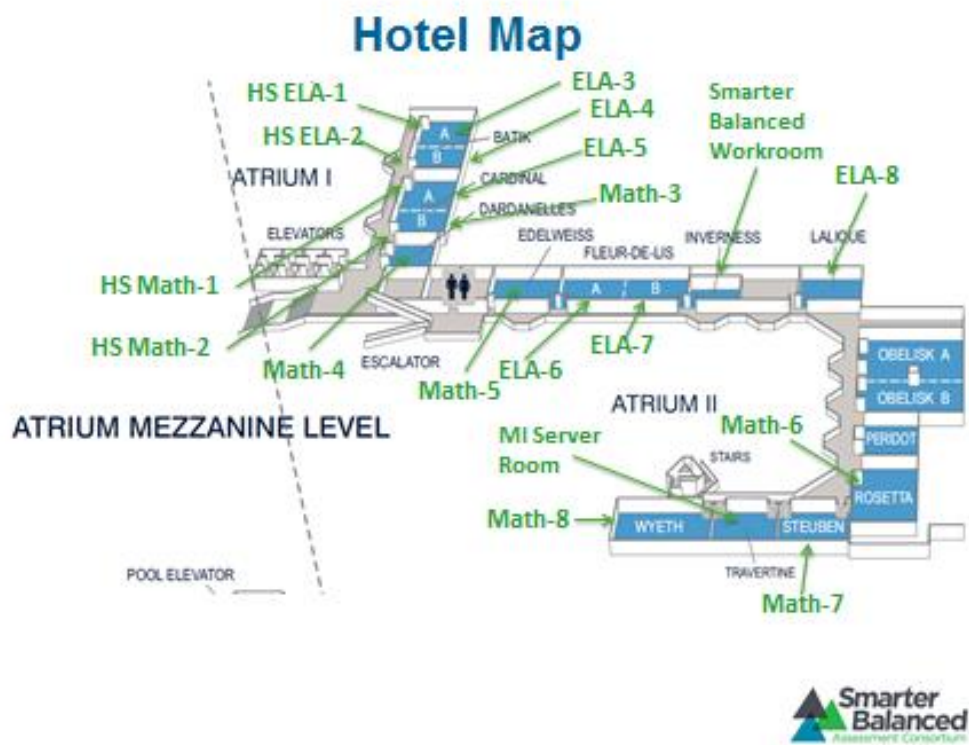


Figure 11. Layout of Meeting Facilities

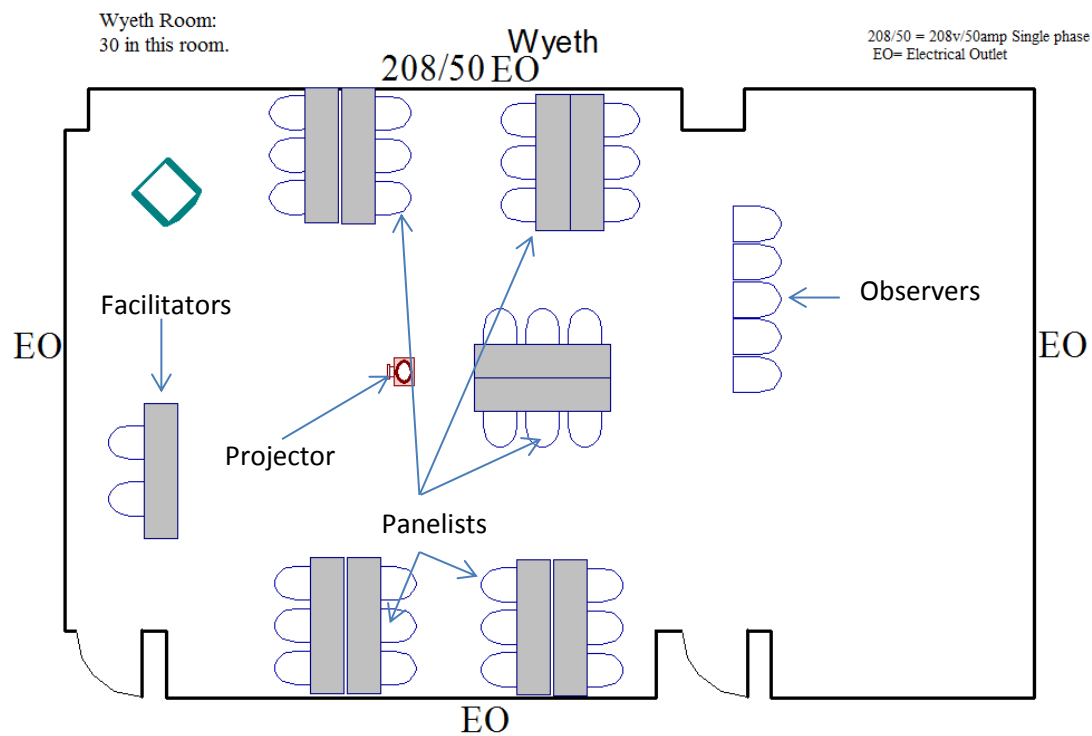


Figure 12. Layout of a Break-Out Room

Transportation. Getting 500 people to Dallas from 22 states required the full-time services of one staff member and a very accommodating travel agent. Some of those 500 people had to fly to Dallas twice, once for achievement level setting and once for cross-grade review. Once in Dallas, panelists needed to get from the airports (Dallas-Fort Worth International and Love Field) to the hotel and back as well as around Dallas. MI provided shuttle service to and from both airports and to nearby restaurants and shopping malls. This last measure contributed to the larger goal of making sure all panelists were comfortable, rested, and alert during achievement level setting. MI provided over 500 airline tickets and over 1,000 shuttle rides.

Lodging. MI staff made hotel reservations for all panelists plus Smarter Balanced staff, MI staff, CTB staff, Hager Sharp staff, observers from Smarter Balanced states, and the auditors. Panelists returning for cross-grade review had two separate reservations. All reservations were for single rooms unless otherwise requested by the panelist. In all, MI made reservations for 2,005 room nights.

Meals. As I have noted elsewhere (Bunch, 2012a), making panelists comfortable and attending to their basic needs are important aspects of successful management of an achievement level setting meeting, particularly one that stretches on for more than a week. MI staff worked closely with hotel staff to plan nutritious breakfasts and lunches as well as afternoon snacks and beverage breaks. Meals and snacks were varied, and individuals with specific dietary needs or restrictions received appropriate meals. All were on their own for dinner, but MI provided reimbursement for dinner meals on site and all meals en route to and from the meeting. In all, MI provided 4,338 meals and 2,169 snacks. Comments on the final evaluation forms about lodging and meals were uniformly very positive.

Material shipment. All achievement level setting was web based, for both online and in-person panelists. While online panelists used their own computing devices and internet connections, in-person panelists needed computers, monitors, and other equipment to do their jobs. MI transported most of this equipment from Durham, North Carolina to Dallas, Texas (see the list on p. 5 above).

The movement of equipment from Durham to Dallas was greatly facilitated by advance inspection of the site, conversations with hotel management and IT staff, and exchange of technical information between the two parties. IT staff at MI assembled and packed all materials, secured bonded transportation, and accompanied the equipment to Dallas. There, they worked with hotel staff to set up each room, run cables, and test all equipment two days prior to the start of the in-person workshop.

Flow. As noted above, there were three waves of panelists for the in-person workshop in Dallas. High school panelists arrived on Sunday, October 12, started work the morning of October 13, and completed Round 3 by noon October 15. Panelists for grades 6-8 arrived on Tuesday, October 14, began work on the morning of October 15, and finished Round 3 by noon October 17. Panelists for grades 3-5 arrived on Thursday, October 16, began work on the morning of October 17, and finished Round 3 by noon on October 19. Table 11 shows the detailed agenda for each wave, including training of table leaders.

Table 11
Detailed Agenda for Each In-Person Workshop

Day	Time	Event
0	4:00 p.m.	Facilitator Training - October 12, 14, 16
	6:00 p.m.	Table Leader Training - October 12, 14, 16
1	7:30 a.m.	Badges, place cards, and room and table assignments
		Continental breakfast
	8:30 a.m.	Opening Session (Carpenter Ballroom)
		Keynote address and charge
		Q&A; introduction of the parking lot concept
	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.	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
	12:30 p.m.	Lunch (Ballroom Lobby). Joe answers Parking Lot questions last 20 minutes.
	1:20 p.m.	Introduction to the Ordered Item Booklet (Carpenter Ballroom)
	1:50 p.m.	OIB training and review (breakout rooms)
	4:50 p.m.	Wrap-up for Day 1. 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
2	7:30 a.m.	Registration opens (Ballroom Lobby)
		Badges, place cards, and room and table assignments
		Continental breakfast
	8:30 a.m.	Orientation to the Bookmark Procedure (Carpenter Ballroom)
	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 a.m.	Reflection questionnaire
	12:00 Noon	Lunch (Ballroom Lobby). Joe answers Parking Lot questions last 20 minutes.

Day	Time	Event
2		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 p.m.	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
3		Facilitator Training
	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.
		Wrap-up for Day 3
	11:30 a.m.	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 Noon	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.
		Panelists may stay for lunch or leave if they have successfully logged out.
		MI/CTB staff discuss Round 3 data with Smarter Balanced staff and auditor, update achievement level setting report.

On-site management. Management of the in-person achievement level setting was a joint responsibility of Smarter Balanced and Measurement Incorporated. However, as prime contractor, MI was ultimately responsible for all activities. We carried out this responsibility through daily briefings with Smarter Balanced staff and the auditors, on-site training and retraining of facilitators and table leaders, and daily observation of panel activities, an activity Peters & Waterman (1982) refer to simply as “management by walking around.” Questions that came up in one panel were quickly discussed among facilitators and shared with other panels if they were relevant. The daily briefings identified problems not only with content but with schedule as well. MI and CTB staff modified the schedule of rounds after the first day of the high school wave, for example, and allowed more time for the completion of Round 1 for all panels while still allowing them to complete the final round on time without being rushed. By clearly defining who was responsible for what, it was possible to respond quickly and effectively to anything that came up. A clear example of this role definition had to do with the ongoing Ebola situation. Hager Sharp was responsible for monitoring all media outlets and crafting messages. Dr. Willhoft was responsible for sharing those messages each morning or during lunch with panelists, observers, and the media. Everyone else supplied pertinent information related to their specific roles.

As Table 11 shows, the flow of activity was carefully orchestrated. Panels received task-focused training, followed by application of that training. Data flowed from panelists’ computers directly to the Bookmark Pro workstations and back again seamlessly. Panel discussions focused specifically on the previous round’s results, always within the context of the achievement level descriptors and what it means to be college/career ready or ready for the next grade’s work. Within panels, on-grade teachers, above-grade teachers, administrators, parents, business leaders, and others contributed their individual perspectives but always within the framework of the Common Core State Standards and ALDs.

Cross-grade review. After Round 3, most panelists went home. Members of the high school and grades 6-8 panels who had agreed to participate in cross-grade review returned on Sunday, October 19. Members of the grades 3-5 panels who had agreed to participate in cross-grade review simply stayed over Sunday night while the rest of their colleagues went home. The cross-grade review consisted of an introduction by Dr. Bunch, with comments from Dr. Willhoft and Dr. Alpert. The orientation focused on expectations, both within and across grades, and democratic principles. Groundrules were as follows:

- Moving bookmarks modestly up or down is reasonable
- Scale score cuts should increase across grades
- New recommendations should align with the ALDs

Panelists were instructed that any cut score could be changed within these guidelines and that the process would be carried out democratically; i.e., each change would require a motion, second, and 2/3 majority vote for passage, following these steps:

- Consider changes across grade-groups that may support articulation
- 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

After the general orientation, each panel started with a review of the high school cut scores and then moved down to the cut scores for grades 3-8 in a systematic way. Panelists had access to all tests in their content area and could easily look up any item and its properties. The facilitators also had access to the items and could project them onto a screen to focus everyone's attention on a single item at one time. In addition, the facilitators had another projector to show Round 3 results in an interactive Excel spreadsheet that would change when any cut score would change, as shown in Figure 13.

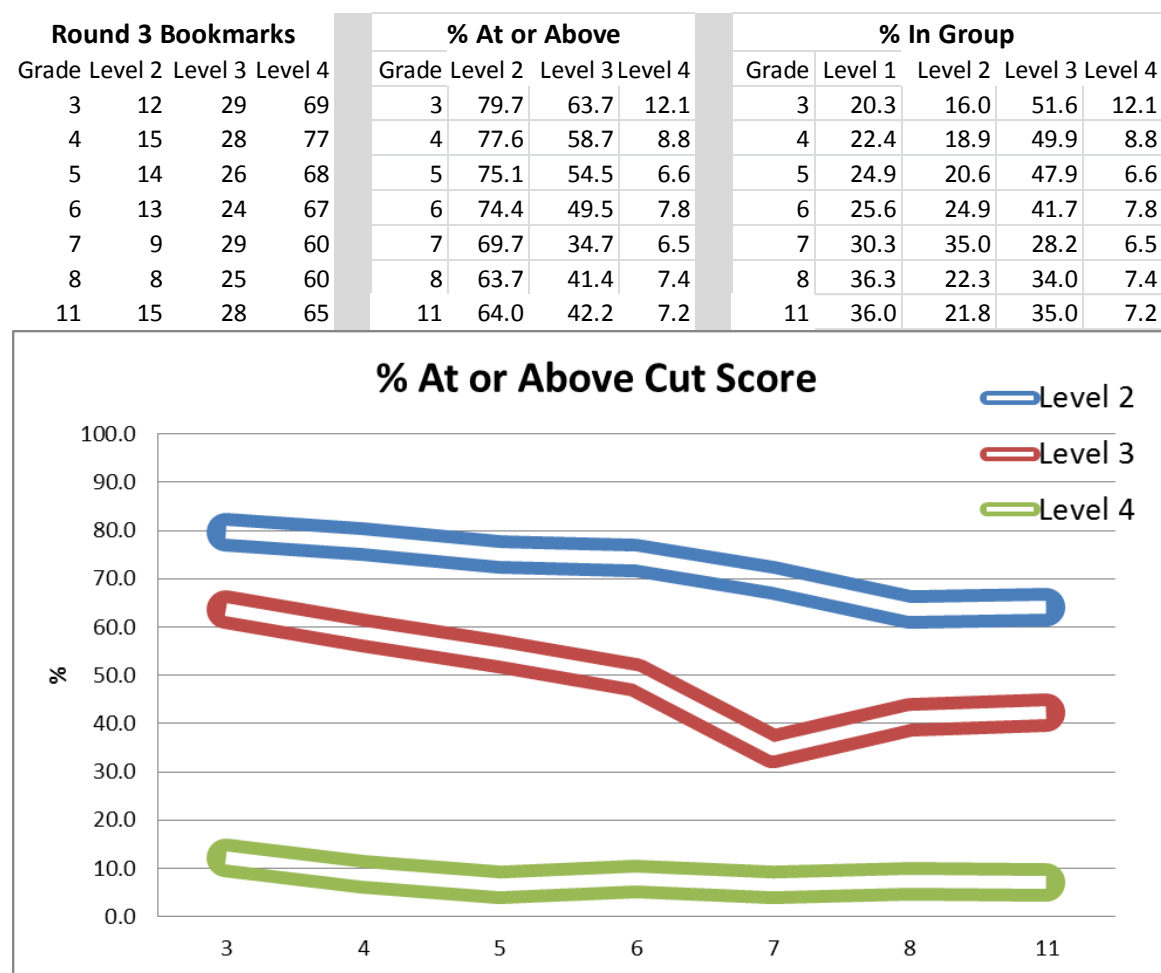


Figure 13. Sample Cross-Grade Review Interactive Spreadsheet

A form of the graphic shown in Figure 13 remained on one screen throughout cross-grade review. Panelists were free to suggest a change to any cut score, and the facilitator would make the appropriate change in the first table (Round 3 Bookmarks), at which time all tables and the graph would update. The ability to see quickly the impact of any change in bookmark placement made the process flow very smoothly. However, changes could not be made simply on the basis of smoothing of the lines in the graph; the person making the motion was required to defend the motion on the basis of the ALDs, the CCCS, and the contents of the test.

Approval by Chiefs. As we moved from the October meetings in Dallas to the November meetings in Chicago and Portland, the delineation and clarification of roles became even more important. MI staff

and our colleagues at CTB and Hager Sharp were responsible for processing all results from the in-person and online activities and preparing summaries for Smarter Balanced to present to the Chiefs. We prepared presentation materials and back-up data for these meetings, but Dr. Willhoft was the lead spokesperson throughout. Between October 20, the final day of cross-grade review, and November 14, the day of the Chiefs' vote, there was considerable exchange of ideas and information between MI and Smarter Balanced, making sure that all data were correct, up to date, and clear.

When Chiefs met in San Diego on November 14, Dr. Willhoft first provided an overview of the process we had carried out, a process the Chiefs had approved at their April 30 meeting in Minneapolis. Dr. Willhoft then provided a context for adopting the recommended cut scores:

1. Honor the work of the panelists who contributed their content expertise to the process
2. Give consideration to external evidence of student readiness for credit-bearing college work: Notably NAGB Grade 12 college preparedness data
3. Maintain a "system perspective" by using information from all grades (and both content areas) to make recommendations that support a coherent system

The Chiefs were then presented with three motions:

1. 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.
2. 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)
3. 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)

Accompanying these motions were tables and graphs depicting the recommended cut scores and impact. These are presented in Figures 14-17.

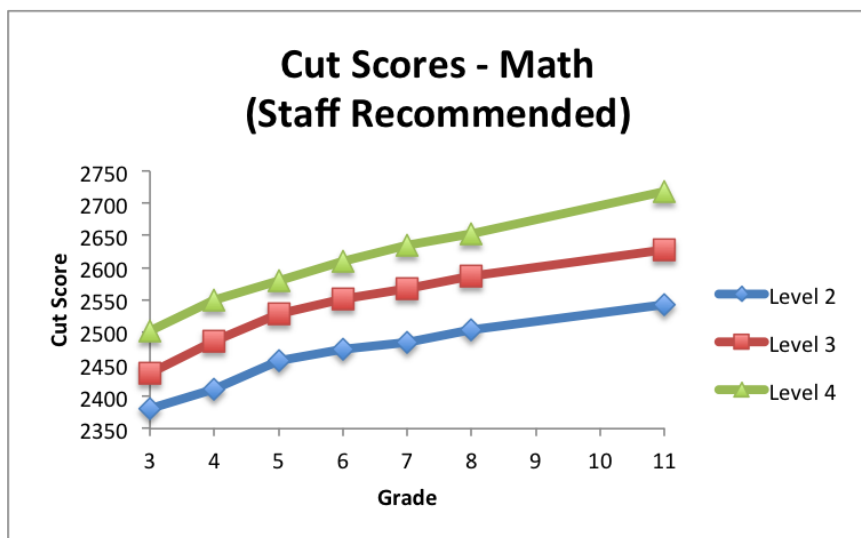


Figure 14. Recommended Mathematics Cut Scores

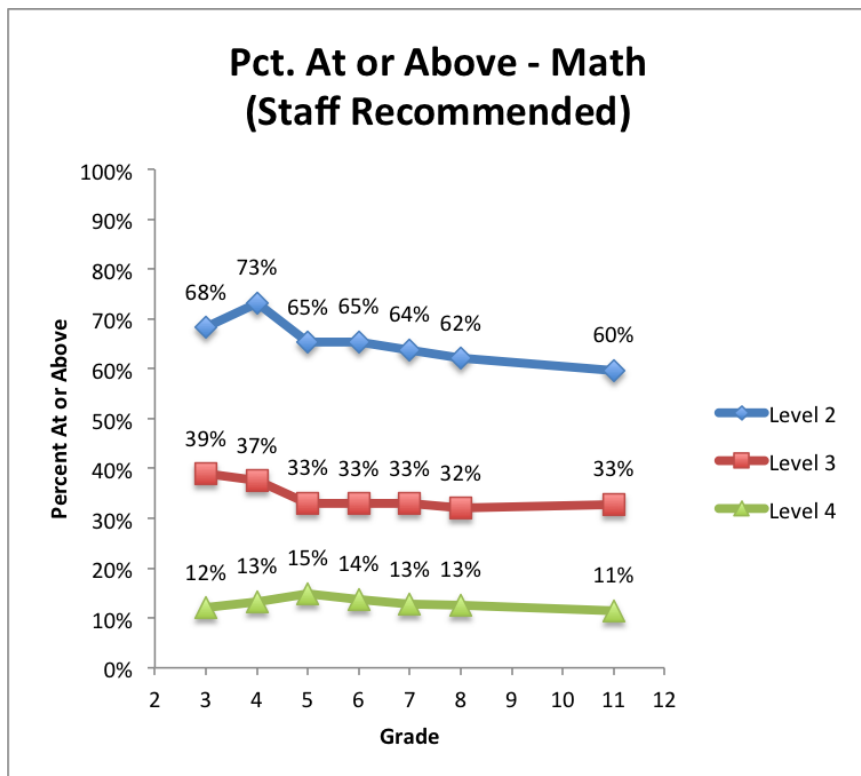


Figure 15. Mathematics Impact Data

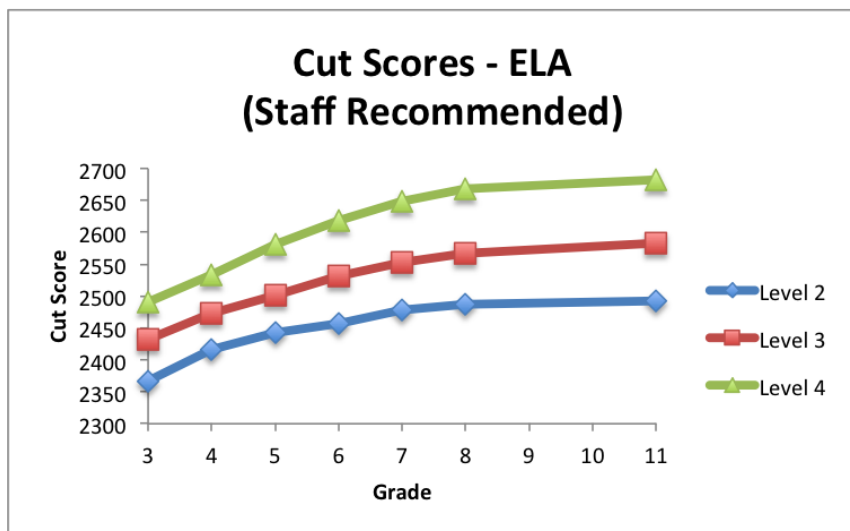


Figure 16. English Language Arts Recommended Cut Scores

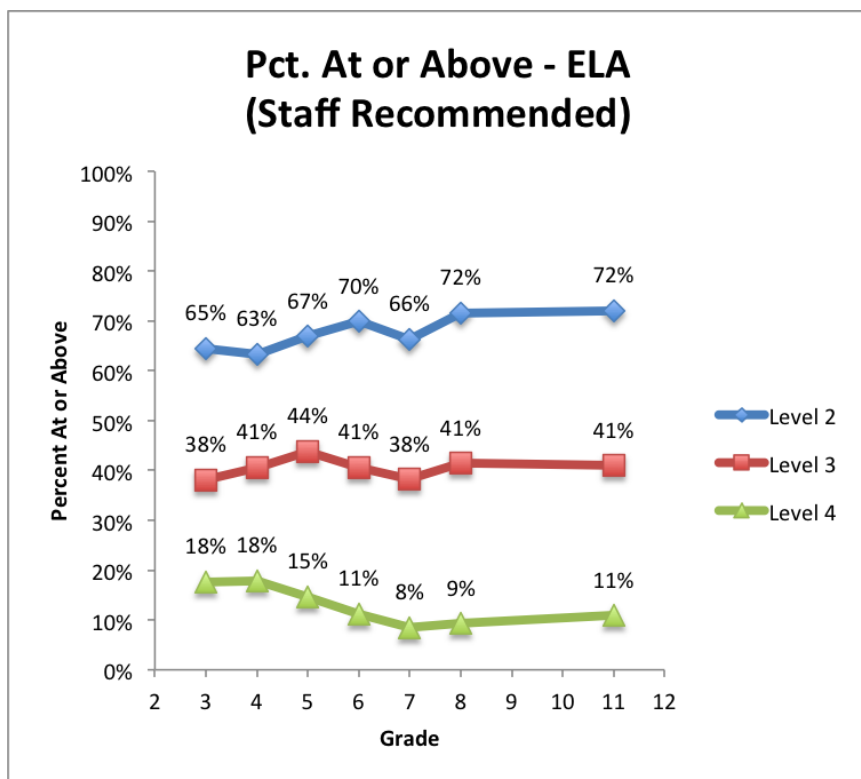


Figure 17. English Language Arts Impact Data

Following a brief discussion, Motion 1 (adoption of the position paper) passed by a vote of 20-0 with no abstentions. Motion 2 (adoption of grade 11 cut scores) passed by a vote of 18-0 with 2 abstentions. Motion 3 (adoption of grades 3-8 cut scores) also passed by a vote of 18-0 with 2 abstentions.

Conclusion

We have moved into a new era of standard setting for K-12 assessments. The moment we applied the phrase “college and career ready” to this process, we moved from documentation of past accomplishments to certification of future expectations. Achievement level setting for the Smarter Balanced assessments occurred in that context, as will that for PARCC and other K-12 assessments in the near future. Within this context, it is important to remember that standard setting, regardless of method, is about preparations, meetings, communications, and logistics. The larger and more complex the task becomes, the more important it will be to pay close attention to all four of these components.

References

- Bunch, M. B. (2012a). Practical issues in standard setting. In G. J. Cizek (Ed.) *Setting Performance Standards: Foundations, Methods, and Innovations* (2nd Ed.). New York, NY: Routledge.
- Bunch, M. B. (2012b). Standard Setting in the Real World. Presession presented at the annual meeting of the National Council on Measurement in Education, Vancouver, British Columbia.
- Bunch, M. B. (2013). Setting Cut Scores on 21st Century Tests. Presession presented at the annual meeting of the National Council on Measurement in Education, San Francisco, California.
- Hager Sharp (2014). Key Insights from Informal Focus Group Among Parents. Internal Working Paper dated April 12, 2014.
- Peters, T. & Waterman, R. H. (1982). *In Search of Excellence*. New York: Harper Collins.

Links to Smarter Balanced Achievement Level Setting Documents

About Achievement Level Setting

<http://www.smarterbalanced.org/achievement-levels/>

Achievement Level Setting Infographic

<http://www.smarterbalanced.org/wordpress/wp-content/uploads/2014/09/Smarter-Balanced-Achievement-Level-Setting-Overview-.pdf>

Auditors' Report

<http://www.smarterbalanced.org/wordpress/wp-content/uploads/2014/11/SBAC-Standard-Setting-Report-.pdf>

Ed Week Article (November 17, 2014)

<http://www.edweek.org/ew/articles/2014/11/17/13sbac.h34.html>

Frequently Asked Questions

<http://www.smarterbalanced.org/wordpress/wp-content/uploads/2014/11/Smarter-Balanced-Achievement-Levels-QA.pdf>

Interpretation and Use of Scores and Achievement Levels

<http://www.smarterbalanced.org/wordpress/wp-content/uploads/2014/11/Interpretation-and-Use-of-Scores.pdf>

Panelist Videos

https://www.youtube.com/watch?v=bW_yGf4BB1E

Reporting Achievement Level Descriptors

<http://www.smarterbalanced.org/wordpress/wp-content/uploads/2015/03/Reporting-ALDs-Final.pdf>.

Smarter Balanced States Approve Achievement Level Recommendations Press Release

<http://www.smarterbalanced.org/news/smarter-balanced-states-approve-achievement-level-recommendations/>

Statements of Support (Auditors, Advisory Panel, and Technical Advisory Committee)

<http://www.smarterbalanced.org/wordpress/wp-content/uploads/2014/11/ALS-Statements-of-Support.pdf>