PARCC

PARCC
Partnership for Assessment of Readiness for College and Careers

# Performance Level Setting Technical Report

**Laurie L. Davis, Ph.D., Eric L. Moyer, Ph.D.**
Pearson
*December 2015*

# Table of Contents

# Chapter 1: Partnership for Assessment of Readiness for College and Careers (PARCC)

This chapter provides an overview of the PARCC assessment program and includes the following sections:

- Goals of the PARCC Assessment Program
- Content Standards Measured by PARCC
- PARCC Performance Levels
- Components of the PARCC Assessment System

## Goals of the PARCC assessment system

The Partnership for Assessment of Readiness for College and Careers (PARCC) is a consortium of states working together since 2010 to develop a next generation assessment system that measures whether students are on track to be successful in college and their careers. The PARCC assessment system was designed to be a more rigorous set of assessments that would serve as an "educational GPS system," assessing students' current performance and provide teachers, schools, students, and parents information as to whether students are on track in their learning and for success after high school. The PARCC assessment system includes assessments in Mathematics and English language arts/literacy, given in grades 3 through 11. The first operational administration of the PARCC assessment system was during the 2014-15 school year, including 11 states and Washington, D.C.

The PARCC assessment system is designed on a new assessment model that includes the following goals:

- Determine whether students are college- and career-ready or "on-track" to succeed in future studies in the same content area
- Generate valid and reliable information to inform instruction and accountability decisions
- Report comparable results across schools, districts, and member states of the PARCC consortium
- Provide tools to assess student learning and support instruction during the school year
- Utilize technology to increase student access and engagement

The following sections provide a high-level description of the PARCC assessment system and how the goals of this new assessment system were met.

## Content Standards Measured by PARCC

The content standards measured by the PARCC assessment system are based on the Common Core State Standards (CCSS) (CCSSO, 2015). The CCSS is a set of standards in mathematics and English language arts/literacy (ELA/L) that outline the knowledge and skills students should gain throughout their education in order to graduate high school prepared to succeed in entry-level college-credit bearing academic courses or training programs. The development of the CCSS was led by the Council of Chief School State Officers (CCSSO) and the National Governors Association (NGA) Center for Best Practices in 2009, as a collaboration between teacher, school chiefs, administrators and other experts.

To support the implementation of the CCSS, PARCC developed the Model Content Frameworks for Mathematics and English Language Arts/Literacy. The PARCC Model Content Frameworks were developed through a process that included content experts in the related content area from the various PARCC states as well as members of the CCSS writing team. The main purpose of the Model Content Frameworks is to provide a frame for the PARCC assessments, by providing claims about what students should know and be able to do for each grade within a content area, leading to college- and career-readiness.

PARCC uses Evidence-Centered Design (ECD) to drive the development of the summative assessments. ECD is a deliberate and systematic approach to assessment development that starts with the claims (or inferences) in the Model Content Framework that describes student proficiency. The evidence that is needed for each grade and subject, stated as student objectives, provide support for the claims, in the Evidence Statement documents for each grade. Items or tasks are then developed that will provide students the optimal opportunity to produce the evidence that would support the claim of student proficiency.

## PARCC Performance Levels

In addition to the new assessment design used for the PARCC assessment system, new performance standards needed to be established. Federal statute requires that any statewide assessment used for accountability purposes include at least three achievement levels. The PARCC assessment classifies student performance into five performance levels that delineate the knowledge and skills that students are able to demonstrate. The five performance levels are:

- Level 5: Exceeded expectations
- Level 4: Met expectations
- Level 3: Approached expectations
- Level 2: Partially met expectations
- Level 1: Did not yet meet expectations

To establish these five performance levels for each PARCC assessment, each assessment needs to have four threshold scores, or performance standards, that distinguish between consecutive performance levels. All states using the PARCC assessment will adopt and report these performance levels, but each

state will independently determine the uses of these performance levels for making decisions about students, schools, and districts. For each performance level, student expectations for the knowledge and skills that would be observed were developed using a process, described in Chapter 4.

The College- and Career-Ready Determination Policy for the PARCC assessment system describe the knowledge, skills and practices that students must demonstrate in ELA/L or mathematics to show that they are able to succeed in entry-level college-credit bearing courses and relevant technical courses. The purpose of this policy is to provide a clear level of academic preparation needed for success in postsecondary courses. Students who earn a CCR determination will be able to enroll directly into entry-level college-credit bearing courses, without the need for taking any remedial courses, at participating institutions of higher education. The level of performance required for a student to earn the CCR determination is set by the PARCC governing board, as one of the performance levels.

## Components of the PARCC Assessment System

The PARCC assessment system was designed to include both summative assessment components and optional assessment tools as shown in Figure 1.1.

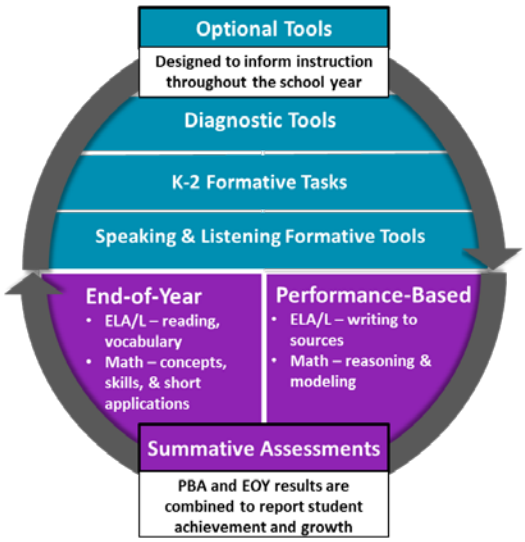



**Figure 1.1:** Components of the PARCC Assessment System

In the 2014-2015 academic year, the summative assessments consisted of two components, the Performance-Based Assessment (PBA) and the End-of-Year Assessment (EOY). The PBA was administered after about 75% of the school year had completed. The EOY was administered after about 90% of the school year had completed. The scores from the PBA and EOY assessments are combined to report a single summative student score for the PARCC assessment. Table 1.1 displays the differences between the content assessed on the PBA and EOY for the ELA/L and mathematics assessments.

The optional tools for the PARCC assessment system were not operational during the 2014-2015 school year. The 2015-2016 school year will be the first year where schools and districts may access these tools.

**Table 1.1:** Contents of the PBA and EOY Assessments for ELA and Mathematics

| | ELA/L | Mathematics |
|---|---|---|
| **Performance-Based Assessment** | • **Test items:** 3 tasks – one each in literary analysis, research simulation, and narrative writing<br>• **Focus:** Writing effectively when analyzing text, using evidence from multiple sources | • **Test items:** 7-10 extended tasks and approx. 10 short tasks<br>• **Focus:** Applying skills, concepts, and understandings to solve multi-step problems requiring reasoning, precision, perseverance, and strategic use of tools |
| **End-of-Year Assessment** | • **Test items:** 3-5 reading passages with 5-8 comprehension and vocabulary questions<br>• **Focus:** Reading for comprehension and understanding vocabulary | • **Test items:** Approx. 35-39 short tasks<br>• **Focus:** Demonstrating conceptual understanding and solving mathematics problems |

The PARCC summative assessment, PBA and EOY, can be administered online and on paper forms. The use of an online administration allowed for the development and administration of technology enhanced item types (TEIs). The use of TEIs provides the opportunity to incorporate more authentic assessments which may increase student engagement in the assessment process.

The unique aspects of the PARCC assessment system required a new approach to setting the performance levels. The following chapters discuss in greater detail the methodology and steps used to establish the PARCC performance levels.

# Chapter 2: Overview of the Performance Level Setting Process

This chapter provides an overview of the performance level setting process used for the PARCC assessments and includes the following sections:

- Goals of Setting Performance Standards
- Evidence-Based Standard Setting
- The PARCC Performance Level Setting Process
- Changes to the Performance Level Setting Process

## Goals of Setting Performance Standards

Once students are administered an assessment, various groups, including students, parents, educators, administrators, and policy makers, want to know how the students performed on the assessment and how to interpret that performance. By establishing performance levels associated with different student performance on the assessment, a frame of reference is developed for interpreting student scores. Setting the level of performance on an assessment sufficient for student performance to be classified into each performance level is one of the most critical steps in developing an assessment program.

For a standards-based assessment, such as the PARCC assessment, performance on the assessment is compared to a set of predefined content standards. The content standards define a set of knowledge and skills the students taking the assessment are expected to demonstrate upon completion of each course or grade level. The performance standards established represent the level of competence students are expected to demonstrate on the assessment to be classified into each performance level.

## Evidence-Based Standard Setting

One of the main objectives of the PARCC assessment system is to provide information to students, parents, educators, and administrators as to whether students are on track in their learning for success after high school, defined as college- and career-readiness. To set performance levels associated with this objective, PARCC used the Evidence-Based Standard Setting (EBSS) method (Beimers, Way, McClarty, & Miles, 2012) for the PARCC Performance Level Setting (PLS) process. The EBSS method is a systematic method for combining various considerations into the process for setting performance levels, including policy considerations, content standards, educator judgment about what students should know and be able to demonstrate, and research to support PARCC's policy goals related to college- and career- readiness. A multi-step process was used to allow a diverse set of stakeholders to consider the interaction of these elements in recommending performance levels for each PARCC assessment.

As illustrated in Figure 2.1, each element of the evidence-based standard-setting approach is described as it relates to the PARCC assessments.

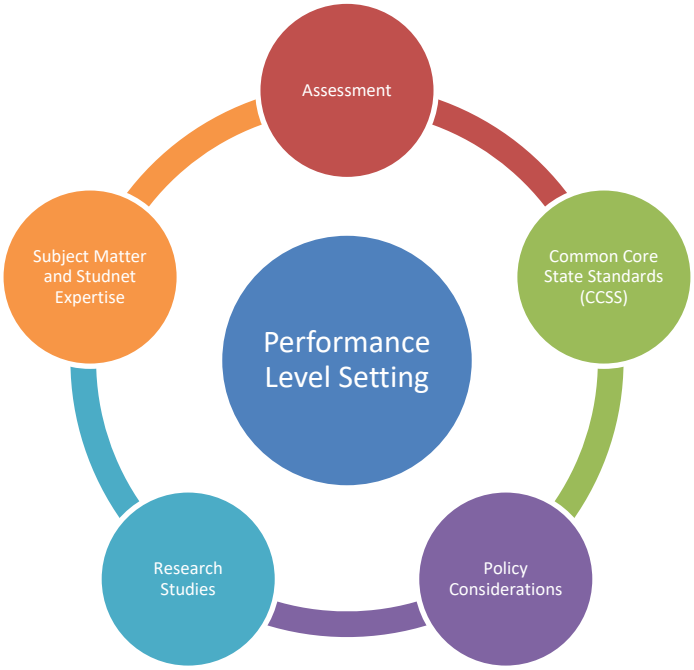



**Figure 2.1**: Critical Elements of the Evidence-Based Standard-Setting Approach

- **Common Core State Standards (CCSS)**: The Common Core State Standards (CCSS) are a set of academic learning standards in mathematics and English language arts/literacy (ELA/L) that outline what a student should know and be able to do at the end of each grade. The standards include the skills and knowledge students graduating from high school need to succeed in college, career, and life. Based on CCSS, PARCC developed several key component of the standard setting process, including the performance labels, policy definitions, and specific performance level descriptors.
- **Assessment:** Each PARCC assessment has been developed to assess the range of knowledge and skills described for each grade and subject or course in the CCSS. Each PARCC assessment is based on the student expectations specified in the PARCC evidence statements and test blueprints.
- **Policy Considerations:** Before performance standards can be established, the purpose and use for an assessment need to be clearly articulated for students, parents, educators, and policy makers. The development of the PARCC CCR Determination Policy and the Policy-Level PLDs detail the policy claims regarding student performance on PARCC assessments which specify the outcomes of interest and the policy goals.
- **Research Studies:** PARCC conducted two research studies in support of their policy goals—the Benchmarking study and the Postsecondary Educators' Judgment (PEJ) study. The purpose of the research studies was to provide evidence which could be used throughout the performance level setting process to keep the PARCC performance levels aligned with the policy goals. The specifics of the research studies are discussed in Chapter 3 of this report.

- **Subject Matter and Student Expertise:** Educators, curriculum specialists, and administrators from elementary, secondary and higher education provided subject matter expertise and experience along with classroom experience for the PLS process. These individuals assisted at various times in the process, from developing the performance labels, policy definitions, and specific performance level descriptors, to making informed judgments for the performance standards.
- **Performance Level Setting:** Within the framework of evidence-based standard setting, an established standard-setting method known as the extended modified Yes-No Angoff method (Plake, Ferdous, Impara, & Buckendahl, 2005) was used to recommend threshold scores for performance standards.

## The PARCC Performance Level Setting Process

PARCC established seven steps within the EBSS process in order to establish performance standards for the PARCC assessments.

**Step 1: Define outcomes of interest and policy goals**
Committees of educators and content experts were convened to recommend performance labels and general policy definitions for each performance level. PARCC then worked with educators to develop specific performance level descriptors for each performance level, describing a reasonable progression of knowledge and skills for each content area and grade, creating an aligned system.

**Step 2: Develop research, data collection, and analysis plans**
Research plans were developed and implemented to collect research evidence to inform the PLS process and support the interpretation of the performance standards. The specific performance level descriptors (PLD) for each PARCC assessment were compared to performance standards for various external assessments with performance standards indicating college- and career-readiness. Additionally, educators from various institutions of higher education were led through a process in which they made judgments about the level of performance on the PARCC Algebra II and ELA/L 11 assessments that would indicate academic readiness in entry-level college-level credit bearing courses within each subject area.

**Step 3: Synthesize the research results**
One of the key elements of the EBSS approach is the incorporation of empirical data in the process to help inform content-based judgments made by the PLS panels and to provide validity evidence in support of the recommended standards. PARCC synthesized and used the results of the research studies in the following ways:

- Prior to the PLS meetings as input to pre-policy panel;
- During the PLS meetings to provide context for panelists and to provide feedback data to panelists following their rounds of judgment, and

- To evaluate the reasonableness of the recommended cut scores afterwards at post-policy meetings

Each of these uses will be described further in steps 4, 5, and 6, respectively.

**Step 4: Conduct pre-policy meeting**
After the research results were synthesized, a pre-policy meeting comprised of stakeholders such as members of the PARCC Governing Board, higher education representatives, state education agency staff, members of state boards of education, and local education officials (such as local superintendents, etc.) was convened to evaluate study results and weigh in on reasonable ranges for the PARCC CCRD (see Chapter 5).

**Step 5: Conduct performance level setting (PLS) meetings with panels**
Committees consisting of K-12 educators and higher education faculty used the performance labels, policy definitions, PLDs, and reasonable ranges set by the policy committee as starting points in their process of making judgments about PARCC items that resulted in recommended threshold scores for each PARCC assessment.

**Step 6: Conduct reasonableness review with post-policy panel**
The recommended threshold scores for each performance level were reviewed for reasonableness of performance level expectations across grades and subject areas by members of the PARCC Governing Board and, for high school, members of the Advisory Committee on College Readiness (ACCR).

**Step 7: Continue to gather evidence in support of standards**
After the PARCC performance standards are adopted, additional evidence will be gathered in support of the reliability and validity of the threshold scores. In particular, there is a research study planned that will examine the external validity of the threshold scores for PARCC CCR performance levels by comparing student performance on the PARCC assessments to student performance on other assessments (e.g. SAT, ACT, etc.) which measure college- and career-readiness.

Members of the PARCC technical advisory committee (TAC) provided input to the PARCC PLS throughout the process. The TAC provided input on design and execution of both the PARCC benchmarking and Post-secondary Educators Judgment (PEJ) studies. Additionally, the TAC reviewed a high level plan for performance level setting and provided input on several key questions related to the process for making item judgments, format of feedback data, development of borderline descriptors, and use of research results within the PLS meeting process. In addition, members of the TAC attended both the dry-run and live PLS meetings as observers to confirm that the process was implemented as expected.

## Changes to the Performance Level Setting Process

Given that the PARCC PLS process was executed over the course of multiple years and paralleled the development of the assessments themselves, adaptation and modification of both the process itself and key elements in the process was necessary to ensure that the resulting performance standards maintained alignment with the intended set of inferences about student knowledge and skills. The primary areas in which modifications occurred were in the performance level descriptors (PLDs) and in the relationship of PARCC performance levels to the PARCC College- and Career-Ready Determination Policy. This section will summarize the changes made at each point in the process. Further discussion of these changes occurs in later sections of the report.

In 2012, PARCC adopted a College- and Career-Ready Determination Policy (CCRD) which defined 5 performance levels with levels 4 and 5 indicating college- and career- readiness. The performance labels for the initial performance levels were defined as:

- Level 5: Distinguished command of CCSS for the grade/course
- Level 4: Strong command of CCSS for the grade/course
- Level 3: Moderate command of CCSS for the grade/course
- Level 2: Partial command of CCSS for the grade/course
- Level 1: Minimal command of CCSS for the grade/course

The initial policy claims and performance level descriptors for each performance level were developed using a process involving educator input which will be discussed in chapter 4.

The College- and Career-Ready Determination Policy (CCRD), as discussed in Chapter 1, defines the performance level that students must meet or exceed to receive the CCRD designation. A student that receives the CCRD designation is exempt from taking remedial courses in the respective subject at participating institutions of higher education. Initially, the CCRD Policy indicated that a student needed to reach performance level 4 or greater to receive the CCRD indicator. This was represented in the policy claim by individuals in performance levels 4 and 5 being academically “prepared” and “well prepared,” respectively, to engage successfully in entry-level college-credit course or further studies in the same subject area.

This initial set of performance labels and policy descriptors which resulted from Step 1 in the EBSS process were used during several successive steps of the PLS process, including:

- Research studies
    - Benchmarking Study
    - Postsecondary Educators Judgment (PEJ) Study
- Pre-policy meeting
- Dry-run Performance Level Setting meeting

Based on the results of the PEJ study and reinforced by the results from the dry-run PLS meetings, an adjustment was made to the level 3 performance label and policy definition. The changes to the level 3

performance label and policy claim were made to reflect a change to the CCRD that students meeting performance level 3 would also receive the CCRD designation. The adjusted performance levels were:

- Level 5: Distinguished command of CCSS for the grade/course
- Level 4: Strong command of CCSS for the grade/course
- Level 3: *Adequate* command of CCSS for the grade/course
- Level 2: Partial command of CCSS for the grade/course
- Level 1: Minimal command of CCSS for the grade/course

The policy claim for performance level 3 was changed from "will likely need academic support" to "likely prepared" to engage successfully in entry-level college-credit courses or in future studies in the same subject. These were the performance labels and policy claims that were used by panelists during the high school and grades 3-8 PLS meetings.

However, during the post-policy review by the PARCC Governing Board, there were additional changes to the performance levels, policy claims, and performance level descriptors. Given the level of student performance on the assessments and the desire to provide a greater amount of differentiated feedback to students performing below the CCR level (as was originally intended), the PARCC Governing Board voted to shift the performance levels to better meet the intended inferences about student performance. Holding the college- and career- ready (or on track) expectations constant, performance levels above this expectation were combined and performance levels below this expectation were expanded to create the final system of performance levels with three below and two above the college- and career-ready (or on track) expectation. Additional details about how this process was conducted, the involvement of PLS educator panelists, and implications for the PLDs are discussed in Chapter 4. The policy labels for the final performance levels are defined as:

- Level 5: *Exceeded Expectations*
- Level 4: *Met Expectations*
- Level 3: *Approached Expectations*
- Level 2: *Partially Met Expectations*
- Level 1: *Did Not Yet Meet Expectations*

Under this final set of performance levels, students who are in Levels 4 or 5 receive the CCRD designation.

Figure 2.2 displays the changes to the performance levels across the PLS process. A further description of the changes is included in Chapter 7.

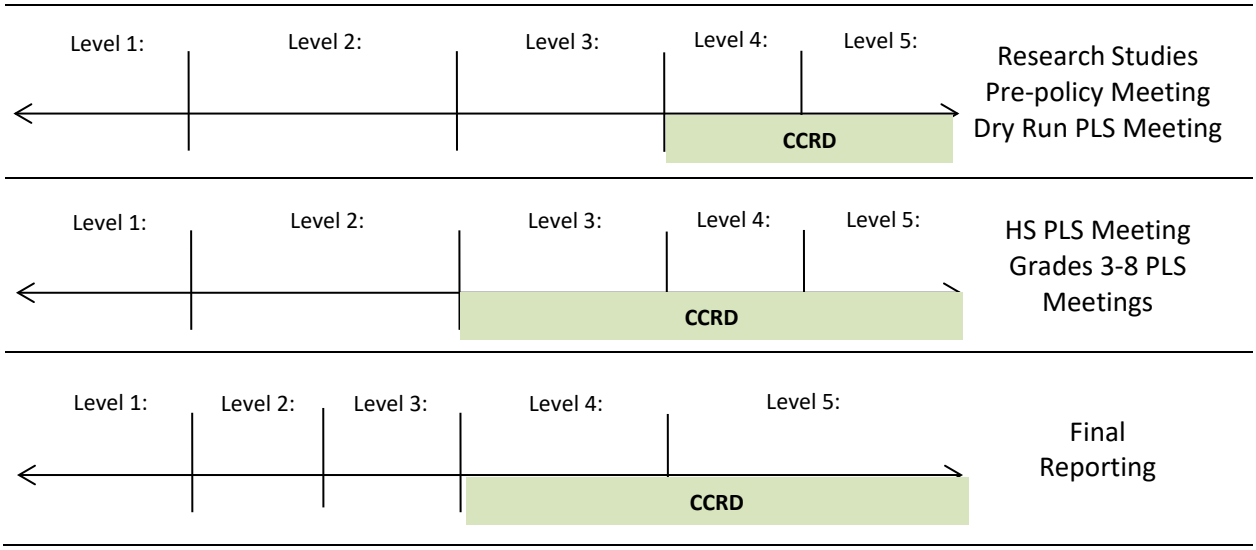



**Figure 2.2:** Changes to Performance Levels through PLS Process

Note: The actual threshold scores for the performance levels were not recommended until after the PLS meetings. The lines are just to represent the ordered nature of the performance level, not any relative position of the threshold scores.

# Chapter 3: Performance Level Descriptors

This chapter provides details about the development of the performance level descriptors (PLD) that were used throughout the PLS process. The sections in this chapter include:

- What are Performance Level Descriptors?
- Performance Level Descriptors and College- and Career-Readiness Determination
- Approach to Performance Level Descriptor Development
- Revision of the Performance Level Descriptors

## What are Performance Level Descriptors?

Performance level descriptors (PLD) are statements that articulate skills, knowledge, and practices that students classified into a particular performance level should be able to demonstrate. Each assessment within the PARCC assessment system, grades 3-8 and high school, has five performance levels. The performance levels range from Level 1, representing the lowest level of student performance, to Level 5, representing the highest level of student performance. PARCC has developed two different kinds of PLDs that were used as part of the PLS process: policy-level PLDs and grade- and subject-specific PLDs.

The PARCC policy-level PLDs were developed as part of the PARCC College- and Career-Ready Determination (CCRD) policy, which was adopted by the PARCC Governing Board and Advisory Committee on College Readiness. The policy-level PLDs describe the educational implications for students who attain a particular performance level on the PARCC assessments.

The policy-level PLDs do not differentiate student performance within each grade level. By contrast, the grade- and subject-specific PLDs indicate the knowledge, skills, and practices that students performing at a given performance level should be able to demonstrate within each content area, mathematics or ELA/L, at each grade level. The following sections will describe the process used to develop PARCC College- and Career-Readiness Determination (CCRD) Policy and the grade- and subject-specific PLDs.

## Performance Level Descriptors and College- and Career-Ready Determination

The College- and Career-Readiness Determination (CCRD) Policy and Policy-Level Performance Level Descriptors are documents that were used as a guidepost for the development of the subject- and grade-specific performance level descriptors (PLDs). The PARCC CCRD policy describes for both ELA/Literacy and mathematics the academic knowledge, skills, and practices students must demonstrate to show that they are able to enter directly into and succeed in entry-level, credit-bearing courses and relevant technical courses in those content areas at two- and four-year public institutions of higher education. Students who attain a CCR determination in either ELA/L or mathematics will have earned direct entry into relevant entry-level, credit-bearing courses without remediation at participating institutions of higher education.

The first draft of the CCRD Policy and policy-level PLDs was reviewed by the PARCC Governing Board and the Advisory Committee on College Readiness (ACCR) in June 2012 and approved for release for public review and comment. Public feedback was collected on the CCRD policy and the policy-level PLDs from various groups, which was used to revise the CCRD policy during Fall 2012. For further information about the public comments concerning the initial draft of the CCRD policy, review the Summary Report of State and Public Feedback from PARCC.

The initial CCRD Policy defined that to earn a College- and Career-Ready Determination in ELA/Literacy or mathematics, a student would need to earn a threshold score for performance level 4 or greater on the grade 11 ELA/Literacy or Algebra II or Integrated Math III assessments, respectively. For all other assessments, a student that earns a threshold score for performance level 4 or greater indicates that the student is academically prepared to engage successfully in further studies in the respective content area.

## Approach to Performance Level Descriptor Development

For every PARCC assessment, performance level descriptors were written for performance Levels 2, 3, 4, and 5. The performance labels for the five performance levels were defined as:

- Level 5: Distinguished command of CCSS for the grade/course
- Level 4: Strong command of CCSS for the grade/course
- Level 3: Moderate command of CCSS for the grade/course
- Level 2: Partial command of CCSS for the grade/course
- Level 1: Minimal command of CCSS for the grade/course

Note that PLDs were not developed for performance Level 1. Each performance level on the PARCC assessments encompasses a wide range of student performance. For the lowest level, Level 1, the performances can include those from students who chose not to answer a majority of the questions, those who answered all but a few questions incorrectly, and those who just missed the cut score for Level 2. Consequently, the most accurate means to describe the performance of a student achieving at Level 1 is to say that this student is one who has not demonstrated the knowledge and skills necessary to achieve Level 2.

A multi-step iterative process was used in developing, reviewing, and approving the PARCC grade- and subject-specific PLDs. Panels were convened, comprised of K-12 educators and curriculum directors and higher education faculty from PARCC Governing States, to develop initial draft PLDs. This step in the process required the use of seven panels, organized by subject and grade span, including:

- Elementary school ELA/L (grades 3-5)
- Middle school ELA/L (grades 6-8)
- High school ELA/L (grades 9-11)
- Elementary school mathematics (grades 3-5)
- Middle school mathematics (grades 6-8)

- High school integrated mathematics (Integrated mathematics I, II, and III)
- High school traditional mathematics (Algebra I, Geometry, Algebra II)

The nine content experts on each committee represented the range of states in PARCC and the range of students within the grade band, including special populations. The meetings were planned so that the high school PLDs were developed first, then the middle school PLDs, and then the elementary school PLDs[1].

Each panel met for three days. Committees started with the highest grade (11, 8, or 5) and worked backwards through the other two grade levels. In addition, each committee focused first on the target performance level (Level 4) in each grade, and then moved to the other levels, 5, 3, and 2, in that order. To generate descriptors, committees were instructed to focus on the "typical" student at each level. Panelists were asked to focus on each claim and then determine the amount of knowledge, skills, and abilities as described by the CCSS needed to meet the rigor established by the policy descriptors. Sample items, draft assessment blueprints, and the task complexity framework were provided to ensure strong alignment with the actual assessment. In addition, the panelists worked with the general content descriptors as a starting point to indicate the expected level of rigor for each claim at each level. Rather than focusing on frequency words like "usually" and "occasionally," however, they were encouraged to focus on the knowledge, skills, and abilities a typical student should exhibit at each level and to vary the levels by accuracy and dimensions in the cognitive complexity frameworks.

After the development committees, the draft PLDs were sent through several review and revision steps, including:

- Review by Educational Testing Service (ETS) and the College Board (CB), focusing primarily on the consistency of the PLDs across level, articulation across grades, and fidelity with the CCSS.
- Review of the revised PLDs by the respective PARCC Operational Working Group (OWG) to address any issues identified by the ETS/CB review and revise the format for public review.
- Review of the revised PLDs, after the OWGs, by the PARCC State Leads and Governing Board, before public release.
    - All high school PLDs were reviewed by the PARCC Advisory Committee on College Readiness for alignment with expectations from participating institutions of higher education (IHE).
- Public review of the revised PLDs which was used by PARCC to refine the PLDs for final review.

The final draft of the grade- and subject-specific PLDs was reviewed and adopted by the PARCC Governing Board and Advisory Committee on College Readiness during a joint session on June 26, 2013.

[1] Due to weather-related events, the middle school PLD panel had to be rescheduled and occurred after the elementary school panel.

For additional information about the development, review, revision, and adoption of the PARCC grade- and subject-specific PLDs, see the full technical report, "PARCC Development of Performance Level Descriptors: Technical Documentation."

## Revision of Performance Level Descriptors

As referenced in Chapter 2, changes were made to both the policy-level PLDs and grade-and subject-specific PLDs at various points in the PLS process to reflect the availability of new information (e.g. PEJ study results (see chapter 4) and PLS dry-run meeting results (see chapter 6)) and to maintain alignment of the PARCC performance levels with the intended set of inferences about student knowledge and skills.

Prior to the performance level setting meetings, changes were made to the level 3 performance label and policy claim to reflect a change to the CCRD that students meeting performance level 3 would also receive the CCRD designation. The label for performance level 3 was changed from *Moderate* to *Adequate*. Additionally, the policy claim for performance level 3 was changed from "will likely need academic support" to "likely prepared" to engage successfully in entry-level college-credit courses or in future studies in the same subject. No changes were made to the grade- and subject- specific PLDs at this juncture. These were the performance labels and policy claims that were used by panelists during the high school and grades 3-8 PLS meetings.

After the performance level setting meetings, additional changes to the PARCC PLDs were undertaken to reflect the shift in performance levels approved by the PARCC governing board (see Chapter 7). The modifications included changes to the performance levels and labels, policy-level PLDs, and the grade- and subject- specific PLDs.

The policy labels for the final performance levels are defined as:

- Level 5: *Exceeded Expectations*
- Level 4: *Met Expectations*
- Level 3: *Approached Expectations*
- Level 2: *Partially Met Expectations*
- Level 1: *Did Not Yet Meet Expectations*

Under this final set of performance levels, students who perform at Levels 4 or 5 receive the CCRD designation, for the Grade 11 ELA/Literacy, Algebra II or Integrated Math III assessments, or designation that the student is "on-track" to engage in further studies in the content area, for all other PARCC assessments.

The shift in performance levels was accomplished using a scale anchoring process which involved two primary steps. In the first step, the top two performance levels, above college- and career-ready (or on-track), were combined into a single performance level and an additional performance level below college-and career- ready (or on track) was created by empirically determining the mid-point between

the existing two levels[2]. In the second step, the performance level descriptors (PLDs) were updated using items which discriminated student performance well at this level to create a PLD aligned with the new empirically determined performance level. At this same time, PLDs for all performance levels were reviewed for consistency and continuity. Members of the original PLS committees from summer 2015 were recruited to participate in this process.

During the fall of 2015, a set of meetings was convened by HumRRO to revise the grade- and subject-specific PLDs to align with the new performance levels, labels, and policy-level PLDs. This work involved two phases and was conducted separately for high school and grades 3-8. In the first phase, HumRRO worked with PARCC to identify a set of individuals familiar with the PARCC grade- and subject-specific PLDs who would make the necessary adjustments to the PLDs to align with the modifications to the performance levels. In the second phase, panelists who had previously participated in the performance level setting meetings were asked to review and provide input on these changes[3]. For more information about the process used during these meetings, please review the technical report for these meetings from PARCC.

[2] Scale score values of 700 and 750 were mapped to the threshold scores for the existing performance levels 2 and 3 using a linear transformation of the underlying IRT theta scale on the reference PLS test form. The midpoint value of 725 (and its associated IRT theta on the reference PLS test form) was used to subdivide performance level 2. As such, scale score values of 700 to 724 now place students into performance level 2, scale score values of 725 to 749 now place students into performance level 3, and scale score values of 750 to 799 now place students into performance level 4.

[3] All PLS meeting panelists were invited to attend conference calls at which time they were informed about the changes to the PARCC performance levels adopted by the PARCC Governing Board, were informed about the revisions to the PLDs, and were invited to volunteer to participate as part of the PLD review process.

# Chapter 4: Research Studies

This chapter provides details about Steps 2 and 3 of the PARCC performance level setting process, which focuses on developing and conducting research studies. The sections in this chapter include

- Use of Research Studies in Performance Level Setting
- Benchmarking Study
- Postsecondary Educators' Judgment (PEJ) Study
- Technical Issues and Caveats

## Use of Research Studies in Performance Level Setting

The PARCC assessment system is designed to be an aligned system of performance standards from grade 3 to high school, based on the Common Core State Standards (CCSS). The results of the PARCC assessment provide indicators of student readiness for college and careers. When performance standards are set with respect to these objectives, empirical evidence validates the use of those standards to describe the likelihood that students will meet future goals, such as success in postsecondary endeavors.

A systematic approach to incorporating empirical evidence into the PLS process was developed, based on the evidence-based standard-setting approach (Beimers, Way, McClarty, & Miles, 2012). The approach involved the following three steps:

- **Determine the research studies to conduct**
  A process was used to determine the research studies in order to gather a sufficient body of validity evidence, but not so much information that the data become difficult to interpret. The result of the process was that two research studies were conducted: The Benchmarking Study and Postsecondary Educators' Judgment Study.
- **Develop data collection plans**
  For each study, an appropriate data collection plan was developed, to inform decision making and report results for the pre-policy meeting.
- **Selecting appropriate analysis methods**
  Each research study served a particular purpose during the PLS process, and presented a unique set of requirements and considerations for quantitative analysis.

The following sections cover the methods and results of each of the research studies implemented.

## Benchmarking Study

The purpose of the PARCC Benchmarking Study was to inform the PLS process by providing one source of empirical evidence to inform the PARCC Level 4 performance standard. This benchmarking study gathered external information that provided estimates of the percentage of graduating high school students that would be considered college- and career-ready, as well as the percentage of students "on-

track" to readiness in earlier grades. Specifically, external information was analyzed to inform performance standards for the grade 11 ELA/literacy, Algebra II, and Integrated Mathematics III assessments, the grade 8 ELA/literacy and mathematics assessments, and the grade 4 ELA/literacy and mathematics assessments.[4]

The study included two sections. The first section was a review of the relevant literature on college- and career-readiness. The literature review provided background and framing for the paper by reviewing relevant statistics from published reports that suggest current levels of college- and career-readiness across the country. The second section was a description of the methods and results of the studies PARCC conducted to directly compare the PARCC assessments and initial Level 4 performance expectations with several other international, national, and state assessments. Results from the published literature and the PARCC comparison study were integrated to form a summary set of recommendations for consideration by the PARCC states and pre-policy committee.

Several different types of statistics have been used to raise awareness about the need to improve the college- and career-readiness of students in the United States, including: 1) referrals and enrollment in remedial courses; 2) college enrollment, graduation rates, and degree attainment; and 3) high school grades, graduation rates, and courses taken. A review of the relevant literature suggests that around 40% of college students enroll in at least one remedial course (with figures ranging from 20% to 60%). However, the enrollment rate is likely to underestimate the number of students who need remediation as more students are referred to developmental coursework than actually enroll. Moreover, about a third of students do not enroll in college after completing high school, and 20% drop out before high-school graduation. These statistics highlight the importance of providing information about students' college readiness and whether they are on track toward being college ready when there is ample time for intervention.

PARCC Level 4 performance was described by 1) a set of Performance Level Descriptors (PLDs) which explains the knowledge, skills, and abilities a student must demonstrate for each grade and content area to reach Level 4, and 2) an empirical definition of college- and career-readiness for high school which suggests students who reach level 4 will have approximately a 0.75 probability of earning a C or better in specific, relevant, entry-level, credit-bearing, college courses[5]. These two definitions of Level 4 – content and empirical – informed comparisons between PARCC and external assessments in the benchmarking study.

The PARCC assessments and Level 4 expectations were compared with international, national, and state assessments. Specific external comparisons included PISA, NAEP, ACT, SAT, the Michigan Merit Exam,

[4] Although the Benchmarking Study included information on external assessments for students in grades 4 and 8, this part of research was ultimately not included in the PLS process because of the changes made to the PARCC performance levels which limited the utility of these comparisons.
[5] This empirical definition of college- and career- readiness as related to the initial PARCC performance level 4 reflects an estimate based on the policy definition associated with the initial performance level. Future research will be needed to validate this definition as it relates to the final CCRD on the PARCC assessments.

and the Virginia End-of-Course exams in high school. For each external assessment, the performance level that best matched PARCC's Level 4 expectations was determined through consideration of either the empirical definition or the content expectations in PARCC's Level 4 PLDs. Once the best matched performance level was determined, the percentage of students reaching that level across the nation and for PARCC states was determined. The contextual information was used to evaluate the performance results and to consider how the study should be emphasized when making summary recommendations. Each external comparison along with the best aligned performance level and percentage of students estimated as reaching that external benchmark are summarized in Table 4.1.

**Table 4.1:** Performance on External Assessments that Best Align to PARCC Performance Level 4

| PARCC Assessment | External Assessment | Best Aligned Performance Level | Percentage Reaching the External Benchmark |
|---|---|---|---|
| Grade 11 ELA/literacy | PISA Literacy | Level 4 | 28% |
| | NAEP Reading | Preparedness/Proficient | 38% |
| | NAEP Writing | Proficient | 27% |
| | ACT Reading | Benchmark = 22 | 44% |
| | ACT English | Benchmark = 18 | 64% |
| | SAT Reading | Benchmark = 500 | 50% |
| | SAT Writing | Benchmark = 500 | 45% |
| | Michigan Reading | Proficient | 53% |
| | Michigan Writing | Proficient | 50% |
| | Virginia Reading | Advanced/College Path | 8% |
| | Virginia Writing | Advanced/College Path | 25% |
| Algebra II and Integrated Math III | PISA Mathematics | Level 4 | 25% |
| | NAEP Mathematics | Preparedness (163) | 39% |
| | ACT Mathematics | Benchmark = 22 | 44% |
| | SAT Mathematics | Benchmark = 500 | 54% |
| | Michigan Mathematics | Proficient | 29% |
| | Virginia Algebra II | Advanced/College Path | 14% |

**Note:** Percentages represent national estimates, when available.

For both ELA/L and mathematics, the data supported a range of 25 to 50 percent of students considered college- and career-ready or on track to readiness based on PARCC's definitions of the initial Level 4 performance. Beyond the general range of 25 to 50 percent, specific, honed-in recommendations were also provided based on a consideration of contextual factors (e.g. similarity to the PARCC assessments, expectations, and student population) relative to external comparison that influenced how much weight that comparison should be given. The specific ranges represent a 10 percentage-point region in which the initial Level 4 performance standard would most likely occur based on the external comparisons. Table 4.2 below shows the recommendations for the specific percentage of students who were expected to reach PARCC Level 4 or higher based on these study results.

**Table 4.2:** Recommended Honed-In Range from Benchmark Study

| | English Language Arts | Mathematics |
|---|---|---|
| **High School (Grade 11 ELA, Algebra II, Integrated Math III)** | 35-45% | 35-45% |

For further information about the PARCC Benchmarking Study, review the full research report for the Benchmarking Study from PARCC.

## Postsecondary Educators' Judgment Study

The purpose of the Postsecondary Educators' Judgment Study was to provide information to support the PLS process, by collecting recommendations from professors and instructors of entry-level college credit bearing courses at institutions of higher education (IHE) regarding the minimum level of performance on PARCC test items that would indicate that students were 'academically-ready' to take and succeed in postsecondary courses.

A sample of instructors from a variety of IHE across the states within the PARCC consortium completed the study, including a total of 100 participants for Algebra II and 90 participants for Grade 11 ELA/L. The sample of individuals that completed the study, shown in Table 4.3, was a representative sample of individuals that were nominated to participate in the study by the Higher Education Leadership Team (HELT).

**Table 4.3:** State Representation of Individuals that Completed the Study

| | Grade 11 ELA/L | | Algebra II | |
|---|---|---|---|---|
| **State** | **N** | **%** | **N** | **%** |
| **AR** | 11 | 12 | 6 | 6 |
| **CO** | 11 | 12 | 7 | 7 |
| **DC** | 0 | 0 | 1 | 1 |
| **IL** | 10 | 11 | 13 | 13 |
| **LA** | 7 | 8 | 6 | 6 |
| **MA** | 16 | 18 | 19 | 19 |
| **MD** | 7 | 8 | 6 | 6 |
| **NJ** | 14 | 16 | 15 | 15 |
| **NM** | 7 | 8 | 13 | 13 |
| **NY** | 3 | 3 | 3 | 3 |
| **OH** | 4 | 4 | 11 | 11 |
| **RI** | 0 | 0 | 0 | 0 |
| ***Total*** | *90* | | *100* | |

The study used an online system to guide participants through a judgmental process, where participants reviewed test items from the PARCC Algebra II or Grade 11 ELA/L assessment and made judgments

about how students taking an entry-level college-credit bearing course that they teach would perform on each item. The participants were guided through the study process using the online system, so they could complete the study at their own pace, including the following steps.

1. Postsecondary Educators' Judgment Study Overview
2. Participation Survey
3. Online Training
4. PARCC Tutorials
5. PARCC Practice Tests
6. Item Judgment Survey
7. Exit Survey

To complete the item judgment survey, participants reviewed the items on both the Performance Based Assessment (PBA) and End-of-Year (EOY) components of a reference form of the assessment. For each item, the participants reviewed the knowledge and skills needed to respond to the item and made judgments about how students taking an entry-level college-credit bearing course, that they teach, would perform on each item. The participants entered their answer to the following question for each item they reviewed on the item judgment survey.

"How many points would a borderline 'academically-ready' student likely earn if they answered this item (or set of items) on their first day of class?"

A student was considered "academically-ready" if they would have at least a 75% probability of earning a C or better in a relevant course, without taking a remediation class, based on PARCC's College- and Career-Ready Determination Policy. The sum of each participant's item-level judgments, including weighting, if applicable, from both the PBA and EOY components was calculated to determine the participant's estimated total raw score recommendation for the borderline academically ready student. The participant judgments were then summarized to provide a set of aggregate results organized by factors such as type of IHE, years of experience, and course taught. This process resembled the judgmental process that educators followed in the performance level setting meetings; however, judgments in the PEJ study were made relative to the definition of academically-ready rather than relative to the specific performance level descriptors (PLDs). Additionally, participants made only one judgment for each item on the test and were led through only one round of the item judgment process.

The mean and median estimated raw scores on each assessment were plotted to create a visual summary of the study results (see Figures 4.1 and 4.2 for Algebra II and ELA/L 11 results, respectively). For each assessment, the results were plotted for the overall group (All) and disaggregated results according to the following factors:

- Type of IHE ("Type"): 2-year (2) or 4-year (4) IHE
- Technical/Non-Technical ("Tech"): Associated with technical program (Y) or non-technical program (N)

**Figure 4.1:** Algebra II Raw Score Range Recommendation

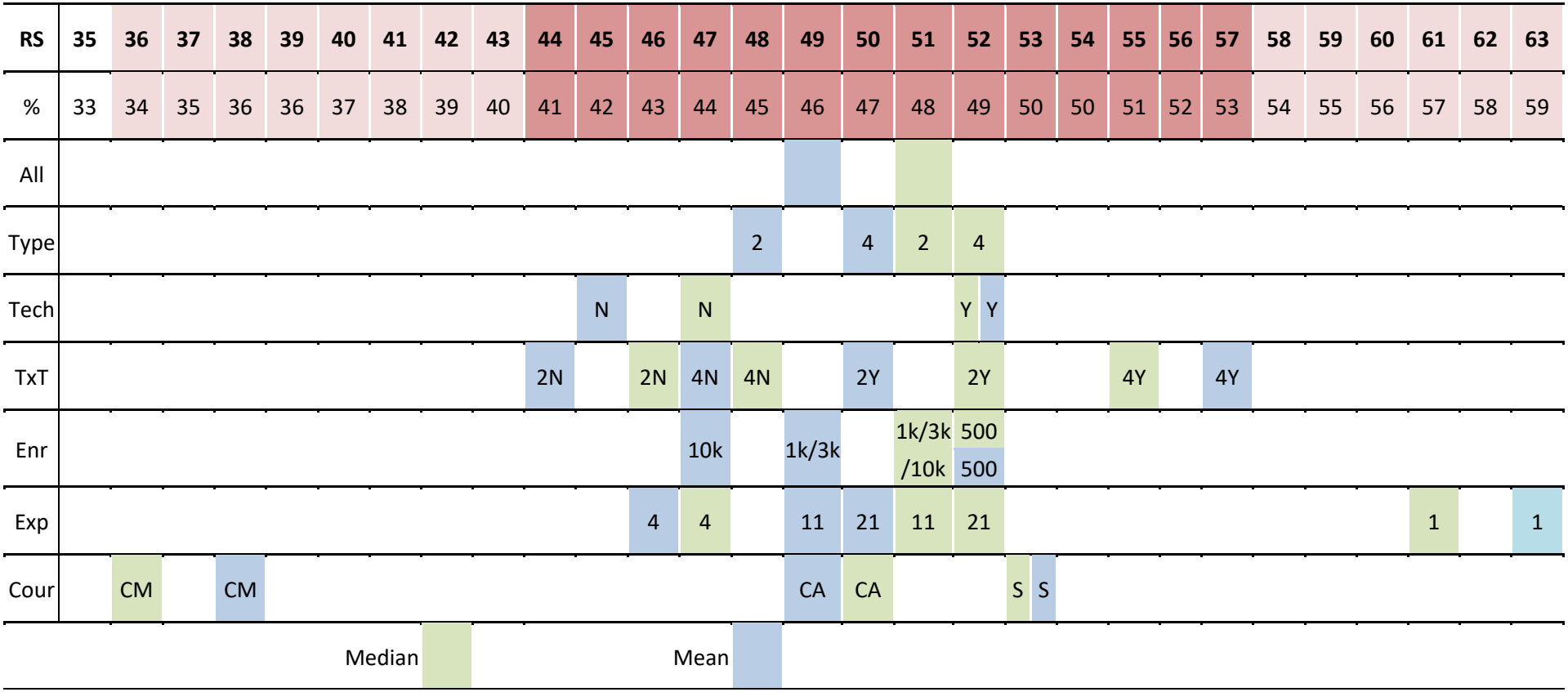
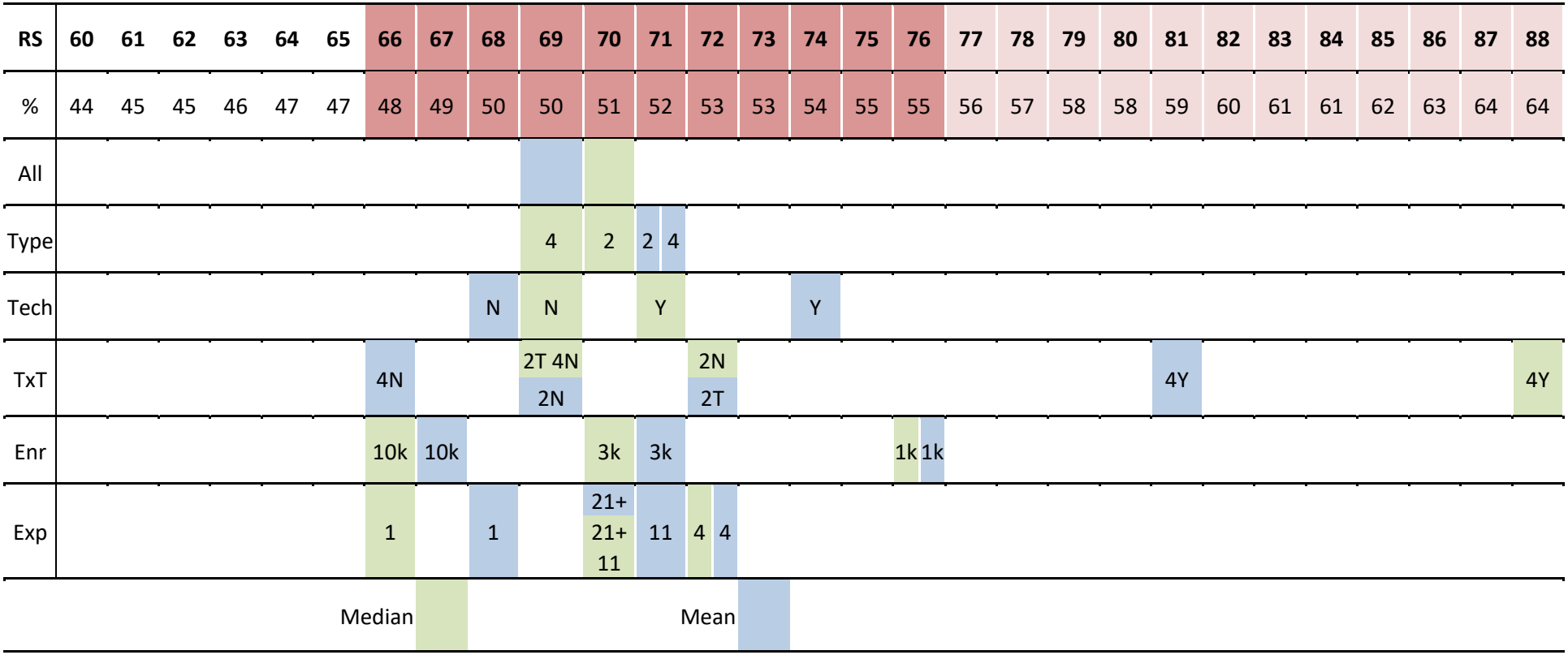


**Figure 4.2:** Grade 11 ELA/L Raw Score Range Recommendation



- Type by Technical ("TxT"): A cross-tabulation of the type of IHE with designation as technical or non-technical.
    - 2-year IHE with technical program (2Y),
    - 2-year IHE with non-technical program (2N),
    - 4-year IHE with technical program (4Y), or
    - 4-year IHE with non-technical program (4N)
- Enrollment ("Enr"): Total student enrollment of 500-999 (500), 1,000-2,999 (1K), 3,000-9,999 (3k), or more than 10,000 (10k)
- Experience ("Exp"): Years of experience in higher education of 1-3 (1), 4-10 (4), 11-20 (11), more than 20 (21+)
- Course Profile ("Cour"): Course profile of College Algebra (CA), Survey of College Math (CM), or Introduction to Statistics (S). Note that for ELA/L 11 there was a single course profile so results are not disaggregated by this group.

Based on the information from the study, two sets of recommendations were created for ranges in which the threshold score for college- and career- readiness could be expected. The "broad" recommended range (indicated by the light pink highlighted range of raw scores in Figures 4.1 and 4.2) encompasses the data points for all participant groups in the study. The "honed-in" range indicated by the dark pink highlighted range of raw scores in Figures 4.1 and 4.2) gives less weight to data points for which the number of participants in the group was relatively small. The recommended ranges for both Algebra II and ELA/L 11 are summarized in Table 4.4.

**Table 4.4:** College and Career Readiness Threshold Score Range Recommendations Summary

| | | Raw Score (Percent Total Points) | |
|---|---|---|---|
| **Subject** | **Range** | **Minimum** | **Maximum** |
| Algebra II | Broad | 36 (34%) | 63 (59%) |
| | Honed-in | 44 (41%) | 57 (53%) |
| Grade 11 ELA/L | Broad | 66 (48%) | 88 (64%) |
| | Honed-in | 66 (48%) | 76 (55%) |

As part of the exit survey, the participants were provided with the opportunity to note which of the PARCC policy level descriptors defined the "academically-ready" student that they had considered while making their item level judgments. The five policy level descriptors for Grade 11 ELA/L and Algebra II were anchored to a Likert-scale, so that some variability in the in the participant judgment could be captured (e.g., a "high-three (3+)" versus and "low-three (3-)"). Participants were instructed to indicate what level of command the academically-ready just-barely C student they considered would

demonstrate with regard to the knowledge, skills, and abilities embodied by the Common Core State Standards by selecting one of the thirteen levels shown below.

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 **Distinguished** | 5- | 4+ | 4 **Strong** | 4- | 3+ | 3 **Moderate** | 3- | 2+ | 2 **Partial** | 2- | 1+ | 1 **Minimal** |

The distributions of the policy level descriptors that the participants selected as part of the exit survey for Algebra II and Grade 11 ELA/L are shown in Figures 4.3 and 4.4, respectively.

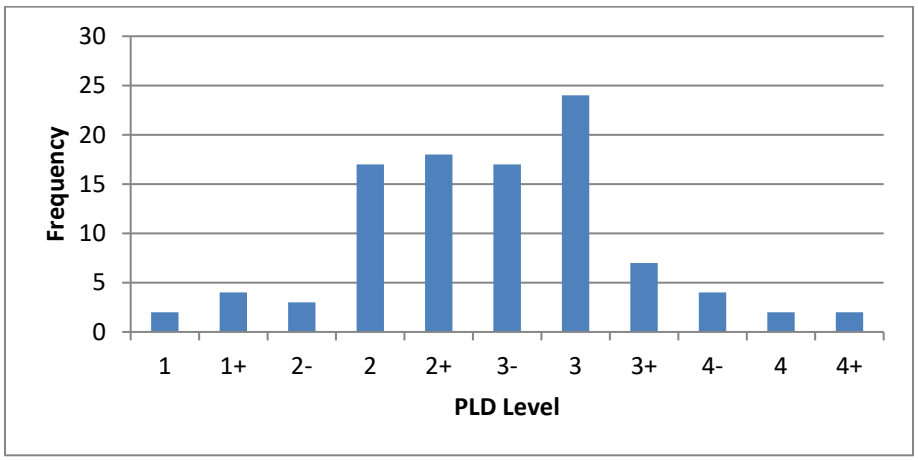



**Figure 4.3:** Distribution of policy level descriptors for all participants of Algebra II study

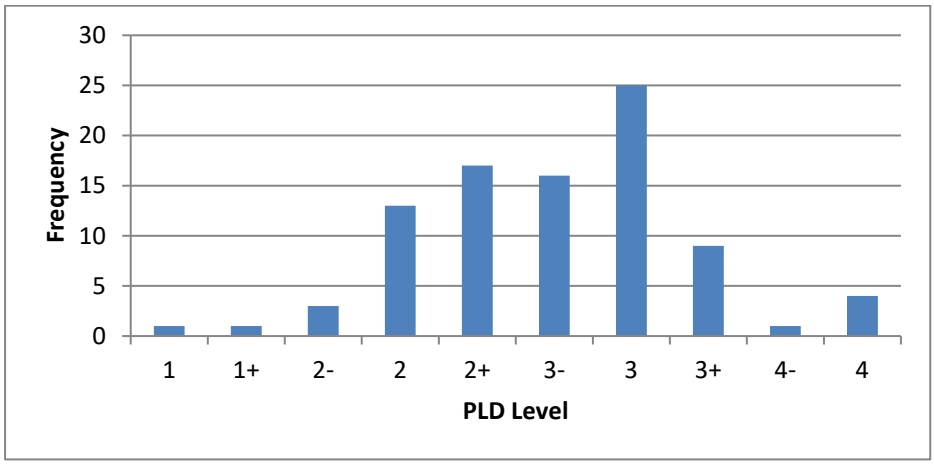



**Figure 4.4:** Distribution of policy level descriptors for all participants completing Grade 11 ELA/L study

The medians of the policy level descriptors selected by the overall participants for Algebra II and Grade 11 ELA/L, representing the knowledge and skills for the borderline 'academically-ready' examinee, were both 3- (Low 3). The modes of the policy level descriptors selected by the overall participants were both 3, for Algebra II and Grade 11 ELA/L. This information played a central role in the decision by PARCC to extend the CCRD label to performance level 3 (see row 2 of Figure 2.1).

For more information about the PARCC Postsecondary Educators' Judgment Study, read the full research report for the Postsecondary Educators' Judgment Study.

## Technical Issues and Caveats

**Benchmarking Study.** When interpreting the results of the study, it is important to note several factors from the contextual analyses. First, many of the external assessments have no stakes associated with them for students. This lack of consequences may result in lower motivation and fewer students performing at the highest levels of the assessment. Second, the PARCC assessments often consist of more constructed-response items and item that require a greater depth of knowledge or higher cognitive complexity level. Third, whenever available, this report includes results from a national sample as well as the states who participated in the PARCC field test or who planned to participate in the PARCC operational assessment. PARCC states tended to perform just as well as – or slightly better than – the nationally reported results. Finally, the PARCC ELA/literacy assessment includes both reading and writing whereas many of the external assessments evaluated these two components separately. Therefore, comparison results are reported separately for reading and writing with the reading results showing greater consistency than the writing results.

Additionally, it is important to note a few limitations. First, the study was based on available data which included some less recent assessments (e.g. NAEP writing from 2011) and some where the population of test takers may not be similar to the PARCC population (e.g., higher performing students taking SAT, fewer students with disabilities and English language learners taking international assessments). Second, the available data needed to be matched to the PARCC Level 4 based on the existing external performance levels. In some cases, it appeared that PARCC Level 4 would fit between two performance levels (e.g., between Proficient and Advanced on NAEP), but student performance is only reported for the existing levels.

**Postsecondary Educators' Judgment Study.** In interpreting these results, it is important to remember that the participants of both studies were asked to make a judgment about the number of points a borderline "academically-ready" student would receive on each item on the assessments, if they were presented the item on the first day of instruction in their course. The judgments that they made were not about the population of students who take these tests in high school. Several participants observed that the students may have forgotten high school content by the first day of taking their course.

> *" I answered these questions based upon the students I see on the first day of my college course. However, I am confident they knew much more when they took Algebra II 2-3 years earlier."* –Participant comment

Other participants observed that students might be expected to continue to build on their math or ELA skills in their 12th grade year. This potential disconnect between the population being judged and the population of students who take the PARCC assessments in high school is important to recognize and deserves consideration in using the study results as one of many inputs into the performance level setting process.

In order for the judgment study to use an open process, so participants could complete the study at their convenience, the test forms used for the study had not yet completed the test construction review process. During the judgment study, several issues were found with items on the test forms as part of the test form review process. Most issues were minor and did not impact the study judgments. However, some changes were made to the items on the Algebra II test forms used for the judgment study to address the issues, including correcting a scoring rubric for an item and correcting an answer option on an item. Although these changes were made early in the study window, a few participants viewed the earlier versions of these items. When looking at the responses of the participants for these items, the data indicated that the participants were following the instructions provided and considered the content needed to respond to the items, rather than focusing on the answer options or scoring rubric provided. Although this should have minimal impact on the results of the study, there is a chance that it could have affected how this small set of participants would have responded to the item.

# Chapter 5: Pre-policy Meeting

This chapter provides details about Step 4 of the PARCC PLS process, which focuses on convening the pre-policy committee and developing "reasonable ranges" based on the PARCC research studies. The sections in this chapter include

- Purpose of Reasonable Ranges
- Purpose and Format of the Pre-policy Committee
- Pre-policy Committee Composition
- Pre-policy Committee Meeting Proceedings
- Reasonable Range Recommendations and Rationale
- Pre-policy Committee Surveys

## Purpose of Reasonable Ranges

In the context of evidence-based standard setting, it is important to understand the extent to which the content judgments of educators are in alignment with what research would suggest relative to the policy goals of the assessment program. Reasonable ranges are intended to provide a range of outcomes that would be considered reasonable relative to the threshold scores for an assessment program. For the PARCC assessments, the reasonable ranges communicated the minimum and maximum percentage of students who may be expected to perform at or above the College and Career Readiness Determination standard in Algebra II, Integrated Math III, and ELA/L 11. This is important to ensure that assessment results will have the intended meaning. Reasonable ranges are intended to be used in conjunction with educator judgment and are not intended to replace educator judgment. In fact, reasonable ranges are not intended to tell educators where the threshold scores should be. Instead, they are intended to identify for educators areas where it simply would not make sense for the threshold scores to fall. If the threshold scores fall outside these reasonable ranges, the meaning of the assessment scores may differ from what was intended.

For PARCC, reasonable ranges relate to the College and Career Readiness Determination (CCRD) and the two research studies (the Benchmarking study and the Postsecondary Educators' Judgment study) that were conducted to inform this indicator. However, research study results do not, by themselves, provide sufficient information to construct the reasonable ranges. To begin with, each study encompasses a variety of different stakeholder perspectives or sources of information which can lead to different conclusions about what would be reasonable. Additionally, the research studies are themselves not perfect and have the types of limitations associated with all research studies relative to scope, participation, timing, and context. As such, the research studies must be first interpreted through a policy lens to identify a range of outcomes which would be considered reasonable. The reasonable ranges recommended by the pre-policy committee were later shared with the high school PLS committees to inform their recommendation of performance standards for the PARCC assessments

## Purpose and Format of the Pre-policy Committee

The purpose for convening the pre-policy committee was to obtain recommendations on the reasonable ranges for college- and career- ready performance standards on the PARCC assessments. The committee included representatives from both K-12 and higher education who served in roles such as Commissioner/Superintendent, Deputy/Assistant Commissioner, State Board Member, Director of Assessment, Director of Academic Affairs, Senior Policy Associate, etc. Committee members reviewed the test purpose and use of the performance standards. They were presented with the methods and results from the PARCC research studies. Using this information and drawing on their policy expertise, the committee was able to provide input into the minimum and maximum percentage of students that would be expected to be classified as college- and career-ready.

## Pre-policy Committee Composition

The pre-policy committee was composed of 22 members who were representatives of K-12 and higher-education from across the PARCC states. Tables 5.1 - 5.6 summarizes the characteristics and experience of the 22 policy committee members. Refer to Appendix 1 for a complete list of the names and positions of the pre-policy committee members.

**Table 5.1:** Gender Distribution of Policy Committee

| Gender | N | % |
|---|---|---|
| Male | 11 | 50 |
| Female | 11 | 50 |

**Table 5.2:** Ethnicity Distribution of Policy Committee

| Ethnicity | N | % |
|---|---|---|
| African American | 1 | 5 |
| Asian or Pacific Islander | 0 | 0 |
| Hispanic | 2 | 9 |
| Multi-racial | 2 | 9 |
| Native American | 0 | 0 |
| White | 17 | 77 |

**Table 5.3:** Years of Experience in Education of Policy Committee Panelists

| Years of Experience | N | % |
|---|---|---|
| None | 0 | 0 |
| 1-5 years | 2 | 9 |
| 6-10 years | 8 | 36 |
| 11-15 years | 3 | 14 |
| 16-20 years | 9 | 41 |
| More than 20 years | 0 | 0 |

**Table 5.4:** Policy Committee's Experience with Roles in Education

| Position | N | % |
|---|---|---|
| Teacher/Instructor/Professor | 18 | 82 |
| School or District Administrator | 11 | 50 |
| State/National Government | 18 | 82 |
| State/National Educational Organization | 5 | 23 |

**Table 5.5:** Policy Committee's Experience with Education Groups

| Population | N | % |
|---|---|---|
| K-12 Education | 22 | 100 |
| Higher Education | 7 | 32 |

**Table 5.6:** Policy Committee's Experience with Student Populations

| Population | N | % |
|---|---|---|
| Students receiving special education services | 14 | 64 |
| Students of low socioeconomic status | 18 | 82 |
| Students who are English language learners | 14 | 64 |
| Students who are receiving general education instruction | 21 | 95 |

## Pre-policy Committee Meeting Proceedings

The pre-policy committee meetings were facilitated by staff from PARCC states, Parcc Inc., and Pearson who were knowledgeable about the PARCC assessment and the PLS meeting process. During the meeting, the pre-policy committee considered and discussed the results of the research studies presented and the policy implications for the PARCC College- and Career-Ready Determination Policy. The full agenda for the pre-policy committee meeting is in Appendix 2. A description of the topics covered in the agenda are described below.

**Welcome and Introductions**

The committee members were introduced and general housekeeping tasks were discussed, including security protocols and reimbursement forms.

**Overview of PARCC Assessment System**

The facilitators provided the committee members with an overview of the PARCC assessment system, including the structure of the assessment system, the assessment performance levels, and the PARCC College- and Career-Ready Determination Policy.

**Overview PARCC Performance Level Setting**

The committee members were taken through the steps of the evidence-based standard-setting process that would be used for the PARCC performance level setting, including the role of the pre-policy committee. A schedule for completing the tasks for the PARCC PLS meetings was also discussed.

**Review of the PARCC Research Studies**

The purpose, methods, and results of the research studies, including the PARCC Benchmarking Study and Postsecondary Educators' Judgment Study, were presented to the committee members. From the PARCC Benchmarking Study, the committee members were presented with a set of ranges representing the percentage of students that would be expected to be college- and career-ready, for Grade 11 ELA/L and High School mathematics, based on PARCC's definition from the PARCC CCRD Policy. From the Postsecondary Educators' Judgment Study, the committee members were presented with a set of ranges representing percentages of total raw scores that college- and career-ready students would be expected to score on the PARCC Algebra II and Grade 11 ELA/L assessments, shown in Table 4.4.

**Committee Discussion and Judgment**

The committee members were organized into table groups with mixed state representation. Within each group, the members examined the results from the research studies and discussed the results in relation to a set of policy questions. The following questions were presented to the committee to discuss in the table groups.

- What percentage of students would you expect to be college- and career-ready?
- Do you expect any differences between English language arts and mathematics?

Each table was provided the opportunity to present the results of their table-level discussion, by providing the range of responses to each of the policy questions.

After a large group discussion around the table groups' responses to the policy questions, the committee members were asked to provide their individual judgments. The committee members provided their individual judgment by answering the following question.

"What is the minimum and maximum percentage of students you expect to be college- and career-ready?"

Individual judgments were collected for each content area, English language arts/literacy and mathematics, separately.

**Evaluations and Closing Remarks**

While the results of the individual judgments were being calculated, there was a whole group discussion about the steps needed by individual states to adopt the performance standards recommended by the PARCC PLS meetings. After this discussion, the committee members were presented the results of the individual judgments and how the recommended reasonable ranges would be used during the PLS meetings. Committee members also completed a process evaluation survey.

## PARCC Reasonable Ranges

Each committee member was asked to provide an individual judgment for the minimum and maximum percentage of students that they would expect to be college- and career-ready. The individual

judgments were made for English language arts/literacy and mathematics, separately. The medians of the minimum and maximum of the individual judgments were calculated to determine the minimum and maximum, respectively, of the reasonable ranges. The statistics for the individual judgments are shown in Table 5.7.

**Table 5.7:** Calculated Statistics for Individual Judgments from Pre-policy Committee

| | Minimum% | | | | Maximum% | | | |
|---|---|---|---|---|---|---|---|---|
| | **Median** | **Mean** | **Min.** | **Max.** | **Median** | **Mean** | **Min.** | **Max.** |
| ELA/L | 20 | 20.15 | 10 | 35 | 35 | 36.50 | 30 | 45 |
| Math | 15 | 15.50 | 5 | 30 | 30 | 30.80 | 20 | 40 |

The final recommended reasonable ranges are shown in Table 5.8.

**Table 5.8:** Final Recommended Ranges by Policy Committee

| | **Minimum %** | **Maximum %** |
|---|---|---|
| English Language Arts/Literacy | 20 | 35 |
| Mathematics | 15 | 30 |

## Pre-policy Committee Surveys

Pre-policy committee members were asked to complete a process evaluation survey at the end of the committee meeting. The purpose of the process evaluation was to collect information about each committee member's experiences in recommending reasonable ranges for the CCRD threshold scores on the PARCC assessments.

The survey was divided into five sections. The first section asked committee members to rate the successfulness of the various components of the policy committee meeting, such as the explanation of the purpose of the meeting and the background and requirements of the PARCC assessment program, the discussion of the research studies, the discussion of the policy questions, and the individual judgments. The second section asked committee members to evaluate the adequacy of the amount of time spent on various elements of the meeting, such as the training, table discussions, and individual judgment tasks. In the third section, committee members were to provide their input on whether they were given adequate opportunities to express their professional opinions about policy questions and reasonable ranges. The fourth section asked committee members whether they were provided adequate opportunities during the meeting to ask questions and interact with their fellow committee members. The fifth section was open-ended so that participants could provide additional comments about the process or their experience as a committee member. The panelists were asked not to include any identifying information on the survey so that the responses would be anonymous.

A summary of the responses to the pre-policy committee process evaluation survey is provided in Appendix 3. The majority of committee members thought that the various components of the meeting

were "successful" or "very successful." The time spent on the research studies, table discussions, and individual judgment tasks was "adequate" to "very adequate." Virtually all committee members responded that they were given adequate opportunities to express their opinions about the research studies and reasonable ranges and to interact with other committee members.

# Chapter 6: Performance Level Setting Committees

This chapter provides details about step 5 of the performance level setting process, which focuses on convening performance level setting meetings. The sections of this chapter include:

- Purpose of Performance Level Setting Meetings
- Dry Run Performance Level Setting Meeting
- Committee Composition and Attendees
- Description of the Performance Level Setting Process
- Data for the Performance Level Setting Meetings
- Meeting Facilitators
- PLS Meeting Proceedings
- Recommended PARCC Threshold Scores

## Purpose of Performance Level Setting Meetings

Performance level setting is based to a large degree on the judgment of educators. Groups of educators make expert recommendations about the level of performance expected for each performance level, based on their experience with different groups of students and knowledge of the assessed content. A specific process or "standard setting method" is used to capture the educator judgments and to translate these into threshold scores for the performance levels. The purpose of the PARCC performance level setting meetings was to gather expert recommendations from groups of educators for the threshold scores that define the performance levels on each PARCC assessment.

Student performance on each of the PARCC assessments is classified into one of five performance levels. Each committee was asked to recommend four threshold scores that would define the boundaries between performance levels 1 through 5. The participants used the following information to make the threshold score recommendations:

- Performance labels and policy definitions
- Grade and subject specific Performance Level Descriptors (PLDs) for each assessment
- Evidence statements for each assessment
- Content of the PARCC assessments

These recommended threshold scores represent the performance level on each assessment that a student would need to meet or exceed to be classified into each performance level.

## Dry-Run Performance Level Setting Meetings

A dry-run of the PARCC PLS meeting process was held for Grade 11 English language arts/literacy (ELA/L) and Algebra II in order to evaluate the implementation of the performance level setting method with

the innovative characteristics of the PARCC assessments. These content areas were selected because they combined all of the various aspects of the PARCC assessments, including the various types of items, scoring rules, and performance level decisions. The dry-run PLS meetings provided the opportunity to implement and evaluate multiple aspects of the operational plan for the actual PLS meeting, including pre-work, meeting materials, data analysis and feedback, and staff and panelist functions. The results of the dry-run PLS meeting were used to implement improvements in the process for the operational PLS meetings.

During the dry-run PLS meetings, panelists were led through the initial planned standard-setting process, facilitated jointly by a content facilitator and process facilitator. After each step in the PLS process, surveys were used to collect participant feedback about the effectiveness of the specific activities involved in each step. After the dry-run PLS meeting, the results of the process evaluations were tabulated and evaluated to identify parts of the process which could be improved for the operational PLS meeting. The results of the dry-run PLS meeting were used to implement improvements in the process for the operational PLS meetings. Specifically, changes made to the PLS process included the following.

- Improved organization of physical materials in the binder and more table space for working with physical materials during the meetings.
- Improved training on assigning full and partial credit for items and tasks.
- Improved online forms for data entry that allowed panelists to submit judgments after each component rather than after each item.
- Improved process for making judgments which allowed panelists to review all performance levels for each item before moving on to the next item.
- Improved structure and visual layout of item-level agreement feedback data to help panelists focus discussion on items with the greatest amount of disagreement in judgments.
- Revision of time allotted for certain process steps (e.g. more time allowed for initial rounds of judgment and table group discussion) to better meet panelist needs.
- Creation of “draft” borderline descriptors to give the panelists a starting point for this task rather than having them start from scratch.

Changes were approved by the TAC prior to implementation. Additional information about the methods and results of the dry-run PLS meetings is available in the full report for the PARCC Dry-Run PLS meetings.

### Committee Composition and Attendees

PARCC solicited nominations for panelists to serve on the performance level setting committees from all states who had administered the PARCC assessments in 2014-2015. Nominations were solicited both from state departments of public education (K-12) and higher education (primarily for participation on the high school panels). Appendix 4 contains materials related to the recruitment of panelists for the PLS meetings, which communicate the criteria used for recruiting panelists. Several panelists were

nominated to participate as part of the PLS meetings due to their experience with prior steps of the PLS process. When selecting panelists, PARCC placed an emphasis on those educators who had content knowledge as well as experience with a variety of student groups and attempted to balance the panels in terms of state representation. Due to various state actions relative to continued participation in the PARCC assessments, some of the panelists originally identified for participation on PLS committees were unable to attend and replacements from other states had to be identified. In some cases these replacements were identified only a couple of weeks before the PLS meetings.

The tables in Appendix 5 summarize the characteristics and experience of the panelists on each PLS committee. These tables provide demographic information about the PLS committee members as well as information about the members' current positions in education, the distributions of panelists across PARCC states, their experience working with various types of student populations, and the types of districts they represent. Panelist response to the gender and ethnicity questions was voluntary.

## Description of the Performance Level Setting Process

For the PARCC PLS meetings, the evidence-based standard-setting method (Beimers, Way, McClarty, & Miles, 2012) was used. This process incorporates elements of the Extended Modified Angoff (Yes/No) standard-setting method (Plake, Ferdous, Impara, & Buckendahl, 2005) with external data to guide meeting participants to recommend performance standards on the PARCC assessments. This method asked panelists to review each item on a reference form of the PARCC assessment and to make the following judgment:

***How many points would a borderline student at each performance level likely earn if they answered the question?***

This extension to the Yes/No standard setting method allowed for incorporation of the multi-point PARCC items by asking educators to evaluate (Yes or No) whether a borderline student would earn the maximum number of points on an item, a lesser number of points on an item, or no points on the item. In the case of a single point or multiple-choice item, this task simplifies to the standard Yes/No method.

The individual item judgments were aggregated to calculate an individual recommended total test score for each performance level, which represented the recommended threshold score for each performance level on the reference PLS test form. Panelists completed three rounds of item judgment. Between the item judgment rounds, the panelists were provided feedback information including data relative to panelist agreement, student performance on the items, and student performance on the test as a whole. High school panelists were shown the pre-policy reasonable ranges prior to making their Round 1 judgments and again as feedback data following each round of judgment.

The materials used for the PLS meetings will be discussed below.

***Moodle.*** A Moodle site was used as the online platform for housing the materials for the PLS meetings and collecting panelist judgments throughout the PLS process. Moodle is an open source e-learning

platform which provides panelists access to the necessary information for completing the PLS meeting. Since some forms of the PARCC assessments were administered through an online environment, TestNav 8, the Moodle site provided panelists access for viewing the secure online items. The Moodle site also provided participants access to online documents which provided background information about the PARCC assessment and practice working in the online environment. Panelists were provided unique user identifications and passwords to provide access to the Moodle site, so they would only have access to the materials for the PLS meetings for which they were participants.

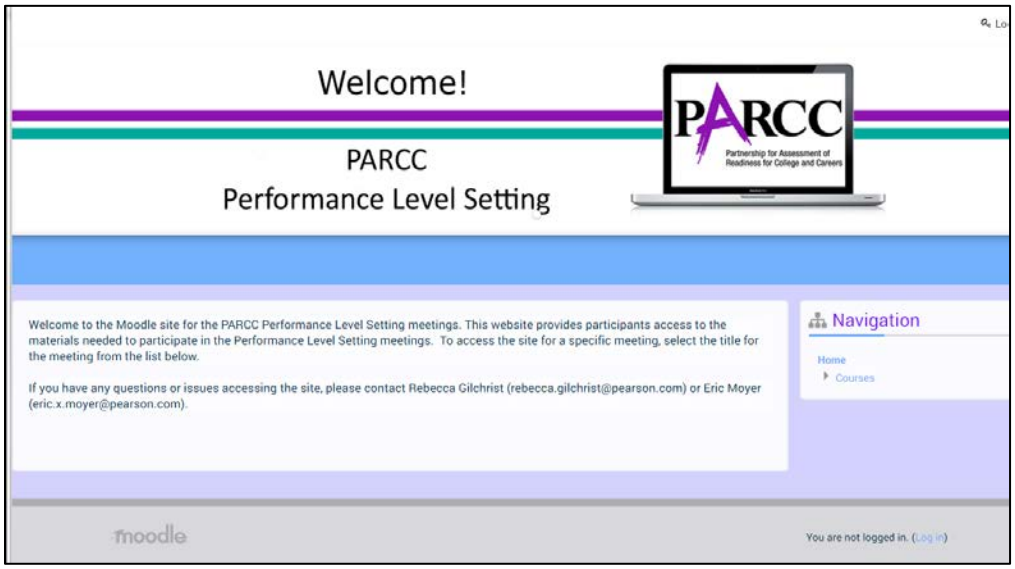



**Figure 6.1**: Example Moodle Home Screen for PARCC PLS Meetings

***Binders.*** In addition to the online resources provided through Moodle, panelists were provided with a meeting binder to organize a variety of hard copy materials they would need to work with throughout the meeting process. These materials included the following.

- Performance Level Descriptors (PLD) for each course
- Evidence Statements for each course
- Scoring Rules
- Test Form Answer Keys
- Item Judgment Record Sheets

The binders were prepared in advance and assigned security numbers. Panelists were required to check out and check in binders each day of the meeting. Panelists were provided additional materials throughout the meeting that they were instructed to insert into their binders.

***Computers.*** Each panelist was provided a laptop computer within their PLS meeting rooms that they could use to access the online resources through the Moodle site. The laptops were Dell latitudes model E5540 with 15.6" screens, standard keyboards with full-size number pad, and an external mouse.

Panelists were not provided with external keyboards, numeric keypads, or external monitors. Panelists were seated in table groups of 3-5 participants at 4 sets of tables in pod configuration, to provide each panelist with enough space to work with the computer and binder materials. The laptops were anchored to the tables through secure cables and power supplies were centrally located in the middle of each table.

***Test Forms.*** The test form used for the PLS meeting is one of the key materials, since for the standard setting method used panelists make judgments on each item on the test form. The results of the individual panelist item judgments are aggregated to create estimated total test scores. The PARCC assessments, during the 2015 administration, were administered using two components, the performance based assessment (PBA) and end-of-year (EOY) assessment. The PBA component was administered after about 75% of the school year has completed. The EOY component was administered after about 90% of the school year has completed. A single form of the operational spring 2015 PBA and EOY assessments was selected for use during the PLS meeting. In all cases the test form that was used as the basis for "Online Accommodated Group 2A" was selected to serve as the performance level setting test form for each test, to allow for the greatest possible inclusion of students receiving accommodations in the performance level setting data. Table 6.1 shows the PBA and EOY form used for the PLS meetings.

**Table 6.1:** PBA and EOY forms used for PLS meetings

| ELA/L | | | Mathematics | | |
|---|---|---|---|---|---|
| **Test** | **PBA** | **EOY** | **Test** | **PBA** | **EOY** |
| Grade 11 | Form 1 | Form 1 | Algebra II | Form 1 | Form 1 |
| Grade 10 | Form 2 | Form 1 | Geometry | Form 1 | Form 1 |
| Grade 9 | Form 5 | Form 1 | Algebra I | Form 1 | Form 1 |
| Grade 8 | Form 5 | Form 1 | Integrated III | Form 1 | Form 1 |
| Grade 7 | Form 2 | Form 1 | Integrated II | Form 1 | Form 1 |
| Grade 6 | Form 5 | Form 4 | Integrated I | Form 2 | Form 1 |
| Grade 5 | Form 2 | Form 1 | Grade 8 | Form 1 | Form 1 |
| Grade 4 | Form 5 | Form 1 | Grade 7 | Form 1 | Form 1 |
| Grade 3 | Form 4 | Form 1 | Grade 6 | Form 1 | Form 1 |
| | | | Grade 5 | Form 1 | Form 1 |
| | | | Grade 4 | Form 1 | Form 1 |
| | | | Grade 3 | Form 1 | Form 1 |

*English Language Arts/Literacy.* The PBA component for the ELA/L assessment included three tasks which addressed narrative writing, literary analysis, and research simulation, respectively. For the narrative writing task, students read one short piece of literature and then answered a small number of reading items to gather ideas to write a narrative story. For the literary analysis task, students read one short piece of literature and one extended piece of literature and answered 4-6 reading comprehension questions. They then wrote an analytic essay analyzing the texts. For the research simulation task, students read one informational text and answered reading items to gather ideas for the writing task.

They then read one or two additional shorter texts, gathered additional ideas by answering more reading items, and then incorporated evidence from multiple texts in an analytic essay. Within each task, student responded to evidence-based selected response (EBSR) items or technology-enhanced constructed response (TECR) items, and completed a prose constructed response (PCR) item, or essay. The EOY component of the ELA/L assessment focused on reading comprehension, including three passages or passage sets. All of the passages in the EOY component were either literary or informational. For each passage or passage set, the students were asked to respond to several items, EBSR and TECR, based on evidence from the text or texts. Table 6.2 shows the total number of items and possible points for each ELA/L test form used during the PLS meetings.

**Table 6.2:** Number of items and total points for each ELA/L assessment

| Grade | PBA | | | | EOY | | Total | |
|---|---|---|---|---|---|---|---|---|
| | EBSR/TECR | | PCR | | EBSR/TECR | | | |
| | Items | Points | Items | Points | Items | Points | Items | Points |
| 3 | 17 | 34 | 3 | 42 | 12 | 24 | 32 | 100 |
| 4 | 20 | 40 | 3 | 42 | 12 | 24 | 35 | 106 |
| 5 | 20 | 40 | 3 | 42 | 12 | 24 | 35 | 106 |
| 6 | 20 | 40 | 3 | 53 | 22 | 44 | 45 | 137 |
| 7 | 20 | 40 | 3 | 53 | 22 | 44 | 45 | 137 |
| 8 | 20 | 40 | 3 | 53 | 22 | 44 | 45 | 137 |
| 9 | 20 | 40 | 3 | 53 | 21 | 42 | 44 | 135 |
| 10 | 20 | 40 | 3 | 53 | 22 | 44 | 45 | 137 |
| 11 | 20 | 40 | 3 | 53 | 22 | 44 | 45 | 137 |

Note: After the PLS committee meetings, items on some test forms were omitted due to content or psychometric reasons. These changes are reflected in Table 7.3.

In the PBA and EOY component, the non-essay items are comprised of two individual items (parts A and B) that are related. Each pair of items is scored together as a single item providing a single score, with a maximum score of 2 points. The score received for each pair of items is defined by scoring rules, which were provided to the panelists for the PLS meeting.

The essay item for each task on the PBA component (called Prose Constructed Response or PCR) was scored on up to three different constructs using a rubric. The PCRs for all three tasks were scored on the "Written Expression" construct and "Knowledge of Language and Convention" construct. The "Comprehension of Key Ideas and Details" construct was only scored for the literary analysis and research simulation tasks. The number of points for each scoring component of the PCR was different for different grade bands, as shown in Table 6.3.

**Table 6.3:** Points for Scoring Components of PCR items

| | Maximum Possible Points | |
|---|---|---|
| | Grades 3-5 | Grades 6-11 |
| **Reading Comprehension*** | 3 | 4 |
| **Written Expression **** | 3 | 4 |

| | | |
|---|---|---|
| **Knowledge of Language and Conventions** | 3 | 3 |

Notes: * - Reading comprehension was not scored as part of the narrative writing task. ** - The written expression score was weighted by 3 when calculating the total score for an item.

*Mathematics.* All math assessment contained three types of items, Type I, Type II and Type III, which vary in the content focus. Type I tasks primarily assess concepts, procedures, skills, and applications. Type II tasks focus on mathematical reasoning and the examinee's ability to justify arguments or critique reasoning. Type III tasks assess mathematical modeling and applications. The PBA assessment includes all three types of tasks while the EOY assessment only includes Type I tasks. The maximum number of points on items ranged from 1 to 6 points. Table 6.4 shows the total number of each type of item and the total possible points on each mathematics test form used during the PLS meetings.

**Table 6.4:** Number of Items and Total Points for each Mathematics Assessment

| | PBA | | | | EOY | | Total | |
|---|---|---|---|---|---|---|---|---|
| **Grade** | **Type I** | **Type II** | **Type III** | **Points** | **Type I** | **Points** | **Items** | **Points** |
| 3 | 10 | 4 | 3 | 38 | 39 | 44 | 56 | 82 |
| 4 | 10 | 4 | 3 | 38 | 36 | 44 | 53 | 82 |
| 5 | 9 | 4 | 3 | 38 | 36 | 44 | 52 | 82 |
| 6 | 10 | 4 | 3 | 38 | 34 | 44 | 51 | 82 |
| 7 | 10 | 4 | 3 | 38 | 33 | 44 | 50 | 82 |
| 8 | 11 | 4 | 3 | 38 | 33 | 44 | 51 | 82 |
| Algebra I | 10 | 4 | 4 | 42 | 35 | 55 | 53 | 97 |
| Geometry | 10 | 4 | 4 | 42 | 34 | 54 | 52 | 96 |
| Algebra II | 10 | 5 | 5 | 52 | 32 | 49 | 52 | 101 |
| Integrated I | 10 | 4 | 4 | 42 | 31 | 51 | 49 | 93 |
| Integrated II | 10 | 4 | 4 | 42 | 34 | 55 | 52 | 97 |
| Integrated III | 10 | 5 | 5 | 52 | 35 | 55 | 55 | 107 |

Note: After the PLS committee meetings, items on some test forms were omitted due to content or psychometric reasons. These changes are reflected in Table 7.3.

## Data for the Performance Level Setting Meetings

Within the separate PBA and EOY administrations, test forms were spiraled across students. This resulted in a relatively large number of form combinations when considering PBA and EOY together. For example, in a situation where there were 6 core forms of PBA and 6 core forms of EOY, there would be 36 different form combinations. Since scaling and equating analyses were still ongoing at the time of the PLS meetings, data used for PLS feedback data (item means, item score point distributions, and impact data) reflect only those students who took the specific form combination reflected in Table 6.1. Additionally, data used for the PLS feedback data represented only those students who took an online test form. Adjustments made through PARCC's mode comparability research were then used to appropriately scale the paper test data to the online test data. Lastly, those students who took an online accommodated version of the test form were represented in the PLS feedback data because the test form selected for PLS contained the same items as the online accommodated form. Because only

one form of the online test was accommodated and students generally took an accommodated form for both the PBA and EOY components of the PARCC assessment, students taking an accommodated form of the test were not spread out across different form combinations. As a result, the sample size of students taking the accommodated form of the test was frequently larger than the sample size of students taking the non-accommodated form of the test. To address this issue, the accommodated and non-accommodated test data sets were proportionally weighted relative to the overall percent of the online testing population (for each grade/content area) who received an accommodated test form. Table 6.5 shows the number of students (from both the accommodated and non-accommodated forms of the test) included in the PLS feedback data.

**Table 6.5:** Number of Students Represented in the PARCC PLS Data

| Math | | | ELA/L | | |
|---|---|---|---|---|---|
| Grade | Non-Accommodated | Accommodated | Grade | Non-Accommodated | Accommodated |
| 3 | 10498 | 33468 | 3 | 12569 | 1961 |
| 4 | 10562 | 32270 | 4 | 13885 | 2804 |
| 5 | 10875 | 31342 | 5 | 15937 | 2797 |
| 6 | 10889 | 33696 | 6 | 13591 | 2855 |
| 7 | 10773 | 31566 | 7 | 15148 | 2809 |
| 8 | 8689 | 26606 | 8 | 15980 | 2632 |
| Algebra I | 5624 | 20575 | 9 | 6251 | 1658 |
| Geometry | 2626 | 6445 | 10 | 4900 | 664 |
| Algebra II | 4463 | 5513 | 11 | 5381 | 477 |
| Integrated I | 6115 | 4557 | | | |
| Integrated II | 2994 | 758 | | | |
| Integrated III | 2224 | 207 | | | |

Item means and item score point distributions (percent of students at each score point on an item) were calculated directly from the proportionally weighted data set. Because of the relatively small number of students who took the specific PLS test form combination (especially for high school math), an additional step was taken to smooth the distribution of student test scores in calculating impact data. The goal was to produce a smoothed distribution of student performance that would more closely resemble the distribution if the full population of student testers was included. Without this step, movement of a threshold score by 1 or 2 points might result in an artificial increase or decrease in student performance that would be misleading feedback to panelists. A loglinear smoothing procedure was conducted for both the overall impact data distribution and subgroup impact data distributions using Proc Genmod through SAS (Moses & vonDavier, no date). Through this method the four moments of the original raw score distribution were preserved while the fluctuations throughout the distribution due to small numbers of students were smoothed out.

### Meeting Facilitators

For each PLS breakout session there were two facilitators assigned, a content facilitator and process facilitator. All PLS meeting facilitators were staff from Pearson or ETS. The content facilitator was responsible for leading the panelists through the sections of the PLS process associated with the development of the test content and the development of the performance level descriptors for the assessment, including:

- Introduction to the PARCC assessment
- Experience the assessment activity
- Scoring rules and scoring the assessment
- Development of the borderline performance level descriptors

The process facilitator was responsible for leading the panelists through the sections of the PLS process associated with the panelists making threshold score recommendations, including:

- Performance level setting training
- Practice item judgment round and discussion
- Item judgment rounds 1, 2, and 3
    - Reviewing the panelists' task
    - Presenting the feedback data to the panelist
    - Facilitating the panelists discussion about the feedback data
- Vertical articulation

Although the PLS process steps were assigned to specific facilitators during the meeting, the facilitators closely collaborated while facilitating the meeting, providing support when needed.

To ensure that all facilitators of the PLS breakout meetings were prepared to lead the discussions associated with setting the performance level standards for the PARCC assessment, each facilitator was required to complete a set of online trainings associated with the PLS meetings. The training sessions were facilitated using a Moodle site created specifically for this purpose, with an overall course created for the facilitator training and different units, within the course, created for each training session. Any materials associated with the training session were housed on the Moodle site, within the specific unit. The actual training was provided as online training, so individuals from various parts of the country participating as facilitators could participate. A recording of the online training was provided on the Moodle site, so facilitators could review the training at their convenience and individuals that could not attend the online training would be able to access the training.

Facilitators were required to attend the training that was associated with the tasks for which they were responsible, but were also able to optionally attend any of the training sessions provided. The training sessions that all facilitators needed to complete included:

- Moodle Bootcamp – an introduction to working within the Moodle platform for the facilitators, since this was the platform used for facilitating the PLS meetings.

- Overview of PARCC – an introduction to the various aspects of the PARCC assessment system, including questions that may come up during the PLS meetings.
- PARCC PLS Meeting Facilitation Parts 1 and 2 – a step by step overview of the PLS process used for setting the PARCC standards, including how to use Moodle within the process.
- Lessons Learned from High School PLS Meetings – the facilitators from the high school PLS meetings were able to share any information learned with the facilitators of the grades 3-8 PLS meetings

The training sessions that the content facilitators were required to complete included:

- Introduction to Performance Level Setting – an overview of the purpose of a PLS meeting and the different elements of standard setting, and how PLS meetings would be used with the PARCC assessment system.
- Developing Borderline Descriptors – a description of the three-step jigsaw cooperative process panelists would be led through to develop the borderline descriptors.

The training sessions that the process facilitators were required to complete included:

- PARCC College and Career Readiness – a review of PARCC's College- and Career-Ready Determination Policy, the research studies performed in support of the PLS meeting, and their use in the PLS process.
- Vertical Articulation – a discussion of the expected flow of events during the vertical articulation meetings and review of available materials for the meetings.

At the end of each training session, a short online quiz was provided for the facilitators to complete, to check that the training sessions were completed by the facilitators.

### Meeting Proceedings

The PLS meetings for the PARCC assessments were conducted during three one-week sessions. The dates of the twelve PLS committee meetings are shown in Table 6.6.

**Table 6.6:** PARCC PLS Committee Meetings and Dates

| Dates | Committees by Subjects and Grades |
|---|---|
| July 27 - 31, 2015 | Algebra I/Integrated Math I |
| | Geometry/Integrated Math II |
| | Algebra II/Integrated Math III |
| | Grade 9 English Language Arts/Literacy |
| | Grade 10 English Language Arts/Literacy |
| | Grade 11 English Language Arts/Literacy |
| August 17 - 21, 2015 | Grades 7 & 8 Mathematics |
| | Grades 7 & 8 English Language Arts/Literacy |
| August 24 - 28, 2015 | Grades 3 & 4 Mathematics |
| | Grades 5 & 6 Mathematics |
| | Grades 3 & 4 English Language Arts/Literacy |
| | Grades 5 & 6 English Language Arts/Literacy |

The organization of the PLS meetings allowed the recommended performance standards from the upper grade-level meetings to be used as part of the feedback for the lower grade-level meetings.

The PLS meetings were facilitated jointly by a content facilitator and a process facilitator. Appendix 6 includes the complete agendas for the High School and Grades 3-8 PLS meetings for mathematics and ELA/L. Table 6.7 shows the high level agenda for the PLS committee meetings.

**Table 6.7:** PLS Committee Meeting Agenda Topics

| | |
|---|---|
| General Session | • Welcome and Introductions<br>• Overview of PARCC Assessment System<br>• Overview of Performance Level Setting Process<br>• Example Performance Level Setting |
| Breakout Sessions | • Introductions<br>• Overview of Item Development<br>• Experience the Assessment*<br>• Review of Scoring Rules<br>• Policy Level and Performance Level Descriptors<br>• College and Career Readiness<br>• Borderline Performance Level Descriptors*<br>• Performance Level Setting Training<br>• Round 1: Judgment and Feedback*<br>• Round 2: Judgment and Feedback*<br>• Round 3: Judgment and Feedback*<br>• Evaluation and Closing Remarks |

Note: *These tasks were repeated for each assessment for which the committee was recommending threshold scores.

In addition, vertical articulation was conducted at the end of each PLS meeting to allow committee members an opportunity to make final adjustments to their recommendations in light of results from the third round of judgment and threshold score recommendations from other committees.

The following will describe the steps used to guide the panelists through the entire PLS process.

**PLS Meeting Pre-work**

The PLS meeting panelists were required to complete a set of activities prior to attending the onsite meetings. The purpose of the assigned pre-work was to expedite the training of the panelists for the PLS meetings, by providing the panelists an opportunity to experience the online testing environment and familiarize themselves with information that would be used throughout the PLS meeting. The pre-work assigned included:

- Tutorials - Participants completed tutorials to introduce them to the online testing environment, TestNav 8, the various tools that students would have access to, and the different types of items they would encounter.
- Practice tests - Participants worked on practice tests for both the PBA and EOY components of the assessments related to the PLS meeting
- Reference materials - Participants reviewed reference materials that would be used throughout the PLS meetings, including evidence statements, performance level descriptors, and the PARCC College- and Career-Ready Determination Policy.
- Non-disclosure agreement (NDA) - panelists were required to read and agree to the provisions contained in a standard non-disclosure agreement. The panelist agreement to the NDA was recorded in Moodle, by participants responding to a single question in an online quiz. Note that for the High School PLS meetings, the NDA was completed as a first step during the onsite meetings rather than as part of the pre-work.

Additionally, the panelists were required to complete an online survey, to collect demographic information. This information was used during the analysis of the panelist responses. The survey collected information about the panelist's current teaching position, teaching experience, school information and personal demographic information.

**General Session**

The purpose of the general session for the PLS meetings was to welcome the members of the PLS committees, provide background information about the PARCC assessment system, and introduce the PLS meeting process. As part of the general session, participants were guided through an example exercise using the PLS process. A single general session took place at the beginning of each set of PLS meetings.

**Breakout Sessions**

After the general session, panelists moved into content-specific breakout sessions for the remainder of the PLS meeting. Within each committee, panelists were divided into groups of 3 to 5 panelists, pre-assigned before the meeting. Each table group consisted of different types of committee members so that there was a blend of expertise at each table. Table leaders were identified to facilitate the discussions and assist in meeting logistics at each table.

**Preparing for the Item Judgment Activity**

To prepare panelists for completing the individual item judgments during the PLS meetings, the panelists received training specific to the item judgment tasks, including completing various activities.

***Experience the Test.*** Panelists were required to experience the specific test form which was used during the PLS meeting by completing each of the items on the test form, as the student would experience the items. The items were accessed through the Moodle site, since it was administered as an online assessment. Since the version of the online system used in the PLS meetings did not store and score participant responses, participants recorded their responses on a separate item response form. During this review, if the panelists identified any issues with items on the test forms, they were asked to record the comments on an Item Comment Form.

Answer key documents were provided to panelists so they could score their responses to the items. For the mathematics and ELA/L assessments, there are unique scoring rules for scoring the different types of items on the assessment. Facilitators trained panelists on these scoring rules and then walked through several examples with them on their respective assessments. The purpose of the scoring rules training was to help the panelists determine the score they would have received on the assessment and to help them evaluate what types of item responses would earn students different point values.

***Borderline Performance Level Descriptors.*** To help inform discussions during the PLS meeting, facilitators reviewed the performance labels, policy definitions, and discussed the meaning and uses of the PARCC College- and Career-Ready Determination Policy. The participants were then directed to review the grade- and subject- specific performance level descriptors (PLDs). The PLDs gave the panelists a common understanding of the knowledge, skills, and abilities students should demonstrate within each performance level for a specific PARCC assessment. When reviewing the PLDs, panelists were asked to discuss the differences between the expectations at the different performance levels.

The panelists were then introduced to differences between a *typical* student and *borderline* student within a performance level, where the borderline student was described as the minimally qualified student to be classified within a particular performance level.

Prior to the PLS meetings, a set of "draft" borderline descriptors for each assessment was created by a set of experts familiar with the performance level descriptors (PLD) for the specific assessment. The “draft” borderline descriptors contained a unique list of knowledge, skills, and abilities that a borderline student would be expected to demonstrate. The participants were guided through a three-step "jig-saw"

review process to revise the "draft" borderline descriptors to develop a final list of descriptors for the borderline student.

Step 1: A representative from each table group was assigned to a performance level group, where they reviewed and discussed the PLDs and the "draft" borderline descriptors for their assigned performance level. The purpose of this step was to have at least one individual from each table group with a deep understanding of each performance level.

Step 2: At the original table groups, the participants reviewed the borderline PLDs and notes from the Step 1 discussion to revise a specific sub-claim or skill area of the "draft" borderline PLDs to describe the borderline students. To make most efficient use of time, each table was assigned one or more sub-claims/skill areas to revise. The participants worked collaboratively within the Moodle site to edit all performance levels for their assigned areas of the "draft" borderline descriptors.

Step 3: The facilitator combined the edited "draft" borderline descriptors from each table group into a single master borderline descriptors document that was displayed for all participants. During a large-group discussion, participants had the opportunity to make edits to the master borderline descriptors document, to represent the level of performance borderline students would be expected to demonstrate for each performance level.

Due to the overlapping relationships of the Integrated Math assessments to the traditional math assessments, the process for creating the borderline descriptors for Integrated Math I, Integrated Math II, and Integrated Math III was slightly different than for the other PARCC tests. "Draft" borderline performance level descriptors were not created for the courses in the Integrated Mathematics series. Instead, the panelists were provided the final master borderline descriptors for Algebra I, Geometry, and Algebra II. During step 2, the panelists used the descriptors from these three documents to create the borderline descriptors for their assigned sub-claims.

A copy of the final master borderline PLD document was printed and provided for each participant to put into their binder.

***Performance Level Setting Training.*** The PLS committee members were provided training on the extended modified Angoff (Yes/No) method (Plake, Ferdous, Impara, & Buckendahl, 2005) and how to use the Moodle site to record their individual item judgments. For each item, the panelists were instructed to view the item within the online testing environment, accessed through Moodle, review the answer key, scoring rules and rubric, and borderline PLD associated with the item. Based on the review of the item and related materials, the panelist would answer the following question:

"How many points would a borderline performance level student likely earn if they answer this item?"

Likely was defined for panelists as 2 out of 3 times. The response to the question for each item was recorded on an item judgment form and within the item judgment survey in the Moodle site. The item

judgments were made for all performance levels for one item before moving on to the next item, starting with performance level 2 and moving up to performance level 5. When the panelists completed making individual item judgments in the item judgment survey, they would submit their responses in the Moodle system (see Figure 6.2).

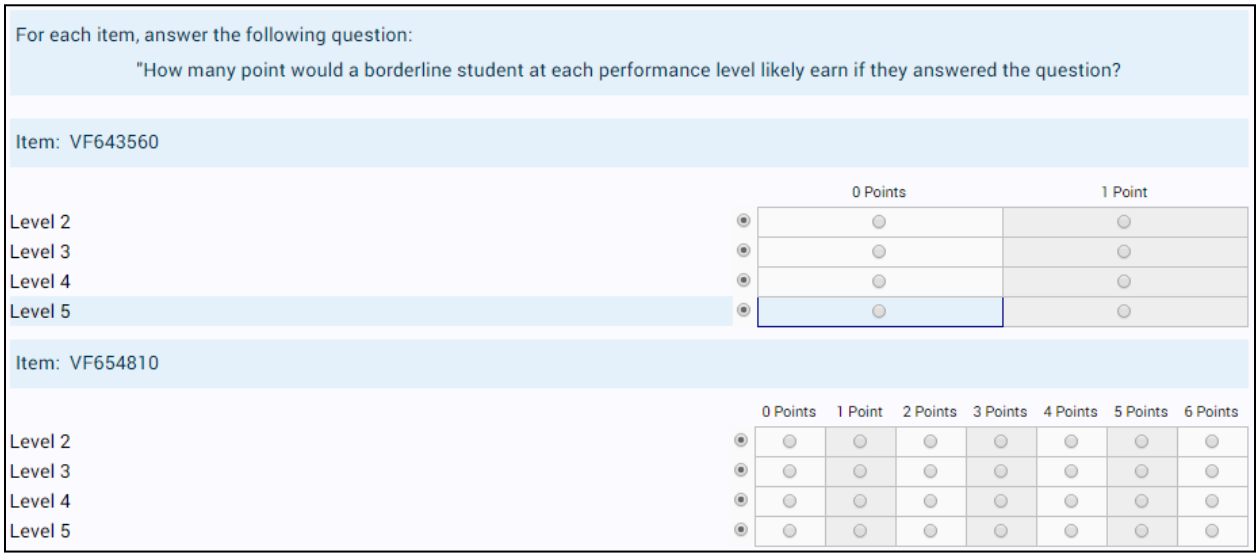



**Figure 6.2**: Example Item Judgment Survey for two items

To provide the panelists practice in making item judgments, the panelists completed a practice item judgment task. The panelists made judgments for all borderline performance levels on a set of practice items, including items from both the PBA and EOY assessments. The practice item set for the mathematics assessments include 8 items, with 4 items from the PBA component and 4 items from the EOY component. The practice item set for the ELA/L assessments included 6 items, with 3 items from the PBA component, including one prose constructed response (PCR), and 3 items from the EOY component. Panelists were asked to complete their judgments independently and without discussion from other panelists. After the panelists completed the practice item judgment activity, a group discussion was used to review the panelist responses and answer any questions that the panelists had about the item judgment process.

If the PLS meeting was planned to set performance levels for two assessments, the panelists only completed training and practice activities for the first assessment. No additional training or practice was conducted for the second assessment as panelists were already experienced with making item judgments at that point in the process.

**Item Judgment Rounds**

After receiving training on the PLS procedure, the committee members participated in three rounds of judgments for each assessment. Within each round, panelists were asked to consider the items in the test form, starting with the PBA component and then the EOY component. Items were sequenced in the same way that students saw them on the operational test forms and were not ordered by difficulty.

Each panelist made a judgment for the Level 2 performance level, followed by judgments for Level 3 performance level, Level 4 performance level, and Level 5 performance level, in this order. The panelists entered their item judgments for each round by completing an item judgment survey in Moodle. Judgments for items on both the PBA and EOY were completed in the same item judgment form. The panelists were provided the opportunity to save or submit their item judgments after completing the item judgments for the PBA and again after completing the item judgments for the EOY.

Before starting each of three item judgment rounds for an assessment, the panelists were required to complete a readiness survey indicating that they understood the task and process used to complete the item judgments. Prior to rounds 2 and 3 the panelists were also asked if they understood the data that was provided during the feedback discussion for the prior round. The panelists had to answer "yes" to all questions on the readiness survey before continuing with the item judgment round. If they answered "no" they were asked to notify a facilitator for additional assistance.

***Round 1 Item Judgment and Feedback***. During the first round of item judgments, panelists made their item judgments primarily based on the content of the test form and the borderline descriptors. After round 1 item judgments, the following types of feedback were provided to the panelists.

- Total Score Level
  - The panelist's individual Round 1 recommended threshold scores for Performance Levels 2, 3, 4, and 5.
  - Table-level Round 1 threshold score recommendations for Performance Levels 2, 3, 4, and 5, including the minimum, maximum, mean, and median.
  - Committee-level Round 1 threshold score recommendations for Performance Levels 2, 3, 4, and 5, including the minimum, maximum, mean, and median.
  - Panelist agreement chart, displaying each committee member's Round 1 threshold score recommendations on a bar graph, for each performance level and adjacent performance levels.
  - For high school assessments, pre-policy reasonable ranges in terms of percent of points on the reference PLS test form. This was compared to the range of panelist threshold scores for performance level 3.
- Item Level
  - Panelist item judgment agreement for each item, showing the percent of panelists that selected each possible score, including the five items with the greatest level of panelist disagreement for each performance level
  - The score point distribution for the items, showing the percent of students that received each of the score points possible, along with the item mean.

An example of committee-level Round 1 feedback can be found in Appendix 7. The panelists had the opportunity to discuss the feedback in their table groups.

For the High School PLS meetings (Grades 9 through 11 ELA/L and high school math sequences), the participants were additionally provided information associated with the reasonable ranges from the pre-policy meeting. For round 1 feedback, this information was expressed as the range of points on the reference PLS test form which would identify a percent of students as college- and career- ready consistent with their reasonable range. This conversion from percent of students (how the pre-policy reasonable ranges were originally expressed) to percent of points was necessary as panelists were not yet provided with information about student performance on the overall test (impact data). The pre-policy reasonable range in terms of percent of points was compared to the range of panelist threshold scores for performance level 3. The participants in the high school mathematics PLS meetings were all presented the reasonable range for the percentage of total points for the Algebra II test, regardless of assessment. The participants in the high school ELA/L PLS meetings were all presented the reasonable range for the percentage of total points for the Grade 11 ELA/L test, regardless of assessment. This decision was made after review of the distribution of student performance on the high school test forms and was intended to support alignment of performance standards across the high school tests throughout the PLS meeting process.

***Round 2 Item Judgment and Feedback.*** For the second round of judgments, panelists made their item judgments based on the first-round feedback, discussion with their table groups, the content in the items and the borderline descriptors. After completing their Round 2 judgments, panelists were provided with the following second-round feedback:

- Total Score Level
    - The panelist's individual Round 2 recommended threshold scores for Performance Levels 2, 3, 4, and 5.
    - Table-level Round 2 threshold score recommendations for Performance Levels 2, 3, 4, and 5, including the minimum, maximum, mean, and median.
    - Committee-level Round 2 threshold score recommendations for Performance Levels 2, 3, 4, and 5, including the minimum, maximum, mean, and median.
    - Panelist agreement chart, displaying each committee member's Round 2 cut score recommendations in a bar chart, for each performance level and adjacent performance levels.
    - For high school assessments, pre-policy reasonable ranges in terms of percent of points on the reference PLS test form. This was compared to the range of panelist threshold scores for performance level 3.
    - Impact data for the assessment based on the committee's Round 2 threshold score recommendations, presented as an overall group and disaggregated as gender, ethnicity, English language learners (ELL), and students with disabilities (SWD).
        - For high school assessments, pre-policy reasonable ranges in terms of the percent of students expected to be classified as college- and career-ready on the PARCC assessments. This was compared to the minimum and maximum

percentage of students in Level 3 or above based on panelists' individual threshold judgments.
    - For grades 7 & 8, the percent of students in Level 3 or above for the High School Assessments in the same subject area were presented. See Appendix 7 for more detail.
    - For grades 3 through 6, the impact data for the grades 7 & 8 assessments in the same subject area were presented. See Appendix 7 for more detail.
- Item Level
    - Panelist item judgment agreement for each item, showing the percent of panelists that selected each possible score, including the five items with the greatest level of panelist disagreement for each performance level

The panelists had the opportunity to discuss the feedback data in their table groups and as an overall group. An example of committee-level feedback after Round 2 item judgments can be found in Appendix 7.

***Round 3 Item Judgment and Feedback.*** During the third round of judgments, participants made their final item judgments based on all of the feedback they received in the first two rounds. The feedback each panelist received after Round 3 included the following:

- Total Score Level
    - Committee-level Round 3 threshold score recommendations for Performance Levels 2, 3, 4, and 5, including the minimum, maximum, mean, and median.
    - Impact data for the assessment based on the committee's Round 3 threshold score recommendations, presented as an overall group and disaggregated as gender, ethnicity, English language learners (ELL), and students with disabilities (SWD).
        - For high school assessments, pre-policy reasonable ranges in terms of the percent of students expected to be classified as college- and career-ready on the PARCC assessments. This was compared to the minimum and maximum percentage of students in Level 3 or above based on panelists' individual threshold judgments.
        - For grades 7 & 8, the percent of students in Level 3 or above for the High School Assessments in the same subject area were presented. See Appendix 7 for more detail.
        - For grades 3 through 6, the impact data for the grades 7 & 8 assessments in the same subject area were presented. See Appendix 7 for more detail.

No feedback data were presented at the item level following Round 3. Panelists were given an opportunity to discuss the Round 3 feedback data. This discussion and the Round 3 results were the primary inputs to the vertical articulation process.

**Vertical Articulation**

***Meeting participants.*** The final activity in which the PARCC PLS panelists participated was the vertical articulation. The purpose of the vertical articulation was to look at the threshold score recommendations that were made across PARCC assessments within a content area and evaluate the reasonableness of these threshold scores. Panelists were shown the impact data resulting from their Round 3 threshold score recommendations across grades/courses within a content area. Recommendations for adjustment to the threshold scores could be made by members of the vertical articulation committee, after reviewing the Round 3 feedback and group discussion. A dynamic spreadsheet was used to show the panelists how changes in the threshold scores would change the impact data across performance levels.

The participants of the high school vertical articulation committee were the table leaders from each of the individual PLS meetings. Vertical articulation meetings were held for the mathematics assessments and ELA/L assessment separately. During the High School mathematics vertical articulation, the panelists reviewed the results from the traditional sequence, Algebra I, Geometry, and Algebra II, and the integrated sequence, Integrated Math I, II, and III, separately.

Because of the logistics of when the Grade 7&8 PLS meetings were held, there was only one committee for each content area. As a result, the vertical articulation for the Grades 7 & 8 PLS meetings was not conducted as a separate meeting, but was integrated into breakout session for these committees. All participants of the PLS meeting were involved in the vertical articulation. During the vertical articulation, the participants were again presented with the results from the grades 7 and 8 assessments on which they had made judgments as well as the percent of students at Level 3 or above on the high school assessments in the same subject area. Participants in the Grade 7 & 8 PLS meetings were only allowed to make changes to the threshold score recommendations for the grade 7 and 8 tests.

The vertical articulation for the Grades 3-6 PLS meetings were separate meetings convened after the PLS meetings were concluded. The participants of the vertical articulation meeting were the table leaders from the individual PLS meetings. Vertical articulation meetings were held for the mathematics assessments and ELA/L assessment separately. During the vertical articulation meetings, the panelists were presented with the final results from the grades 3 - 8 assessments, for the respective subject. Participants in the Grade 3-6 vertical articulation meeting were only allowed to make changes to the threshold score recommendations for the grade 3-6 tests.

***Meeting process.*** The vertical articulation process involved three steps.

- PLD Review Activity
- Review and Discussion of Cross-Grade Impact Data Charts
- Recommend Changes to Threshold Scores

Participants began the vertical articulation process by reviewing the PLDs for the set of grade levels they were reviewing as well as PLDs for the higher grade levels (for grades 7 & 8, this was the grade 9 ELA/L,

Algebra I, and Integrated Math I PLDs; for grades 3-6 this was the grades 7 & 8 PLDs). Panelists reviewed the PLDs independently with instructions to look for differences across grades. They then discussed differences as a table group. Finally, based on these discussions, panelists completed a survey within Moodle which asked them to provide their expectations for student performance for each pair of adjacent grade levels (see Figure 6.3). Aggregate results for the group were recorded on a flip chart for reference in the next step of the process. Areas where a majority of panelists indicated "much greater than" or "much less than" were noted.

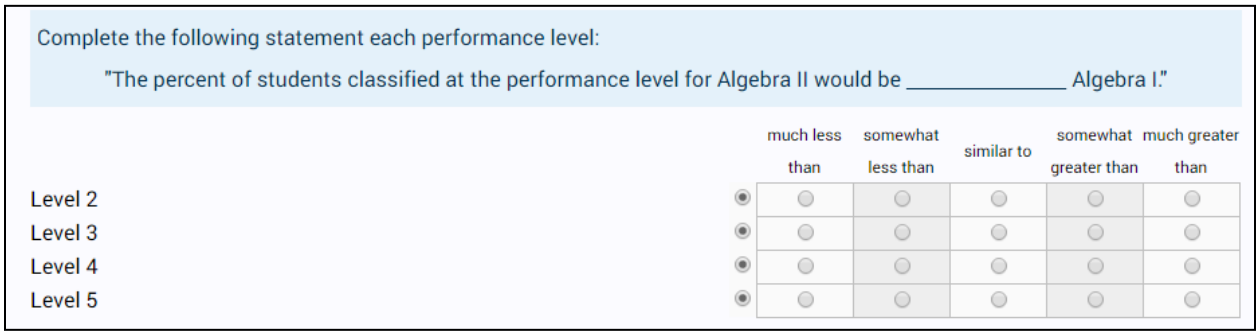



**Figure 6.3:** Sample Moodle Survey Question for PLD Review Activity

In the next step of the vertical articulation process, panelists were shown cross-grade impact data charts reflecting the Round 3 results of all grade levels completed that week as well as any higher grade level impact information, as appropriate. The group discussed how these results looked across grades relative to the expectations they had generated in the PLD review activity—noting areas where their expectations based on the PLDs differed from the actual impact data.

Lastly, panelists had the opportunity to recommend changes to the Round 3 threshold scores if they felt that the pattern of impact data across grade levels was inconsistent for reasons that they could not attribute to the test content, the PLDs, or the students. These changes were made directly at the threshold score level and did not involve item level judgments. The range of individual panelists' threshold scores from Round 3 as well as the committee medians from Rounds 1 and 2 were used as a guide when evaluating how much change would be reasonable to make. Panelists were very cognizant of honoring the work their committees had done and were very judicious in making changes such that the number and magnitude of changes was limited to only those changes necessary to support articulation across grade levels. Facilitators used an interactive spreadsheet to show panelists how changes to the threshold scores would change the pattern of impact data across grades and performance levels.

**Process Evaluation Survey**

At the end of the PLS meeting, panelists were asked to complete a process evaluation survey in the Moodle site. The purpose of the survey was to collect information about each panelist's experience in recommending threshold scores for the PARCC assessments. The survey asked participants to provide feedback on the following:

1. The level of success of the various components of the meeting.
2. The usefulness of the activities conducted during the meeting.
3. The adequacy of the various components of the meeting.
4. The adequacy of opportunities to ask questions, etc. the meeting.
5. How confident participants were that the threshold scores accurately reflected student performance at each performance level.
6. The level of support the participants had for the threshold scores for each performance level.
7. Whether committee members thought that their judgments and opinions were treated with respect by facilitators and fellow panelists.

All participants were also allowed to provide any additional information concerning their evaluation of the process of the PLS meeting through an open response question.

Although there was some variation, overall, the PLS process was positively perceived by the panelists. Most committee members thought that the various components of the meeting were "Successful" or "Very Successful" and that the activities conducted during the meeting were either "Useful" or "Very Useful." Additionally, a large majority of panelists indicated that they were confident or very confident in the threshold scores and said that they either moderately supported or strongly supported the threshold scores. The full results of the process evaluation surveys are in Appendix 8.

### Recommended PARCC Threshold Scores

During each item judgment round, the panelists provided a judgment for each item representing the number of points that they expected a borderline student at each performance level would likely receive, if they were administered the item. The sum of the item judgments across all items on both the PBA and EOY components, including weighting, if applicable, was calculated for each performance level. The calculated sum for the performance level represented the recommended threshold score the panelist expected an examinee would need to meet or exceed to be classified into that performance level.

During the PLS process, it was expected that there would be variation between panelist threshold score recommendations for each performance level. To determine a single threshold score recommendation for a performance level for a committee, the threshold score recommendations for a performance level were averaged across panelists. Specifically, the median score from a set of panelists' threshold score recommendations was used to determine the recommended threshold score for the group.

The recommended threshold score recommendations from the three item judgment rounds and the vertical articulation, represented as raw scores, are shown in Appendix 9. The summary statistics for the threshold scores for each performance level for Rounds 1, 2, and 3 from each PLS meeting are shown in Appendix 10. The panelist agreement data for performance levels 2, 3, 4 and 5 for Rounds 1 and 2 from each PLS meeting are shown in Appendix 11. The estimated impact data after round 3 and vertical articulation for each PLS meeting are shown in Appendix 12. Table 6.8 shows the threshold scores as

raw score and percentage of maximum total raw score for each performance level after vertical articulation for each subject and course.

**Table 6.8:** Threshold Scores for Each Performance Level after Vertical Articulation

| | Level 2 | | Level 3 | | Level 4 | | Level 5 | |
|---|---|---|---|---|---|---|---|---|
| **Subject** | **Raw** | **%** | **Raw** | **%** | **Raw** | **%** | **Raw** | **%** |
| Grade 3 ELA/L | 17 | 17 | 40 | 40 | 67 | 67 | 86 | 86 |
| Grade 4 ELA/L | 19 | 18 | 48 | 45 | 74 | 70 | 93 | 88 |
| Grade 5 ELA/L | 15 | 14 | 44 | 42 | 77 | 73 | 98 | 92 |
| Grade 6 ELA/L | 25 | 18 | 66 | 48 | 98 | 72 | 123 | 90 |
| Grade 7 ELA/L | 25 | 18 | 61 | 45 | 89 | 65 | 115 | 84 |
| Grade 8 ELA/L | 27 | 20 | 63 | 46 | 96 | 70 | 115 | 84 |
| Grade 9 ELA/L | 20 | 15 | 53 | 39 | 86 | 64 | 114 | 84 |
| Grade 10 ELA/L | 29 | 21 | 62 | 45 | 92 | 67 | 111 | 81 |
| Grade 11 ELA/L | 19 | 14 | 51 | 37 | 85 | 62 | 112 | 82 |
| Grade 3 Math | 15 | 18 | 33 | 40 | 56 | 68 | 71 | 87 |
| Grade 4 Math | 15 | 18 | 37 | 45 | 63 | 77 | 77 | 94 |
| Grade 5 Math | 12 | 15 | 34 | 41 | 58 | 71 | 68 | 83 |
| Grade 6 Math | 11 | 13 | 37 | 45 | 63 | 77 | 74 | 90 |
| Grade 7 Math | 8 | 10 | 26 | 32 | 54 | 66 | 68 | 83 |
| Grade 8 Math | 11 | 13 | 26 | 32 | 55 | 67 | 73 | 89 |
| Algebra I | 9 | 9 | 26 | 27 | 53 | 55 | 75 | 77 |
| Geometry | 9 | 9 | 28 | 29 | 56 | 58 | 76 | 79 |
| Algebra II | 12 | 12 | 31 | 31 | 62 | 61 | 82 | 81 |
| Integrated Math I | 12 | 13 | 30 | 32 | 57 | 61 | 72 | 77 |
| Integrated Math II | 9 | 9 | 27 | 28 | 47 | 48 | 67 | 69 |
| Integrated Math III | 12 | 11 | 30 | 28 | 62 | 58 | 85 | 79 |

# Chapter 7: Review and Approval of Performance Levels

This chapter details the process used for reviewing performance standards for reasonableness. The sections in this chapter include the following:

- Purpose and Process for Reasonableness Review
- Data for Reasonableness Review
- Reasonableness Review Results
- Final Approval of Recommended Performance Levels

## Purpose and Process for Reasonableness Review

After educator committees recommended threshold scores for the performance levels for the PARCC assessments, PARCC conducted a reasonableness review of the recommended threshold score recommendations across content areas and made adjustments as appropriate. The reasonableness review process following the PLS meetings was intended to ensure that performance standards contribute to a cohesive assessment system. The reasonableness review of the recommended threshold scores from the high school PLS meetings and the grades 3-8 PLS meetings were held separately. The reasonableness review of the threshold scores from the high school PLS meeting was completed before the grades 3-8 PLS meetings, so the grades 3-8 PLS meetings could refer to the results from the high school PLS meetings. Participants in the reasonableness review process were members of the PARCC Governing Board and, for high school, members of the Advisory Committee on College Readiness (ACCR) from each PARCC state. A full list of participants in the reasonableness review process is provided in Appendix 13. The reasonableness review process involved two components—a set of webinars in which PARCC Governing Board and ACCR members could review the results of the PLS meetings, and an in-person meeting in which adjustments to the threshold scores were considered.

The webinars were conducted the week prior to the in-person meetings and included a review of the following:

- Overview of the Evidence Based Standard Setting Process
- Summary of the PLS meeting process
- Summary of panelist information (participation by state and role)
- The pre-policy reasonable ranges and key data points from the PARCC Benchmarking study
- The impact data associated with the recommended threshold scores following vertical articulation
- The median threshold scores from all 3 judgment rounds as well as the recommended threshold scores following vertical articulation

During the in-person review meetings, PARCC Governing Board and ACCR members additionally considered the following:

- Summary of process evaluation surveys from PLS meetings

- The recommended threshold scores from vertical articulation
- The standard error of judgment from Round 3 judgments
- Updated impact data with the threshold scores from vertical articulation applied across all forms of the online PARCC tests

For the in-person meetings, participants first reviewed results from the ELA/L tests and considered any changes to the threshold scores from those tests before repeating the process with the mathematics tests. For high school, participants jointly reviewed results of the traditional and integrated math tests. Participants were given a brief training on the standard error of judgment (SEJs) including what it represented in terms of a measure of variability of PLS panelist judgments, how it was calculated, and how it should be used in considering changes to the recommended threshold scores. Participants were instructed that changes within 2 SEJs were considered to be consistent with the recommendations of the PLS panels.

In the interest of time, a set of options was prepared in advance of the in-person meetings for making adjustments using the SEJ information in key areas of interest based on discussions during the webinars. However, the SEJ information was available for all threshold scores and participants were able to propose adjustments to any of the set of threshold scores as they felt appropriate. An interactive spreadsheet (very similar to the one used during vertical articulation of the PLS meetings) was used during the reasonableness review to show participants how changes in the threshold scores would change the percent of students in each performance level. Adjustments to the threshold scores were made using a three-step process.

- In the first step, areas of interest (e.g. Level 3 thresholds for high school mathematics; Level 5 threshold for grade 9 ELA/L, etc.) for making adjustments were identified.
- In the second step, each area of interest was evaluated one at a time.
    - What is the current impact data based on the recommended threshold scores?
    - What is the current percent of points associated with each threshold score on the PLS reference form?
    - What would these look like if we moved the threshold by 1 SEJ (up or down)? By 2 SEJ?
- In the third step, the impact to the other areas of interest as well to the adjacent performance levels was evaluated.

### Data for Reasonableness Review

Following the conclusion of the PLS meetings, the threshold scores on the PLS reference forms from vertical articulation were converted to “theta” values representing the underlying score on the Item Response Theory metric. These theta values were then identified on all online forms of the test and student performance classified into one of the five performance levels based upon these values. The

percent of students in each performance level was then aggregated across the set of online[6] PARCC test forms. Table 7.1 provides the sample size for each grade and subject of the data used in reasonableness review.

**Table 7.1:** Number of Students Represented in the Reasonableness Review Impact Data

| Mathematics | | ELA/L | |
|---|---|---|---|
| **Grade** | **Sample Size** | **Grade** | **Sample Size** |
| 3 | 447812 | 3 | 379596 |
| 4 | 472514 | 4 | 479513 |
| 5 | 489139 | 5 | 496405 |
| 6 | 510819 | 6 | 512159 |
| 7 | 501762 | 7 | 513712 |
| 8 | 410072 | 8 | 504867 |
| Algebra I | 401521 | 9 | 342461 |
| Geometry | 183221 | 10 | 244302 |
| Algebra II | 172609 | 11 | 160746 |
| Integrated I | 26751 | | |
| Integrated II | 11373 | | |
| Integrated III | 7281 | | |

Students were excluded from the reasonableness review data if they did not meet the PARCC attemptedness criterion (at least one item attempted on each of the two components). Additionally, students who had irregularities such as system or interface issues identified by the states were also excluded. Finally, for students with duplicate records, the highest score was used. Impact data shown during the reasonableness review meetings included the same breakouts as for the PLS meetings (overall, by gender, by ethnicity, and by special population). In addition, a breakout of overall student performance by PARCC state was included—though the names of individual states were masked during the meeting proceedings. Individual states were able to obtain information on student performance within their own states prior to the in-person meeting.

The Standard Error of Judgment (SEJ) was based on Round 3 judgments for each PLS meeting. For each threshold score, the SEJ was calculated as the standard deviation of panelist individual threshold scores in Round 3 divided by the square root of the number of panelists. The SEJ takes into account the variability of judgments across panelists and accounts for the number of panelists providing judgments. SEJs are presented in terms of the number of raw score points on the PLS reference form. Table 7.2 shows the SEJs for each threshold score for the PARCC high school and 3-8 tests.

[6] Evaluation of mode comparability (comparison of online to paper test performance) was still ongoing at the time of the reasonableness review meetings so student performance on the paper forms of the test could not be included in the reasonableness review datasets.

**Table 7.2:** Round 3 SEJ Values for the PARCC High School Tests

| | Level 2 | Level 3 | Level4 | Level 5 |
|---|---|---|---|---|
| Grade 3 ELA/L | 0.6 | 1.0 | 1.4 | 1.8 |
| Grade 4 ELA/L | 0.8 | 1.2 | 1.1 | 1.0 |
| Grade 5 ELA/L | 0.7 | 0.9 | 0.9 | 1.0 |
| Grade 6 ELA/L | 1.9 | 2.3 | 2.1 | 1.1 |
| Grade 7 ELA/L | 1.5 | 2.1 | 1.1 | 1.7 |
| Grade 8 ELA/L | 1.2 | 1.7 | 1.7 | 1.5 |
| Grade 9 ELA/L | 1.6 | 2.3 | 2.0 | 1.9 |
| Grade 10 ELA/L | 1.7 | 1.9 | 1.7 | 1.9 |
| Grade 11 ELA/L | 1.4 | 1.8 | 1.4 | 2.0 |
| Grade 3 Math | 0.9 | 1.6 | 0.9 | 0.7 |
| Grade 4 Math | 1.5 | 1.8 | 1.0 | 0.6 |
| Grade 5 Math | 1.4 | 1.5 | 1.6 | 1.2 |
| Grade 6 Math | 0.7 | 2.1 | 1.3 | 1.2 |
| Grade 7 Math | 0.8 | 1.3 | 1.4 | 0.9 |
| Grade 8 Math | 1.1 | 1.9 | 1.9 | 1.1 |
| Algebra I | 0.6 | 2.1 | 1.9 | 1.3 |
| Geometry | 0.7 | 1.0 | 1.3 | 1.1 |
| Algebra II | 1.1 | 2.6 | 2.9 | 2.1 |
| Integrated Math I | 1.5 | 2.4 | 1.8 | 1.1 |
| Integrated Math II | 0.4 | 0.8 | 0.8 | 0.7 |
| Integrated Math III | 0.6 | 1.7 | 1.9 | 2.0 |

As the PLS meeting process was occurring, psychometric data analysis using Item Response Theory (IRT) was also ongoing. This process identified a small sub-set of items which were included on the PLS reference forms reviewed by the committees, but which were ultimately omitted from scoring. In preparation for the reasonableness review meetings, these items were omitted from the total possible number of points on the PLS reference forms and committee threshold recommendations (in terms of raw score points on the PLS reference form) were adjusted proportionally. In some cases this adjustment resulted in a shift of the raw score value associated with a threshold score on the PLS reference form; however, the underlying “theta” value on the IRT scale associated with each threshold score was maintained. Table 7.3 shows the original and updated number of points for impacted test forms as well as any changes to the threshold scores which occurred as a result of this adjustment. If a content/grade is not listed, then there was no change to the test form.

**Table 7.3** Changes to the Total Points and Threshold Scores Prior to Reasonableness Review

| Subject/Course | Total # of Points PLS | Total # of Points RR | Impact to Threshold Scores | | | |
|---|---|---|---|---|---|---|
| | | | Level 2 | Level 3 | Level 4 | Level 5 |
| Grade 7 ELA/L | 137 | 135 | 25 to 24 | 61 to 60 | 89 to 88 | 115 to 114 |
| Integrated Math I | 93 | 92 | N/C | N/C | N/C | N/C |
| Integrated Math II | 97 | 95 | N/C | N/C | N/C | N/C |
| Integrated Math III | 107 | 100 | N/C | N/C | 62 to 61 | 85 to 84 |

Note: NC means that there was no change to the threshold score for the associated performance level

## Reasonableness Review Results

Table 7.4 shows the changes that were made to the performance level threshold scores for each content area based on the reasonableness review. In total, 11 out of 84 threshold scores (13%) were adjusted.

**Table 7.4:** Reasonableness Review Recommendations

| Subject | Recommended Changes to Performance Levels | | | |
|---|---|---|---|---|
| | Level 2 | Level 3 | Level 4 | Level 5 |
| Grade 3 ELA/L | N/C | N/C | N/C | N/C |
| Grade 4 ELA/L | N/C | 48 to 50 | 74 to 76 | N/C |
| Grade 5 ELA/L | N/C | N/C | N/C | N/C |
| Grade 6 ELA/L | N/C | 66 to 64 | N/C | N/C |
| Grade 7 ELA/L | N/C | N/C | N/C | N/C |
| Grade 8 ELA/L | N/C | N/C | N/C | N/C |
| Grade 9 ELA/L | N/C | N/C | N/C | N/C |
| Grade 10 ELA/L | N/C | N/C | N/C | N/C |
| Grade 11 ELA/L | N/C | N/C | N/C | N/C |
| Grade 3 Math | N/C | N/C | N/C | N/C |
| Grade 4 Math | N/C | N/C | N/C | N/C |
| Grade 5 Math | N/C | N/C | N/C | N/C |
| Grade 6 Math | N/C | N/C | N/C | N/C |
| Grade 7 Math | N/C | N/C | N/C | N/C |
| Grade 8 Math | N/C | N/C | N/C | N/C |
| Algebra I | N/C | 31 to 28 | N/C | N/C |
| Geometry | N/C | 28 to 27 | N/C | N/C |
| Algebra II | N/C | 26 to 22 | N/C | N/C |
| Integrated Math I | 12 to 11 | 30 to 28 | N/C | N/C |
| Integrated Math II | 9 to 11 | 27 to 26 | N/C | N/C |
| Integrated Math III | N/C | 30 to 28 | N/C | N/C |

Note: N/C means that there was no change to the threshold score for the associated performance level

The impact data based on the threshold scores after the reasonableness review can be found in Appendix 12. Table 7.5 shows the threshold scores after the reasonableness review as raw score and percentage of the maximum total raw score for each performance level for each subject and course.

**Table 7.5:** Threshold Scores for Each Performance Level after Reasonableness Review

| | Level 2 | | Level 3 | | Level 4 | | Level 5 | |
|---|---|---|---|---|---|---|---|---|
| **Subject** | **Raw** | **%** | **Raw** | **%** | **Raw** | **%** | **Raw** | **%** |
| Grade 3 ELA/L | 17 | 17 | 40 | 40 | 67 | 67 | 86 | 86 |
| Grade 4 ELA/L | 19 | 18 | 50 | 47 | 76 | 72 | 93 | 88 |
| Grade 5 ELA/L | 15 | 14 | 44 | 42 | 77 | 73 | 98 | 92 |
| Grade 6 ELA/L | 25 | 18 | 64 | 47 | 98 | 72 | 123 | 90 |
| Grade 7 ELA/L | 24 | 18 | 60 | 44 | 88 | 65 | 114 | 84 |
| Grade 8 ELA/L | 27 | 20 | 63 | 46 | 96 | 70 | 115 | 84 |
| Grade 9 ELA/L | 20 | 15 | 53 | 39 | 86 | 64 | 114 | 84 |
| Grade 10 ELA/L | 29 | 21 | 62 | 45 | 92 | 67 | 111 | 81 |
| Grade 11 ELA/L | 19 | 14 | 51 | 37 | 85 | 62 | 112 | 82 |
| Grade 3 Math | 15 | 18 | 33 | 40 | 56 | 68 | 71 | 87 |
| Grade 4 Math | 15 | 18 | 37 | 45 | 63 | 77 | 77 | 94 |
| Grade 5 Math | 12 | 15 | 34 | 41 | 58 | 71 | 68 | 83 |
| Grade 6 Math | 11 | 13 | 37 | 45 | 63 | 77 | 74 | 90 |
| Grade 7 Math | 8 | 10 | 26 | 32 | 54 | 66 | 68 | 83 |
| Grade 8 Math | 11 | 13 | 26 | 32 | 55 | 67 | 73 | 89 |
| Algebra I | 9 | 9 | 22 | 23 | 53 | 55 | 75 | 77 |
| Geometry | 9 | 9 | 27 | 28 | 56 | 58 | 76 | 79 |
| Algebra II | 12 | 12 | 28 | 28 | 62 | 61 | 82 | 81 |
| Integrated Math I | 11 | 12 | 28 | 30 | 57 | 62 | 72 | 78 |
| Integrated Math II | 11 | 12 | 26 | 27 | 47 | 49 | 67 | 71 |
| Integrated Math III | 12 | 12 | 28 | 28 | 61 | 61 | 84 | 84 |

## Final Approval of Performance Levels

PARCC's original goal in developing five performance levels and a College- and Career-Ready Determination policy had been to better evaluate student proficiency and provide better information to inform instruction and student support. The rationale for five performance levels (rather than four or three) was to allow for three levels *below* the CCRD level in order to better differentiate among levels and provide instructional supports to students working towards college- and career- readiness. Although the change to include the top three performance levels (rather than the top two levels) in the definition of college- and- career readiness was supported by research and evidence from the dry-run PLS meetings, it resulted in only two levels below the CCRD level. Members of the PARCC Governing Board and ACCR were concerned that this did not provide sufficient or useful differentiation for families, teachers, and schools in the range of the tests where most students were performing. Additionally, the very low percentage of students in performance Level 5 called into question the utility of maintaining this as a separate performance level.

As a result, the PARCC Governing Board and ACCR voted to shift the PARCC performance levels to better meet the intended inferences about student performance and to provide better differentiation to inform feedback to families, teachers, and schools. The overarching goal in making this shift was to be true to the data and the work of the PLS committees, but to provide more useful information about student performance and growth. Figure 7.1 shows how the performance level shifts were made. Holding the college- and career- ready (or on track) expectations constant, the following changes were made:

- performance Levels 4 and 5 were collapsed to create a new "performance Level 5,"
- performance Level 3 was renamed to "performance Level 4," and
- performance Level 2 was divided to create two new performance levels (new "performance Levels 2 and 3")[7].

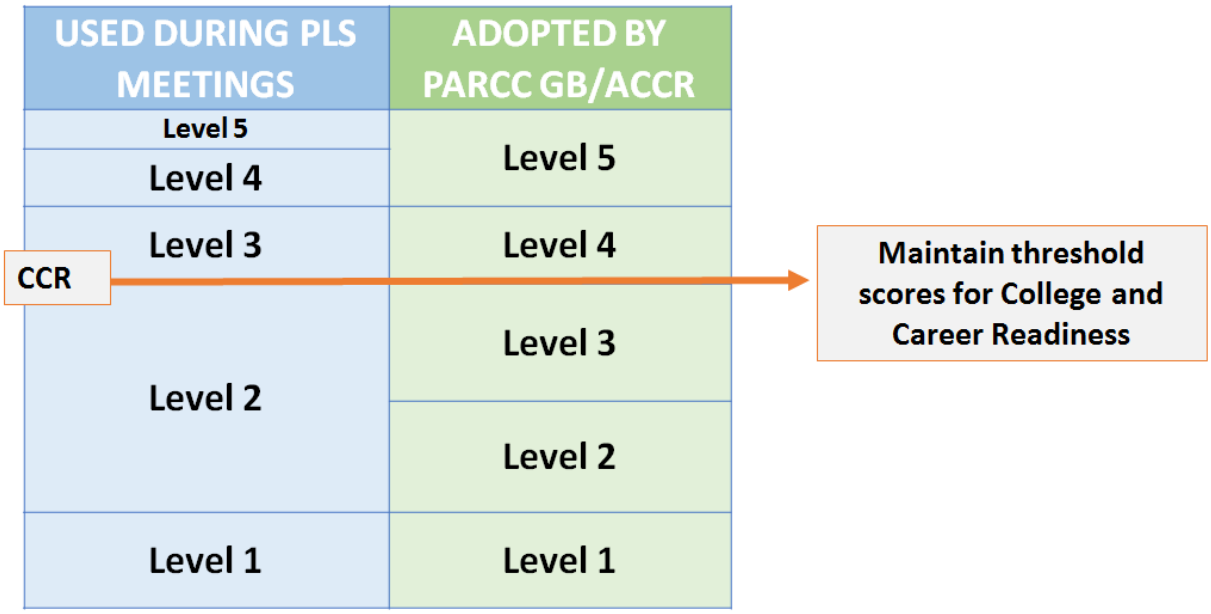



**Figure 7.1:** Illustration of Performance Level Shifts

This resulted in a final system of performance levels with three below and two above the college- and career- ready (or on track) expectation as per the original intentions of the PARCC assessment program. Discussion of how the PLDs were revised to reflect these changes is provided in Chapter 2. The impact data based on the final threshold scores can be found in Appendix 12. Table 7.6 shows the final threshold scores as raw score and percentage of the maximum total raw score for each performance level for each subject and course.

[7] Performance Level 2 was divided in half empirically using scale score values. The scale score mid-point of the original Level 2 was used as the threshold score for the new Level 3.

**Table 7.6:** Final Threshold Scores for Each Performance Level

| | Level 2 | | Level 3 | | Level 4 | | Level 5 | |
|---|---|---|---|---|---|---|---|---|
| **Subject** | **Raw** | **%** | **Raw** | **%** | **Raw** | **%** | **Raw** | **%** |
| Grade 3 ELA/L | 17 | 17 | 28 | 28 | 40 | 40 | 67 | 67 |
| Grade 4 ELA/L | 19 | 18 | 33 | 31 | 50 | 47 | 76 | 72 |
| Grade 5 ELA/L | 15 | 14 | 28 | 26 | 44 | 42 | 77 | 73 |
| Grade 6 ELA/L | 25 | 18 | 43 | 31 | 64 | 47 | 98 | 72 |
| Grade 7 ELA/L | 24 | 18 | 40 | 30 | 60 | 44 | 88 | 65 |
| Grade 8 ELA/L | 27 | 20 | 43 | 31 | 63 | 46 | 96 | 70 |
| Grade 9 ELA/L | 20 | 15 | 34 | 25 | 53 | 39 | 86 | 64 |
| Grade 10 ELA/L | 29 | 21 | 44 | 32 | 62 | 45 | 92 | 67 |
| Grade 11 ELA/L | 19 | 14 | 33 | 24 | 51 | 37 | 85 | 62 |
| Grade 3 Math | 15 | 18 | 24 | 29 | 33 | 40 | 56 | 68 |
| Grade 4 Math | 15 | 18 | 26 | 32 | 37 | 45 | 63 | 77 |
| Grade 5 Math | 12 | 15 | 21 | 26 | 34 | 41 | 58 | 71 |
| Grade 6 Math | 11 | 13 | 22 | 27 | 37 | 45 | 63 | 77 |
| Grade 7 Math | 8 | 10 | 15 | 18 | 26 | 32 | 54 | 66 |
| Grade 8 Math | 11 | 13 | 18 | 22 | 26 | 32 | 55 | 67 |
| Algebra I | 9 | 9 | 15 | 15 | 22 | 23 | 53 | 55 |
| Geometry | 9 | 9 | 16 | 17 | 27 | 28 | 56 | 58 |
| Algebra II | 12 | 12 | 19 | 19 | 28 | 28 | 62 | 61 |
| Integrated Math I | 11 | 12 | 18 | 20 | 28 | 30 | 57 | 62 |
| Integrated Math II | 11 | 12 | 17 | 18 | 26 | 27 | 47 | 49 |
| Integrated Math III | 12 | 12 | 19 | 19 | 28 | 28 | 61 | 61 |

## Appendix 1: Pre-Policy Committee Members

A complete list of all members of the pre-policy committee, including their names and positions at the time of the pre-policy meeting is provided below. Additionally, a summary of the information about the members of the pre-policy committee is provided.

| Name | Position | State |
|---|---|---|
| Mike Hernandez | Deputy Commissioner Arkansas | AR |
| Ian Macgillivary | Director of Academic Affairs | CO |
| Will Morton | Director of Assessment Administration, Colorado Department of Education | CO |
| Jeffrey Noel | Assistant superintendent data, accountability, and research | DC |
| Naomi Watson | Deputy Chief Data and Strategy (DCPS) | DC |
| Rashida Young | Senior Manager, Equity and Fidelity Team (DC public charter school board) | DC |
| Dan Cullen | Deputy Director for Academic Affairs | IL |
| Susie Morrison | Deputy Supt/Chief Education Officer, Illinios | IL |
| Susan Lane | Sr. Assistant Commissioner P-16 Alignment MA Deputy of Higher Education | MA |
| Mitchell D. Chester | Commissioner of Ed | MA |
| Dr. Mary Kay Finan | MD State Board of Education Member | MD |
| Jack R. Smith | Deputy Supt CAO Maryland State Depatment of Education | MD |
| Richard Baliko | NAEP State Coordinator/PARCC Program Manager | MS |
| Walt Drane | State Assessment Director | MS |
| Bari Erlichson | Asst Commisioner NJDOE | NJ |
| Leighann Lenti | Deputy Secretary NMPED | NM |
| Justlyn Overby | Asst Director of Assessment for NMPED | NM |
| Rick (Patrick) Scott | NM PARCC Higher Education Lead | NM |
| Sarah Wickham | Senior Policy Advisor, Ohio Depatment of Ed | OH |
| Jim Wright | Dircetor Office of Curriculum and Assessment ODE | OH |
| Andrea Castaneda | Chief, Finance & Operation @ RI Dept of Ed | RI |
| Katherine Sipala | Superintendent of Schools (RI); President of School Supt's Association | RI |

## Appendix 2: Pre-Policy Committee Agenda

***Pre-Policy Meeting: Wednesday, April 8, 2015***

**8:00-8:30 – Breakfast**

8:30-8:45 – Welcome and Meeting Purpose (PARCC, Inc. staff / State lead)

8:45-9:00 – Introductions

9:00-10:00 – Overview of PARCC assessment system (State leads)

**10:00-10:15 – Break**

10:15 -11:00 – Overview of standard setting process (Pearson)

11:00-11:45 – Presentation of PARCC Benchmarking Study (Pearson)

**11:45-12:30 – Lunch**

12:30-1:15 – Presentation of Educator Judgment Study (Pearson)

1:15-1:25 – Introduction to policy judgment activity (Pearson)

1:25-1:50 – Table discussion: What percentage of students would you expect to be college- and career-ready?

1:50-2:15 – Whole group discussion: What percentage of students would you expect to be college- and career-ready?

2:15-2:40 – Individual judgments for the minimum and maximum percentage of students expected to be college- and career-ready in English language arts and mathematics

**2:40-3:00 – Break**

3:00-4:30 – Process for standards adoption (PARCC, Inc. staff / State lead)

4:30-4:45 – Feedback from judgment activity (Pearson)

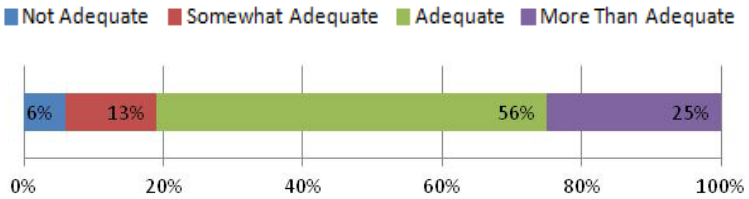
4:45-5:00 – Closing remarks (PARCC, Inc. staff / State lead)

## Appendix 3: Pre-Policy Committee Process Evaluation Summary

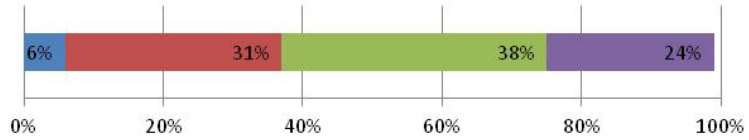
1. Check the column below that best reflects your opinion about the level of success of the various components of the meeting in which you have just participated. The activities were designed to help you both understand the process and be supportive of the recommendations made by the committee.

**2.** How adequate were the following elements of the meeting?

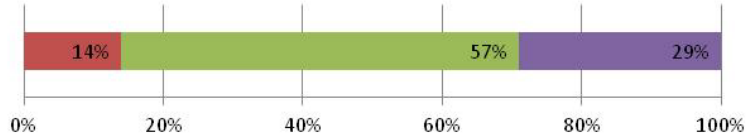
a. Amount of time spent reviewing studies




b. Amount of time spent in table discussions

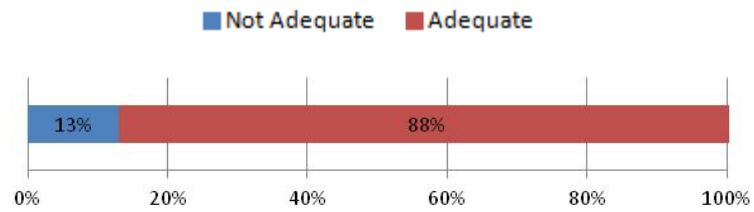



c. Amount of time spent in individual judgments

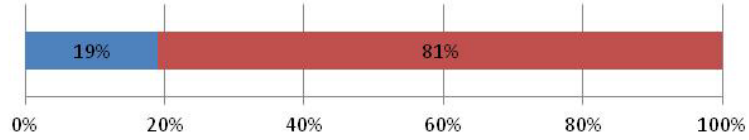



**3.** Did you have adequate opportunities during the session to express your professional opinions about the following elements?

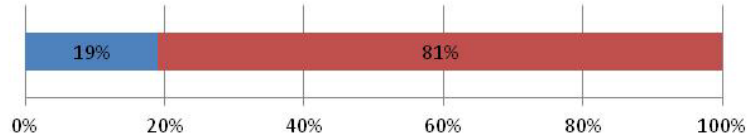
a. Benchmark Study




b. Postsecondary Educators' Judgment Study

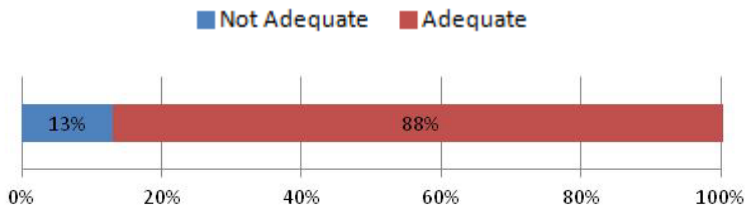



c. Reasonable percentages of students college- and career-ready

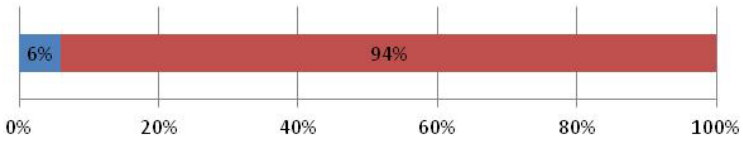


**4.** Did you have adequate opportunities during the session regarding the following elements?

a. Ask questions about the studies




b. Interact with your fellow committee members



# Appendix 4: Performance Level Setting Committee Recruitment

**PARCC PLS Committee Recruitment Memo**

## I. Overview

In summer 2015, PARCC will be conducting three performance level setting events to develop recommended threshold scores for each performance level for its 21 summative assessments. Twelve grade-span panels will meet for five days each according to the following schedule:

| Grade/Course Panel | Meeting Dates |
|---|---|
| ELA/literacy grades 9-11<br>Mathematics high school assessments | July 27 – 31, 2015 |
| ELA/literacy & Mathematics grades 7-8 | August 17 – 21, 2015 |
| ELA/literacy & Mathematics grades 3-6 | August 24 – 28, 2015 |

PARCC will also conduct a dry run performance level setting for Algebra II/Integrated Mathematics III and ELA/Literacy grade 11 from **May 4-7, 2015**. Each content area will include a panel of 10 members. The purpose of the dry run is to examine and evaluate the performance level setting process to be conducted during the summer meetings, and to produce recommendations for improvement. Tiger Team members have suggested that panelists identified to participate in the dry run be different than those who participate in the summer meetings in order to maximize opportunities to involve educators from the states.

## II. Panel Composition

For the summer 2015 meetings, PARCC will recruit 12 performance level setting panels with 20 members each. The panels will be composed of K-12 educators, postsecondary faculty, members of the business community, and community members with appropriate content expertise for their assigned grade span and subject area. The tables below list the assigned grade levels for each panel, and proposed composition of each panel. The demographic composition of each panel should reflect the diversity of PARCC states, and include educators with experience with special populations, including English learners and students with disabilities.

**PARCC Performance Level Setting Panels**

| Mathematics | ELA/ Literacy |
|---|---|
| **1.** Grades 3-4 | **7.** Grades 3-4 |
| **2.** Grades 5-6 | **8.** Grades 5-6 |
| **3.** Grades 7-8 | **9.** Grades 7-8 |
| **4.** Algebra I / Integrated I | **10.** Grade 9 |
| **5.** Geometry / Integrated II | **11.** Grade 10 |

| | |
|---|---|
| **6.** Algebra II / Integrated III | **12.** Grade 11 |

| Panel(s) | Composition |
|---|---|
| **Grades 3-4**<br>**Grades 5-6**<br>**Grades 7-8** | • 18-19 K-12 educators (classroom teachers, school/district level curriculum leads, administrators)<br>• 1-2 other stakeholders (e.g., school of education faculty, business, community) |
| **Grade 9 ELA**<br>**Grade 10 ELA**<br>**Alg. I/Int. I**<br>**Geo./Int. II** | • 14 K-12 educators (classroom teachers, school/district level curriculum leads, administrators)<br>• 4-5 postsecondary faculty (representing 2-year, 4-year, CTE)<br>• 1-2 other stakeholders (e.g., business, community) |
| **Grade 11 ELA**<br>**Alg. II/Int. III** | • 8 K-12 educators (classroom teachers, school/district level curriculum leads, administrators)<br>• 2 high school technical/vocational educators<br>• 6 postsecondary faculty (representing 2-year, 4-year, CTE)<br>• 3-4 other stakeholders (e.g., business, community) |

## III. Benefits for Panelists

The PARCC performance level setting meetings are a unique opportunity for educators to engage professionally with their peers from across the PARCC states. Participants will leave the meetings with a better understanding of how assessment threshold scores are determined, and of the expectations for student performance on the PARCC assessments.

Participants who are not being paid by their employers during the performance level setting meetings will receive a stipend of $150 per day. Classroom educators who receive their salary over a 12-month period will receive a stipend if school is not in session. Reimbursement of $100 per day is available for school districts or universities that will be paying for substitute coverage during the meetings.

## IV. Suggested Panelists

Panelists selected for the performance level setting events should have extensive content knowledge in their subject area and be familiar with the Common Core State Standards. For the middle and high school panels, it is preferable that participants hold a degree in their content area. States may consider naming:

- Current or former general education teachers
- Postsecondary faculty responsible for teaching first-year courses

- State agency content experts
- District-level curriculum directors or subject-area coaches
- Teachers of the year

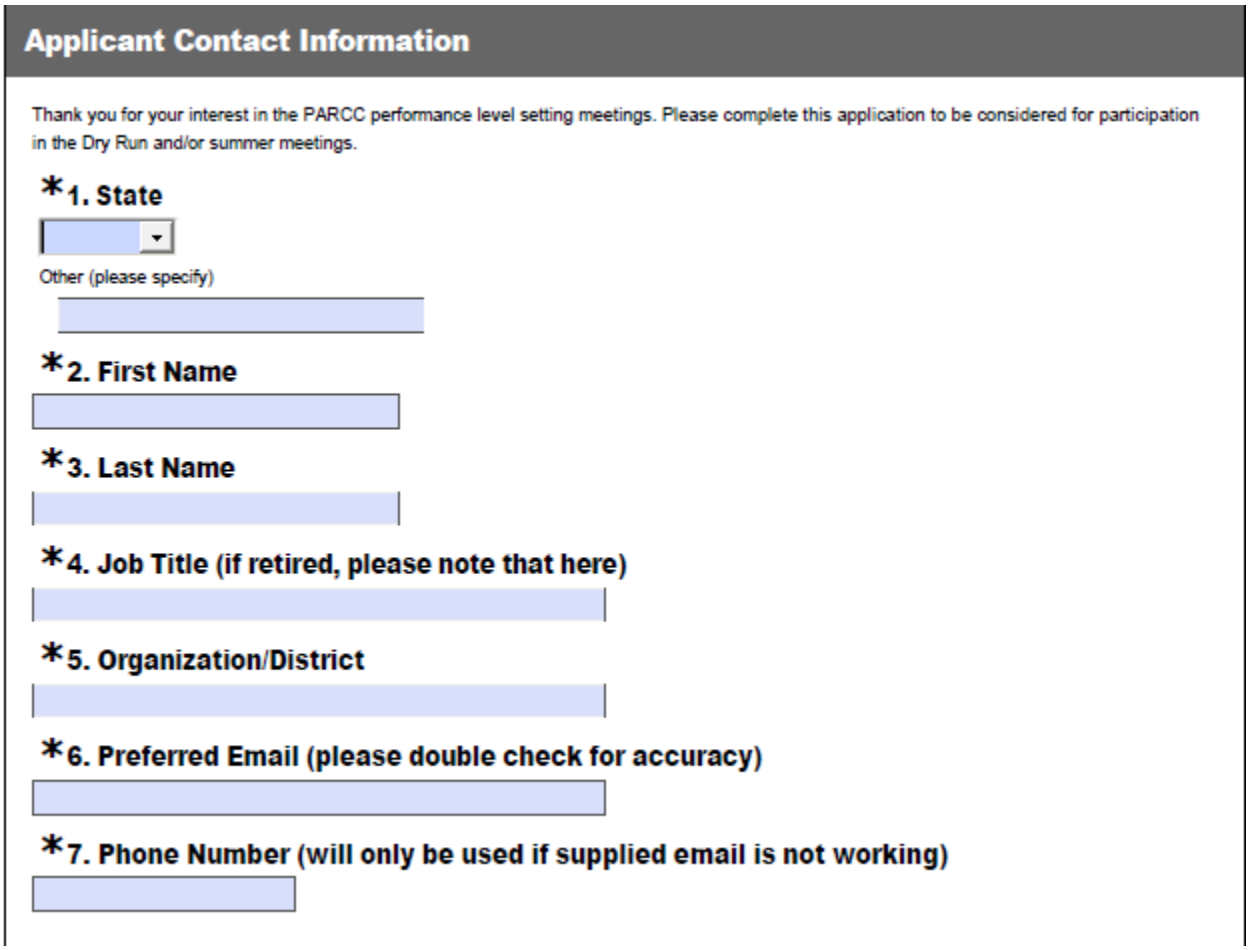
## V. Recruitment Process

States will identify panel nominees and back-ups, and distribute recruitment information to these individuals. Each state will have approximately 20 panelists across the grade levels and subject areas. States may choose to distribute materials to pre-selected nominees, or to a wider distribution of potential participants. State Leads should keep in mind that one goal of the performance level setting panels is to have balanced representation on each of the 12 panels, so recruitment should target educators from diverse backgrounds at varying grade levels.

The following documents have been developed to help in communicating with potential participants:

- **Recruitment Cover Letter**
  This overview document can be sent to nominees as a high-level explanation of the performance level setting opportunity. It directs recipients to the FAQ or a state contact for additional information. The "nominated panelist" version is meant for states that are targeting specific nominees. The "general audience" version is meant for states that may be sending out information via a wider distribution list.
- **FAQ for Potential Panelists**
  This document provides detailed information to nominees about the performance level setting process, meeting dates, and how to apply.
- **Application Survey (**https://www.surveymonkey.com/r/PARCC_PLS_320**)**
  Applicants are asked to describe their experience in education, including their familiarity with PARCC and the Common Core State Standards, and to identify their grade span and content area assignment preferences. The Performance Level Setting Tiger Team and Parcc Inc. will review the pool of nominees and suggest final assignments, pending state approval. The application will close at **5:00 PM ET** on **March 13**.

**PARCC PLS Panel Application**



## Applicant Contact Information

Thank you for your interest in the PARCC performance level setting meetings. Please complete this application to be considered for participation in the Dry Run and/or summer meetings.

***1. State**

Other (please specify)

***2. First Name**

***3. Last Name**

***4. Job Title (if retired, please note that here)**

***5. Organization/District**

***6. Preferred Email (please double check for accuracy)**

***7. Phone Number (will only be used if supplied email is not working)**

## Assignment Preferences

***8. Which week(s) are you interested in and available to participate? Please check all that apply.**

- May 4 - May 7 Dry Run (ELA/L grade 11 and Algebra II/Integrated III only)
- July 27 - July 31 Performance Level Setting (All high school assessments)
- August 17 - August 21 Performance Level Setting (Grades 7-8)
- August 24 - August 28 Performance Level Setting (Grades 3-4 and 5-6)

Additional Comments

***9. Preferred Content Area**

- English Language Arts/Literacy
- Mathematics

***10. Preferred Grade Span**

- Grades 3-4
- Grades 5-6
- Grades 7-8
- ELA/L Grade 9
- ELA/L Grade 10
- ELA/L Grade 11
- Algebra I / Integrated Math I
- Geometry / Integrated Math II
- Algebra II / Integrated Math III

**11. Would you be willing and available to serve on two panels, from July 27-31 (Grade 9 or Algebra I/Integrated Math I) and August 17-21 (Grades 7-8)?**

- Yes
- No

**12. Second Choice Grade Span**

- Grades 3-4
- Grades 5-6
- Grades 7-8
- ELA/L Grade 9
- ELA/L Grade 10
- ELA/L Grade 11
- Algebra I / Integrated Math I
- Geometry / Integrated Math II
- Algebra II / Integrated Math III

**13. Second Choice Grade Span**

- Grades 3-4
- Grades 5-6
- Grades 7-8
- ELA/L Grade 9
- ELA/L Grade 10
- ELA/L Grade 11
- Algebra I / Integrated Math I
- Geometry / Integrated Math II
- Algebra II / Integrated Math III

## Applicant Experience

***14. Which category best describes your current role?**

- K-12 classroom educator
- K-12 administrator
- K-12 curriculum coordinator or subject-area coach
- 2-year or technical college faculty member
- 2-year or technical college administrator
- 4-year college or university faculty member
- 4-year college or university administrator
- State education agency employee
- Member of the business community
- Other

Other (please specify)

**15. How long have you served in your current capacity?**

- 0-1 years
- 1-4 years
- 4-7 years
- 7-10 years
- 10+ years

## PARCC Performance Level Setting Panel Application

**16. Which category or categories best describe former role(s) you have held?**

- K-12 classroom educator
- K-12 administrator
- K-12 curriculum coordinator or subject-area coach
- 2-year or technical college faculty member
- 2-year or technical college administrator
- 4-year college or university faculty member
- 4-year college or university administrator
- State education agency employee
- Member of the business community

Other (please specify)

**17. How long did you serve in those roles?**

***18. Please indicate your level of experience working with the following populations of students.**

| | No Experience | Somewhat Experienced | Experienced | Very Experienced |
|---|---|---|---|---|
| Students with disabilities | | | | |
| English learners | | | | |
| Gifted students | | | | |
| Alternative education | | | | |

Other (please specify the populations)

***19. Please indicate the geographic setting(s) in which you have experience.**

- Rural
- Suburban
- Urban

Other (please specify)

***20. Please rate your familiarity with the following aspects of PARCC.**

| | Not at all familiar | Somewhat familiar | Familiar | Very familiar |
|---|---|---|---|---|
| PARCC assessment design | | | | |
| PARCC performance level descriptors | | | | |
| College- and Career-Ready Determination Policy | | | | |
| PARCC sample items | | | | |
| PARCC practice tests | | | | |

**21. Briefly describe your experience with PARCC.**

***22. Please rate your familiarity with the Common Core State Standards.**

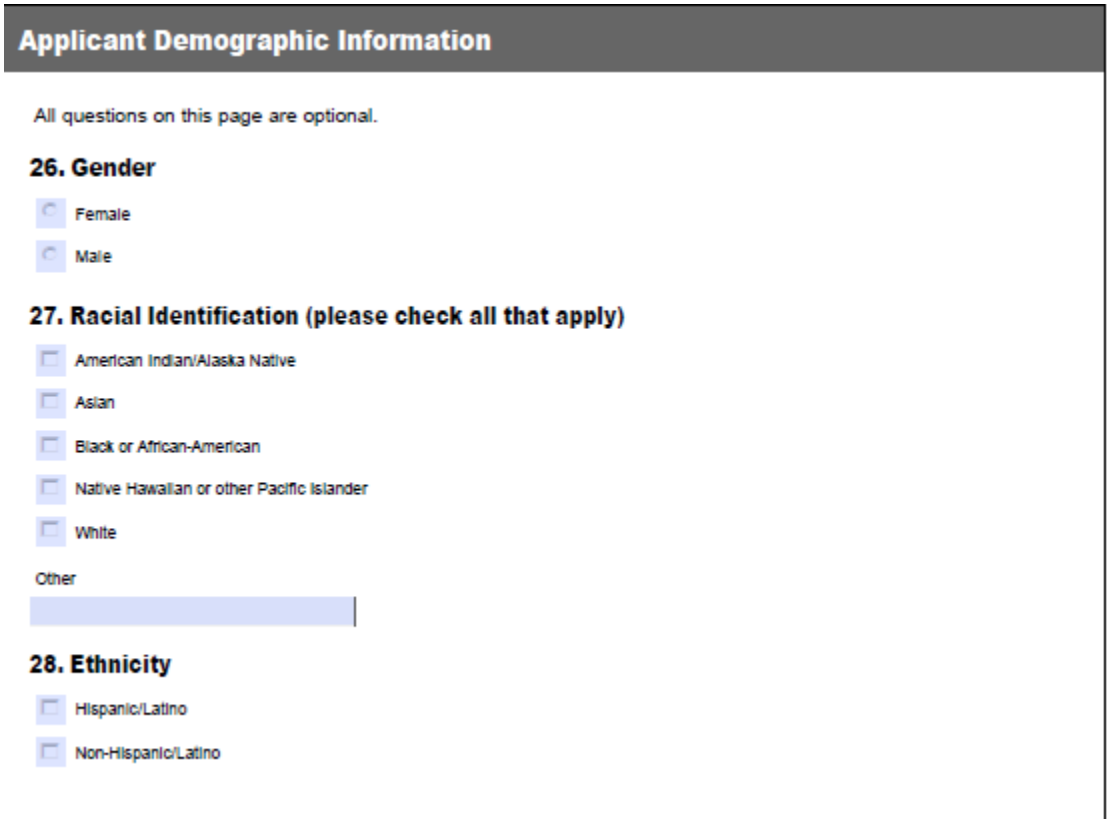
| Not at all familiar | Somewhat familiar | Familiar | Very familiar |
|---|---|---|---|
| | | | |

**23. Briefly describe your experience working with the Common Core State Standards.**

***24. Please indicate your level of experience with the following aspects of student assessment.**

| | No Experience | Somewhat Experienced | Experienced | Very Experienced |
|---|---|---|---|---|
| Developing student assessment materials | | | | |
| Scoring student assessment responses | | | | |
| Analyzing student assessment results | | | | |

**25. Additional Comments**



## Applicant Demographic Information

All questions on this page are optional.

**26. Gender**

- Female
- Male

**27. Racial Identification (please check all that apply)**

- American Indian/Alaska Native
- Asian
- Black or African-American
- Native Hawaiian or other Pacific Islander
- White

Other

**28. Ethnicity**

- Hispanic/Latino
- Non-Hispanic/Latino

# Appendix 5: Performance Level Setting Committee Composition

As part of the pre-work for the PLS meeting, each panelist was required to complete a Panelist Information Survey. The results of that self-reported survey were used to calculate the composition of each committee based on various demographic variables.

## High School Performance Level Setting Committees

In which state do you work?

| State | ELA/L | | | |
|---|---|---|---|---|
| | Grade 9 | Grade 10 | Grade 11 | Total |
| **AR** | 0 | 1 | 2 | **3** |
| **CO** | 2 | 4 | 2 | **8** |
| **DC** | 1 | 1 | 1 | **3** |
| **IL** | 1 | 1 | 1 | **3** |
| **MA** | 2 | 2 | 2 | **6** |
| **MD** | 2 | 3 | 3 | **8** |
| **MS** | 1 | 1 | 1 | **3** |
| **NJ** | 2 | 1 | 1 | **4** |
| **NM** | 2 | 2 | 2 | **6** |
| **OH** | 2 | 0 | 1 | **3** |
| **RI** | 2 | 1 | 1 | **4** |
| **National** | 0 | 0 | 0 | **0** |
| ***Total*** | ***17*** | ***17*** | ***17*** | ***51*** |

| State | Math | | | |
|---|---|---|---|---|
| | Algebra I /Math I | Geometry /Math II | Algebra II /Math III | Total |
| **AR** | 1 | 1 | 1 | **3** |
| **CO** | 2 | 1 | 3 | **6** |
| **DC** | 2 | 1 | 2 | **5** |
| **IL** | 1 | 2 | 2 | **5** |
| **MA** | 1 | 2 | 2 | **5** |
| **MD** | 2 | 1 | 2 | **5** |
| **MS** | 0 | 0 | 0 | **0** |
| **NJ** | 1 | 1 | 2 | **4** |
| **NM** | 2 | 2 | 1 | **5** |
| **OH** | 1 | 3 | 3 | **7** |
| **RI** | 2 | 1 | 2 | **5** |
| **National** | 2 | 2 | 0 | **4** |
| ***Total*** | ***17*** | ***17*** | ***20*** | ***54*** |

What is your current position?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | Grade 9 | Grade 10 | Grade 11 | Total |
| **Teacher (K-12)** | 12 | 9 | 6 | **27** |
| **Teacher (Higher Education)** | 0 | 4 | 4 | **8** |
| **Administrator** | 2 | 3 | 3 | **8** |
| **Other** | 3 | 1 | 4 | **8** |
| ***Total*** | ***17*** | ***17*** | ***17*** | ***51*** |

| Position | Math | | | |
|---|---|---|---|---|
| | Algebra I /Math I | Geometry /Math II | Algebra II /Math III | Total |
| **Teacher (K-12)** | 8 | 8 | 10 | **26** |
| **Teacher (Higher Education)** | 3 | 2 | 5 | **10** |
| **Administrator** | 2 | 2 | 2 | **6** |
| **Other** | 4 | 5 | 3 | **12** |
| ***Total*** | ***17*** | ***17*** | ***20*** | ***54*** |

Do you currently teach the course?

| | ELA | | |
|---|---|---|---|
| | Grade 9 | Grade 10 | Grade 11 |
| **Yes** | 6 | 5 | 3 |
| **No** | 11 | 12 | 14 |

| | Math | | | | | |
|---|---|---|---|---|---|---|
| | Algebra I /Math I | | Geometry /Math II | | Algebra II /Math III | |
| | Algebra I | Integrated Math I | Geometry | Integrated Math II | Algebra II | Integrated Math III |
| **Yes** | 8 | 0 | 8 | 0 | 6 | 0 |
| **No** | 9 | 17 | 9 | 17 | 14 | 20 |

Note: There were several panelists who indicated that they had prior experience with teaching integrated math courses, although they did not currently teach an integrated math course.

For which population do you have educational experience?

| | ELA | | |
|---|---|---|---|
| | Grade 9 | Grade 10 | Grade 11 |
| **K-12 Education** | 17 | 15 | 16 |
| **Higher Education** | 3 | 6 | 9 |

| | Math | | |
|---|---|---|---|
| | Algebra I /Math I | Geometry /Math II | Algebra II /Math III |
| **K-12 Education** | 15 | 15 | 20 |
| **Higher Education** | 5 | 2 | 9 |

For which position(s) within education do you have experience?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | Grade 9 | Grade 10 | Grade 11 | Total |
| **Teacher/Instructor/Professor** | 16 | 17 | 17 | **50** |
| **Department Head or Administrator** | 4 | 7 | 11 | **22** |
| **School or District Administrator** | 2 | 4 | 4 | **10** |
| **Educational Organization** | 2 | 4 | 6 | **12** |

| Position | Math | | | |
|---|---|---|---|---|
| | Algebra I /Math I | Geometry /Math II | Algebra II /Math III | Total |
| **Teacher/Instructor/Professor** | 17 | 17 | 20 | **54** |
| **Department Head or Administrator** | 11 | 10 | 15 | **36** |
| **School or District Administrator** | 3 | 4 | 3 | **10** |
| **Educational Organization** | 3 | 5 | 3 | **11** |

For which of the following populations do you have educational experience with?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | Grade 9 | Grade 10 | Grade 11 | Total |
| **Students receiving special education services** | 17 | 15 | 13 | **45** |
| **Students of low socioeconomic status** | 17 | 16 | 16 | **49** |
| **Students who are English language learners** | 12 | 16 | 13 | **41** |

| | | | | |
|---|---|---|---|---|
| **Students who are receiving general education instruction** | 17 | 16 | 15 | **48** |

| Position | Math | | | |
|---|---|---|---|---|
| | **Algebra I /Math I** | **Geometry /Math II** | **Algebra II /Math III** | **Total** |
| **Students receiving special education services** | 16 | 15 | 15 | **46** |
| **Students of low socioeconomic status** | 14 | 15 | 16 | **45** |
| **Students who are English language learners** | 9 | 13 | 12 | **34** |
| **Students who are receiving general education instruction** | 16 | 17 | 20 | **53** |

How many years of professional experience in education do you have?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | **Grade 9** | **Grade 10** | **Grade 11** | **Total** |
| **1-5 years** | 1 | 0 | 1 | **2** |
| **6-10 years** | 2 | 5 | 3 | **10** |
| **11-15 years** | 4 | 6 | 4 | **14** |
| **16-20 years** | 3 | 4 | 3 | **10** |
| **More than 20 years** | 7 | 2 | 6 | **15** |
| ***Total*** | ***17*** | ***17*** | ***17*** | ***51*** |

| Position | Math | | | |
|---|---|---|---|---|
| | **Algebra I /Math I** | **Geometry /Math II** | **Algebra II /Math III** | **Total** |
| **1-5 years** | 0 | 1 | 1 | **2** |
| **6-10 years** | 2 | 2 | 0 | **4** |
| **11-15 years** | 5 | 4 | 8 | **17** |
| **16-20 years** | 7 | 3 | 2 | **12** |
| **More than 20 years** | 3 | 7 | 9 | **19** |
| ***Total*** | ***17*** | ***17*** | ***20*** | ***54*** |

What is the highest degree you have completed?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | **Grade 9** | **Grade 10** | **Grade 11** | **Total** |

| | | | | |
|---|---|---|---|---|
| **Bachelors degree** | 6 | 1 | 2 | **9** |
| **Masters degree** | 10 | 14 | 10 | **34** |
| **Doctoral degree** | 1 | 2 | 5 | **8** |
| ***Total*** | ***17*** | ***17*** | ***17*** | ***51*** |

| Position | Math | | | |
|---|---|---|---|---|
| | **Algebra I /Math I** | **Geometry /Math II** | **Algebra II /Math III** | **Total** |
| **Bachelors degree** | 2 | 2 | 3 | **6** |
| **Masters degree** | 10 | 14 | 12 | **36** |
| **Doctoral degree** | 5 | 1 | 5 | **11** |
| ***Total*** | ***17*** | ***17*** | ***20*** | ***54*** |

What is your gender?

| Position | ELA/L | | |
|---|---|---|---|
| | **Grade 9** | **Grade 10** | **Grade 11** |
| **Male** | 6 | 3 | 3 |
| **Female** | 11 | 12 | 13 |
| ***Total*** | ***17*** | ***15*** | ***16*** |

| Position | Math | | |
|---|---|---|---|
| | **Algebra I /Math I** | **Geometry /Math II** | **Algebra II /Math III** |
| **Male** | 6 | 5 | 6 |
| **Female** | 11 | 11 | 11 |
| ***Total*** | ***17*** | ***16*** | ***17*** |

What is your ethnicity?

| Position | ELA/L | | |
|---|---|---|---|
| | **Grade 9** | **Grade 10** | **Grade 11** |
| **Hispanic or Latino** | 0 | 1 | 1 |
| **Not Hispanic or Latino** | 15 | 10 | 15 |
| ***Total*** | ***15*** | ***11*** | ***16*** |

| Position | Math | | |
|---|---|---|---|
| | **Algebra I /Math I** | **Geometry /Math II** | **Algebra II /Math III** |
| **Hispanic or Latino** | 0 | 0 | 1 |

| | | | |
|---|---|---|---|
| **Not Hispanic or Latino** | 14 | 16 | 15 |
| ***Total*** | ***14*** | ***16*** | ***16*** |

What is your race?

| Position | ELA/L | | |
|---|---|---|---|
| | Grade 9 | Grade 10 | Grade 11 |
| **Asian** | 0 | 0 | 0 |
| **Black or African American** | 1 | 2 | 3 |
| **White** | 14 | 11 | 13 |
| ***Total*** | ***15*** | ***13*** | ***16*** |

| Position | Math | | |
|---|---|---|---|
| | Algebra I /Math I | Geometry /Math II | Algebra II /Math III |
| **Asian** | 2 | 0 | 0 |
| **Black or African American** | 4 | 1 | 3 |
| **White** | 9 | 15 | 16 |
| ***Total*** | ***15*** | ***16*** | ***19*** |

Do you currently work in a school district?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | Grade 9 | Grade 10 | Grade 11 | Total |
| **Yes** | 15 | 12 | 13 | **40** |
| **No** | 2 | 5 | 4 | **11** |
| ***Total*** | ***17*** | ***17*** | ***17*** | ***51*** |

| Position | Math | | | |
|---|---|---|---|---|
| | Algebra I /Math I | Geometry /Math II | Algebra II /Math III | Total |
| **Yes** | 12 | 14 | 15 | **41** |
| **No** | 5 | 3 | 5 | **13** |
| ***Total*** | ***17*** | ***17*** | ***20*** | ***54*** |

Which word best describes the size of the district where you work?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | Grade 9 | Grade 10 | Grade 11 | Total |
| **Small** | 6 | 5 | 4 | 15 |
| **Medium** | 5 | 4 | 6 | 15 |
| **Large** | 4 | 3 | 3 | 10 |
| ***Total*** | ***15*** | ***12*** | ***13*** | ***40*** |

| Position | Math | | | |
|---|---|---|---|---|
| | Algebra I /Math I | Geometry /Math II | Algebra II /Math III | Total |
| **Small** | 6 | 6 | 1 | 13 |
| **Medium** | 3 | 5 | 7 | 15 |
| **Large** | 3 | 3 | 7 | 13 |
| ***Total*** | ***12*** | ***14*** | ***15*** | ***41*** |

Which word best describes the type of district where you work?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | Grade 9 | Grade 10 | Grade 11 | Total |
| **Rural** | 6 | 5 | 5 | 16 |
| **Metropolitan** | 3 | 4 | 2 | 9 |
| **Suburban** | 6 | 3 | 6 | 15 |
| ***Total*** | ***15*** | ***12*** | ***13*** | ***40*** |

| Position | Math | | | |
|---|---|---|---|---|
| | Algebra I /Math I | Geometry /Math II | Algebra II /Math III | Total |
| **Rural** | 7 | 5 | 3 | 15 |
| **Metropolitan** | 3 | 3 | 4 | 10 |
| **Suburban** | 2 | 6 | 8 | 16 |
| ***Total*** | ***12*** | ***14*** | ***15*** | ***41*** |

Which word best describes the socioeconomic status of the district where you work?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | Grade 9 | Grade 10 | Grade 11 | Total |
| **Low** | 8 | 9 | 8 | 25 |
| **Moderate** | 6 | 3 | 4 | 13 |
| **High** | 1 | 0 | 1 | 2 |
| ***Total*** | ***15*** | ***12*** | ***13*** | ***40*** |

| Position | Math | | | |
|---|---|---|---|---|
| | Algebra I /Math I | Geometry /Math II | Algebra II /Math III | Total |
| **Low** | 6 | 7 | 7 | 20 |
| **Moderate** | 4 | 5 | 8 | 17 |
| **High** | 2 | 2 | 0 | 4 |
| ***Total*** | ***12*** | ***14*** | ***15*** | ***41*** |

## Grades 3-8 Performance Level Setting Committees

In which state do you work?

| State | ELA/L | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **AR** | 1 | 0 | 0 | **1** |
| **CO** | 2 | 2 | 3 | **7** |
| **DC** | 1 | 3 | 2 | **6** |
| **IL** | 1 | 2 | 2 | **5** |
| **MA** | 2 | 2 | 2 | **6** |
| **MD** | 2 | 1 | 2 | **5** |
| **MS** | 1 | 1 | 2 | **4** |
| **NJ** | 2 | 0 | 2 | **4** |
| **NM** | 1 | 2 | 2 | **5** |
| **OH** | 0 | 1 | 0 | **1** |
| **RI** | 2 | 2 | 1 | **5** |
| **National** | 0 | 0 | 0 | **0** |
| ***Total*** | ***15*** | ***16*** | ***18*** | ***49*** |

| State | Math | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **AR** | 0 | 1 | 2 | **3** |
| **CO** | 2 | 3 | 1 | **6** |
| **DC** | 1 | 2 | 2 | **5** |
| **IL** | 2 | 2 | 2 | **6** |
| **MA** | 3 | 1 | 2 | **6** |
| **MD** | 1 | 1 | 3 | **5** |
| **MS** | 2 | 0 | 2 | **4** |
| **NJ** | 2 | 2 | 2 | **6** |
| **NM** | 1 | 2 | 1 | **4** |
| **OH** | 1 | 0 | 1 | **2** |
| **RI** | 2 | 2 | 2 | **6** |
| **National** | 0 | 0 | 0 | **0** |
| ***Total*** | ***17*** | ***16*** | ***20*** | ***53*** |

What is your current position?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **Teacher (K-12)** | 6 | 8 | 7 | **21** |
| **Teacher (Higher Education)** | 1 | 0 | 0 | **1** |
| **Administrator** | 5 | 3 | 3 | **11** |
| **Other** | 3 | 5 | 8 | **16** |
| ***Total*** | ***15*** | ***16*** | ***18*** | ***49*** |

| Position | Math | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **Teacher (K-12)** | 11 | 9 | 12 | **32** |
| **Teacher (Higher Education)** | 0 | 0 | 2 | **2** |
| **Administrator** | 3 | 3 | 1 | **7** |
| **Other** | 3 | 4 | 5 | **12** |
| ***Total*** | ***17*** | ***16*** | ***20*** | ***53*** |

Do you currently teach the course?

| | ELA | | | | | |
|---|---|---|---|---|---|---|
| | Grades 3/4 | | Grades 5/6 | | Grades 7/8 | |
| | Grade 3 | Grade 4 | Grade 5 | Grade 6 | Grade 7 | Grade 8 |
| **Yes** | 2 | 3 | 6 | 5 | 3 | 5 |
| **No** | 13 | 12 | 10 | 11 | 15 | 13 |

| | Math | | | | | |
|---|---|---|---|---|---|---|
| | Grades 3/4 | | Grades 5/6 | | Grades 7/8 | |
| | Grade 3 | Grade 4 | Grade 5 | Grade 6 | Grade 7 | Grade 8 |
| **Yes** | 6 | 6 | 4 | 2 | 4 | 6 |
| **No** | 11 | 11 | 12 | 14 | 16 | 14 |

For which population do you have educational experience?

| | ELA | | |
|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 |
| **K-12 Education** | 15 | 16 | 18 |
| **Higher Education** | 4 | 3 | 4 |

| | Math | | |
|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 |
| **K-12 Education** | 17 | 16 | 20 |
| **Higher Education** | 1 | 3 | 3 |

For which position(s) within education do you have experience?

| | ELA/L | | | |
|---|---|---|---|---|
| **Position** | **Grades 3/4** | **Grades 5/6** | **Grades 7/8** | **Total** |
| **Teacher/Instructor/Professor** | 15 | 16 | 18 | **49** |
| **Department Head or Administrator** | 7 | 8 | 13 | **28** |
| **School or District Administrator** | 5 | 4 | 7 | **16** |
| **Educational Organization** | 2 | 2 | 5 | **9** |

| | Math | | | |
|---|---|---|---|---|
| **Position** | **Grades 3/4** | **Grades 5/6** | **Grades 7/8** | **Total** |
| **Teacher/Instructor/Professor** | 17 | 16 | 20 | **53** |
| **Department Head or Administrator** | 5 | 6 | 9 | **20** |
| **School or District Administrator** | 3 | 4 | 3 | **10** |
| **Educational Organization** | 2 | 3 | 1 | **6** |

For which of the following populations do you have educational experience with?

| | ELA/L | | | |
|---|---|---|---|---|
| **Position** | **Grades 3/4** | **Grades 5/6** | **Grades 7/8** | **Total** |
| **Students receiving special education services** | 14 | 14 | 15 | **43** |
| **Students of low socioeconomic status** | 14 | 16 | 18 | **48** |
| **Students who are English language learners** | 13 | 12 | 13 | **38** |
| **Students who are receiving general education instruction** | 15 | 16 | 17 | **48** |

| Position | Math | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **Students receiving special education services** | 15 | 16 | 19 | **50** |
| **Students of low socioeconomic status** | 16 | 14 | 19 | **49** |
| **Students who are English language learners** | 11 | 9 | 12 | **32** |
| **Students who are receiving general education instruction** | 17 | 15 | 19 | **51** |

How many years of professional experience in education do you have?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **1-5 years** | 1 | 1 | 0 | **2** |
| **6-10 years** | 4 | 3 | 1 | **8** |
| **11-15 years** | 3 | 5 | 3 | **11** |
| **16-20 years** | 3 | 2 | 3 | **8** |
| **More than 20 years** | 4 | 5 | 11 | **20** |
| ***Total*** | ***15*** | ***16*** | ***18*** | ***49*** |

| Position | Math | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **1-5 years** | 3 | 1 | 1 | **5** |
| **6-10 years** | 4 | 2 | 3 | **9** |
| **11-15 years** | 1 | 6 | 4 | **11** |
| **16-20 years** | 3 | 3 | 7 | **13** |
| **More than 20 years** | 6 | 4 | 5 | **15** |
| ***Total*** | ***17*** | ***16*** | ***20*** | ***53*** |

What is the highest degree you have completed?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **Bachelors degree** | 0 | 3 | 1 | **4** |
| **Masters degree** | 13 | 12 | 15 | **40** |
| **Doctoral degree** | 2 | 1 | 2 | **5** |
| ***Total*** | ***15*** | ***16*** | ***18*** | ***49*** |

| Position | Math | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **Bachelors degree** | 4 | 3 | 4 | **11** |
| **Masters degree** | 12 | 13 | 15 | **40** |
| **Doctoral degree** | 1 | 0 | 1 | **2** |
| ***Total*** | ***17*** | ***16*** | ***20*** | ***53*** |

What is your gender?

| Position | ELA/L | | |
|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 |
| **Male** | 5 | 2 | 2 |
| **Female** | 10 | 13 | 16 |
| ***Total*** | ***15*** | ***15*** | ***18*** |

| Position | Math | | |
|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 |
| **Male** | 2 | 3 | 3 |
| **Female** | 15 | 13 | 16 |
| ***Total*** | ***17*** | ***16*** | ***19*** |

What is your ethnicity?

| Position | ELA/L | | |
|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 |
| **Hispanic or Latino** | 0 | 2 | 3 |
| **Not Hispanic or Latino** | 13 | 13 | 14 |
| ***Total*** | ***13*** | ***15*** | ***17*** |

| Position | Math | | |
|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 |
| **Hispanic or Latino** | 1 | 2 | 0 |
| **Not Hispanic or Latino** | 16 | 13 | 15 |
| ***Total*** | ***17*** | ***15*** | ***15*** |

What is your race?

| Position | ELA/L | | |
|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 |
| **Asian** | 0 | 0 | 0 |
| **Black or African American** | 1 | 1 | 4 |
| **White** | 14 | 14 | 13 |
| ***Total*** | ***15*** | ***15*** | ***17*** |

| Position | Math | | |
|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 |
| **Asian** | 0 | 0 | 0 |
| **Black or African American** | 1 | 2 | 4 |
| **White** | 16 | 14 | 15 |
| ***Total*** | ***17*** | ***16*** | ***19*** |

Do you currently work in a school district?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **Yes** | 14 | 16 | 16 | **46** |
| **No** | 1 | 0 | 2 | **3** |
| ***Total*** | ***15*** | ***16*** | ***18*** | ***49*** |

| Position | Math | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **Yes** | 15 | 16 | 16 | **47** |
| **No** | 2 | 0 | 4 | **6** |
| ***Total*** | ***17*** | ***16*** | ***20*** | ***53*** |

Which word best describes the size of the district where you work?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **Small** | 2 | 3 | 7 | **12** |
| **Medium** | 7 | 7 | 5 | **19** |
| **Large** | 5 | 6 | 4 | **15** |
| ***Total*** | ***14*** | ***16*** | ***16*** | ***46*** |

| Position | Math | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **Small** | 4 | 7 | 6 | **17** |
| **Medium** | 8 | 7 | 5 | **20** |
| **Large** | 3 | 2 | 5 | **10** |
| ***Total*** | ***15*** | ***16*** | ***16*** | ***47*** |

Which word best describes the type of district where you work?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **Rural** | 4 | 1 | 3 | **8** |
| **Metropolitan** | 5 | 6 | 6 | **17** |
| **Suburban** | 5 | 9 | 7 | **21** |
| ***Total*** | ***14*** | ***16*** | ***16*** | ***46*** |

| Position | Math | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **Rural** | 5 | 5 | 5 | **15** |
| **Metropolitan** | 3 | 3 | 4 | **10** |
| **Suburban** | 7 | 8 | 7 | **22** |
| ***Total*** | ***15*** | ***16*** | ***16*** | ***47*** |

Which word best describes the socioeconomic status of the district where you work?

| Position | ELA/L | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **Low** | 9 | 8 | 10 | **27** |
| **Moderate** | 4 | 8 | 5 | **17** |
| **High** | 1 | 0 | 1 | **2** |
| ***Total*** | ***14*** | ***16*** | ***16*** | ***46*** |

| Position | Math | | | |
|---|---|---|---|---|
| | Grades 3/4 | Grades 5/6 | Grades 7/8 | Total |
| **Low** | 7 | 9 | 6 | **22** |
| **Moderate** | 8 | 7 | 8 | **23** |
| **High** | 0 | 0 | 2 | **2** |
| ***Total*** | ***15*** | ***16*** | ***16*** | ***47*** |

# Appendix 6: Performance Level Setting Meeting Agendas

## High School PLS Meeting Agenda - Mathematics

**PARCC Performance Level Setting– Detailed Agenda**
**High School Math Meetings**
**July 27-31, 2015**

**DAY 1**

*Breakfast* (8:00 – 8:30am)

**General Session**

- Welcome (8:30 – 9:00am)
  - Who is in the room? (Introductions of key meeting staff and facilitators)
  - Why are we here? (Meeting Purpose)
    - What are we trying to accomplish?
    - What will we be doing this week?
  - Overview of PARCC
- Overview and Purpose of Performance Level Setting (9:00 – 10:00am)
  - What is Performance Level Setting?
  - Basic Vocabulary
  - Evidence Based Standard Setting
  - Overview of the Item Level Judgment Task
  - Try it out! Pop-culture assessment example

*Break (10:00 – 10:15am)* – Move to Breakout Session Rooms

**Breakout Session**

- Breakout Session Introductions (10:15 – 10:45am)
  - Introductions (panelists, facilitators, table leaders)
  - Orientation to Materials
  - Binder Check-out
  - Login to Moodle (password resets if needed)
  - Security discussion
- Experience the Assessment (PBA) Test 1 (10:45 – 11:45am)
  - Testing times for PARCC
  - How PARCC items are developed
  - Overview of Math/ELA item types
  - Panelists take the PBA test form

*Lunch (11:45 – 12:30pm)*

- Scoring the PARCC Assessments (12:30 – 1:15pm)

  - Facilitators review scoring rules for PARCC Items
  - Orientation to Answer Key/Rubric Documents
- Scoring Your PBA Assessment (1:15 – 1:45pm)
  - Panelists score their responses to the PBA Assessment
- Experience the Assessment (EOY) Test 1 (1:45 – 2:45pm)
  - Panelists take the EOY test form

*Break (2:45 – 3:00pm)*

- Scoring Your EOY Assessment (3:00 – 3:45pm)
  - Facilitators review scoring rules for EOY Items
  - Panelists score their responses to the EOY Assessment
- Review and Discuss PLDs (included in pre-work) Test 1 (3:45 – 4:30pm)
  - Panelists discuss PLDs with their table groups noting key differences between performance levels
- College and Career Readiness Discussion (4:30 – 5:00pm)
  - Review of PARCC CCR Definitions and Policies
  - Facilitators discuss PARCC Benchmarking , PEJ studies, and pre-policy meeting
  - Facilitators share historical data points on college and career readiness

**DAY 2 (Breakout Session)**

*Breakfast* (8:00 – 8:30am)

- Borderline Descriptor Training (8:30 – 9:00am)
  - Typical vs. Borderline Students
  - Working with Draft Borderline Descriptors
- Borderline Student Table Discussion (9:00 – 9:45am)
  - A representative from each table is assigned a borderline level (2, 3, 4, or 5)
  - Borderline level groups meet first to discuss the draft descriptors for their assigned level.
    - How well do these describe the borderline student as we envision them?
- Editing Discussion at Original Tables (9:45 – 10:30am)
  - Table groups reconvene and discuss what they learned about each borderline student group
  - Then work as a table group to edit assigned descriptors

*Break* (10:30 – 10:45am)—Facilitators copy table descriptors into master document

- Group Discussion of Borderline Descriptors (10:45 – 11:15am)
  - Facilitator reviews compiled descriptors from each table with group
- Performance Level Setting Training (11:15 – 12:00pm)
  - Process for making item judgments
  - Definition of likely
  - Judgments that make sense and those that don't
  - Ceiling and floor judgments
  - Recording item judgments

*Lunch* (12:00– 12:45pm) – Print Borderline Descriptors

- Practice Judgment Task (12:45 – 1:30pm)
  - Panelists work independently to make judgments for practice set items
- Discuss Practice Task (1:30 – 2:00pm)
  - Group discussion of judgments, challenges, and points of confusion
  - Relationship between item judgments and threshold scores

*Break* (2:00 – 2:15pm)

- Round 1 Judgments PBA Test 1(2:15 – 3:45pm)
  - Round 1 Readiness Form
  - Panelists work independently to make judgments for PBA items
- Round 1 Judgments EOY Test 1 (3:45 – 5:15pm)
  - Panelists work independently to make judgments for EOY items

**DAY 3 (Breakout Session)**

*Breakfast* (8:00 – 8:30am)

- Round 1 feedback Test 1 (8:30 – 9:00am)
  - Item Level
    - Item means & score point distributions
    - Item level panelist agreement data
  - Test Level
    - Individual threshold scores recommendations
    - Table threshold scores recommendations
    - Group threshold score recommendations
    - Pre-policy ranges—Percent of points
- Table Discussion: Round 1 Feedback Test 1 (9:00 – 9:45am)
  - Panelists discuss item mean and score point distributions
  - Panelists discuss test level agreement
  - Panelists discuss items flagged for greatest disagreement
  - Panelists discuss additional items of interest
- Round 2 Judgments (PBA and EOY) Test 1 (9:45 – 11:45am)
  - Round 2 Readiness Form
  - Panelists revise item judgments based on feedback data and discussions

*Lunch (11:45 – 12:30pm)* – Data analyzed during this time

- Round 2 feedback (Panelist Agreement Data) Test 1 (12:30 –12:45pm)
  - Item Level
    - Item means & Score point distributions
    - Item level panelist agreement data
  - Test Level
    - Individual threshold scores recommendations

        - Table threshold scores recommendations
        - Group threshold score recommendations
        - Pre-policy ranges—Percent of points
- Round 2 feedback table discussions Test 1 (12:45 – 1:15pm)
    - Panelists discuss test level agreement
    - Panelists discuss items flagged for greatest disagreement
    - Panelists discuss additional items of interest
- Whole group discussion of agreement results Test 1 (1:15 – 1:30pm)
- Round 2 Feedback (Impact Data) Test 1 (1:30 – 1:45pm)
    - Overall
    - By student subgroup
    - Pre-policy ranges—Percent of students
- Whole group discussion of impact data results Test 1 (1:45 – 2:00pm)
- Round 3 Judgments Test 1 (2:00 – 3:00pm)
    - Round 3 Readiness Form
    - Panelists revise item judgments based on feedback data and discussions

*Break* (3:00 – 3:30pm) – Data analyzed during this time

- Round 3 Feedback Test 1 (3:30 – 3:45pm)
    - Group threshold score recommendations
    - Impact data (overall and for subgroups)
    - Pre-policy ranges—Percent of points
    - Pre-policy ranges—Percent of students
- Experience the Assessment (PBA) Test 2 (3:45 – 4:45pm)
    - Panelists take the PBA test form
- Scoring Your PBA Assessment Test 2 (4:45 – 5:15pm)
    - Panelists score their responses to the PBA Assessment

**DAY 4 (Breakout Session)**

*Breakfast* (8:00 – 8:30am)

- Experience the Assessment (EOY) Test 2 (8:30 – 9:30am)
    - Panelists take the EOY test form
- Scoring Your EOY Assessment Test 2 (9:30 – 10:00am)
    - Panelists score their responses to the EOY Assessment

*Break* (10:00 – 10:15am)

- Review and Discuss PLDs Test 2 (included in pre-work) (10:15 – 10:45am)
    - Panelists discuss PLDs with their table groups noting key differences between performance levels
- Discuss Traditional Math Borderline Descriptors (10:45 – 11:15am)
    - Review borderline descriptors from the other two traditional math tests
    - Compare to the borderline descriptors developed for test 1 (traditional math)
- Develop Integrated Math Borderline Descriptors (11:15 – 12:00pm)

- Table groups create borderline descriptors for test 2 (integrated math)for assigned sub-claim pulling from and modifying statements from the traditional math

*Lunch* (12:00– 12:45pm)—Facilitators copy table descriptors into master document

- Group Discussion of Borderline Descriptors (12:45 – 1:15pm)—Print Borderline Descriptors
  - Facilitator review compiled descriptors from each table with group
- Round 1 Judgments PBA Test 2 (1:15 – 2:15pm)
  - Round 1 Readiness Form
  - Panelists work independently to make judgments for PBA items
- Round 1 Judgments EOY Test 2 (2:15 – 3:15pm)
  - Panelists work independently to make judgments for EOY items

*Break* (3:15– 3:45pm) – Data analyzed during this time

- Round 1 feedback Test 2 (3:45 – 4:00pm)
  - Item Level
    - Item means & Score point distributions
    - Item level panelist agreement data
  - Test Level
    - Individual threshold scores recommendations
    - Table threshold scores recommendations
    - Group threshold score recommendations
    - Pre-policy ranges—Percent of points
- Table Discussion: Round 1 Feedback Test 2 (4:00 – 4:30pm)
  - Panelists discuss item mean and score point distributions
  - Panelists discuss test level agreement
  - Panelists discuss items flagged for greatest disagreement
  - Panelists discuss additional items of interest
- Round 2 Judgments (PBA and EOY) Test 2 (4:30 – 5:30pm)
  - Round 2 Readiness Form
  - Panelists revise item judgments based on feedback data and discussions

**DAY 5**

**Breakout Session**

*Breakfast* (8:00 – 8:30am)

- Round 2 feedback (Panelist Agreement Data) Test 2 (8:30 –8:45am)
  - Item Level
    - Item means & score point distributions
    - Item level panelist agreement data
  - Test Level
    - Individual threshold scores recommendations
    - Table threshold scores recommendations

    - Group threshold score recommendations
    - Pre-policy ranges—Percent of points
- Round 2 feedback table discussions Test 2 (8:45 – 9:30am)
    - Panelists discuss test level agreement
    - Panelists discuss items flagged for greatest disagreement
    - Panelists discuss additional items of interest
- Whole group discussion of agreement results Test 2 (9:30 – 9:45am)
- Round 2 Feedback (Impact Data) Test 2(9:45 – 10:00am)
    - Overall
    - By student subgroup
    - Pre-policy ranges—Percent of students
- Whole group discussion of impact data results Test 2 (10:00 – 10:15am)
- Round 3 Judgments Test 2 (10:15 – 11:15am)
    - Round 3 Readiness Form
    - Panelists revise item judgments based on feedback data and discussions

*Break* (11:15 –11:45am) – Data analyzed during this time

- Round 3 Feedback Test 2 (11:45 – 12:00pm)
    - Group threshold score recommendations
    - Impact data (overall and for subgroups)
    - Pre-policy ranges—Percent of points
    - Pre-policy ranges—Percent of students
- Next Steps and Close Out (12:00 – 12:15pm)
    - Secure materials check-in
    - Thank you for participating!

*Lunch* (12:15 – 1:00pm)

**Vertical Articulation** (Table Leaders from each panel reconvene)

- Introductions and session purpose (1:00 –1:15pm)
- Cross-grade PLD review (1:15 –2:15pm)
- Cross-grade impact data review (2:15 –2:45pm)
- Recommend changes to Round 3 results (2:45 –3:30pm)

## High School PLS Meeting Agenda - English Language Arts/Literacy (ELA/L)

**PARCC Performance Level Setting– Detailed Agenda**
**High School ELA Meetings**
**July 27-30, 2015**

**DAY 1**

*Breakfast* (8:00 – 8:30am)

**General Session**

- Welcome (8:30 – 9:00am)
  - Who is in the room? (Introductions of key meeting staff and facilitators)
  - Why are we here? (Meeting Purpose)
    - What are we trying to accomplish?
    - What will we be doing this week?
  - Overview of PARCC
- Overview and Purpose of Performance Level Setting (9:00 – 10:00am)
  - What is Performance Level Setting?
  - Basic Vocabulary
  - Evidence Based Standard Setting
  - Overview of the Item Level Judgment Task
  - Try it out! Pop-culture assessment example

*Break (10:00 – 10:15am)* – Move to Breakout Session Rooms

**Breakout Session**

- Breakout Session Introductions (10:15 – 10:45am)
  - Introductions (panelists, facilitators, table leaders)
  - Orientation to Materials
  - Binder Check-out
  - Login to Moodle (password resets if needed)
  - Security discussion
- Experience the Assessment (PBA) (10:45 – 12:15pm)
  - Testing times for PARCC
  - How PARCC items are developed
  - Overview of Math/ELA item types
  - Panelists take the PBA test form

*Lunch (12:15 – 1:00pm)*

- Scoring the PARCC Assessments (1:00 – 1:45pm)
  - Facilitators review scoring rules for PARCC Items
  - Orientation to Answer Key/Rubric Documents
- Scoring your PBA Assessment (1:45 – 2:15pm)
  - Panelists score their responses to the PBA Assessment

*Break (2:15 – 2:30pm)*

- Experience the Assessment (EOY) (2:30 – 4:00pm)
  - Panelists take the EOY test form
- Scoring your EOY Assessment (4:00 – 4:45pm)
  - Panelists score their responses to the EOY Assessment

**DAY 2 (Breakout Session)**

*Breakfast* (8:00 – 8:30am)

- Review and Discuss PLDs (included in pre-work) (8:30 – 9:15am)
  - Panelists discuss PLDs with their table groups noting key differences between performance levels (hardcopy in panelist folders and available through Moodle site)
- Borderline Descriptor Training (9:15 – 9:30am)
  - Typical vs. Borderline Students
  - Working with Draft Borderline Descriptors
- Borderline Student Table Discussion (9:30 – 10:30am)
  - A representative from each table is assigned a borderline level (2, 3, 4, or 5)
  - Borderline level groups meet first to discuss the draft descriptors for their assigned level.
    - How well do these describe the borderline student as we envision them?

*Break (10:30 – 10:45am)*

- Editing Discussion at Original Tables (10:45 – 11:45am)
  - Table groups reconvene and discuss what they learned about each borderline student group
  - Then work as a table group to edit assigned descriptors

*Lunch* (11:45 – 12:30pm)—Facilitators copy table descriptors into master document

- Group Discussion of Borderline Descriptors (12:30 – 1:15pm)
  - Facilitator reviews compiled descriptors from each table with group
- College and Career Readiness Discussion (1:15 – 1:45pm)
  - Review of PARCC CCR Definitions and Policies
  - Facilitators discuss PARCC Benchmarking , PEJ studies, and pre-policy meeting
  - Facilitators share historical data points on college and career readiness
- Performance Level Setting Training (1:45 – 2:30pm)
  - Process for making item judgments
  - Definition of likely
  - Judgments that make sense and those that don't
  - Ceiling and floor judgments
  - Recording item judgments

*Break* (2:30– 2:45pm)

- Practice Judgment Task (2:45 – 3:30pm)
  - Panelists work independently to make judgments for practice set items
- Discuss Practice Task (3:30 – 4:00pm)
  - Group discussion of judgments, challenges, and points of confusion
  - Relationship between item judgments and threshold scores

**DAY 3 (Breakout Session)**

*Breakfast* (8:00 – 8:30am)

- Round 1 Judgments PBA (8:30 – 10:00am)
  - Round 1 Readiness Form
  - Panelists work independently to make judgments for PBA items
- Round 1 Judgments EOY(10:00 – 11:30am)
  - Panelists work independently to make judgments for EOY items

*Lunch* (11:30 – 12:15am) – Data analyzed during this time

- Round 1 feedback (12:15 – 12:45pm)
  - Item Level
    - Item means & score point distributions
    - Item level panelist agreement data
  - Test Level
    - Individual threshold scores recommendations
    - Table threshold scores recommendations
    - Group threshold score recommendations
    - Pre-Policy reasonable ranges (# of Points)
- Table Discussion: Round 1 Feedback (12:45 – 1:30pm)
  - Panelists discuss item mean and score point distributions
  - Panelists discuss test level agreement
  - Panelists discuss items flagged for greatest disagreement
  - Panelists discuss additional items of interest
- Round 2 Judgments (PBA and EOY) (1:30 – 3:00pm)
  - Round 2 Readiness Form
  - Panelists revise item judgments based on feedback data and discussions

*Break (3:00 – 3:30pm)* – Data analyzed during this time

- Round 2 feedback (Panelist Agreement Data) (3:30 – 3:45pm)
  - Item Level
    - Item level panelist agreement data
  - Test Level
    - Individual threshold scores recommendations
    - Table threshold scores recommendations
    - Group threshold score recommendations
    - Pre-Policy reasonable ranges (# of Points)
- Round 2 feedback table discussions (3:45 – 4:30pm)
  - Panelists discuss test level agreement
  - Panelists discuss items flagged for greatest disagreement
  - Panelists discuss additional items of interest
- Whole group discussion of agreement results (4:30 – 5:00pm)

**DAY 4**

**Breakout Session**

*Breakfast* (8:00 – 8:30am)

- Round 2 Feedback (Impact Data) (8:30 – 8:45am)
  - Overall
  - By student subgroup
  - Pre-Policy reasonable ranges (% Students)
- Whole group discussion of impact data results (8:45 – 9:15am)
- Round 3 Judgments (9:15 – 10:45am)
  - Round 3 Readiness Form
  - Panelists revise item judgments based on feedback data and discussions

*Break* (10:45 – 11:15am) – Data analyzed during this time

- Round 3 Feedback (11:15 – 11:30am)
  - Group threshold score recommendations
  - Impact data (overall and for subgroups)
  - Pre-Policy reasonable ranges (# Points and % Students)

- Next Steps and Close Out (11:30 – 12:00pm)
  - Secure materials check-in
  - Thank you for participating!

*Lunch* (12:00 – 1:00pm)

**Vertical Articulation** (Table Leaders from each panel reconvene)

- Introductions and session purpose (1:00 –1:15pm)
- Cross-grade PLD review (1:15 –2:15pm)
- Cross-grade impact data review (2:15 –2:45pm)
- Recommend changes to Round 3 results (2:45 –3:30pm)

## Grades 3-8 PLS Meeting Agenda

**PARCC Performance Level Setting– Detailed Agenda**
**August 17-21, 2015 Grades 7/8**
**August 24-28, 2015 Grades 3/4 & 5/6**

**DAY 1**

*Breakfast* (8:00 – 8:30am)

**General Session**

- Welcome (8:30 – 9:00am)
  - Who is in the room? (Introductions of key meeting staff and facilitators)

    - Why are we here? (Meeting Purpose)
        - What are we trying to accomplish?
        - What will we be doing this week?
    - Overview of PARCC
- Overview and Purpose of Performance Level Setting (9:00 – 10:00am)
    - What is Performance Level Setting?
    - Basic Vocabulary
    - Evidence Based Standard Setting
    - Overview of the Item Level Judgment Task
    - Try it out! Pop-culture assessment example

*Break (10:00 – 10:15am)* – Move to Breakout Session Rooms

**Breakout Session**

- Breakout Session Introductions (10:15 – 10:45am)
    - Introductions (panelists, facilitators, table leaders)
    - Orientation to Materials
    - Binder Check-out
    - Login to Moodle (password resets if needed)
    - Security discussion
- Experience the Assessment (PBA) Test 1 (10:45 – 11:45am)
    - Testing times for PARCC
    - How PARCC items are developed
    - Overview of Math/ELA item types
    - Panelists take the PBA test form

*Lunch (11:45 – 12:30pm)*

- Scoring the PARCC Assessments (12:30 – 1:15pm)
    - Facilitators review scoring rules for PARCC Items
    - Orientation to Answer Key/Rubric Documents
- Scoring Your PBA Assessment (1:15 – 1:45pm)
    - Panelists score their responses to the PBA Assessment
- Experience the Assessment (EOY) Test 1 (1:45 – 2:45pm)
    - Panelists take the EOY test form

*Break (2:45 – 3:00pm)*

- Scoring Your EOY Assessment (3:00 – 3:45pm)
    - Facilitators review scoring rules for EOY Items
    - Panelists score their responses to the EOY Assessment
- Review and Discuss PLDs (included in pre-work) Test 1 (3:45 – 4:30pm)
    - Panelists discuss PLDs with their table groups noting key differences between performance levels
- College and Career Readiness Discussion (4:30 – 5:00pm)
    - Review of PARCC CCR Definitions and Policies
    - Facilitators discuss PARCC Benchmarking , PEJ studies, and pre-policy meeting

- Facilitators share historical data points on college and career readiness

**DAY 2 (Breakout Session)**

*Breakfast* (8:00 – 8:30am)

- Borderline Descriptor Training (8:30 – 9:00am)
  - Typical vs. Borderline Students
  - Working with Draft Borderline Descriptors
- Borderline Student Table Discussion (9:00 – 9:45am)
  - A representative from each table is assigned a borderline level (2, 3, 4, or 5)
  - Borderline level groups meet first to discuss the draft descriptors for their assigned level.
    - How well do these describe the borderline student as we envision them?
- Editing Discussion at Original Tables (9:45 – 10:30am)
  - Table groups reconvene and discuss what they learned about each borderline student group
  - Then work as a table group to edit assigned descriptors

*Break* (10:30 – 10:45am)—Facilitators copy table descriptors into master document

- Group Discussion of Borderline Descriptors (10:45 – 11:15am)
  - Facilitator review compiled descriptors from each table with group
- Performance Level Setting Training (11:15 – 12:00pm)
  - Process for making item judgments
  - Definition of likely
  - Judgments that make sense and those that don't
  - Ceiling and floor judgments
  - Recording item judgments

*Lunch* (12:00– 12:45pm) – Print Borderline Descriptors

- Practice Judgment Task (12:45 – 1:30pm)
  - Panelists work independently to make judgments for practice set items
- Discuss Practice Task (1:30 – 2:00pm)
  - Group discussion of judgments, challenges, and points of confusion
  - Relationship between item judgments and threshold scores

*Break* (2:00 – 2:15pm)

- Round 1 Judgments PBA Test 1(2:15 – 3:45pm)
  - Round 1 Readiness Form
  - Panelists work independently to make judgments for PBA items
- Round 1 Judgments EOY Test 1 (3:45 – 5:15pm)
  - Panelists work independently to make judgments for EOY items

**DAY 3 (Breakout Session)**

*Breakfast* (8:00 – 8:30am)

- Round 1 feedback Test 1 (8:30 – 9:00am)
  - Item Level
    - Item means & score point distributions
    - Item level panelist agreement data
  - Test Level
    - Individual threshold scores recommendations
    - Table threshold scores recommendations
    - Group threshold score recommendations
- Table Discussion: Round 1 Feedback Test 1 (9:00 – 9:45am)
  - Panelists discuss item mean and score point distributions
  - Panelists discuss test level agreement
  - Panelists discuss items flagged for greatest disagreement
  - Panelists discuss additional items of interest
- Round 2 Judgments (PBA and EOY) Test 1 (9:45 – 11:45am)
  - Round 2 Readiness Form
  - Panelists revise item judgments based on feedback data and discussions

*Lunch (11:45 – 12:30pm)* – Data analyzed during this time

- Round 2 feedback (Panelist Agreement Data) Test 1 (12:30 –12:45pm)
  - Item Level
    - Item means & Score point distributions
    - Item level panelist agreement data
  - Test Level
    - Individual threshold scores recommendations
    - Table threshold scores recommendations
    - Group threshold score recommendations
- Round 2 feedback table discussions Test 1 (12:45 – 1:15pm)
  - Panelists discuss test level agreement
  - Panelists discuss items flagged for greatest disagreement
  - Panelists discuss additional items of interest
- Whole group discussion of agreement results Test 1 (1:15 – 1:30pm)
- Round 2 Feedback (Impact Data) Test 1 (1:30 – 1:45pm)
  - Overall
  - By student subgroup
  - Comparison to higher grade level results
- Whole group discussion of impact data results Test 1 (1:45 – 2:00pm)
- Round 3 Judgments Test 1 (2:00 – 3:00pm)
  - Round 3 Readiness Form
  - Panelists revise item judgments based on feedback data and discussions

*Break* (3:00 – 3:30pm) – Data analyzed during this time

- Round 3 Feedback Test 1 (3:30 – 3:45pm)
  - Group threshold score recommendations

  - Impact data (overall and for subgroups)
  - Comparison to higher grade level results
- Experience the Assessment (PBA) Test 2 (3:45 – 4:45pm)
  - Panelists take the PBA test form
- Scoring Your PBA Assessment Test 2 (4:45 – 5:15pm)
  - Panelists score their responses to the PBA Assessment

**DAY 4 (Breakout Session)**

*Breakfast* (8:00 – 8:30am)

- Experience the Assessment (EOY) Test 2 (8:30 – 9:30am)
  - Panelists take the EOY test form
- Scoring Your EOY Assessment Test 2 (9:30 – 10:00am)
  - Panelists score their responses to the EOY Assessment

*Break* (10:00 – 10:15am)

- Review and Discuss PLDs Test 2 (included in pre-work) (10:15 – 10:45am)
  - Panelists discuss PLDs with their table groups noting key differences between performance levels
- Borderline Student Table Discussion (10:45 – 11:15am)
  - A representative from each table is assigned a borderline level (2, 3, 4, or 5)
  - Borderline level groups meet first to discuss the draft descriptors for their assigned level.
    - How well do these describe the borderline student as we envision them?
- Editing Discussion at Original Tables (11:15 – 12:00pm)
  - Table groups reconvene and discuss what they learned about each borderline student group
  - Then work as a table group to edit assigned descriptors

*Lunch* (12:00– 12:45pm)—Facilitators copy table descriptors into master document

- Group Discussion of Borderline Descriptors (12:45 – 1:15pm)—Print Borderline Descriptors
  - Facilitator review compiled descriptors from each table with group
- Round 1 Judgments PBA Test 2 (1:15 – 2:15pm)
  - Round 1 Readiness Form
  - Panelists work independently to make judgments for PBA items
- Round 1 Judgments EOY Test 2 (2:15 – 3:15pm)
  - Panelists work independently to make judgments for EOY items

*Break* (3:15– 3:45pm) – Data analyzed during this time

- Round 1 feedback Test 2 (3:45 – 4:00pm)
  - Item Level
    - Item means & Score point distributions
    - Item level panelist agreement data
  - Test Level
    - Individual threshold scores recommendations

        - Table threshold scores recommendations
        - Group threshold score recommendations
- Table Discussion: Round 1 Feedback Test 2 (4:00 – 4:30pm)
    - Panelists discuss item mean and score point distributions
    - Panelists discuss test level agreement
    - Panelists discuss items flagged for greatest disagreement
    - Panelists discuss additional items of interest
- Round 2 Judgments (PBA and EOY) Test 2 (4:30 – 5:30pm)
    - Round 2 Readiness Form
    - Panelists revise item judgments based on feedback data and discussions

**DAY 5**

**Breakout Session**

*Breakfast* (8:00 – 8:30am)

- Round 2 feedback (Panelist Agreement Data) Test 2 (8:30 –8:45am)
    - Item Level
        - Item means & score point distributions
        - Item level panelist agreement data
    - Test Level
        - Individual threshold scores recommendations
        - Table threshold scores recommendations
        - Group threshold score recommendations
- Round 2 feedback table discussions Test 2 (8:45 – 9:30am)
    - Panelists discuss test level agreement
    - Panelists discuss items flagged for greatest disagreement
    - Panelists discuss additional items of interest
- Whole group discussion of agreement results Test 2 (9:30 – 9:45am)
- Round 2 Feedback (Impact Data) Test 2(9:45 – 10:00am)
    - Overall
    - By student subgroup
    - Comparison to higher grade level results
- Whole group discussion of impact data results Test 2 (10:00 – 10:15am)
- Round 3 Judgments Test 2 (10:15 – 11:15am)
    - Round 3 Readiness Form
    - Panelists revise item judgments based on feedback data and discussions

*Break* (11:15 –11:45am) – Data analyzed during this time

- Round 3 Feedback Test 2 (11:45 – 12:00pm)
    - Group threshold score recommendations
    - Impact data (overall and for subgroups)
    - Comparison to higher grade level results
- Next Steps and Close Out (12:00 – 12:15pm)

- Secure materials check-in
- Thank you for participating!

*Lunch* (12:15 – 1:00pm)

**Vertical Articulation** (Grade 3-6: Table Leaders from each panel reconvene; Grade 7-8: All panelists reconvene)

- Introductions and session purpose (1:00 –1:15pm)
- Cross-grade PLD review (1:15 –2:15pm)
- Cross-grade impact data review (2:15 –2:45pm)
- Recommend changes to Round 3 results (2:45 –3:30pm)

## Appendix 7: Example of Performance Level Setting Feedback Data

This appendix provides examples of the feedback data which was provided to panelists after each round of item-level judgments.

After each item judgment round, panelists were provided with the threshold scores that would have resulted from their own item judgments. During the High School meetings, the individual threshold score feedback also included the percent of total score scores for each performance level, to assist with the interpretation of the reasonable range feedback. Additionally, they were provided with the summary statistics for the threshold score recommendations for each performance level. Figure A6.1 shows an example table presenting the summary statistics for the entire committee from a meeting. The panelists were also provided similar information for all table groups, following Rounds 1 and 2.

| Label | N | Mean | Median | Min | Max |
|---|---|---|---|---|---|
| Level 2 | 17 | 14.65 | 14 | 5.00 | 28.00 |
| Level 3 | 17 | 35.00 | 35 | 21.00 | 56.00 |
| Level 4 | 17 | 58.41 | 58 | 43.00 | 80.00 |
| Level 5 | 17 | 73.06 | 74 | 58.00 | 82.00 |

**Figure A6.1:** Summary of Threshold Score Recommendations Feedback Provided

The distribution of the threshold score recommendations for each performance level was displayed as a frequency bar graph of recommendations from the entire committee, after Rounds 1 and 2. Figure A6.2 shows an example threshold score recommendation distribution for performance level 4. Similar graphs were also displayed for performance level 2, 3, and 5 threshold score recommendation distributions. Additionally, graphs which showed the frequency distribution for each adjacent pair of performance levels (e.g. level 2 with level 3, level 3 with level 4, and level 4 with level 5) were provided to illustrate areas of overlap (e.g. some panelists threshold scores for level 4 were lower than some panelists threshold scores for level 3, etc.).

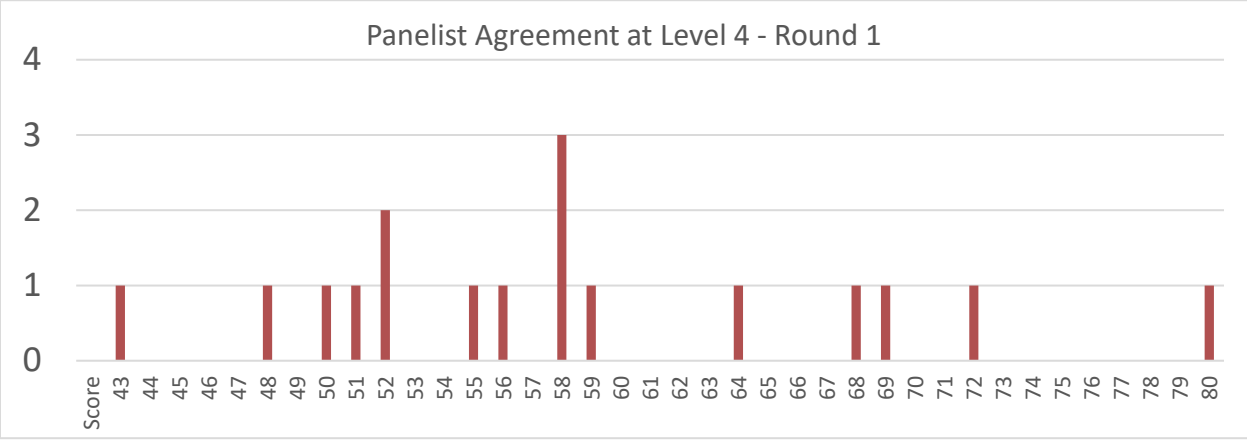



**Figure A6.2:** Display of Panelists Threshold Score Distribution

Panelists were also provided feedback about agreement between panelist judgments at the item level for all items for borderline levels 2 through 5. Figure A6.3 shows an example of what this feedback looked like for 2 ELA items for borderline level 4. The data represent the percent of panelists who indicated 0 to 4 points for each item (or trait for the PCR item) on their item judgment forms. In this example, 70% of panelists thought students at the borderline of Level 4 would get 2 out of 2 points on the first item, 20% thought they would get 1 out of 2 points, and 10% thought they would get 0 out of 2 points. For each borderline level, the five items with the greatest level of panelist disagreement were specifically identified, by listing these items on a separate cover sheet for this information and highlighting these items in the table. Panelists were asked to discuss this information with their table groups after Round 1 and with both their table groups and the whole committee after Round 2.

| Item | ES | Max Points | Borderline Level 4 | | | | |
|---|---|---|---|---|---|---|---|
| | | | 0 | 1 | 2 | 3 | 4 |
| 8098_A | RL 11.1.2 | 2 | 10% | 20% | 70% | -- | -- |
| 8699 | RL 11.1.2 | | | | | | |
| | Reading Comprehension | 4 | 0% | 20% | 30% | 40% | 10% |
| | Written Expression | 4 | 0% | 0% | 30% | 60% | 10% |
| | Knowledge of Language and Conventions | 3 | 0% | 10% | 60% | 30% | -- |

**Figure A6.3:** Example of Item Level Panelist Agreement Data for Borderline Level 4

After the Round 2 and 3 item judgments, panelists were provided the expected impact on student classifications into performance levels, based on the current recommended threshold scores for each performance level. Figure A6.4 displays an example impact data graph for all students and disaggregated by gender. Similar data was provided for students disaggregated by ethnicity, for English Language Learners (ELL), and students with disabilities (SWD). During the high school meetings, panelists were presented with the just the impact data for the assessment related to their meeting. For the grades 7 & 8 PLS meetings, panelists were presented impact data for assessments related to their meeting as well as the percent of students classified as "Level 3 and above" on the high school assessments[8]. For the grades 3 through 6 meetings, panelists were presented the impact data for assessments related to their meeting along with the impact data from the related grades 7 & 8

[8] This information reflected a broader set of student data than was shared during the high school PLS meetings as it included student performance across all online test forms (the same data that were shared with the PARCC Governing Board). Additionally, because of the Governing Board's decision to shift the High School performance levels, a decision was made not to present the full set of impact data across all 5 performance levels. Since the percent of students college- and career- ready or on-track was held constant despite the shift in performance levels, the "Level 3 and above" information presented accurate and actionable information for panelists.

assessments, based on the recommended threshold scores from the grade 7/8 vertical articulation meetings.

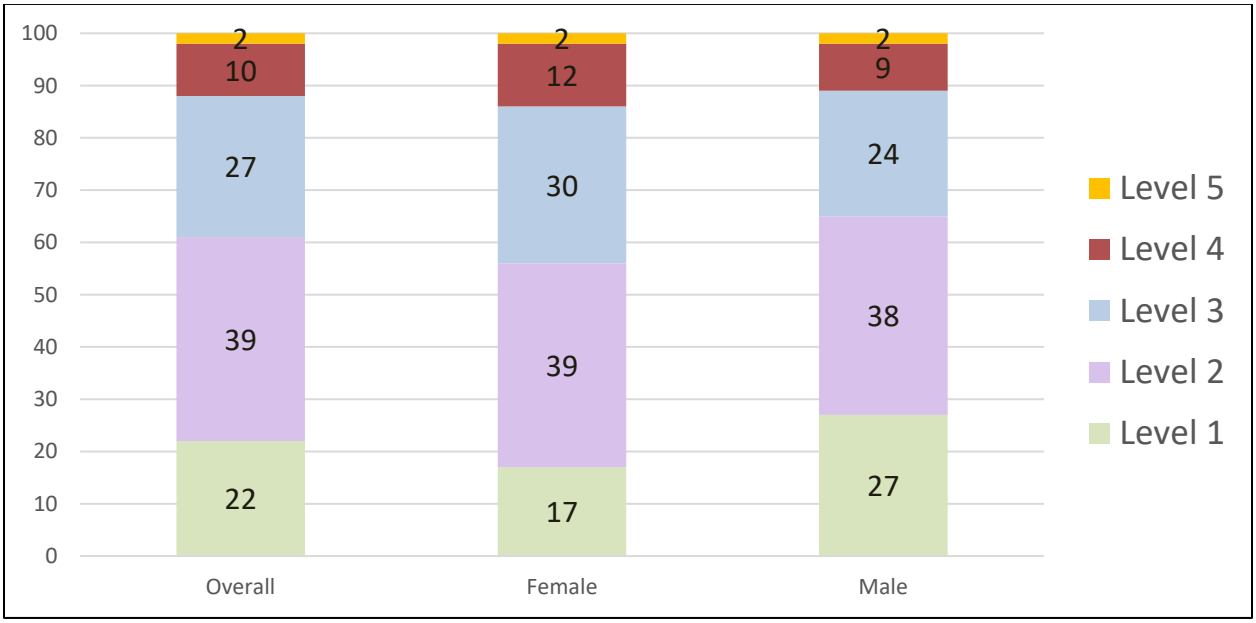



**Figure A6.4:** Estimated Impact Data for Total Group and by Gender Based on Threshold Score Recommendations

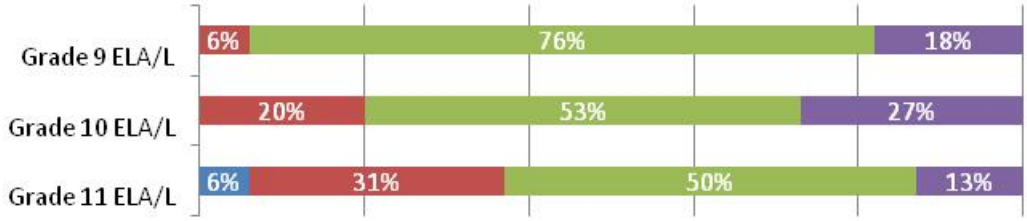
# Appendix 8: Performance Level Setting Process Evaluation Summary

## High School English Language Arts/Literacy (ELA/L)

**Question 1:** Select the response that best reflects your opinion about the level of success of the various components of the meeting in which you have just participated. The activities were designed to help you both understand the performance level setting process and be supportive of the recommendations made by the committee.

Not Successful | Partially Successful | Successful | Very Successful

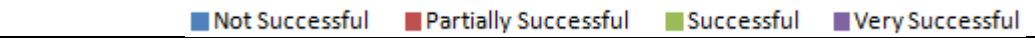
Meeting pre-work




Introduction to the process of setting performance level standards


Discussion of the performance labels and policy definitions

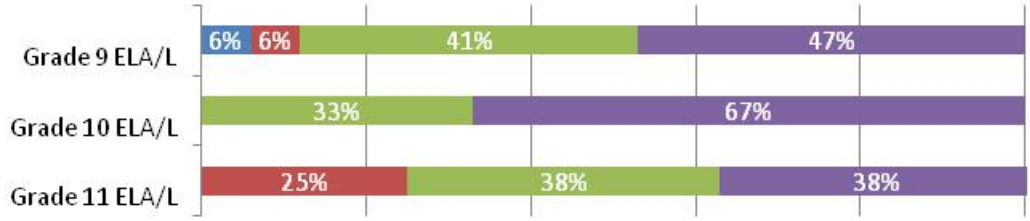

Discussion of the scoring of the items on the assessment




Taking the actual assessment(s)

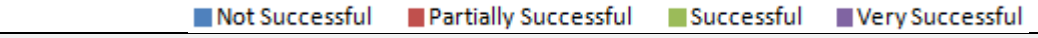

Overview of the item judgment process


Development of the borderline performance level descriptors



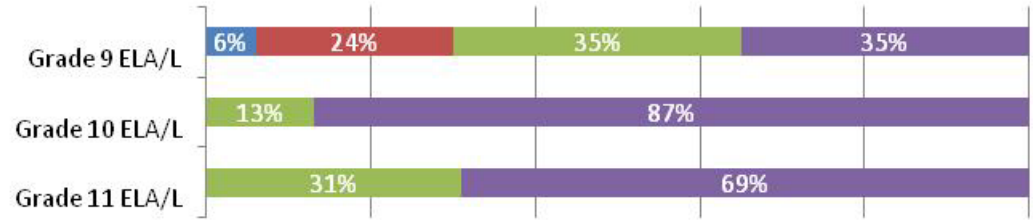

Practice exercise for the item judgment process


Panelist agreement provided after each item judgment round



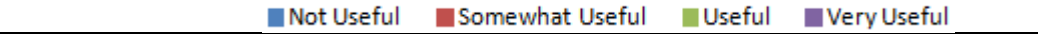

Item difficulty data provided after item judgment rounds 2 and 3


Student performance (impact) data provided after item judgment rounds 2 and 3



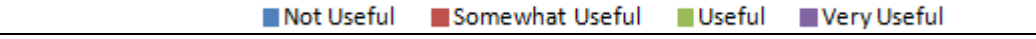

Discussion after each item judgment round


**Question 2:** How useful do you feel the following activities or information were in assisting you to make your recommendations?




Performance level descriptors (PLDs) and evidence tables


Training in the item judgment method



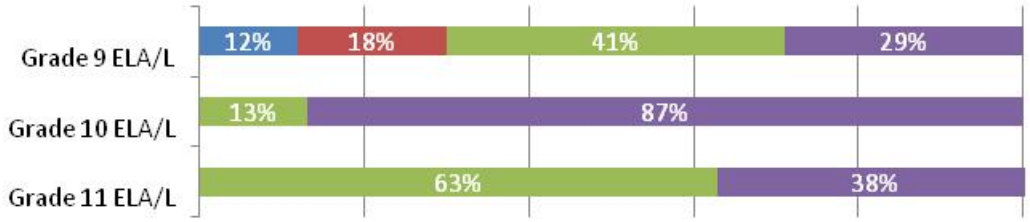

Borderline performance level descriptors

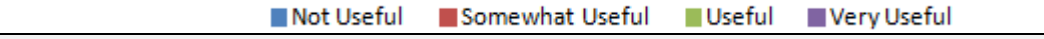

Panelist agreement data provided after Round 1


Item means and score distributions


Panelist agreement data provided after Round 2

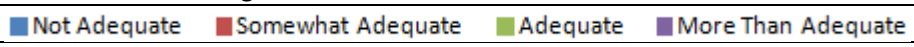

Student performance (impact) data






Discussion after each round


**Question 3:** How adequate were the following elements of the session?

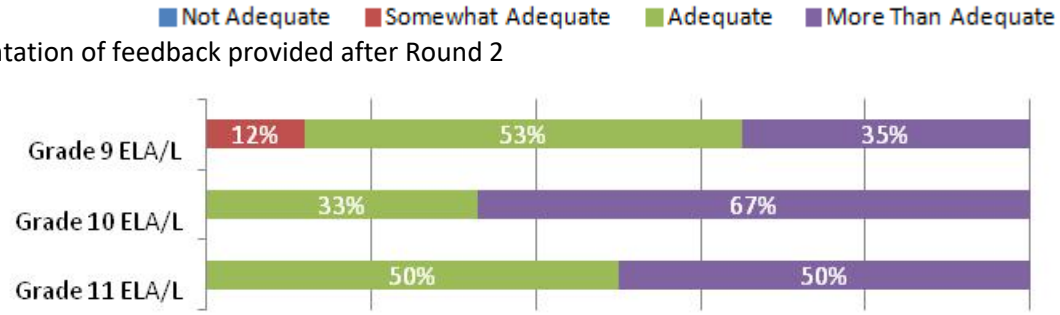



Training provided on the item judgment process

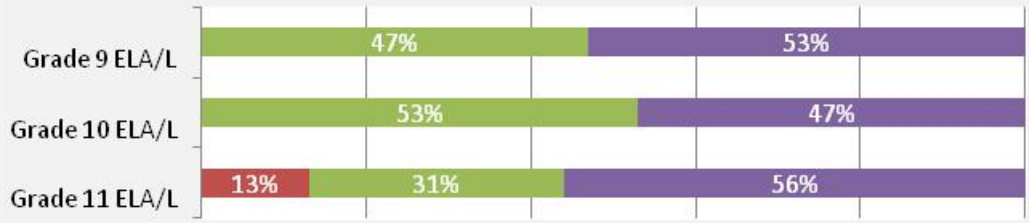

Amount of time spent training on the item judgment process


Time spent creating and discussing the borderline performance level descriptors


Visual presentation of feedback provided after Round 1




Visual presentation of feedback provided after Round 2

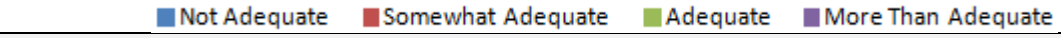



Facilities used for the session


Work space in table groups during meeting


Moodle site for accessing materials and making judgments

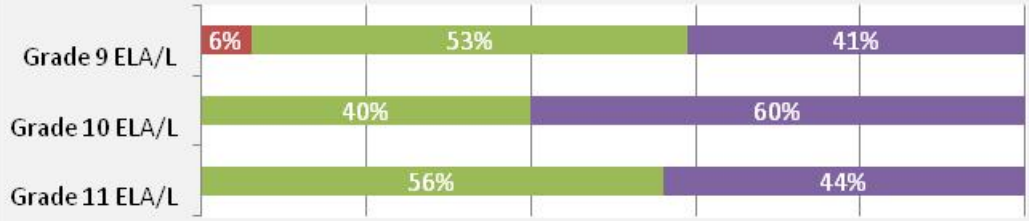

Computers used during meeting



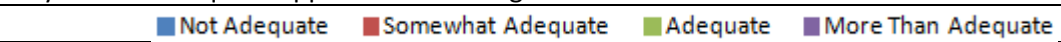

Materials provided in the binder


Total amount of time in breakout groups to make judgments


Number of rounds for the judgments

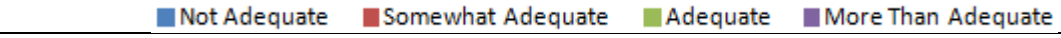



**Question 4:** Did you have adequate opportunities during the session to...




Express your opinions about student performance levels

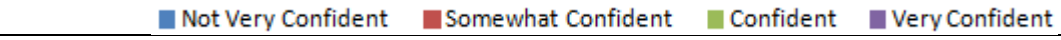

Ask questions about the standards and how they will be used




Ask questions about the process of making item judgments

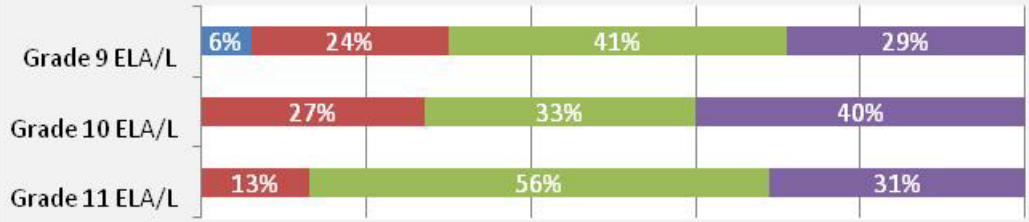

Interaction with your fellow panelists


**Question 5:** In applying the performance level setting method, you were asked to recommend scores (separating five performance levels) for student performance on the PARCC assessments. How confident do you feel that the threshold scores based on Round 3 judgments represent appropriate levels of student performance?

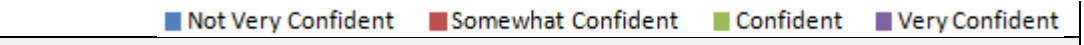



Level 2: Partial command


Level 3: Adequate command



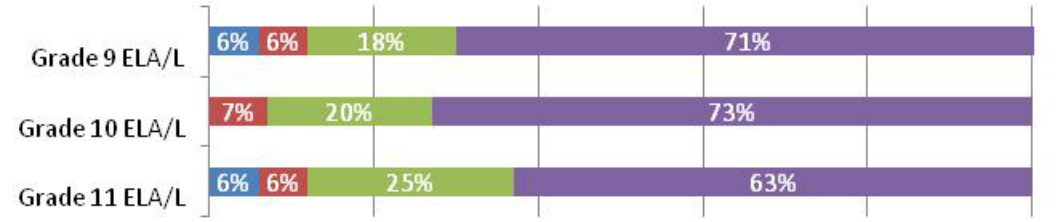

Level 4: Strong command




Level 5: Distinguished command


**Question 6:** To what degree do you support the threshold score based on Round 3 judgments for each performance level?


Level 2: Partial command

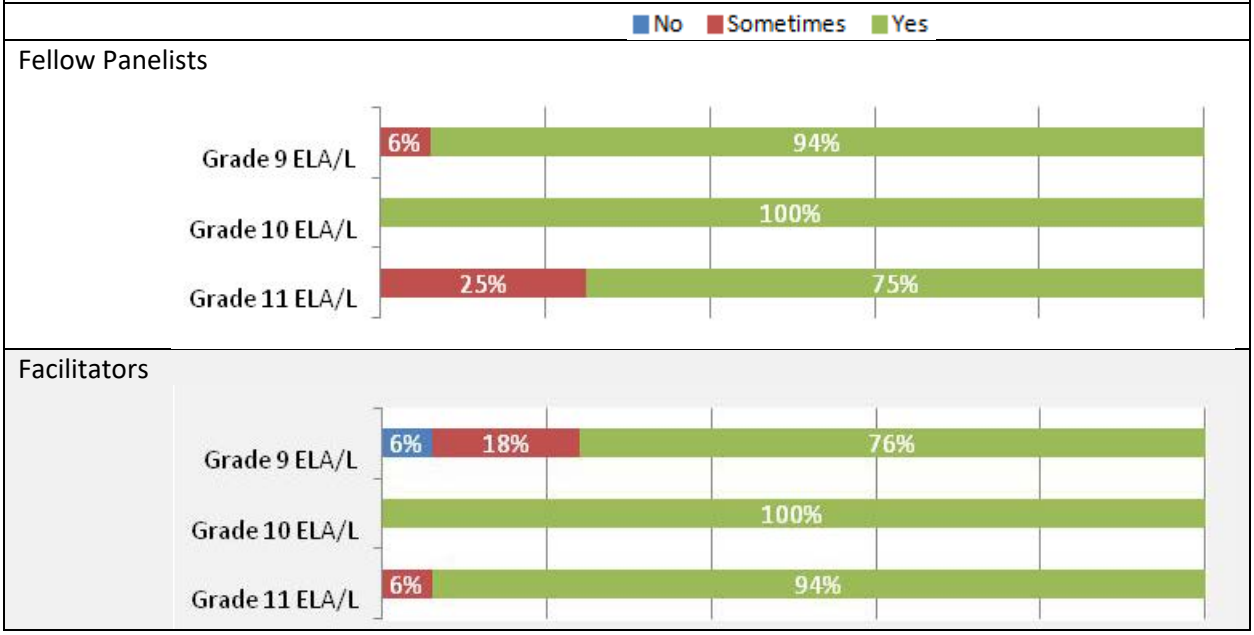



Level 3: Adequate command


Level 4: Strong command

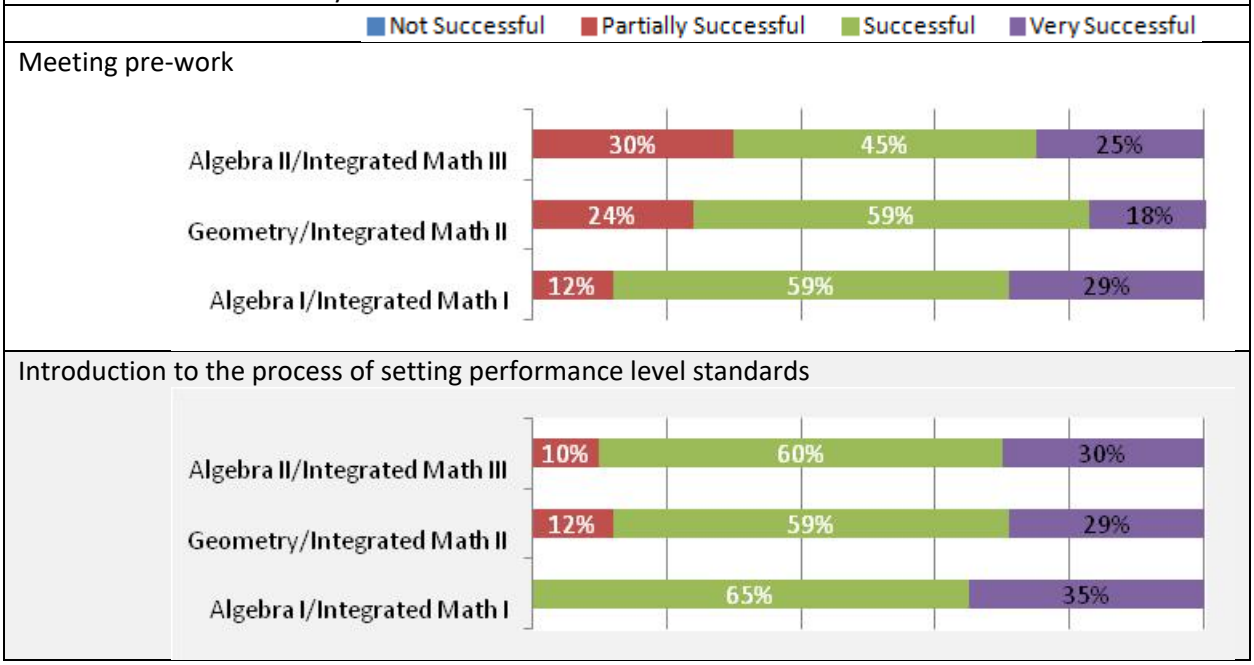

Level 5: Distinguished command

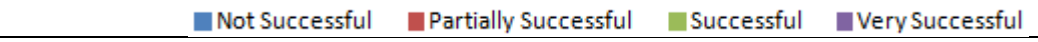
**Question 7:** Do you believe your opinions and judgments were treated with respect by...




## High School Mathematics

**Question 1:** Select the response that best reflects your opinion about the level of success of the various components of the meeting in which you have just participated. The activities were designed to help you both understand the performance level setting process and be supportive of the recommendations made by the committee.





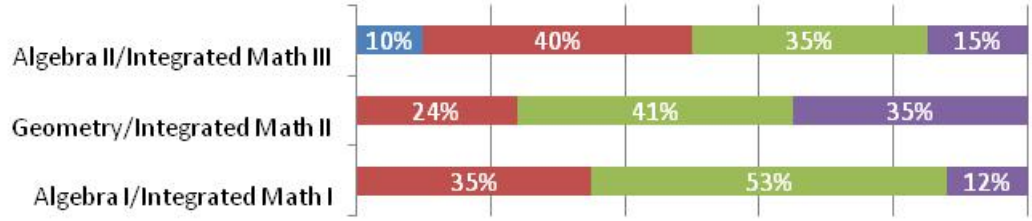

Discussion of the performance labels and policy definitions

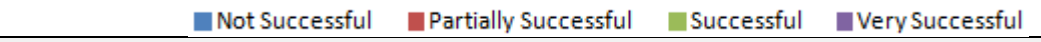

Discussion of the scoring of the items on the assessment


Taking the actual assessment(s)


Overview of the item judgment process


Development of the borderline performance level descriptors



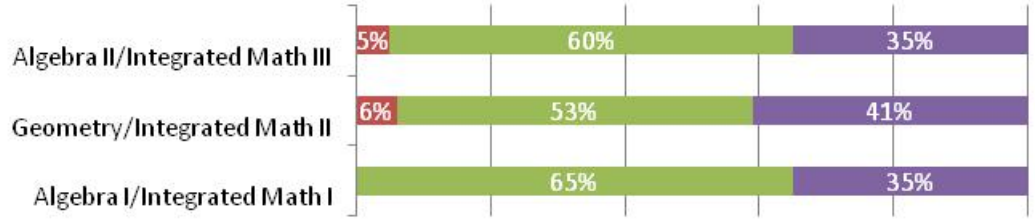



Practice exercise for the item judgment process


Panelist agreement provided after each item judgment round

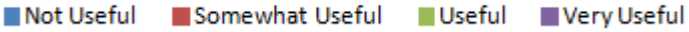

Item difficulty data provided after item judgment rounds 2 and 3


Student performance (impact) data provided after item judgment rounds 2 and 3



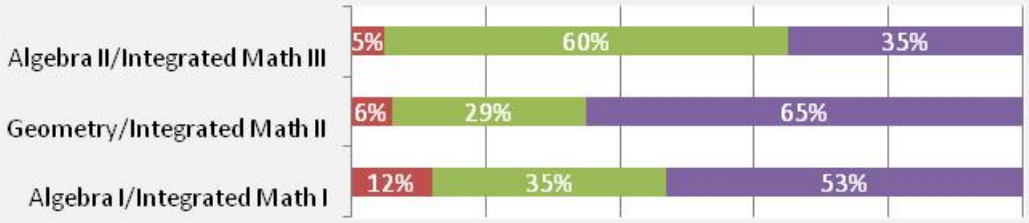

Discussion after each item judgment round

**Question 2:** How useful do you feel the following activities or information were in assisting you to make your recommendations?




Performance level descriptors (PLDs) and evidence tables

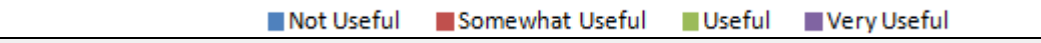

Training in the item judgment method




Borderline performance level descriptors


Panelist agreement data provided after Round 1


Item means and score distributions

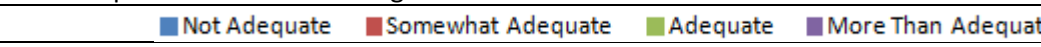



Panelist agreement data provided after Round 2


Student performance (impact) data

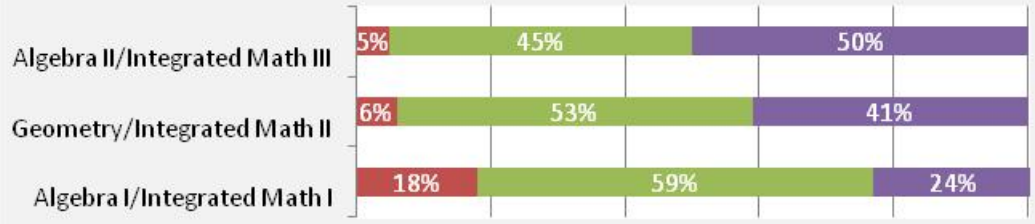

Discussion after each round

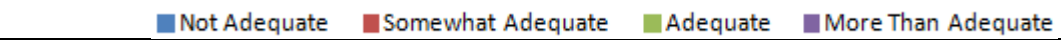

**Question 3:** How adequate were the following elements of the session?




Training provided on the item judgment process

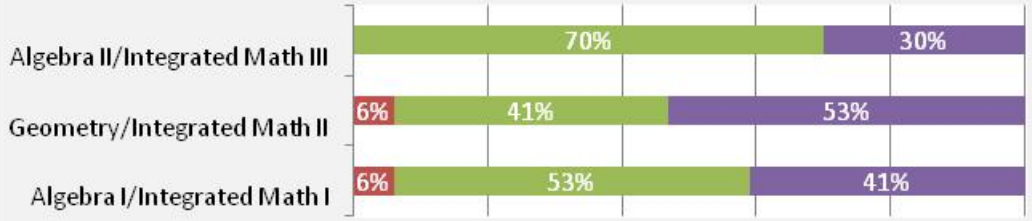

Amount of time spent training on the item judgment process





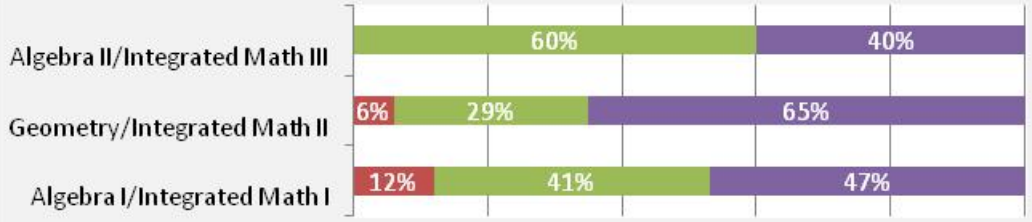

Time spent creating and discussing the borderline performance level descriptors


Visual presentation of feedback provided after Round 1




Visual presentation of feedback provided after Round 2

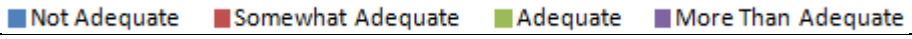

Facilities used for the session




Work space in table groups during meeting


Not Adequate
Somewhat Adequate
Adequate
More Than Adequate
Moodle site for accessing materials and making judgments
Algebra II/Integrated Math III
30%
70%
Geometry/Integrated Math II
24%
76%
Algebra I/Integrated Math I
12%
18%
71%
Computers used during meeting
Algebra II/Integrated Math III
25%
75%
Geometry/Integrated Math II
18%
82%
Algebra I/Integrated Math I
6%
35%
59%
Materials provided in the binder
Algebra II/Integrated Math III
25%
75%
Geometry/Integrated Math II
24%
76%
Algebra I/Integrated Math I
6%
35%
59%
Total amount of time in breakout groups to make judgments
Algebra II/Integrated Math III
10%
35%
55%
Geometry/Integrated Math II
24%
29%
47%
Algebra I/Integrated Math I
24%
53%
24%
Number of rounds for the judgments
Algebra II/Integrated Math III
60%
40%
Geometry/Integrated Math II
41%
59%
Algebra I/Integrated Math I
6%
76%
18%

**Question 4:** Did you have adequate opportunities during the session to...

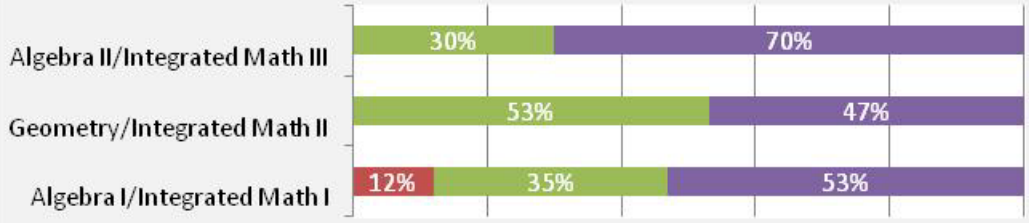



Express your opinions about student performance levels

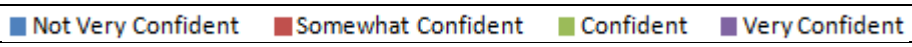

Ask questions about the standards and how they will be used


Ask questions about the process of making item judgments

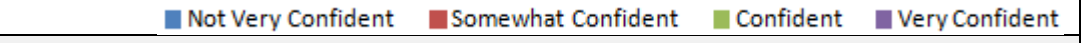

Interaction with your fellow panelists




**Question 5:** In applying the performance level setting method, you were asked to recommend scores (separating five performance levels) for student performance on the PARCC assessments. How confident do you feel that the threshold scores based on Round 3 judgments represent appropriate levels of student performance?

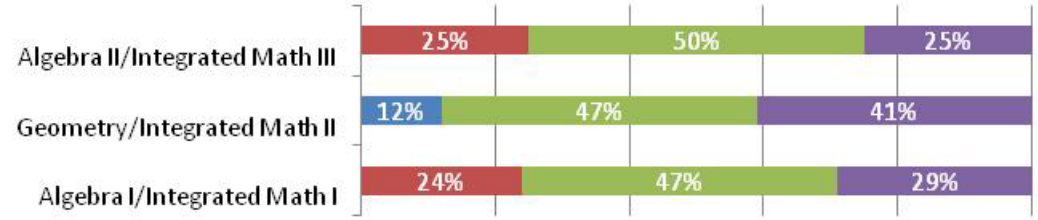



Level 2: Partial command



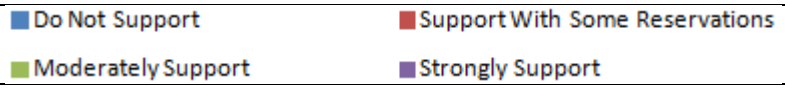

Level 3: Adequate command


Level 4: Strong command




Level 5: Distinguished command


**Question 6:** To what degree do you support the threshold score based on Round 3 judgments for each performance level?



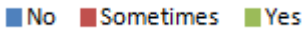

Level 2: Partial command


Level 3: Adequate command

Level 4: Strong command

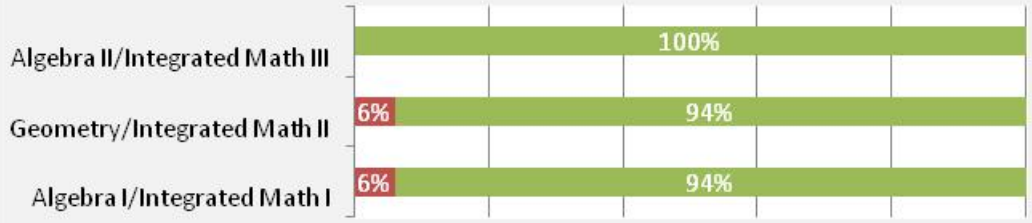

Level 5: Distinguished command


**Question 7:** Do you believe your opinions and judgments were treated with respect by...




Fellow Panelists


Facilitators

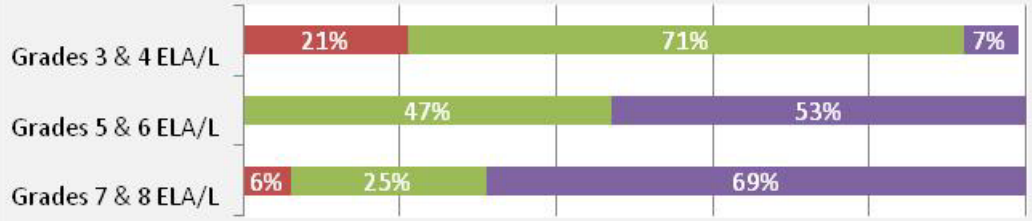


## Grades 3-8 English Language Arts/Literacy

**Question 1:** Select the response that best reflects your opinion about the level of success of the various components of the meeting in which you have just participated. The activities were designed to help you both understand the performance level setting process and be supportive of the recommendations made by the committee.

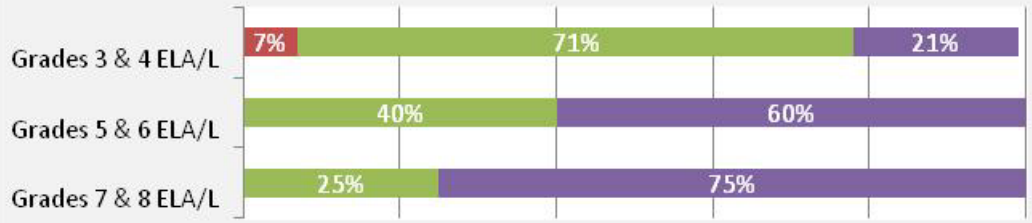
Meeting pre-work


Introduction to the process of setting performance level standards

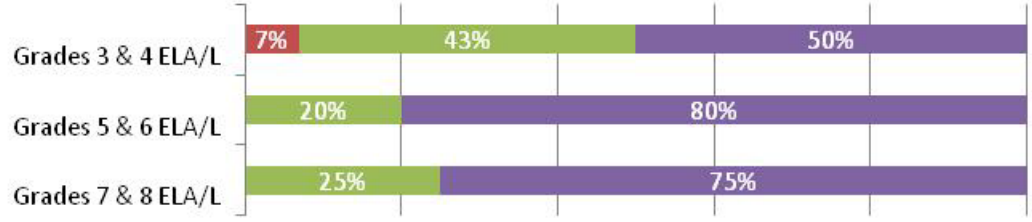


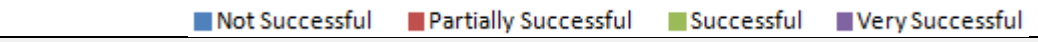

Discussion of the performance labels and policy definitions


Discussion of the scoring of the items on the assessment




Taking the actual assessment(s)





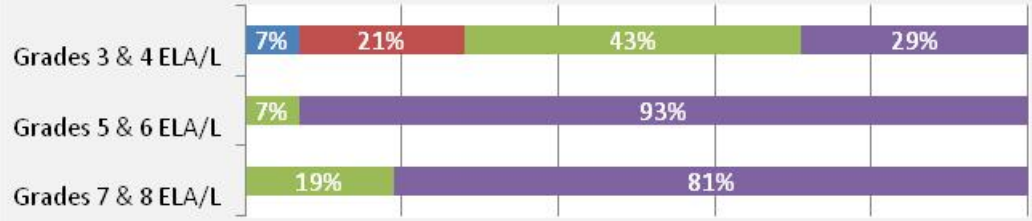

Overview of the item judgment process

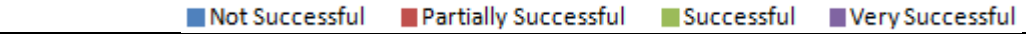

Development of the borderline performance level descriptors


Practice exercise for the item judgment process


Panelist agreement provided after each item judgment round


Item difficulty data provided after item judgment rounds 2 and 3

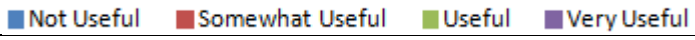





Student performance (impact) data provided after item judgment rounds 2 and 3


Discussion after each item judgment round

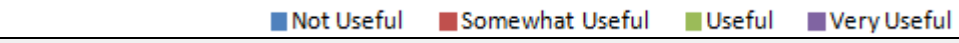

**Question 2:** How useful do you feel the following activities or information were in assisting you to make your recommendations?



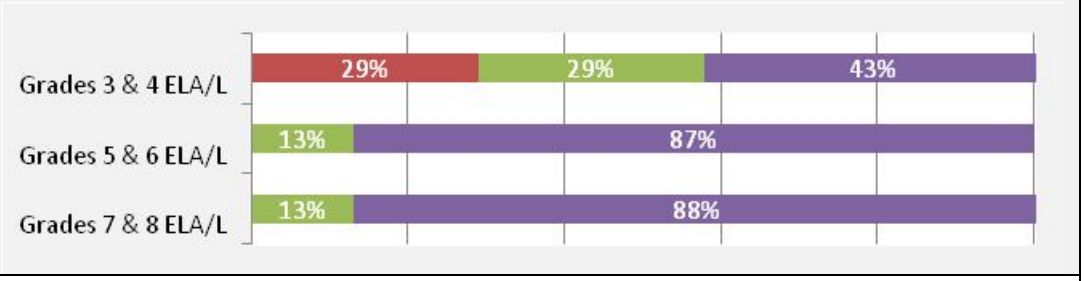

Performance level descriptors (PLDs) and evidence tables


Training in the item judgment method


Borderline performance level descriptors

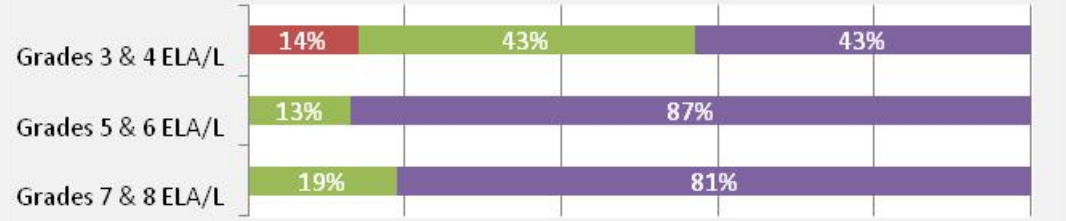



Panelist agreement data provided after Round 1



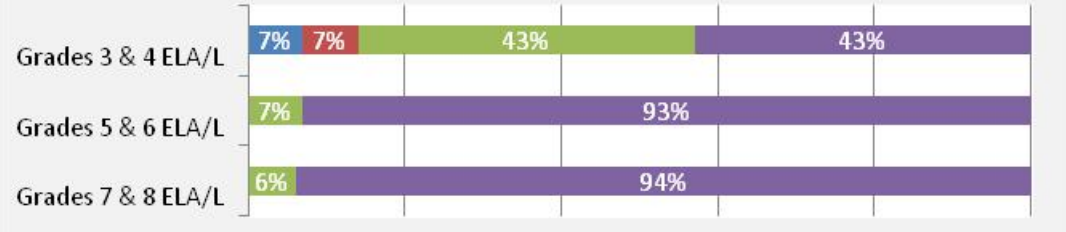

Item means and score distributions


Panelist agreement data provided after Round 2

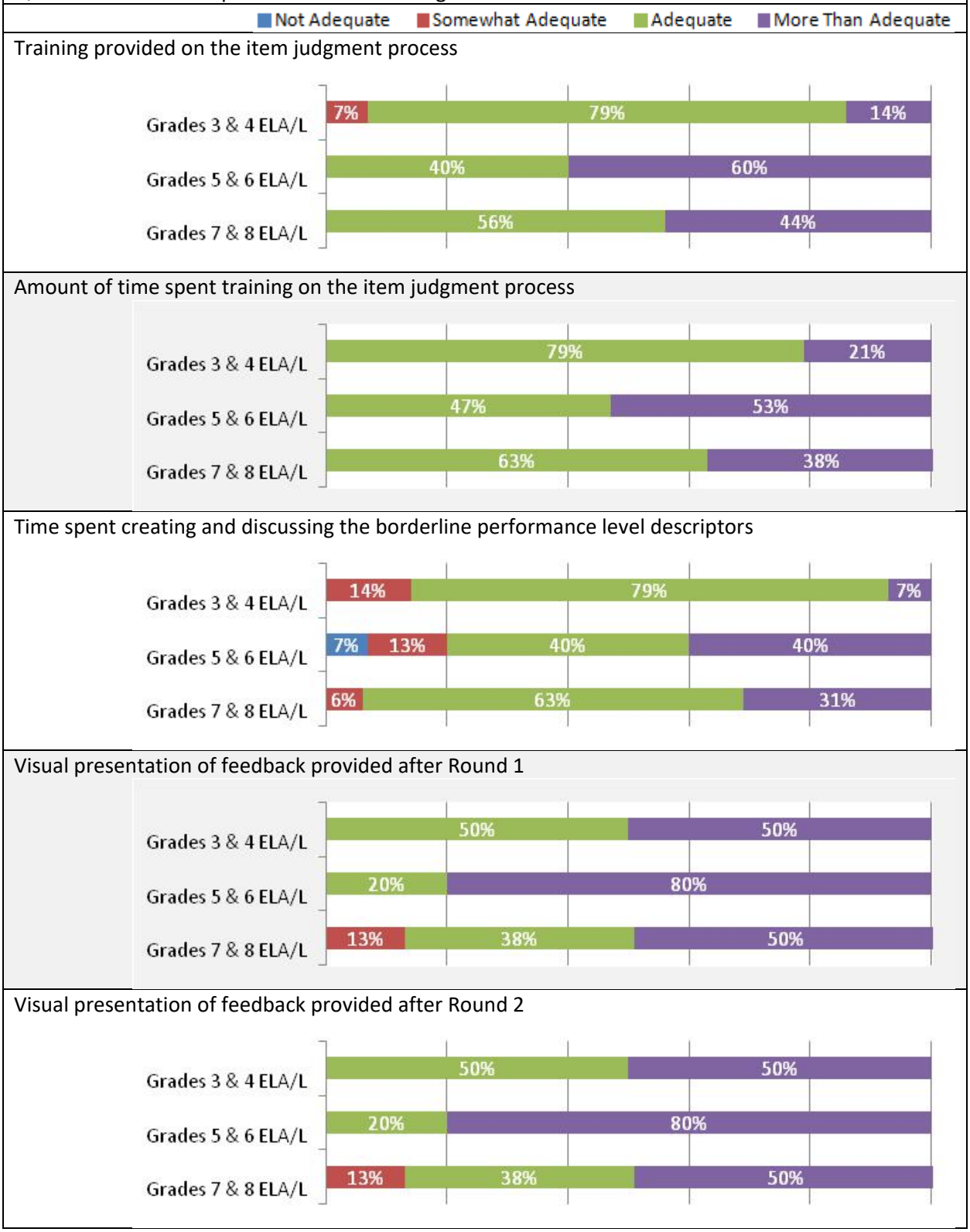


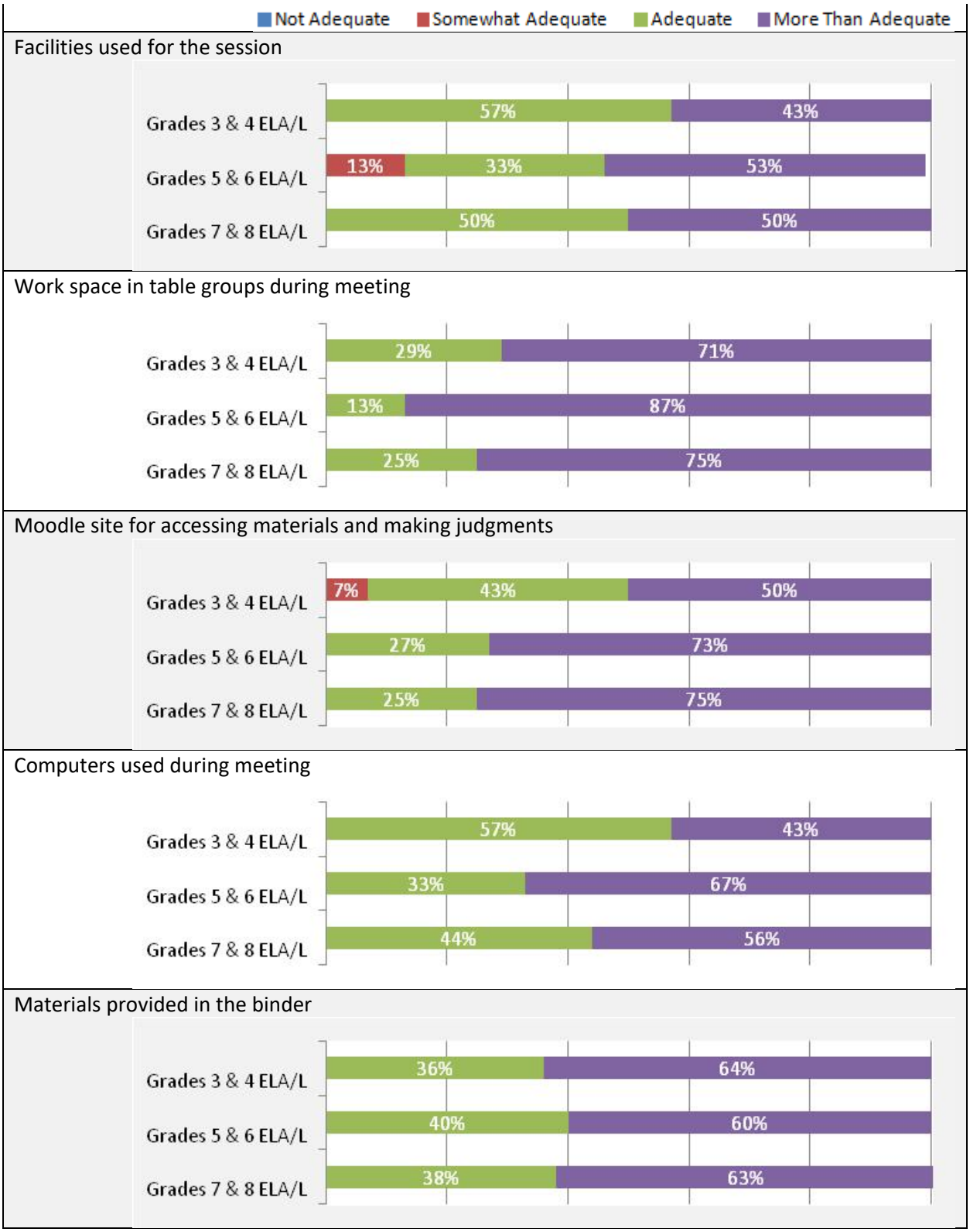

Student performance (impact) data

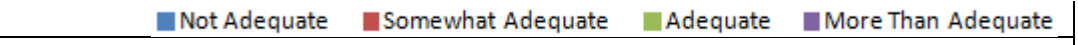

Discussion after each round



**Question 3:** How adequate were the following elements of the session?



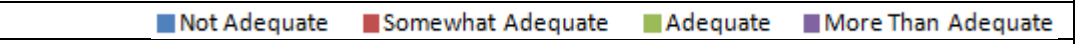

Not Adequate
Somewhat Adequate
Adequate
More Than Adequate
Facilities used for the session
Grades 3 & 4 ELA/L
57%
43%
Grades 5 & 6 ELA/L
13%
33%
53%
Grades 7 & 8 ELA/L
50%
50%
Work space in table groups during meeting
Grades 3 & 4 ELA/L
29%
71%
Grades 5 & 6 ELA/L
13%
87%
Grades 7 & 8 ELA/L
25%
75%
Moodle site for accessing materials and making judgments
Grades 3 & 4 ELA/L
7%
43%
50%
Grades 5 & 6 ELA/L
27%
73%
Grades 7 & 8 ELA/L
25%
75%
Computers used during meeting
Grades 3 & 4 ELA/L
57%
43%
Grades 5 & 6 ELA/L
33%
67%
Grades 7 & 8 ELA/L
44%
56%
Materials provided in the binder
Grades 3 & 4 ELA/L
36%
64%
Grades 5 & 6 ELA/L
40%
60%
Grades 7 & 8 ELA/L
38%
63%




Total amount of time in breakout groups to make judgments


Number of rounds for the judgments


**Question 4:** Did you have adequate opportunities during the session to...



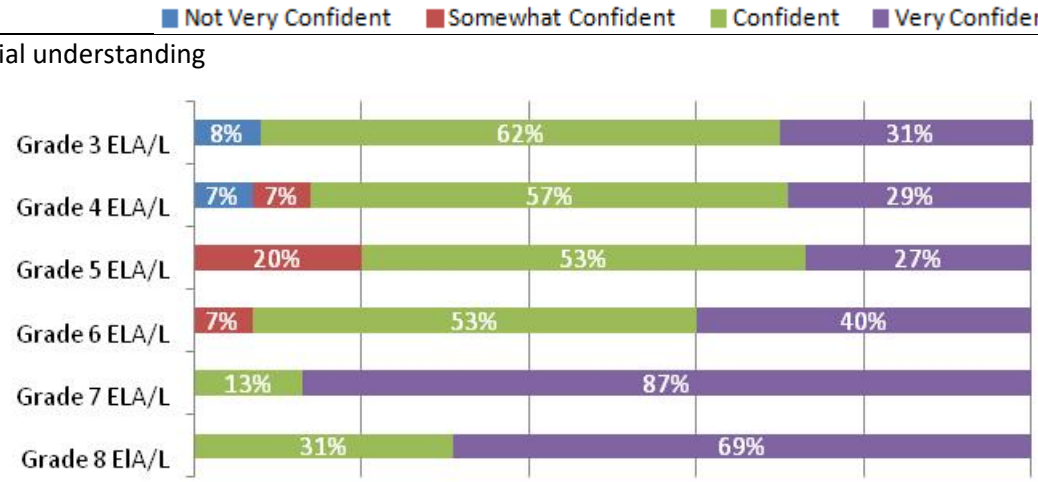

Express your opinions about student performance levels


Ask questions about the standards and how they will be used

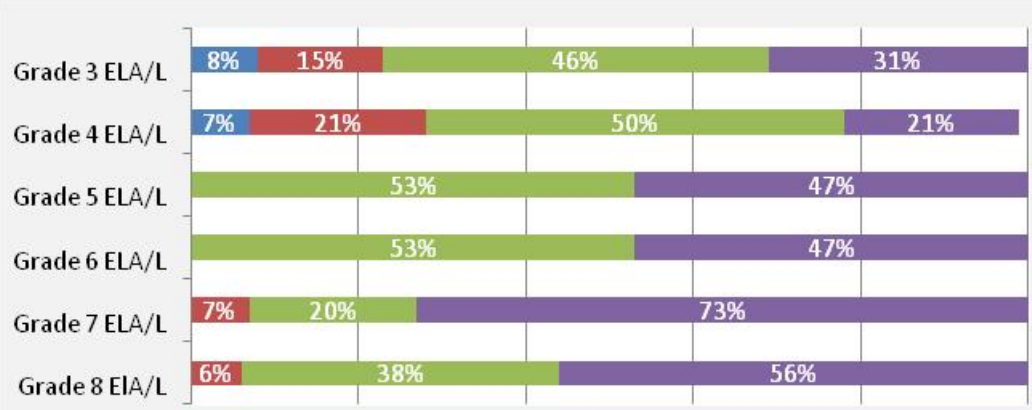

Ask questions about the process of making item judgments


Interact with your fellow panelists

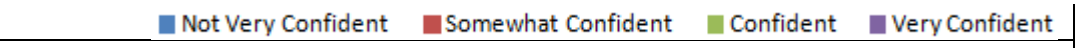
**Question 5:** In applying the performance level setting method, you were asked to recommend scores (separating five performance levels) for student performance on the PARCC assessments. How confident do you feel that the threshold scores based on Round 3 judgments represent appropriate levels of student performance?




Level 2: Partial understanding

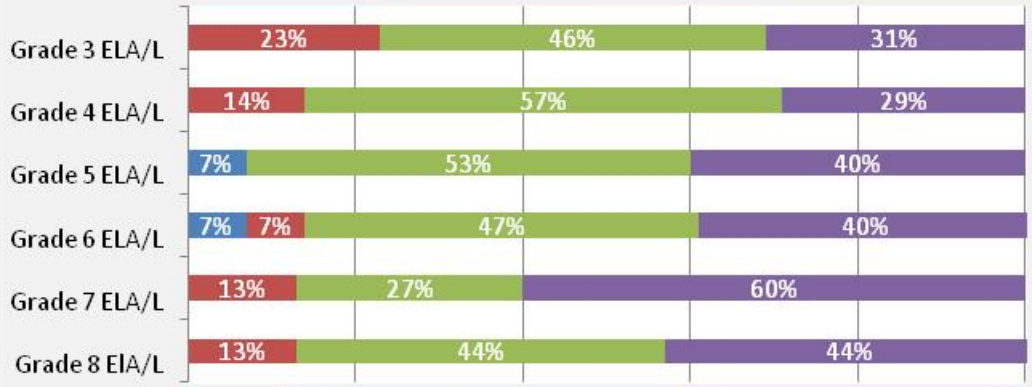



Level 3: Adequate understanding




Level 4: Strong understanding


Level 5: Distinguished understanding




**Question 6:** To what degree do you support the threshold score base on Round 3 judgments for each performance level?

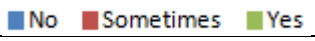

Level 2: Partial understanding


Level 3: Adequate understanding


Level 4: Strong understanding


Level 5: Distinguished understanding

**Question 7:** Do you believe your opinions and judgments were treated with respect by...

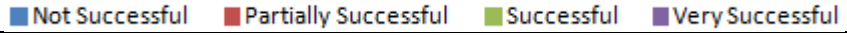



Fellow Panelists


Facilitators

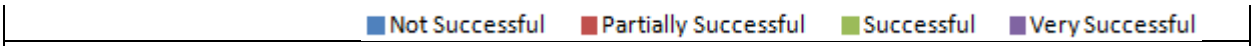

## Grades 3-8 Mathematics

**Question 1:** Select the response that best reflects your opinion about the level of success of the various components of the meeting in which you have just participated. The activities were designed to help you both understand the performance level setting process and be supportive of the recommendations made by the committee.




Meeting pre-work


Introduction to the process of setting performance level standards



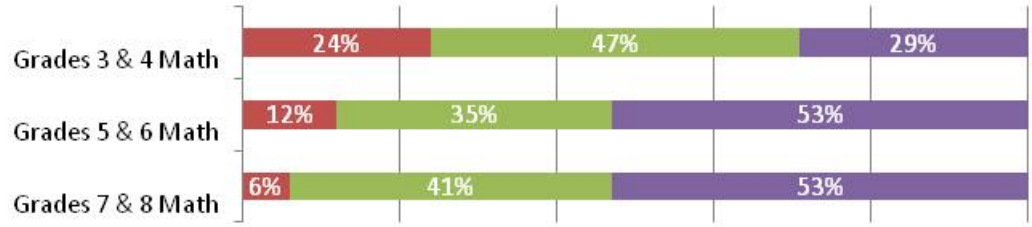

Discussion of the performance labels and policy definitions

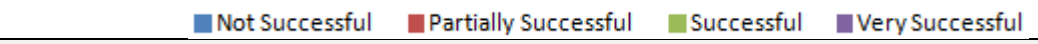

Discussion of the scoring of the items on the assessment


Taking the actual assessment(s)


Overview of the item judgment process


Development of the borderline performance level descriptors





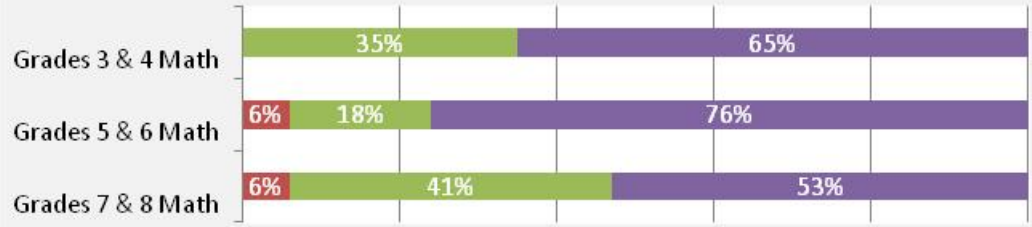

Practice exercise for the item judgment process


Panelist agreement provided after each item judgment round

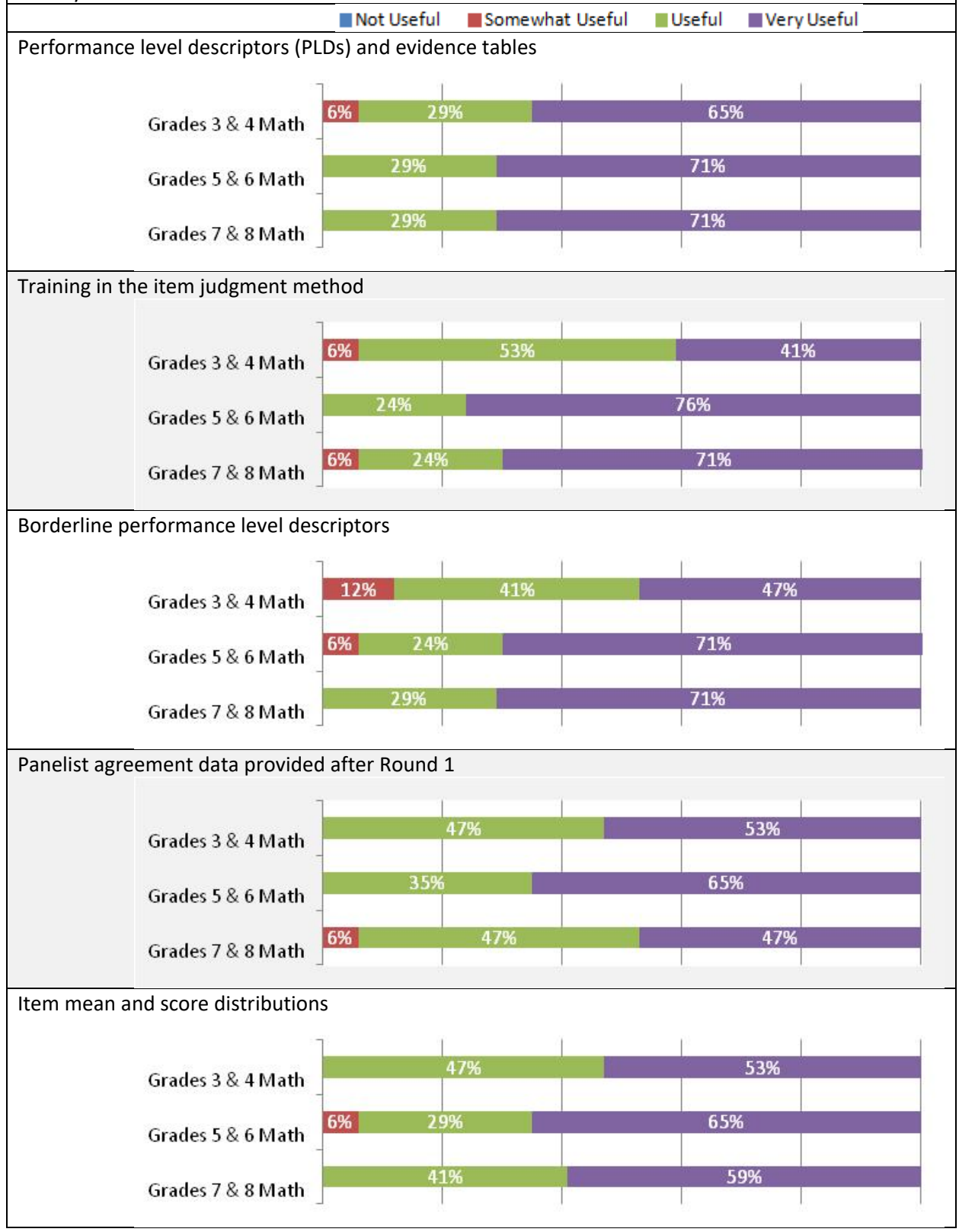

Item difficulty data provided after item judgment rounds 2 and 3

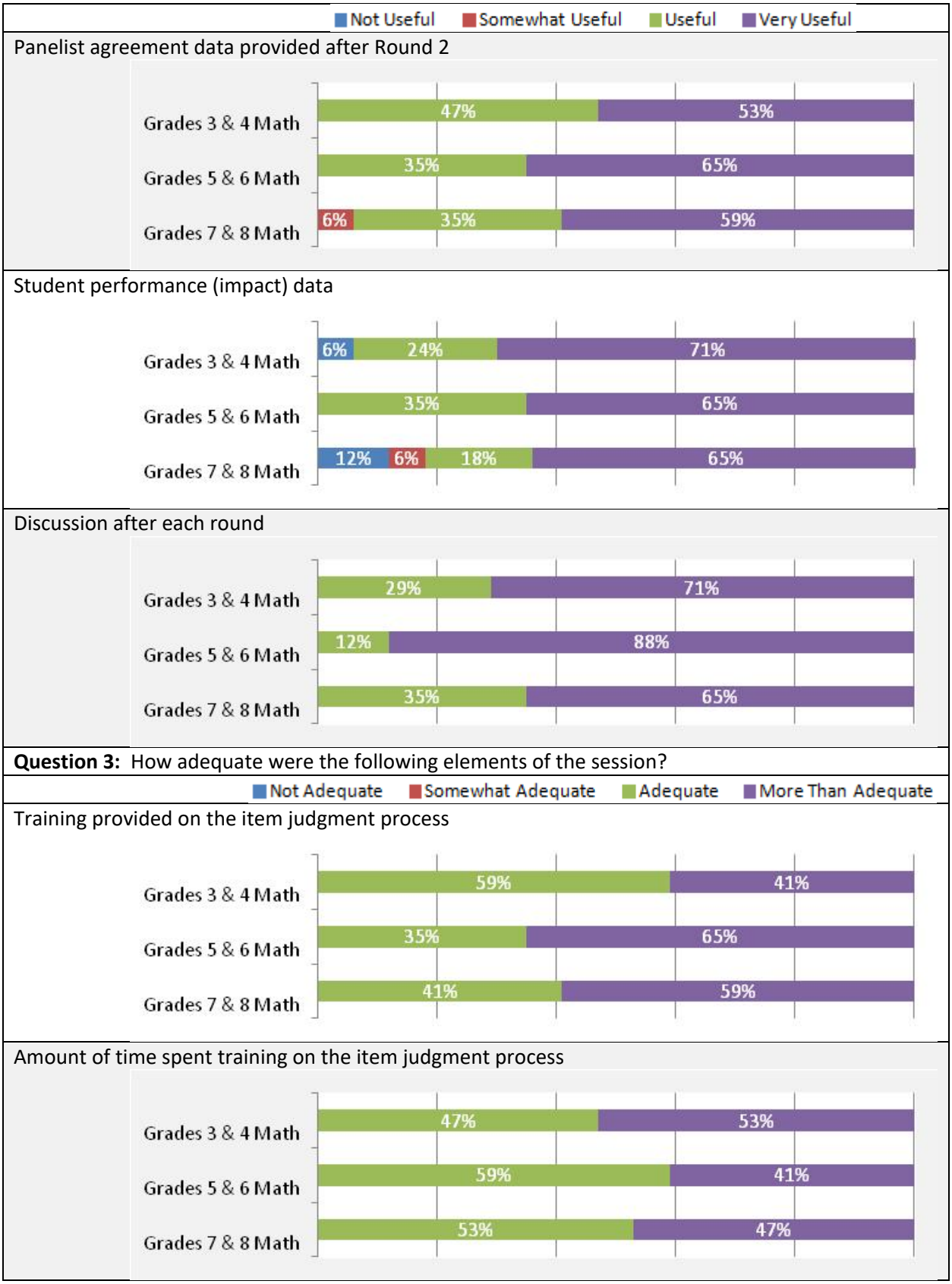

Student performance (impact) data provided after item judgment rounds 2 and 3

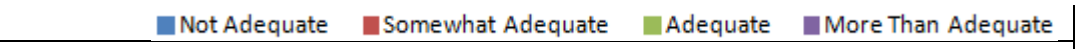

Discussion after each item judgment round



**Question 2:** How useful do you feel the following activities or information were in assisting you to make your recommendations?




Not Useful
Somewhat Useful
Useful
Very Useful
Panelist agreement data provided after Round 2
Grades 3 & 4 Math
47%
53%
Grades 5 & 6 Math
35%
65%
Grades 7 & 8 Math
6%
35%
59%
Student performance (impact) data
Grades 3 & 4 Math
6%
24%
71%
Grades 5 & 6 Math
35%
65%
Grades 7 & 8 Math
12%
6%
18%
65%
Discussion after each round
Grades 3 & 4 Math
29%
71%
Grades 5 & 6 Math
12%
88%
Grades 7 & 8 Math
35%
65%
Question 3: How adequate were the following elements of the session?
Not Adequate
Somewhat Adequate
Adequate
More Than Adequate
Training provided on the item judgment process
Grades 3 & 4 Math
59%
41%
Grades 5 & 6 Math
35%
65%
Grades 7 & 8 Math
41%
59%
Amount of time spent training on the item judgment process
Grades 3 & 4 Math
47%
53%
Grades 5 & 6 Math
59%
41%
Grades 7 & 8 Math
53%
47%



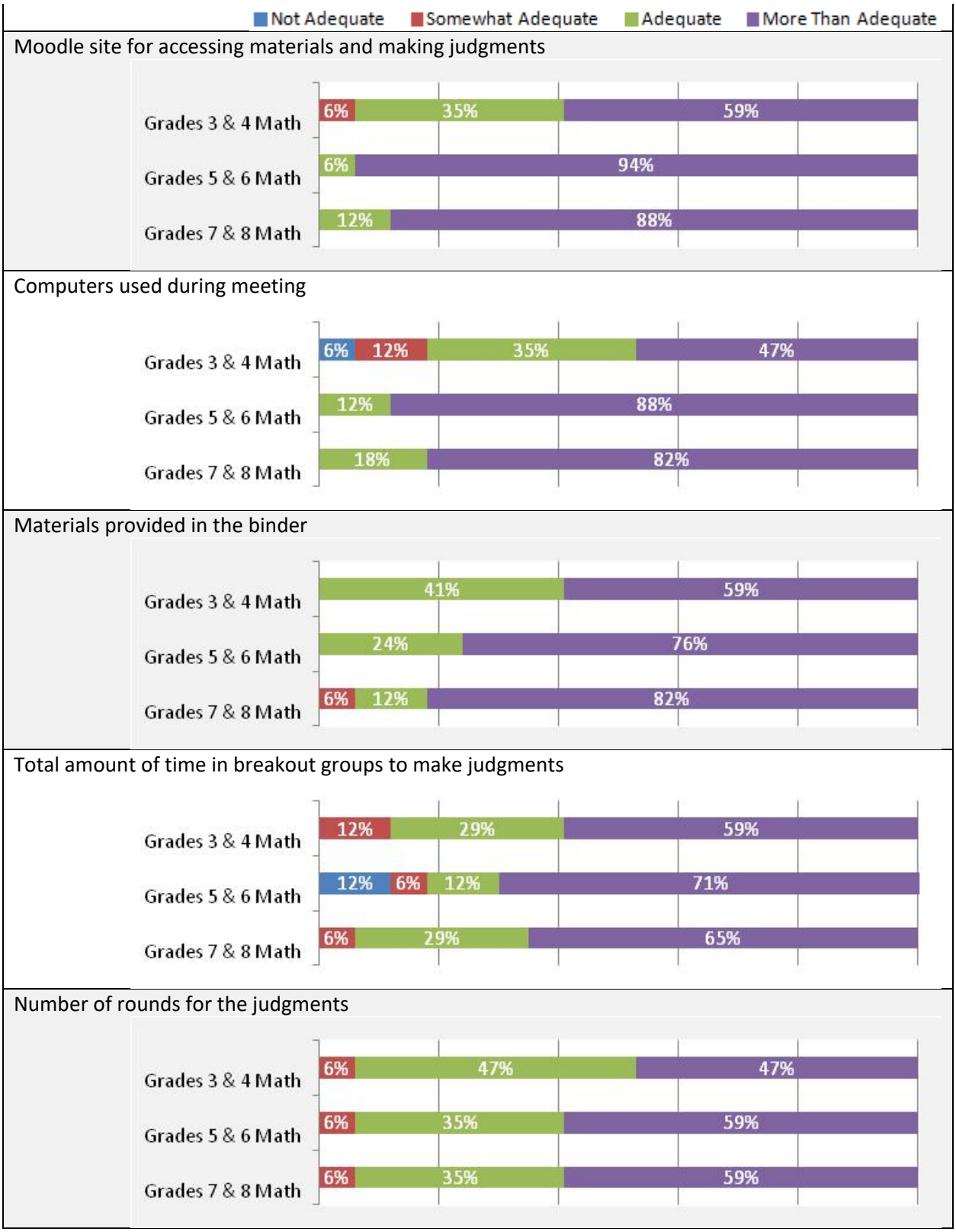

Time spent creating and discussing the borderline performance level descriptors


Visual presentation of feedback provided after Round 1

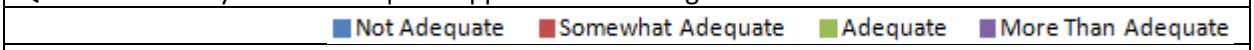

Visual presentation of feedback provided after Round 2


Facilities used for the session


Work space in table groups during meeting


Not Adequate
Somewhat Adequate
Adequate
More Than Adequate
Moodle site for accessing materials and making judgments
Grades 3 & 4 Math
6%
35%
59%
Grades 5 & 6 Math
6%
94%
Grades 7 & 8 Math
12%
88%
Computers used during meeting
Grades 3 & 4 Math
6%
12%
35%
47%
Grades 5 & 6 Math
12%
88%
Grades 7 & 8 Math
18%
82%
Materials provided in the binder
Grades 3 & 4 Math
41%
59%
Grades 5 & 6 Math
24%
76%
Grades 7 & 8 Math
6%
12%
82%
Total amount of time in breakout groups to make judgments
Grades 3 & 4 Math
12%
29%
59%
Grades 5 & 6 Math
12%
6%
12%
71%
Grades 7 & 8 Math
6%
29%
65%
Number of rounds for the judgments
Grades 3 & 4 Math
6%
47%
47%
Grades 5 & 6 Math
6%
35%
59%
Grades 7 & 8 Math
6%
35%
59%

**Question 4:** Did you have adequate opportunities during the session to...



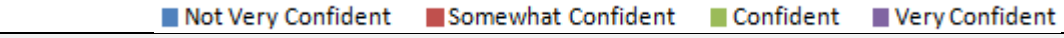

Express your opinions about student performance levels


Ask questions about the standards and how they will be used

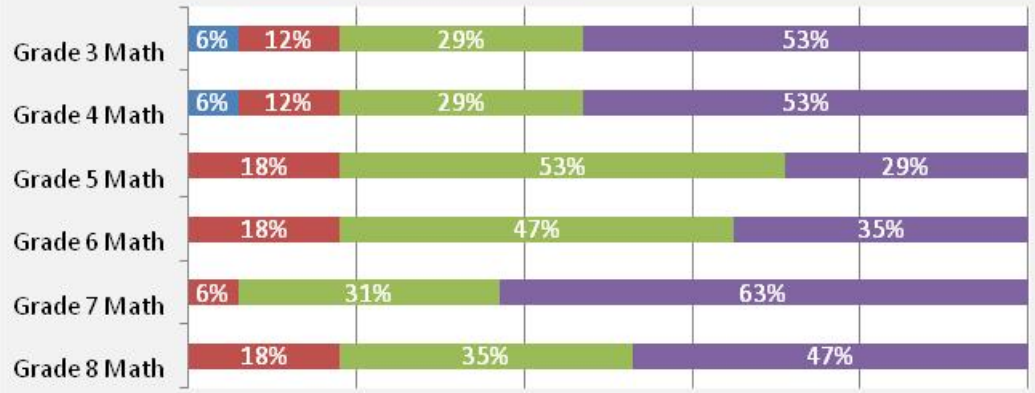

Ask questions about the process of making item judgments


Interaction with your fellow panelists

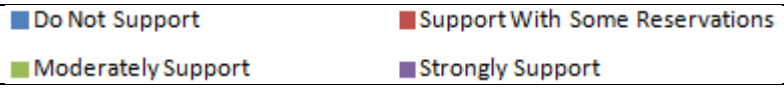
**Question 5:** In applying the performance level setting method, you were asked to recommend scores (separating five performance levels) for student performance on the PARCC assessments. How confident do you feel that the threshold scores based on Round 3 judgments represent appropriate levels of student performance?



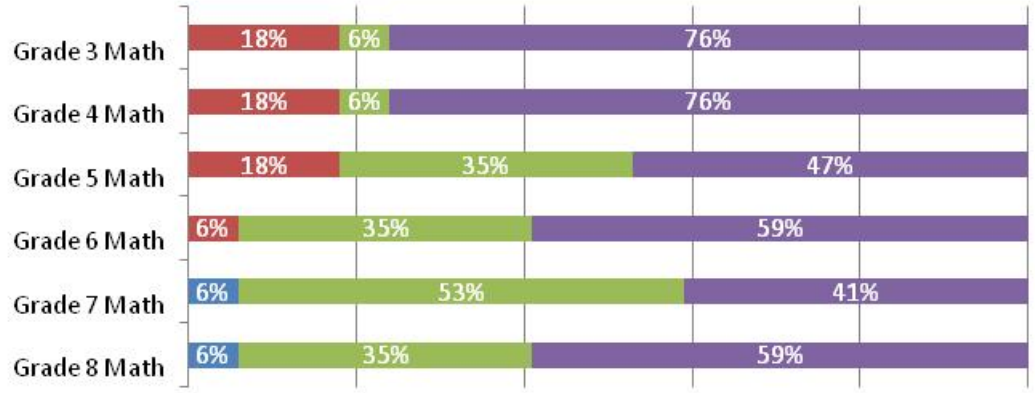

Level 5: Distinguished understanding



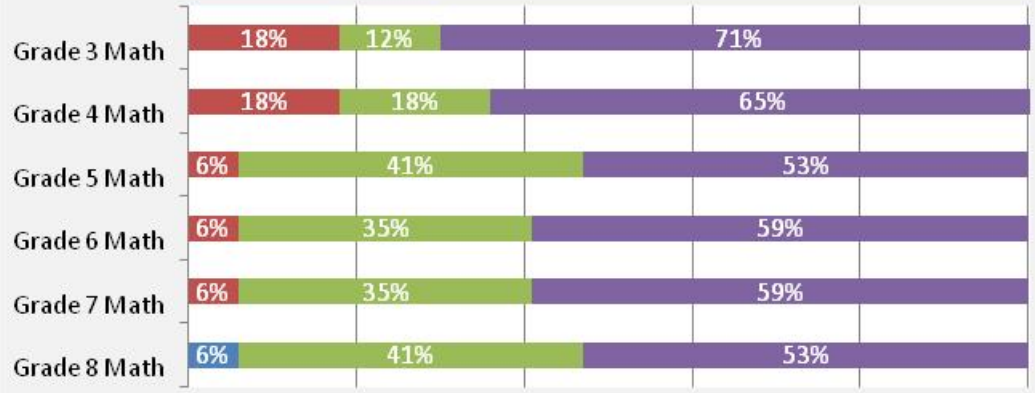

**Question 6:** To what degree do you support the threshold score base on Round 3 judgments for each performance level?

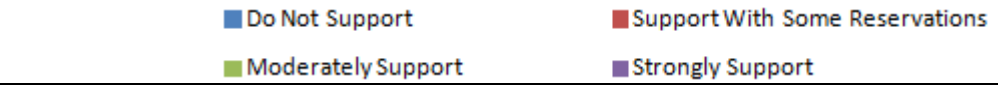



Level 2: Partial understanding

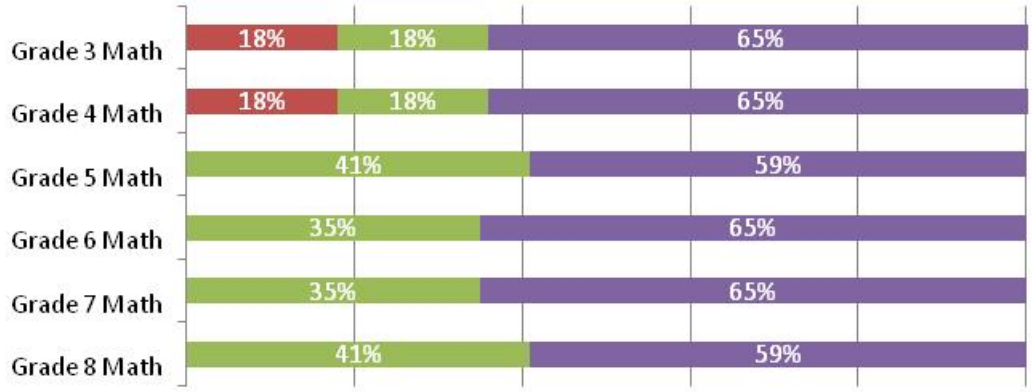



Level 3: Adequate understanding

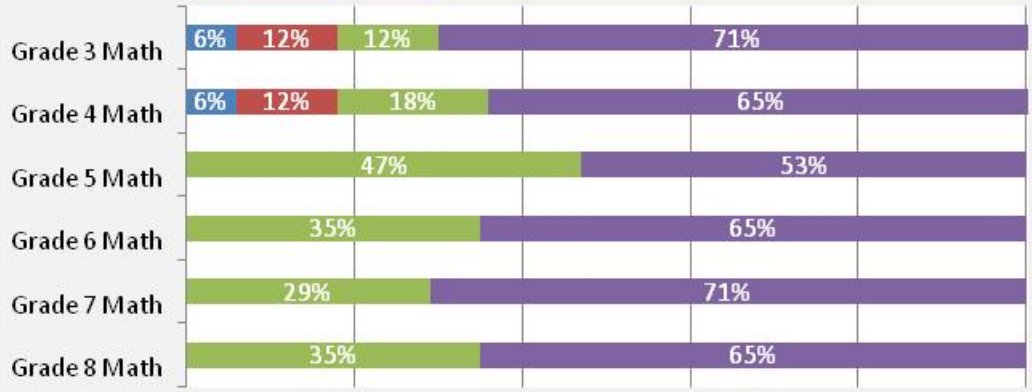




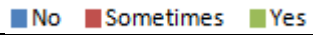

Level 4: Strong understanding




Level 5: Distinguished understanding

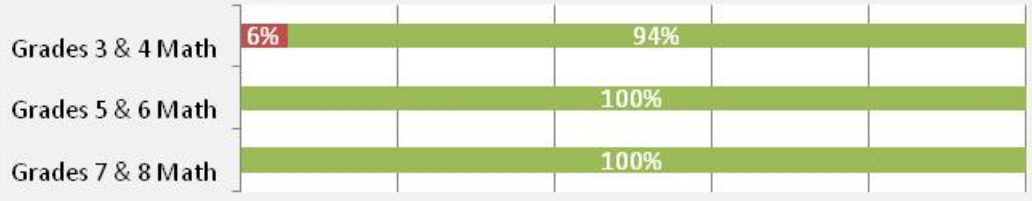



**Question 7:** Do you believe your opinions and judgments were treated with respect by...




Fellow Panelists


Facilitators



# Appendix 9: Summary of Threshold Score Recommendations

This appendix presents a summary of the recommended threshold scores for each performance level for each assessment at various times within the PLS process, including:

- PLS committee round 1 item judgments
- PLS committee round 2 item judgments
- PLS committee round 3 item judgments
- Vertical articulation

The recommended threshold scores are based on the combined Performance Based Assessment (PBA) and End-of-Year (EOY) form used for the PLS process. Since there were some instances in which changes were made to the form used between the PLS meeting and the governing board meeting, the reasonableness review results and final threshold scores are reported in Table 7.5 and 7.6, respectively.

## English Language Arts/Literacy (ELA/L)

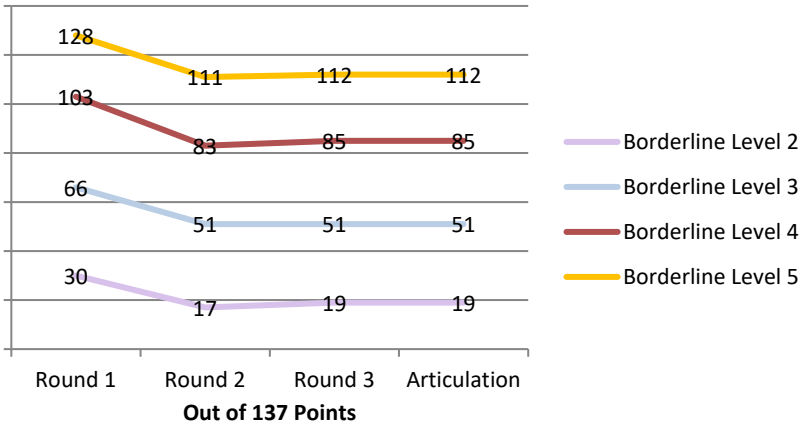



**Figure A8.1:** Grade 11 ELA/L Threshold Scores Across Judgments

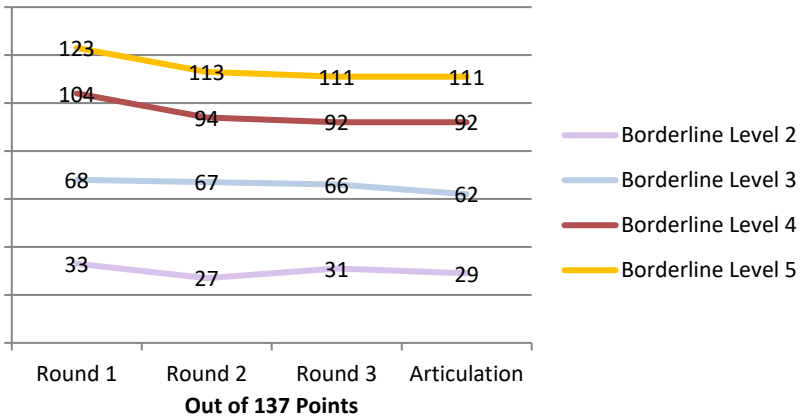



**Figure A8.2:** Grade 10 ELA/L Threshold Scores Across Judgments

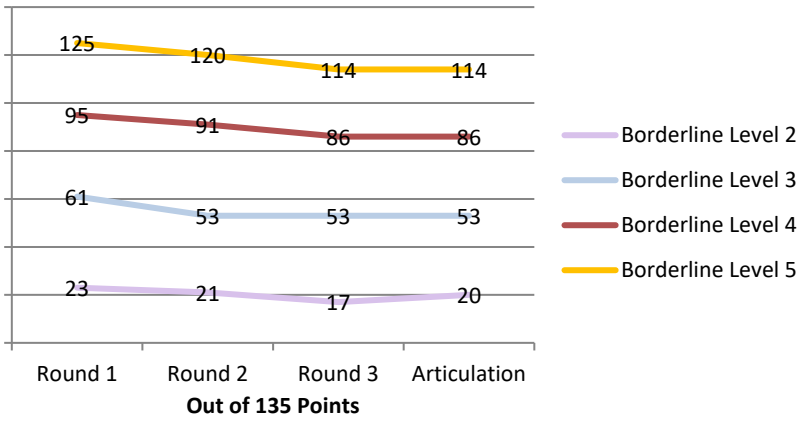



**Figure A8.3:** Grade 9 ELA/L Threshold Scores Across Judgments

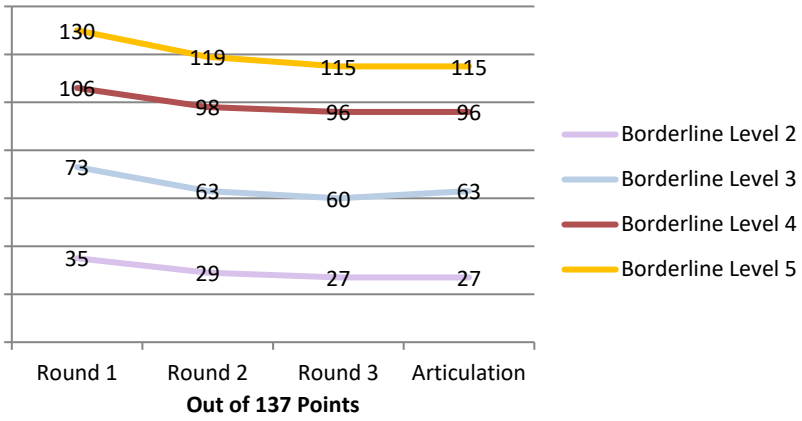



**Figure A8.4:** Grade 8 ELA/L Threshold Scores Across Judgments

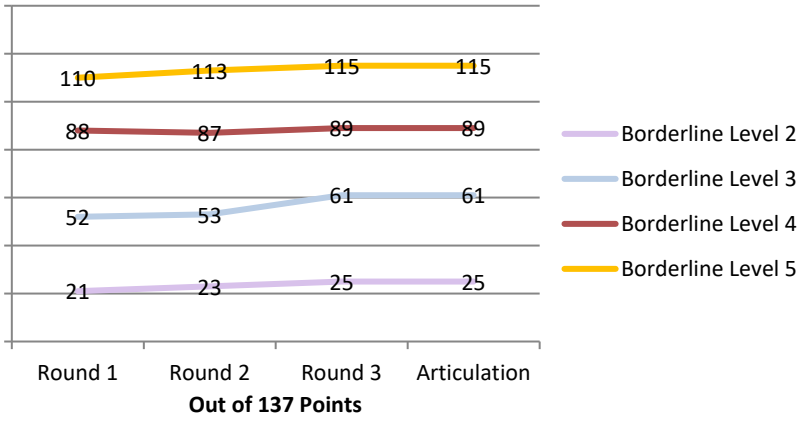



**Figure A8.5:** Grade 7 ELA/L Threshold Scores Across Judgments

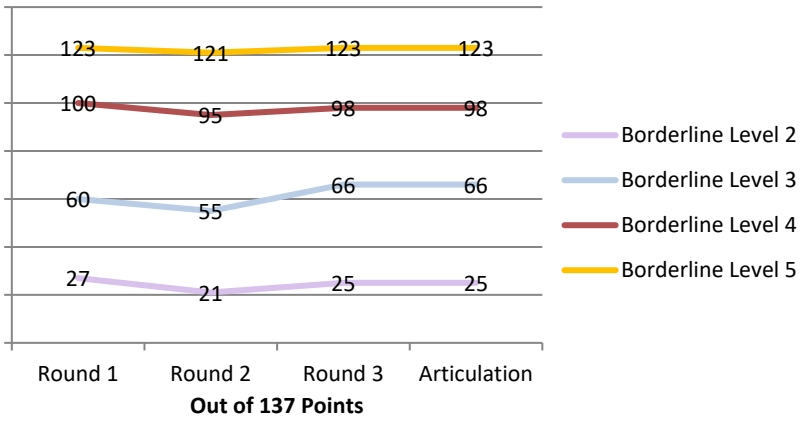



**Figure A8.6:** Grade 6 ELA/L Threshold Scores Across Judgments

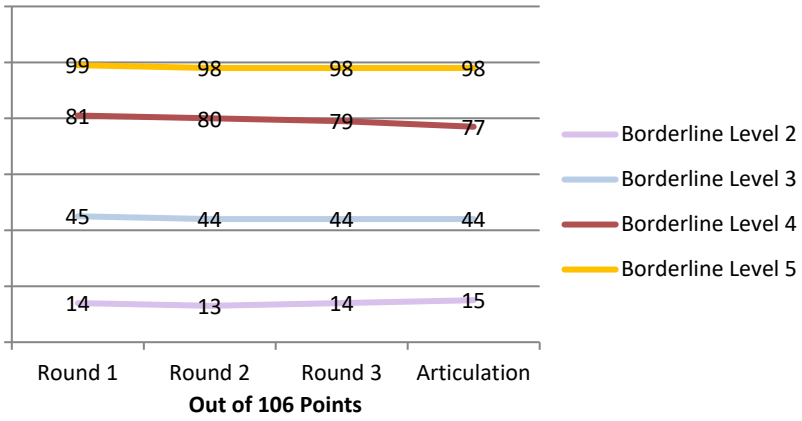



**Figure A8.7:** Grade 5 ELA/L Threshold Scores Across Judgments

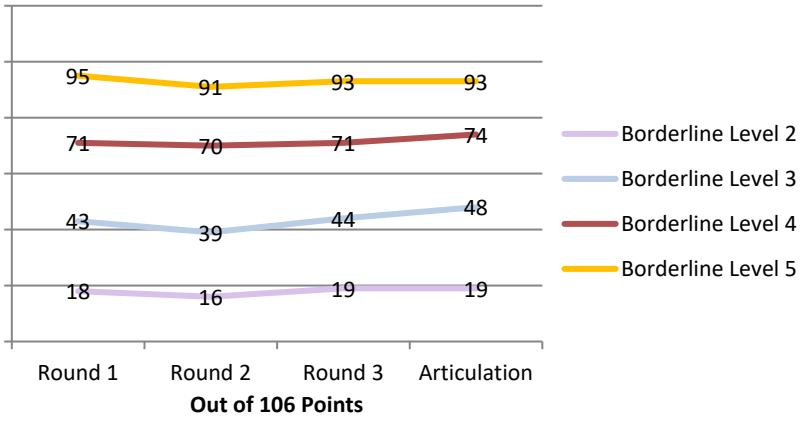



**Figure A8.8:** Grade 4 ELA/L Threshold Scores Across Judgments

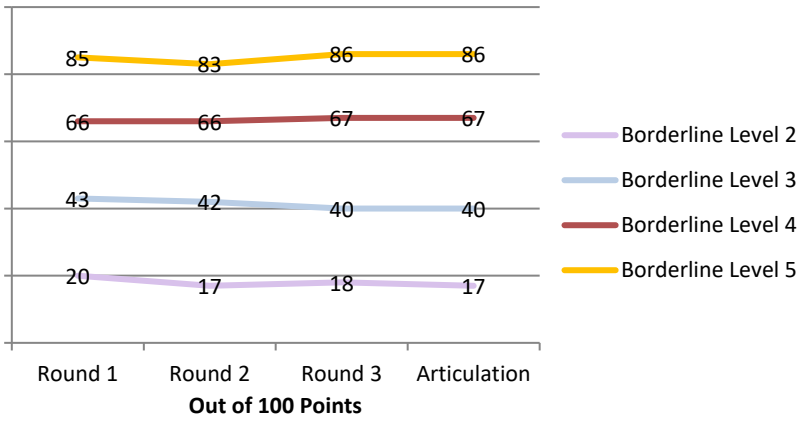



**Figure A8.9:** Grade 3 ELA/L Threshold Scores Across Judgments

## Mathematics

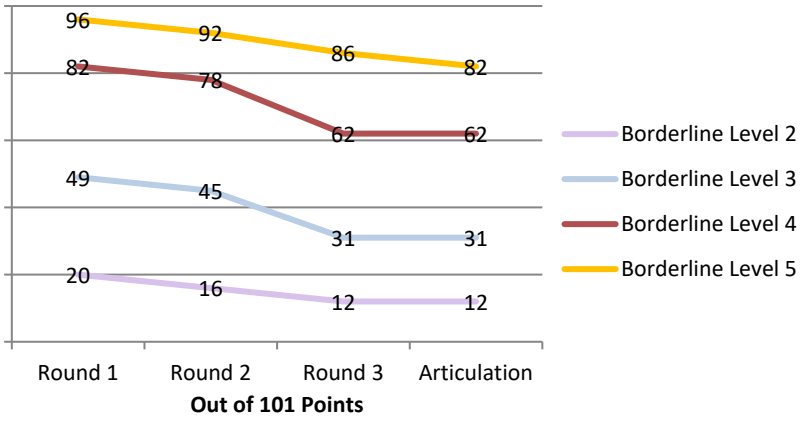



**Figure A8.10:** Algebra II Threshold Scores Across Judgments

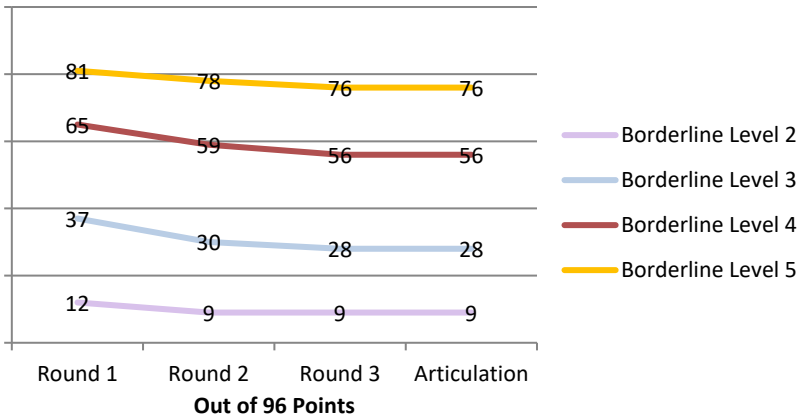



**Figure A8.11:** Geometry Threshold Scores Across Judgments

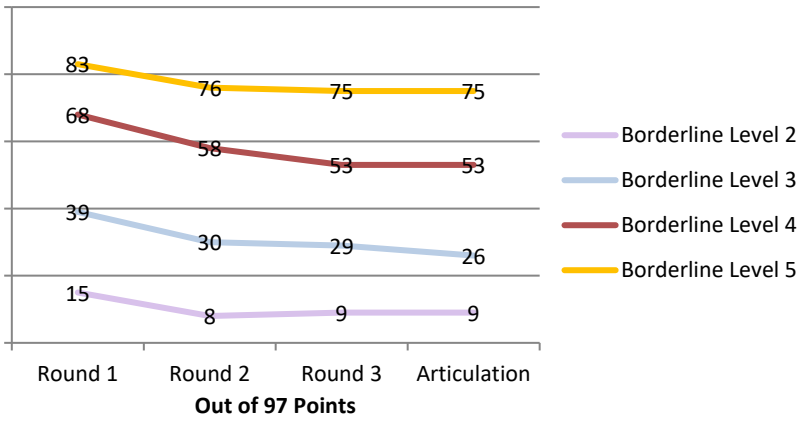



**Figure A8.12:** Algebra I Threshold Scores Across Judgments

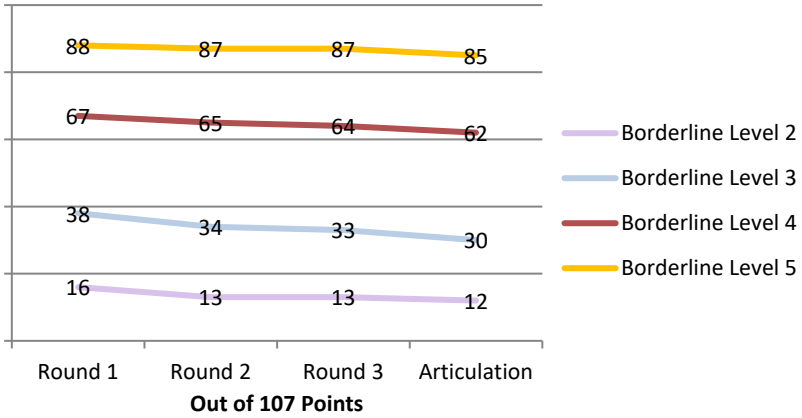



**Figure A8.13:** Integrated Math III Threshold Scores Across Judgments

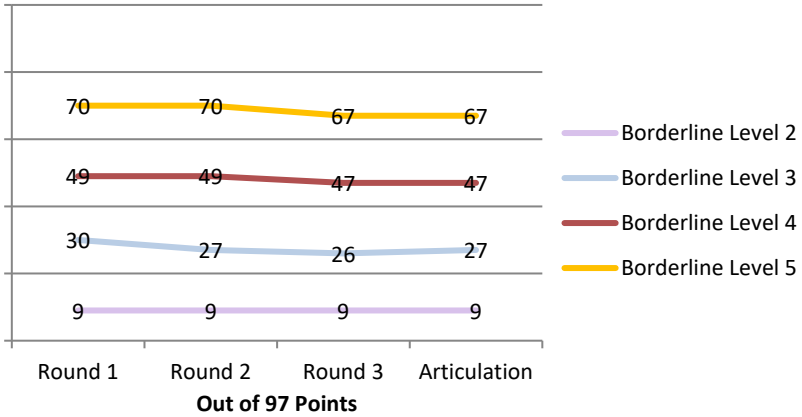



**Figure A8.14:** Integrated Math II Threshold Scores Across Judgments

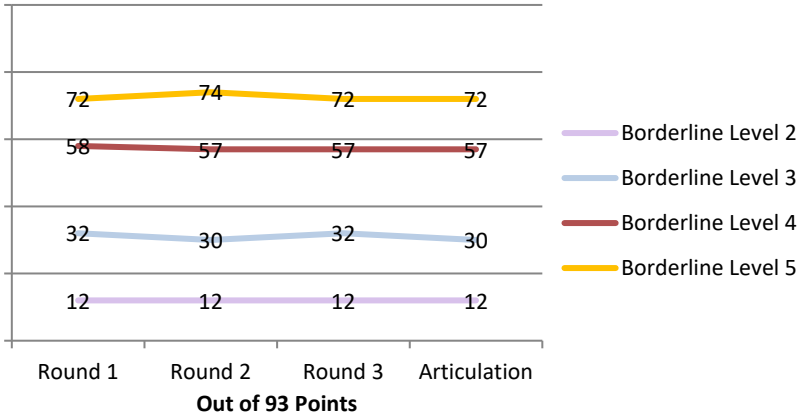



**Figure A8.15:** Integrated Math I Threshold Scores Across Judgments

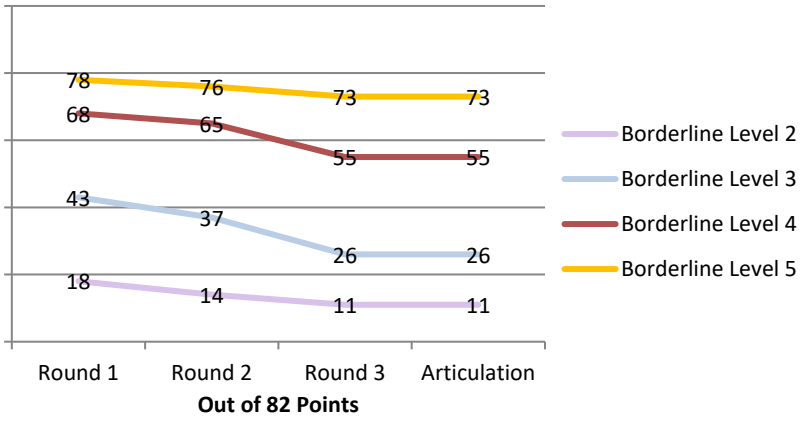



**Figure A8.16:** Grade 8 Mathematics Threshold Scores Across Judgments

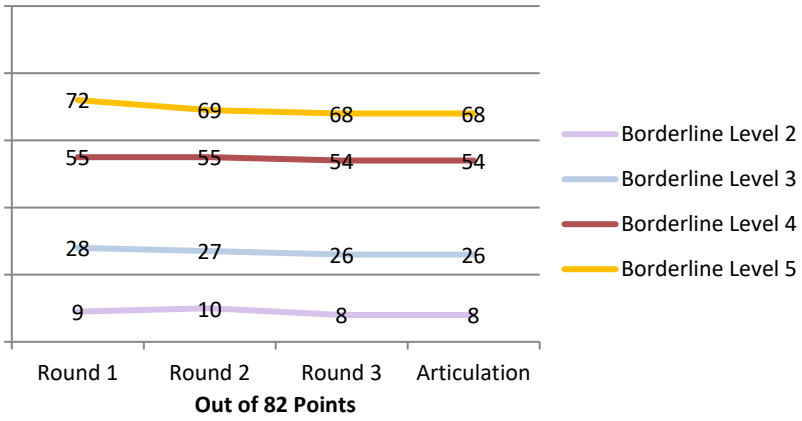



**Figure A8.17:** Grade 7 Mathematics Threshold Scores Across Judgments

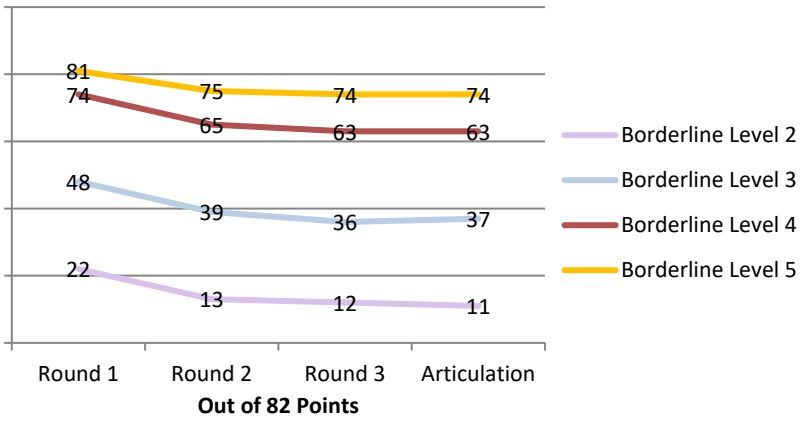



**Figure A8.18:** Grade 6 Mathematics Threshold Scores Across Judgments

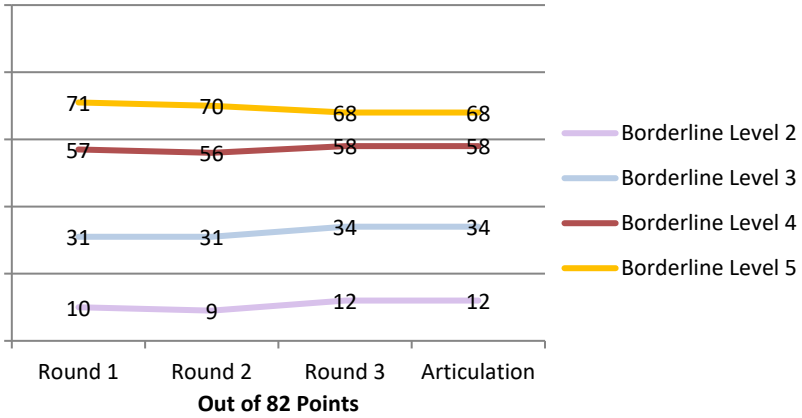



**Figure A8.19:** Grade 5 Mathematics Threshold Scores Across Judgments

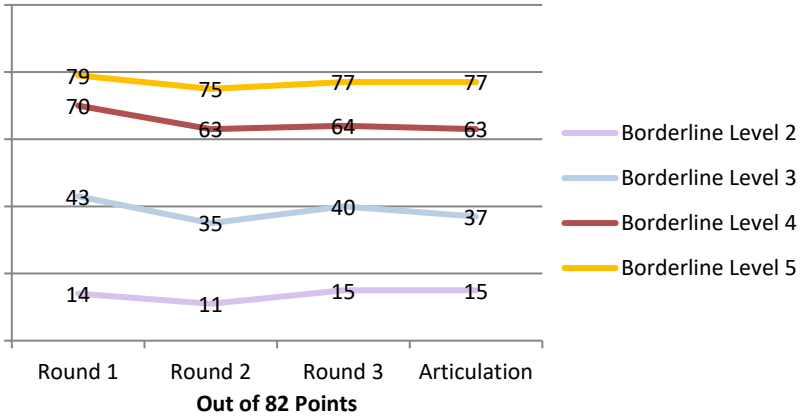



**Figure A8.20:** Grade 4 Mathematics Threshold Scores Across Judgments

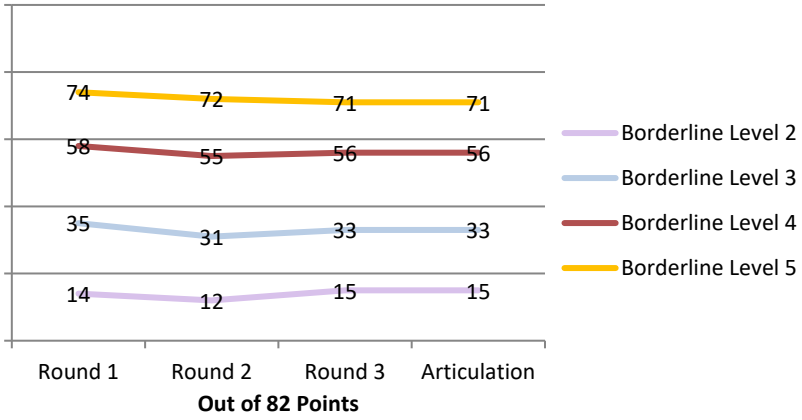



**Figure A8.21:** Grade 3 Mathematics Threshold Scores Across Judgments

# Appendix 10: Summary of Performance Level Setting Panelists' Judgments

## English Language Arts/Literacy (ELA/L)

| Grade 11 English Language Arts/Literacy | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 17 | 31.65 | 30 | 11 | 69 |
| Round 2 | 17 | 15.94 | 17 | 0 | 29 |
| Round 3 | 17 | 19.88 | 19 | 12 | 31 |
| *Performance Level 3* | | | | | |
| Round 1 | 17 | 67.65 | 66 | 43 | 97 |
| Round 2 | 17 | 47.59 | 51 | 15 | 82 |
| Round 3 | 17 | 51.65 | 51 | 41 | 63 |
| *Performance Level 4* | | | | | |
| Round 1 | 17 | 101.00 | 103 | 76 | 120 |
| Round 2 | 17 | 80.65 | 83 | 56 | 107 |
| Round 3 | 17 | 85.47 | 85 | 73 | 96 |
| *Performance Level 5* | | | | | |
| Round 1 | 17 | 124.24 | 128 | 97 | 137 |
| Round 2 | 17 | 109.18 | 111 | 79 | 134 |
| Round 3 | 17 | 111.59 | 112 | 99 | 128 |

| Grade 10 English Language Arts/Literacy | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 17 | 32.00 | 33 | 12 | 49 |
| Round 2 | 17 | 27.76 | 27 | 13 | 48 |
| Round 3 | 17 | 29.18 | 31 | 15 | 45 |
| *Performance Level 3* | | | | | |
| Round 1 | 17 | 68.59 | 68 | 39 | 96 |
| Round 2 | 17 | 61.59 | 67 | 36 | 73 |
| Round 3 | 17 | 63.88 | 66 | 46 | 75 |
| *Performance Level 4* | | | | | |
| Round 1 | 17 | 100.53 | 104 | 81 | 118 |
| Round 2 | 17 | 91.88 | 94 | 70 | 103 |
| Round 3 | 17 | 91.71 | 92 | 77 | 102 |
| *Performance Level 5* | | | | | |
| Round 1 | 17 | 120.06 | 123 | 103 | 135 |
| Round 2 | 17 | 111.18 | 113 | 82 | 124 |
| Round 3 | 17 | 112.24 | 111 | 103 | 130 |

| Grade 9 English Language Arts/Literacy | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 17 | 25.00 | 23 | 12 | 43 |
| Round 2 | 17 | 19.65 | 21 | 1 | 34 |
| Round 3 | 17 | 18.53 | 17 | 9 | 36 |
| *Performance Level 3* | | | | | |
| Round 1 | 17 | 61.41 | 61 | 43 | 78 |
| Round 2 | 17 | 52.12 | 53 | 14 | 70 |
| Round 3 | 17 | 52.24 | 53 | 30 | 69 |
| *Performance Level 4* | | | | | |
| Round 1 | 17 | 96.12 | 95 | 83 | 109 |
| Round 2 | 17 | 88.12 | 91 | 52 | 101 |
| Round 3 | 17 | 84.06 | 86 | 64 | 98 |
| *Performance Level 5* | | | | | |
| Round 1 | 17 | 122.88 | 125 | 101 | 133 |
| Round 2 | 17 | 116.53 | 120 | 98 | 128 |
| Round 3 | 17 | 114.29 | 114 | 101 | 125 |

| Grade 8 English Language Arts/Literacy | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 18 | 35.22 | 35 | 18 | 60 |
| Round 2 | 18 | 27.28 | 29 | 16 | 37 |
| Round 3 | 18 | 25.78 | 27 | 15 | 32 |
| *Performance Level 3* | | | | | |
| Round 1 | 18 | 74.89 | 73 | 58 | 97 |
| Round 2 | 18 | 61.00 | 63 | 41 | 73 |
| Round 3 | 18 | 59.00 | 60 | 48 | 70 |
| *Performance Level 4* | | | | | |
| Round 1 | 18 | 107.56 | 106 | 93 | 125 |
| Round 2 | 18 | 97.78 | 98 | 83 | 108 |
| Round 3 | 18 | 93.89 | 96 | 81 | 106 |
| *Performance Level 5* | | | | | |
| Round 1 | 18 | 127.83 | 130 | 116 | 137 |
| Round 2 | 18 | 119.56 | 119 | 100 | 132 |
| Round 3 | 18 | 115.22 | 115 | 102 | 124 |

| Grade 7 English Language Arts/Literacy | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 17 | 23.71 | 21 | 2 | 56 |
| Round 2 | 17 | 23.71 | 23 | 18 | 39 |
| Round 3 | 16 | 23.06 | 25 | 11 | 30 |
| *Performance Level 3* | | | | | |
| Round 1 | 17 | 52.76 | 52 | 35 | 86 |
| Round 2 | 17 | 56.12 | 53 | 40 | 74 |
| Round 3 | 16 | 60.88 | 61 | 50 | 77 |
| *Performance Level 4* | | | | | |
| Round 1 | 17 | 86.12 | 88 | 68 | 112 |
| Round 2 | 17 | 88.29 | 87 | 75 | 101 |
| Round 3 | 16 | 89.81 | 89 | 83 | 98 |
| *Performance Level 5* | | | | | |
| Round 1 | 17 | 109.59 | 110 | 97 | 125 |
| Round 2 | 17 | 112.47 | 113 | 94 | 127 |
| Round 3 | 16 | 113.38 | 115 | 101 | 123 |

| Grade 6 English Language Arts/Literacy | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 15 | 25.93 | 27 | 14 | 37 |
| Round 2 | 15 | 21.93 | 21 | 9 | 34 |
| Round 3 | 15 | 25.20 | 25 | 15 | 39 |
| *Performance Level 3* | | | | | |
| Round 1 | 15 | 61.40 | 60 | 50 | 79 |
| Round 2 | 15 | 54.27 | 55 | 38 | 69 |
| Round 3 | 15 | 62.73 | 66 | 47 | 80 |
| *Performance Level 4* | | | | | |
| Round 1 | 15 | 99.53 | 100 | 73 | 111 |
| Round 2 | 15 | 92.33 | 95 | 68 | 105 |
| Round 3 | 15 | 98.13 | 98 | 83 | 108 |
| *Performance Level 5* | | | | | |
| Round 1 | 15 | 123.60 | 123 | 108 | 137 |
| Round 2 | 15 | 119.80 | 121 | 106 | 132 |
| Round 3 | 15 | 123.07 | 123 | 113 | 132 |

| Grade 5 English Language Arts/Literacy | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 15 | 16.20 | 14 | 6 | 36 |
| Round 2 | 15 | 14.80 | 13 | 8 | 28 |
| Round 3 | 15 | 13.53 | 14 | 9 | 17 |
| *Performance Level 3* | | | | | |
| Round 1 | 15 | 47.20 | 45 | 33 | 78 |
| Round 2 | 15 | 46.47 | 44 | 36 | 66 |
| Round 3 | 15 | 43.93 | 44 | 38 | 50 |
| *Performance Level 4* | | | | | |
| Round 1 | 15 | 81.00 | 81 | 69 | 93 |
| Round 2 | 15 | 80.53 | 80 | 73 | 92 |
| Round 3 | 15 | 78.33 | 79 | 72 | 84 |
| *Performance Level 5* | | | | | |
| Round 1 | 15 | 97.60 | 99 | 88 | 106 |
| Round 2 | 15 | 98.13 | 98 | 90 | 103 |
| Round 3 | 15 | 97.40 | 98 | 91 | 103 |

| Grade 4 English Language Arts/Literacy | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 15 | 17.27 | 18 | 6 | 25 |
| Round 2 | 15 | 15.40 | 16 | 3 | 24 |
| Round 3 | 15 | 19.07 | 19 | 14 | 24 |
| *Performance Level 3* | | | | | |
| Round 1 | 15 | 42.87 | 43 | 31 | 55 |
| Round 2 | 15 | 38.73 | 39 | 11 | 52 |
| Round 3 | 15 | 43.47 | 44 | 36 | 51 |
| *Performance Level 4* | | | | | |
| Round 1 | 15 | 70.80 | 71 | 60 | 65 |
| Round 2 | 15 | 66.40 | 70 | 34 | 76 |
| Round 3 | 15 | 70.33 | 71 | 61 | 78 |
| *Performance Level 5* | | | | | |
| Round 1 | 15 | 95.53 | 95 | 83 | 106 |
| Round 2 | 15 | 88.80 | 91 | 56 | 98 |
| Round 3 | 15 | 92.93 | 93 | 83 | 99 |

| Grade 3 English Language Arts/Literacy | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 14 | 18.86 | 20 | 9 | 25 |
| Round 2 | 14 | 16.57 | 17 | 11 | 22 |
| Round 3 | 13 | 17.46 | 18 | 12 | 21 |
| *Performance Level 3* | | | | | |
| Round 1 | 14 | 40.43 | 43 | 26 | 52 |
| Round 2 | 14 | 39.57 | 42 | 26 | 46 |
| Round 3 | 13 | 39.15 | 40 | 30 | 46 |
| *Performance Level 4* | | | | | |
| Round 1 | 14 | 66.00 | 66 | 55 | 78 |
| Round 2 | 14 | 64.79 | 66 | 51 | 74 |
| Round 3 | 13 | 65.77 | 67 | 56 | 72 |
| *Performance Level 5* | | | | | |
| Round 1 | 14 | 83.93 | 85 | 72 | 97 |
| Round 2 | 14 | 82.64 | 83 | 71 | 92 |
| Round 3 | 13 | 83.62 | 86 | 71 | 96 |

## Mathematics

| Algebra II | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 20 | 20.25 | 20 | 4 | 36 |
| Round 2 | 20 | 15.65 | 16 | 6 | 29 |
| Round 3 | 20 | 11.00 | 12 | 0 | 19 |
| *Performance Level 3* | | | | | |
| Round 1 | 20 | 49.95 | 49 | 30 | 74 |
| Round 2 | 20 | 42.80 | 45 | 25 | 52 |
| Round 3 | 20 | 31.90 | 31 | 5 | 49 |
| *Performance Level 4* | | | | | |
| Round 1 | 20 | 81.05 | 82 | 70 | 96 |
| Round 2 | 20 | 74.35 | 78 | 51 | 88 |
| Round 3 | 20 | 62.40 | 62 | 29 | 80 |
| *Performance Level 5* | | | | | |
| Round 1 | 20 | 95.15 | 96 | 87 | 100 |
| Round 2 | 20 | 90.60 | 92 | 77 | 99 |
| Round 3 | 20 | 83.60 | 86 | 66 | 96 |

| Geometry | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 17 | 13.12 | 12 | 8 | 23 |
| Round 2 | 17 | 10.00 | 9 | 5 | 18 |
| Round 3 | 17 | 9.71 | 9 | 7 | 18 |
| *Performance Level 3* | | | | | |
| Round 1 | 17 | 38.18 | 37 | 28 | 46 |
| Round 2 | 17 | 30.18 | 30 | 18 | 52 |
| Round 3 | 17 | 27.18 | 28 | 19 | 34 |
| *Performance Level 4* | | | | | |
| Round 1 | 17 | 66.41 | 65 | 56 | 83 |
| Round 2 | 17 | 58.00 | 59 | 37 | 74 |
| Round 3 | 17 | 56.24 | 56 | 46 | 67 |
| *Performance Level 5* | | | | | |
| Round 1 | 17 | 81.59 | 81 | 73 | 91 |
| Round 2 | 17 | 76.29 | 78 | 52 | 87 |
| Round 3 | 17 | 74.24 | 76 | 64 | 79 |

| Algebra I | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 17 | 18.00 | 15 | 4 | 58 |
| Round 2 | 17 | 7.53 | 8 | 0 | 21 |
| Round 3 | 17 | 8.29 | 9 | 3 | 12 |
| *Performance Level 3* | | | | | |
| Round 1 | 17 | 42.29 | 39 | 25 | 63 |
| Round 2 | 17 | 28.00 | 30 | 3 | 42 |
| Round 3 | 17 | 28.65 | 29 | 11 | 47 |
| *Performance Level 4* | | | | | |
| Round 1 | 17 | 69.53 | 68 | 52 | 84 |
| Round 2 | 17 | 53.18 | 58 | 20 | 71 |
| Round 3 | 17 | 55.24 | 53 | 44 | 70 |
| *Performance Level 5* | | | | | |
| Round 1 | 17 | 83.41 | 83 | 73 | 93 |
| Round 2 | 17 | 72.35 | 76 | 40 | 85 |
| Round 3 | 17 | 75.53 | 75 | 65 | 85 |

| Integrated Math III | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 20 | 16.15 | 16 | 9 | 28 |
| Round 2 | 19 | 12.11 | 13 | 2 | 16 |
| Round 3 | 19 | 12.32 | 13 | 7 | 16 |
| *Performance Level 3* | | | | | |
| Round 1 | 20 | 37.20 | 38 | 22 | 49 |
| Round 2 | 19 | 32.42 | 34 | 11 | 43 |
| Round 3 | 19 | 30.84 | 33 | 17 | 40 |
| *Performance Level 4* | | | | | |
| Round 1 | 20 | 67.00 | 67 | 54 | 81 |
| Round 2 | 19 | 62.84 | 65 | 46 | 73 |
| Round 3 | 19 | 61.16 | 64 | 45 | 72 |
| *Performance Level 5* | | | | | |
| Round 1 | 20 | 88.55 | 88 | 75 | 100 |
| Round 2 | 19 | 84.11 | 87 | 66 | 95 |
| Round 3 | 19 | 83.16 | 87 | 64 | 92 |

| Integrated Math II | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 15 | 9.67 | 9 | 5 | 17 |
| Round 2 | 17 | 9.88 | 9 | 5 | 20 |
| Round 3 | 17 | 8.76 | 9 | 5 | 13 |
| *Performance Level 3* | | | | | |
| Round 1 | 15 | 28.53 | 30 | 19 | 38 |
| Round 2 | 17 | 28.06 | 27 | 23 | 37 |
| Round 3 | 17 | 26.35 | 26 | 22 | 33 |
| *Performance Level 4* | | | | | |
| Round 1 | 15 | 50.53 | 49 | 40 | 58 |
| Round 2 | 17 | 48.65 | 49 | 42 | 59 |
| Round 3 | 17 | 47.06 | 47 | 43 | 57 |
| *Performance Level 5* | | | | | |
| Round 1 | 15 | 69.07 | 70 | 60 | 77 |
| Round 2 | 17 | 68.06 | 70 | 61 | 76 |
| Round 3 | 17 | 66.82 | 67 | 63 | 72 |

| Integrated Math I | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 17 | 11.88 | 12 | 1 | 25 |
| Round 2 | 17 | 11.12 | 12 | 0 | 22 |
| Round 3 | 17 | 11.29 | 12 | 2 | 22 |
| *Performance Level 3* | | | | | |
| Round 1 | 17 | 32.53 | 32 | 7 | 61 |
| Round 2 | 17 | 31.18 | 30 | 13 | 48 |
| Round 3 | 17 | 31.71 | 32 | 14 | 48 |
| *Performance Level 4* | | | | | |
| Round 1 | 17 | 58.18 | 58 | 33 | 78 |
| Round 2 | 17 | 58.65 | 57 | 48 | 70 |
| Round 3 | 17 | 57.65 | 57 | 45 | 72 |
| *Performance Level 5* | | | | | |
| Round 1 | 17 | 72.06 | 72 | 46 | 85 |
| Round 2 | 17 | 73.71 | 74 | 67 | 80 |
| Round 3 | 17 | 73.00 | 72 | 64 | 82 |

| Grade 8 Mathematics | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 20 | 17.05 | 18 | 5 | 33 |
| Round 2 | 20 | 13.85 | 14 | 5 | 31 |
| Round 3 | 20 | 11.80 | 11 | 5 | 22 |
| *Performance Level 3* | | | | | |
| Round 1 | 20 | 41.25 | 43 | 21 | 58 |
| Round 2 | 20 | 36.50 | 37 | 22 | 55 |
| Round 3 | 20 | 29.10 | 26 | 15 | 43 |
| *Performance Level 4* | | | | | |
| Round 1 | 20 | 65.80 | 68 | 45 | 76 |
| Round 2 | 20 | 62.90 | 65 | 51 | 74 |
| Round 3 | 20 | 56.10 | 55 | 42 | 70 |
| *Performance Level 5* | | | | | |
| Round 1 | 20 | 77.80 | 78 | 71 | 82 |
| Round 2 | 20 | 75.90 | 76 | 69 | 80 |
| Round 3 | 20 | 71.55 | 73 | 62 | 80 |

| Grade 7 Mathematics | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 19 | 10.21 | 9 | 5 | 23 |
| Round 2 | 18 | 9.50 | 10 | 4 | 21 |
| Round 3 | 18 | 9.28 | 8 | 6 | 20 |
| *Performance Level 3* | | | | | |
| Round 1 | 19 | 30.37 | 28 | 17 | 50 |
| Round 2 | 18 | 27.94 | 27 | 17 | 40 |
| Round 3 | 18 | 26.67 | 26 | 18 | 38 |
| *Performance Level 4* | | | | | |
| Round 1 | 19 | 55.05 | 55 | 38 | 66 |
| Round 2 | 18 | 53.00 | 55 | 36 | 62 |
| Round 3 | 18 | 52.33 | 54 | 37 | 60 |
| *Performance Level 5* | | | | | |
| Round 1 | 19 | 70.37 | 72 | 58 | 77 |
| Round 2 | 18 | 67.83 | 69 | 58 | 75 |
| Round 3 | 18 | 67.33 | 68 | 59 | 73 |

| Grade 6 Mathematics | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 17 | 23.12 | 22 | 10 | 45 |
| Round 2 | 17 | 13.29 | 13 | 7 | 22 |
| Round 3 | 17 | 11.71 | 12 | 7 | 18 |
| *Performance Level 3* | | | | | |
| Round 1 | 17 | 49.12 | 48 | 31 | 66 |
| Round 2 | 17 | 39.47 | 39 | 26 | 50 |
| Round 3 | 17 | 37.29 | 36 | 26 | 54 |
| *Performance Level 4* | | | | | |
| Round 1 | 17 | 73.06 | 74 | 65 | 82 |
| Round 2 | 17 | 64.12 | 65 | 47 | 79 |
| Round 3 | 17 | 62.41 | 63 | 49 | 69 |
| *Performance Level 5* | | | | | |
| Round 1 | 17 | 80.18 | 81 | 73 | 82 |
| Round 2 | 17 | 74.41 | 75 | 58 | 81 |
| Round 3 | 17 | 73.59 | 74 | 59 | 78 |

| Grade 5 Mathematics | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 17 | 11.00 | 10 | 5 | 23 |
| Round 2 | 17 | 10.12 | 9 | 3 | 21 |
| Round 3 | 16 | 11.88 | 12 | 4 | 23 |
| *Performance Level 3* | | | | | |
| Round 1 | 17 | 32.00 | 31 | 20 | 50 |
| Round 2 | 17 | 31.12 | 31 | 19 | 43 |
| Round 3 | 16 | 33.25 | 34 | 23 | 45 |
| *Performance Level 4* | | | | | |
| Round 1 | 17 | 56.82 | 57 | 44 | 73 |
| Round 2 | 17 | 56.71 | 56 | 49 | 64 |
| Round 3 | 16 | 57.38 | 58 | 49 | 72 |
| *Performance Level 5* | | | | | |
| Round 1 | 17 | 70.53 | 71 | 63 | 79 |
| Round 2 | 17 | 70.82 | 70 | 64 | 79 |
| Round 3 | 16 | 69.69 | 68 | 64 | 80 |

| Grade 4 Mathematics | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 17 | 15.71 | 14 | 2 | 32 |
| Round 2 | 17 | 10.94 | 11 | 2 | 17 |
| Round 3 | 17 | 14.47 | 15 | 6 | 31 |
| *Performance Level 3* | | | | | |
| Round 1 | 17 | 42.59 | 43 | 22 | 59 |
| Round 2 | 17 | 36.00 | 35 | 16 | 52 |
| Round 3 | 17 | 40.88 | 40 | 26 | 52 |
| *Performance Level 4* | | | | | |
| Round 1 | 17 | 68.94 | 70 | 59 | 79 |
| Round 2 | 17 | 62.88 | 63 | 49 | 72 |
| Round 3 | 17 | 64.71 | 64 | 57 | 72 |
| *Performance Level 5* | | | | | |
| Round 1 | 17 | 78.65 | 79 | 69 | 82 |
| Round 2 | 17 | 75.35 | 75 | 69 | 80 |
| Round 3 | 17 | 75.71 | 77 | 72 | 80 |

| Grade 3 Mathematics | | | | | |
|---|---|---|---|---|---|
| | **Count** | **Mean** | **Median** | **Minimum** | **Maximum** |
| *Performance Level 2* | | | | | |
| Round 1 | 17 | 14.65 | 14 | 5 | 28 |
| Round 2 | 17 | 13.29 | 12 | 8 | 23 |
| Round 3 | 17 | 14.18 | 15 | 9 | 21 |
| *Performance Level 3* | | | | | |
| Round 1 | 17 | 35.00 | 35 | 21 | 56 |
| Round 2 | 17 | 31.76 | 31 | 19 | 49 |
| Round 3 | 17 | 32.88 | 33 | 24 | 49 |
| *Performance Level 4* | | | | | |
| Round 1 | 17 | 58.41 | 58 | 43 | 80 |
| Round 2 | 17 | 55.71 | 55 | 49 | 66 |
| Round 3 | 17 | 56.41 | 56 | 51 | 65 |
| *Performance Level 5* | | | | | |
| Round 1 | 17 | 73.06 | 74 | 58 | 82 |
| Round 2 | 17 | 70.76 | 72 | 65 | 78 |
| Round 3 | 17 | 71.53 | 71 | 67 | 79 |

# Appendix 11: Performance Level Setting Panelists' Agreement Data

## English Language Arts/Literacy (ELA/L)

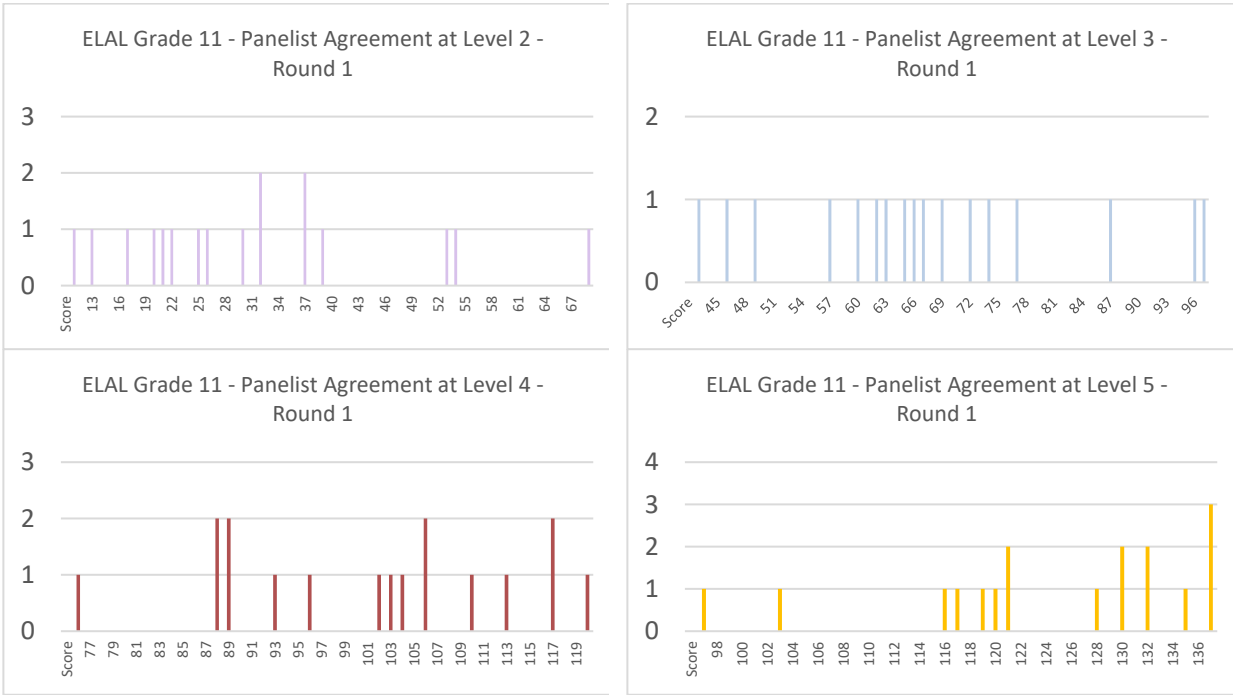



**Figure A10.1:** Grade 11 ELA/L - Round 1

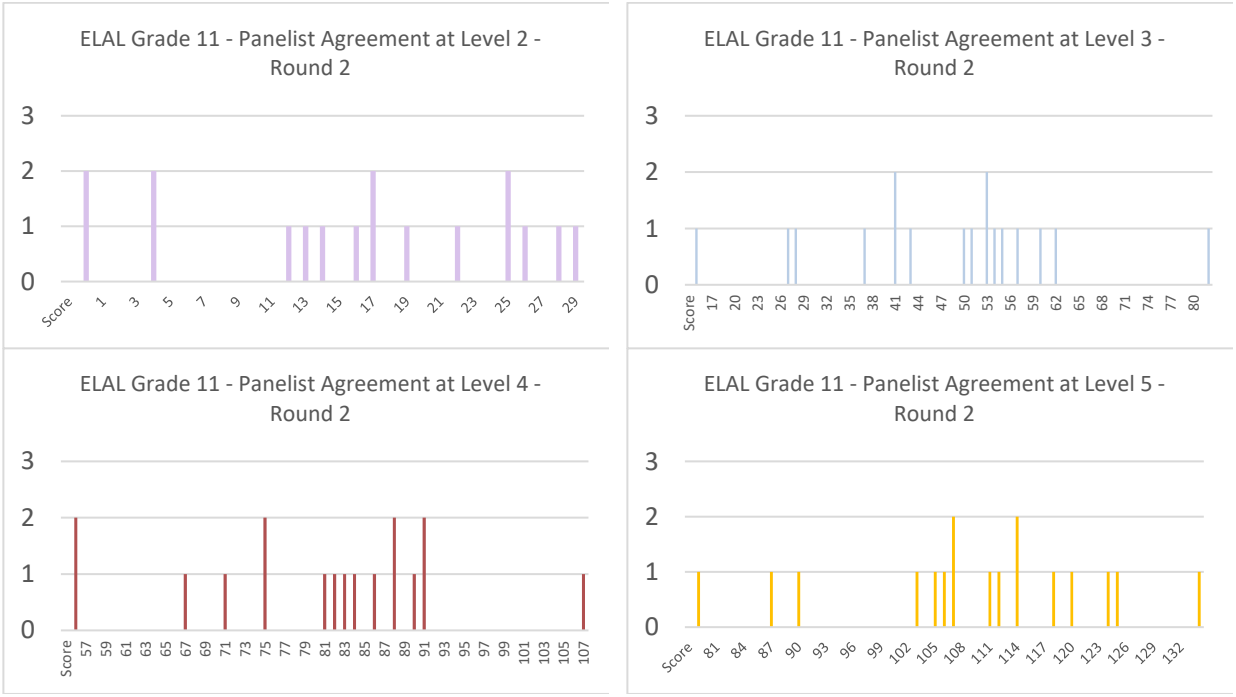



**Figure A10.2:** Grade 11 ELA/L - Round 2

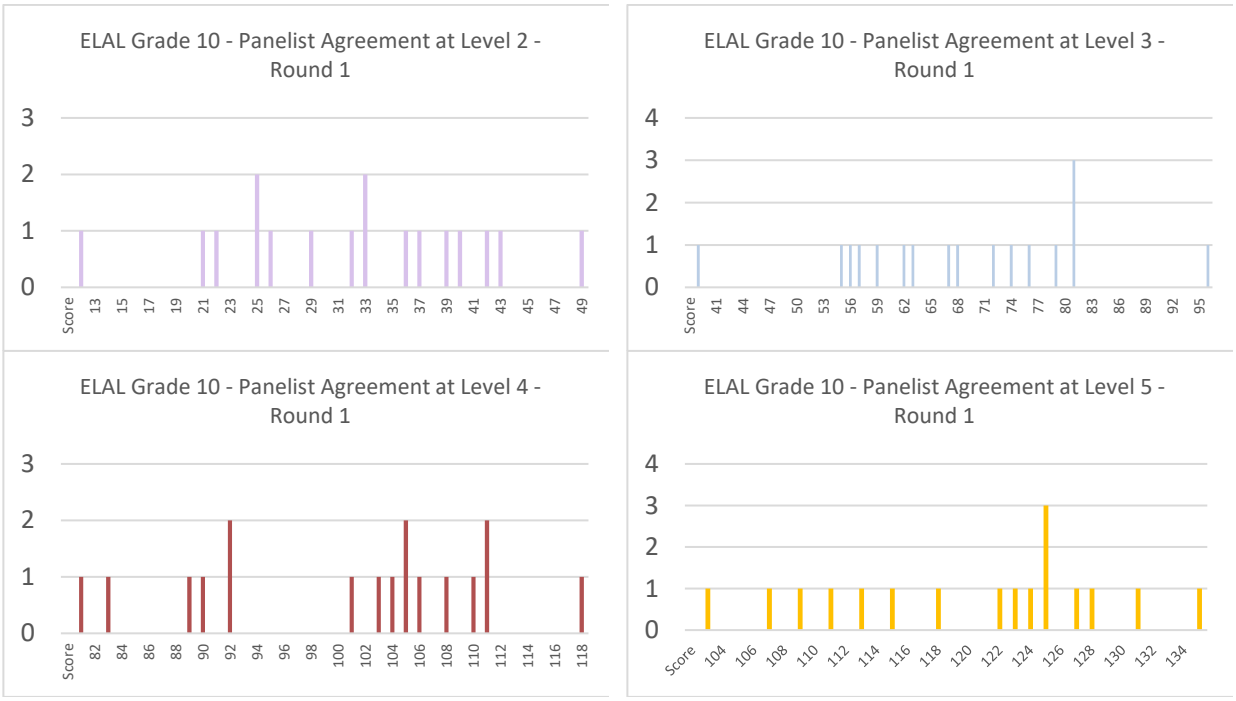



**Figure A10.3:** Grade 10 ELA/L - Round 1

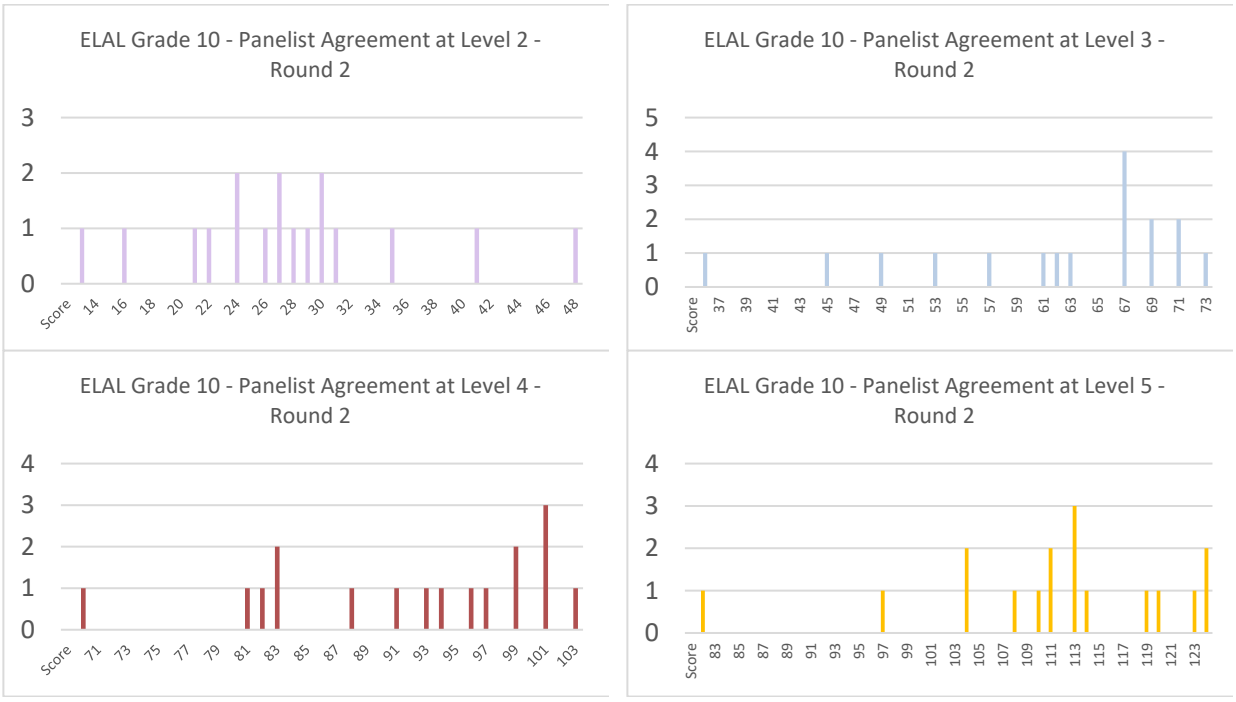



**Figure A10.4:** Grade 10 ELA/L - Round 2

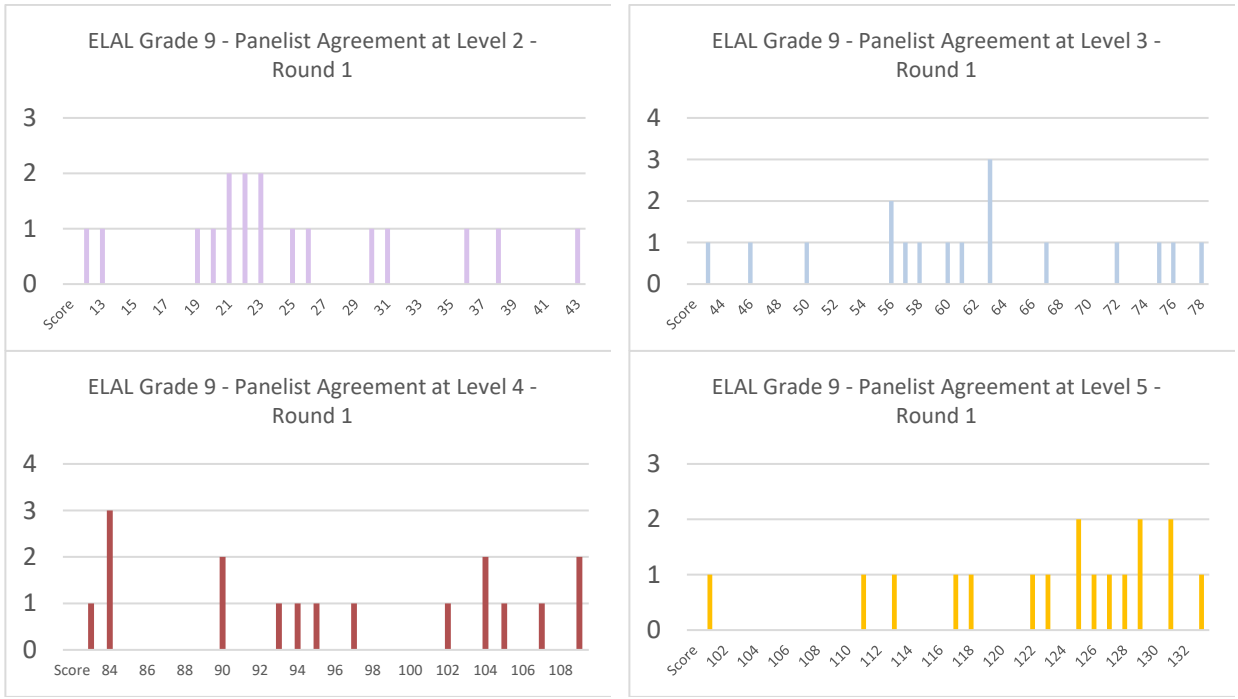



**Figure A10.5:** Grade 9 ELA/L - Round 1

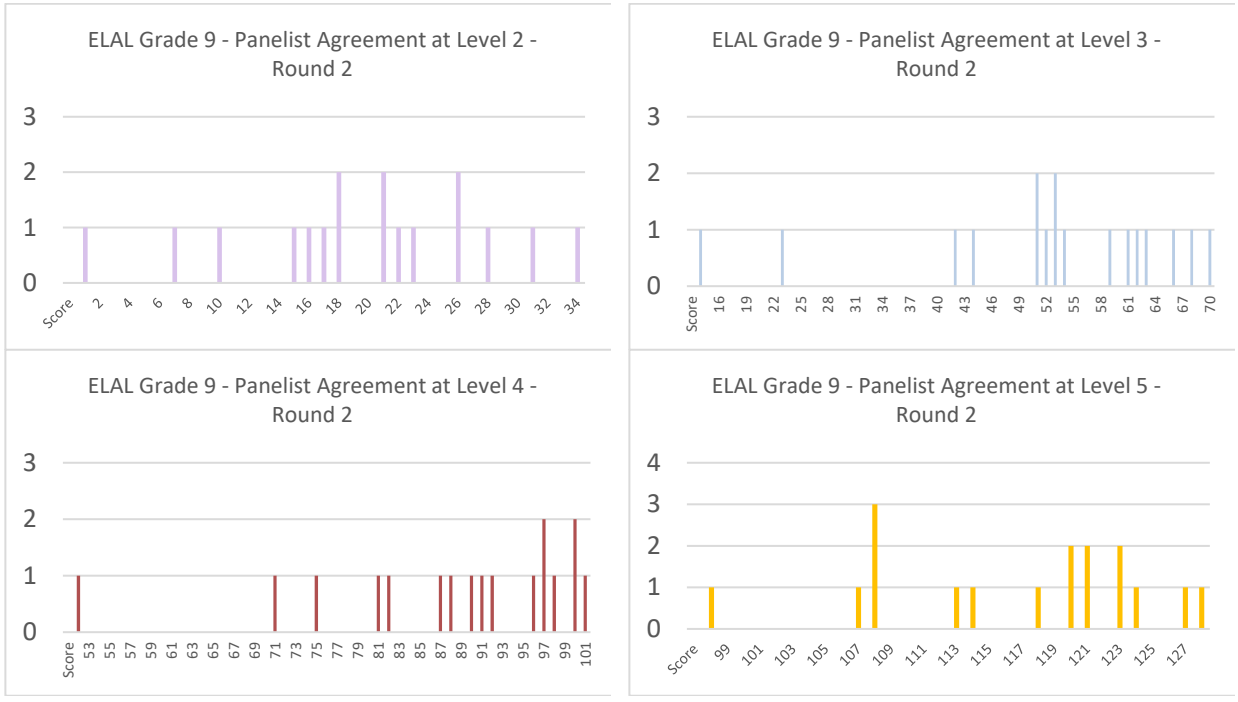



**Figure A10.6:** Grade 9 ELA/L - Round 2

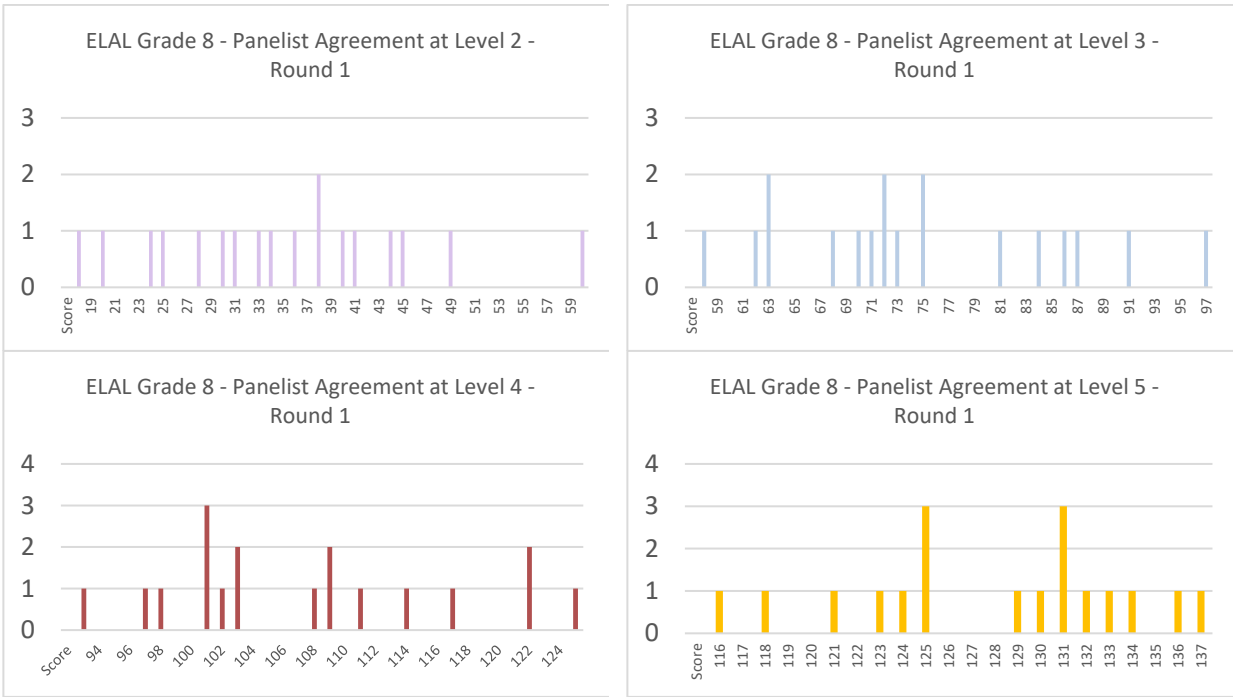



**Figure A10.7:** Grade 8 ELA/L - Round 1

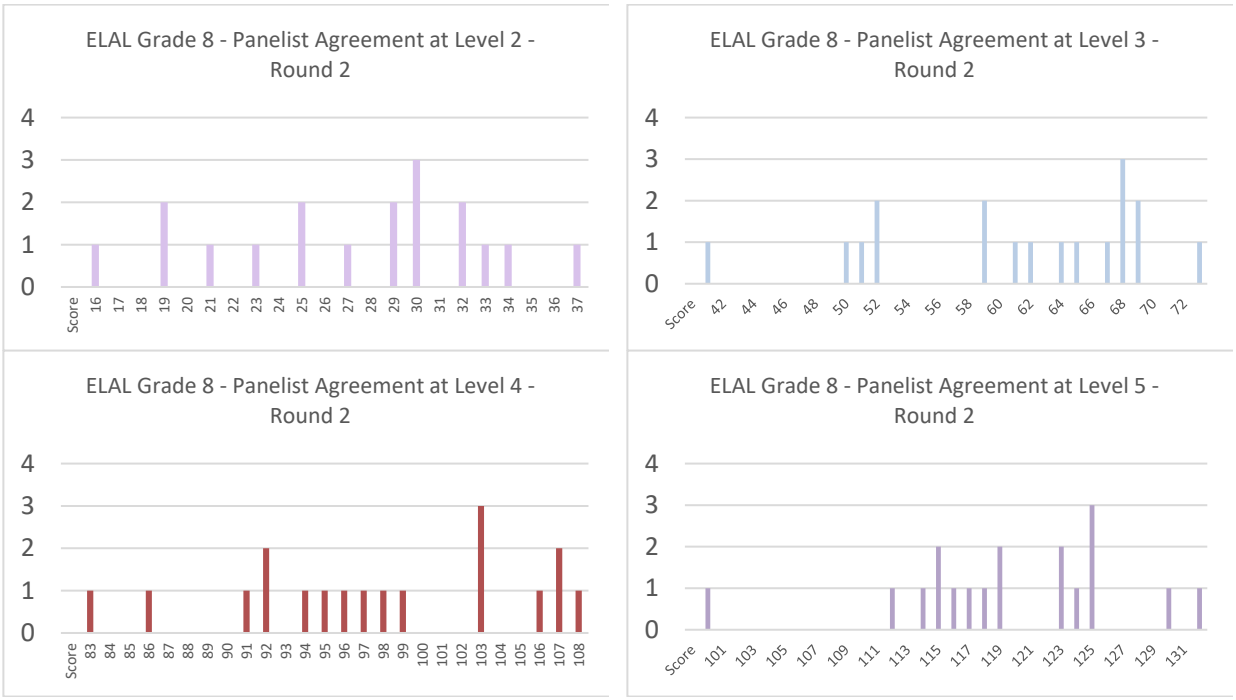



**Figure A10.8:** Grade 8 ELA/L - Round 2

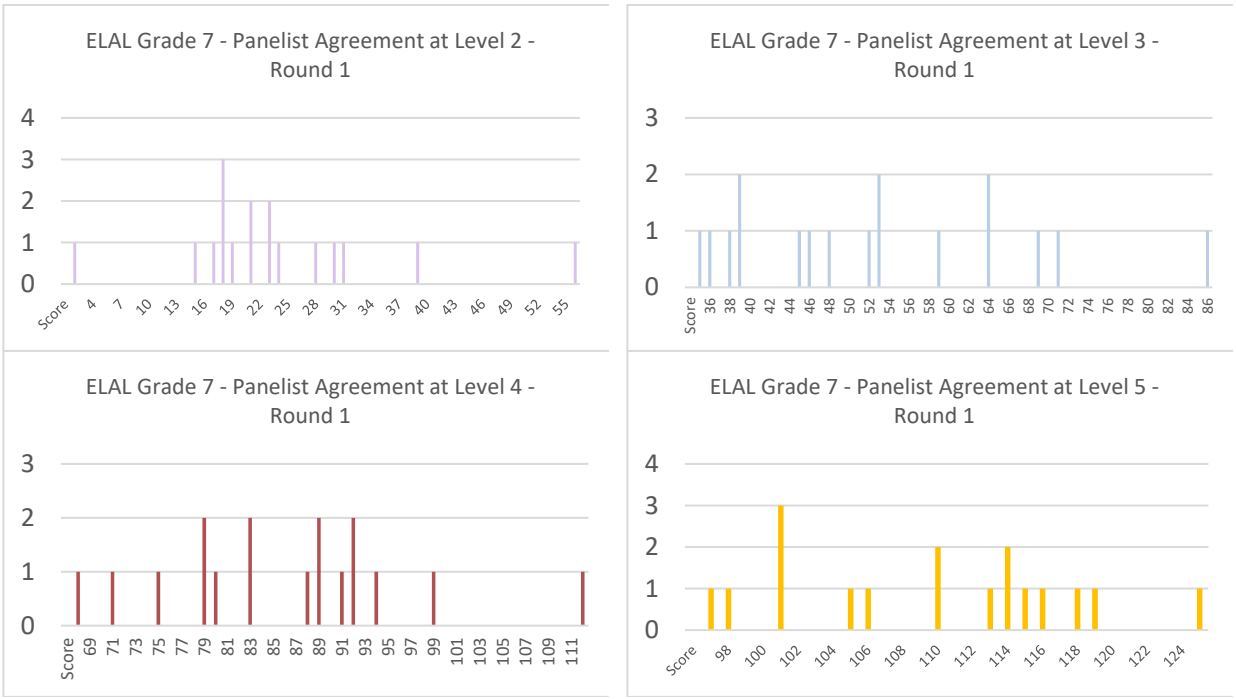



**Figure A10.9:** Grade 7 ELA/L - Round 1

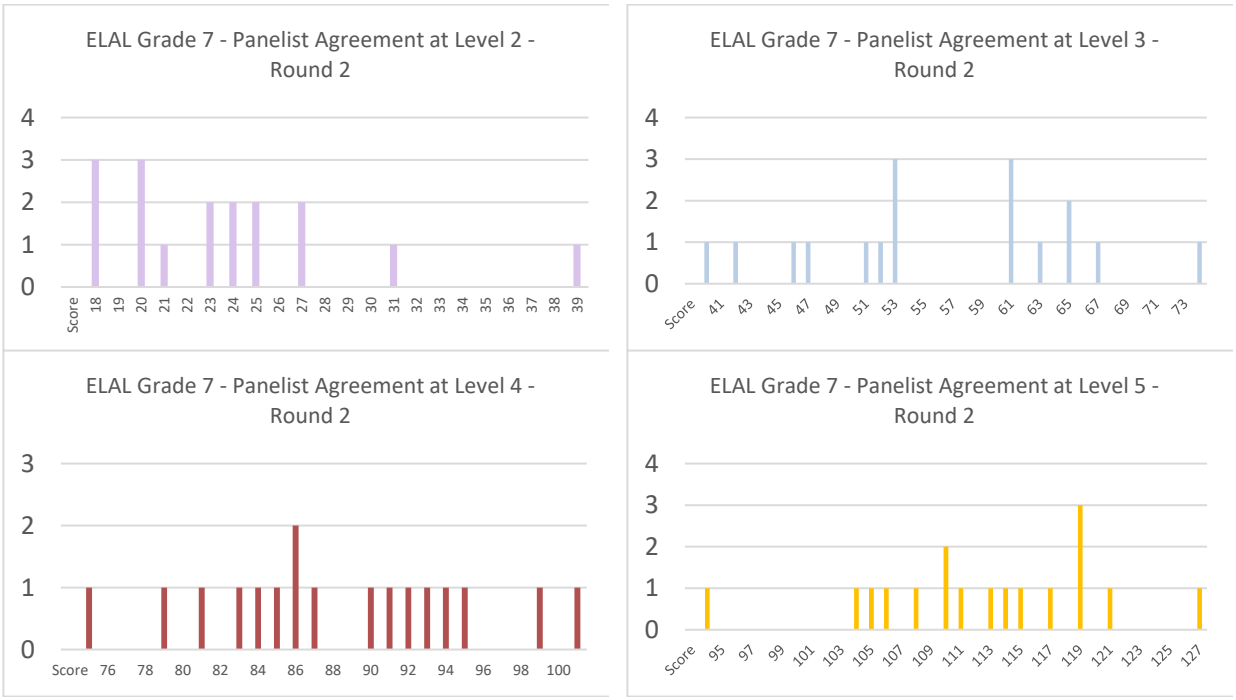



**Figure A10.10:** Grade 7 ELA/L - Round 2

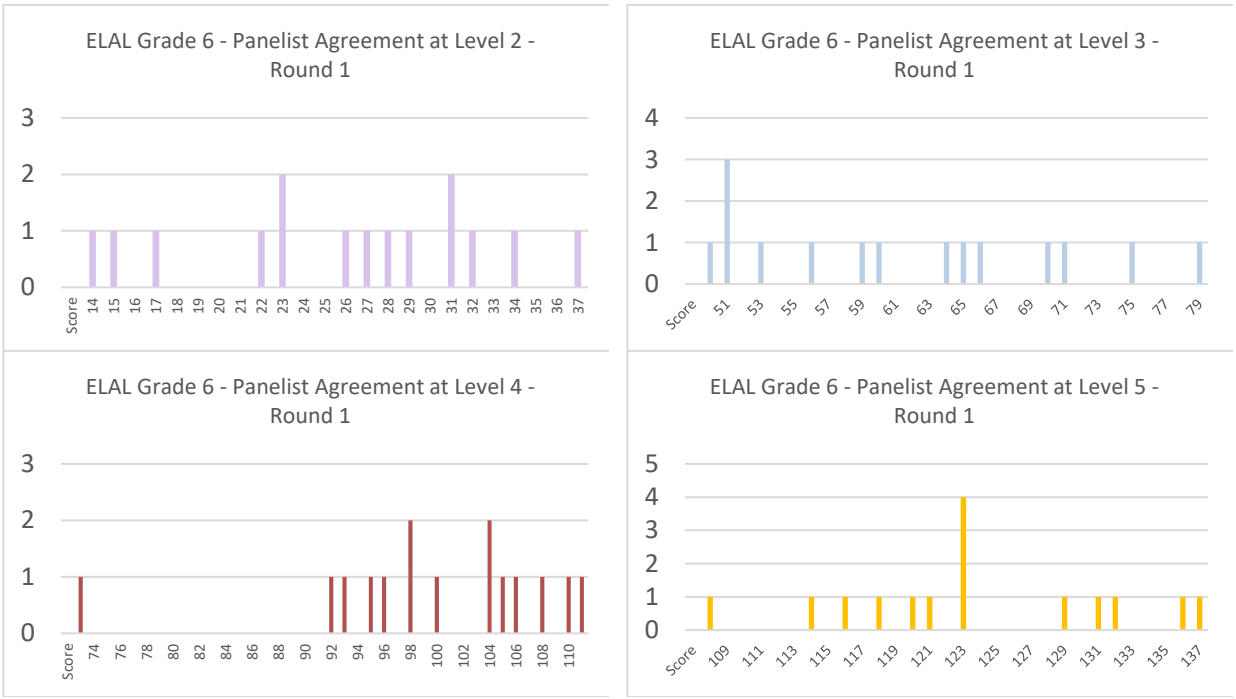



**Figure A10.11:** Grade 6 ELA/L - Round 1

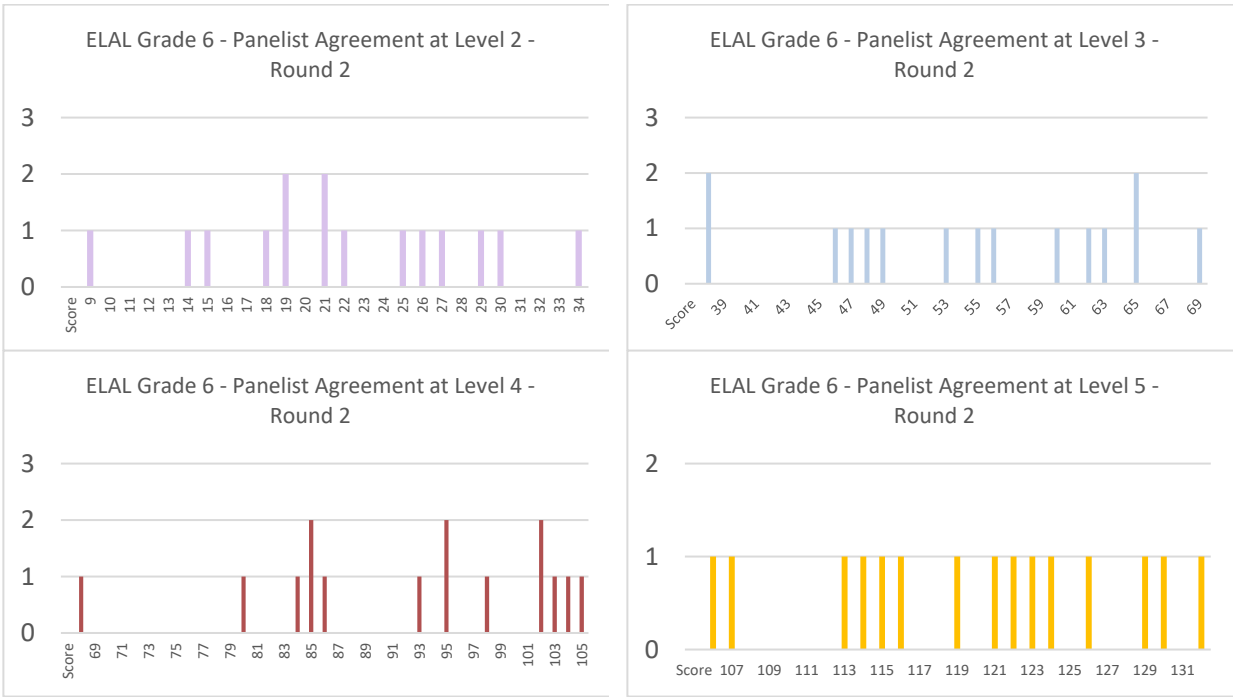



**Figure A10.12:** Grade 6 ELA/L - Round 2

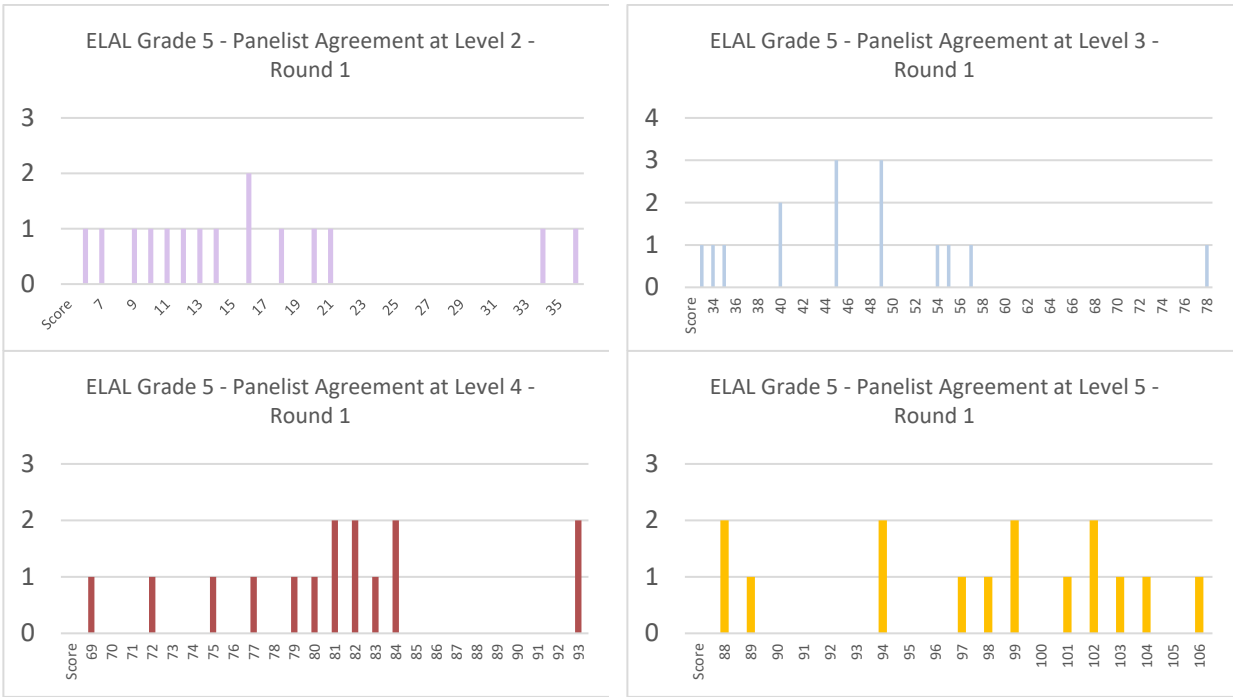



**Figure A10.13:** Grade 5 ELA/L - Round 1

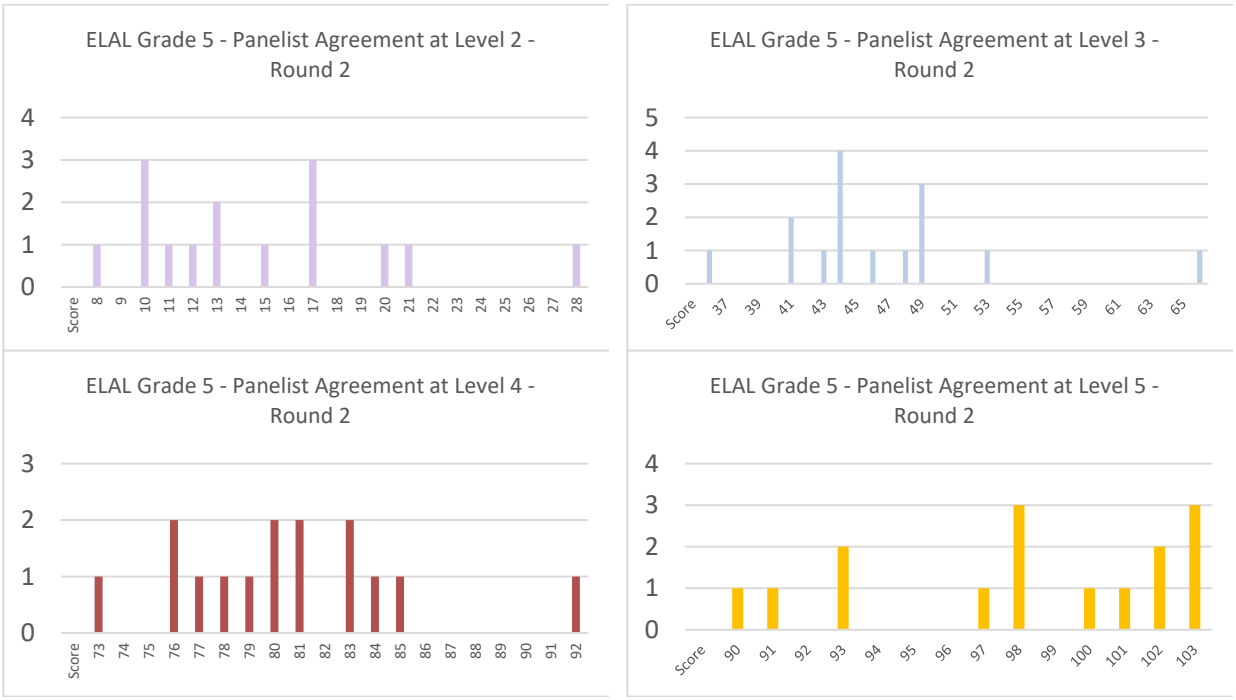



**Figure A10.14:** Grade 5 ELA/L - Round 2

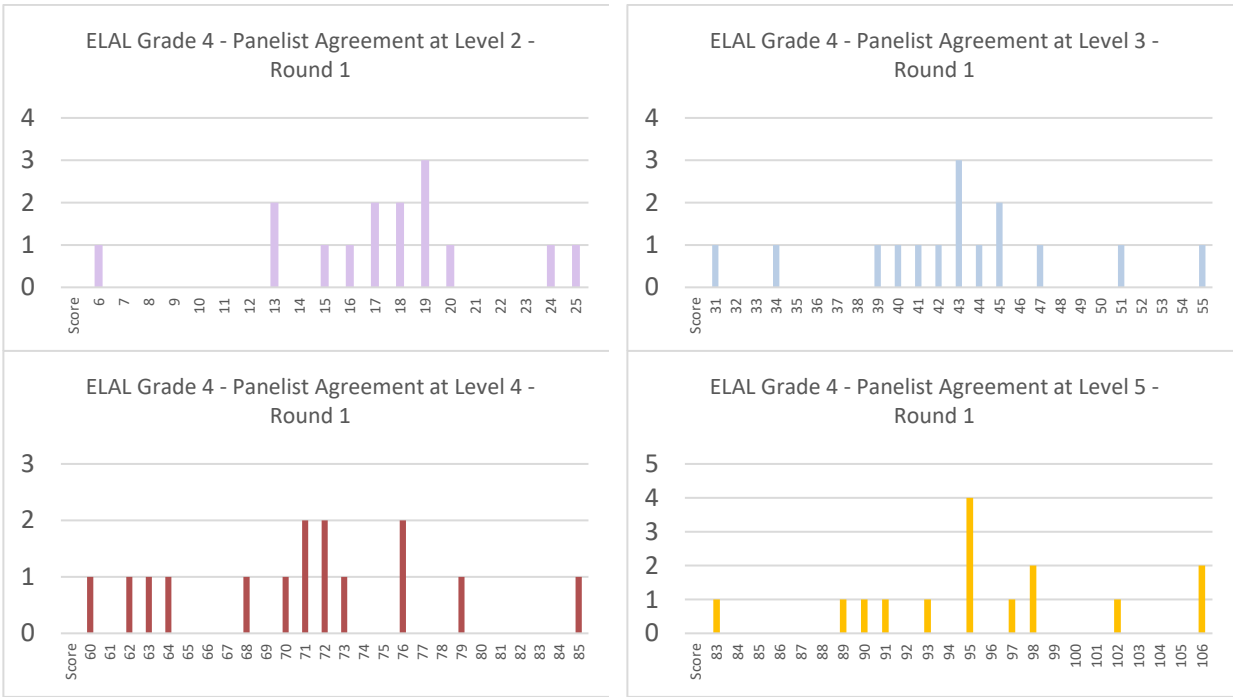



**Figure A10.15:** Grade 4 ELA/L - Round 1

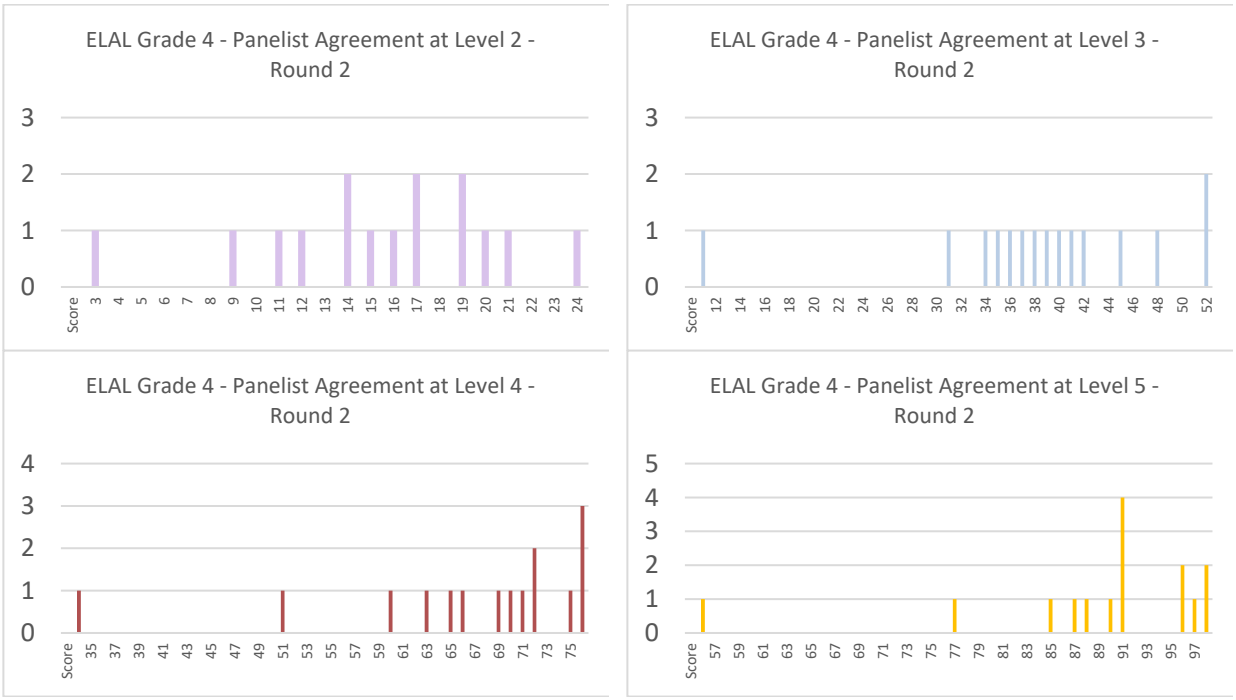



**Figure A10.16:** Grade 4 ELA/L - Round 2

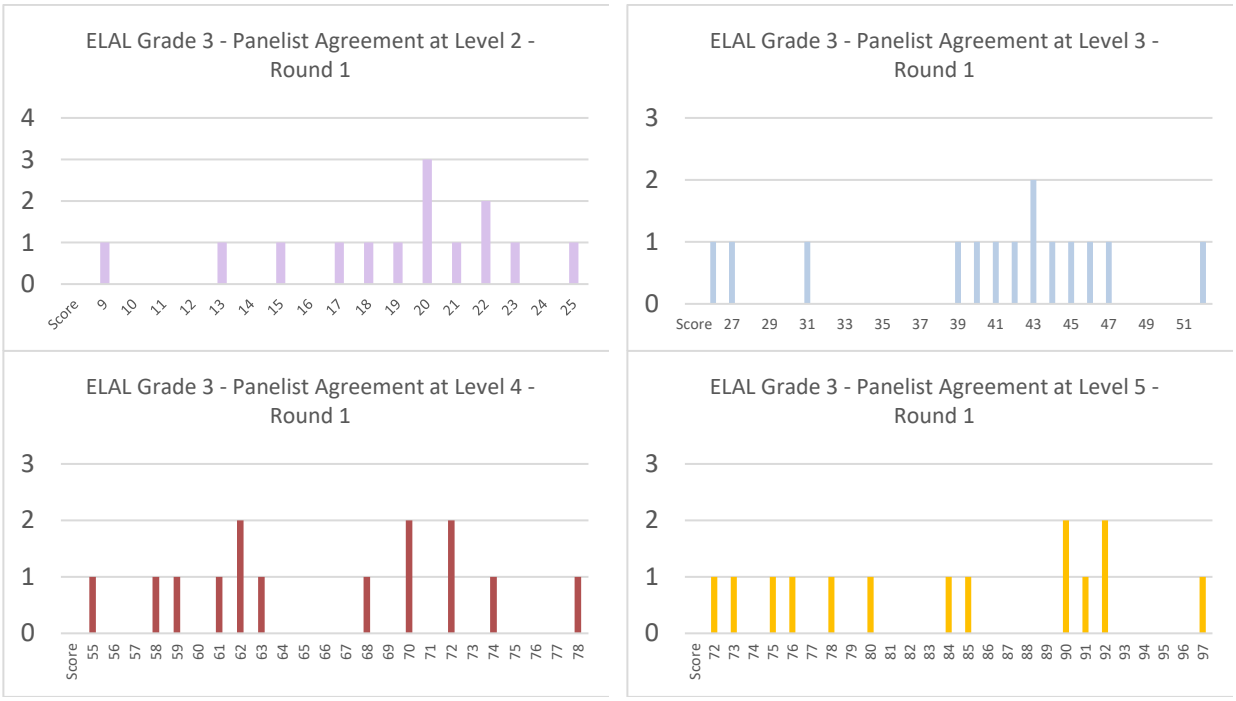



**Figure A10.17:** Grade 3 ELA/L - Round 1

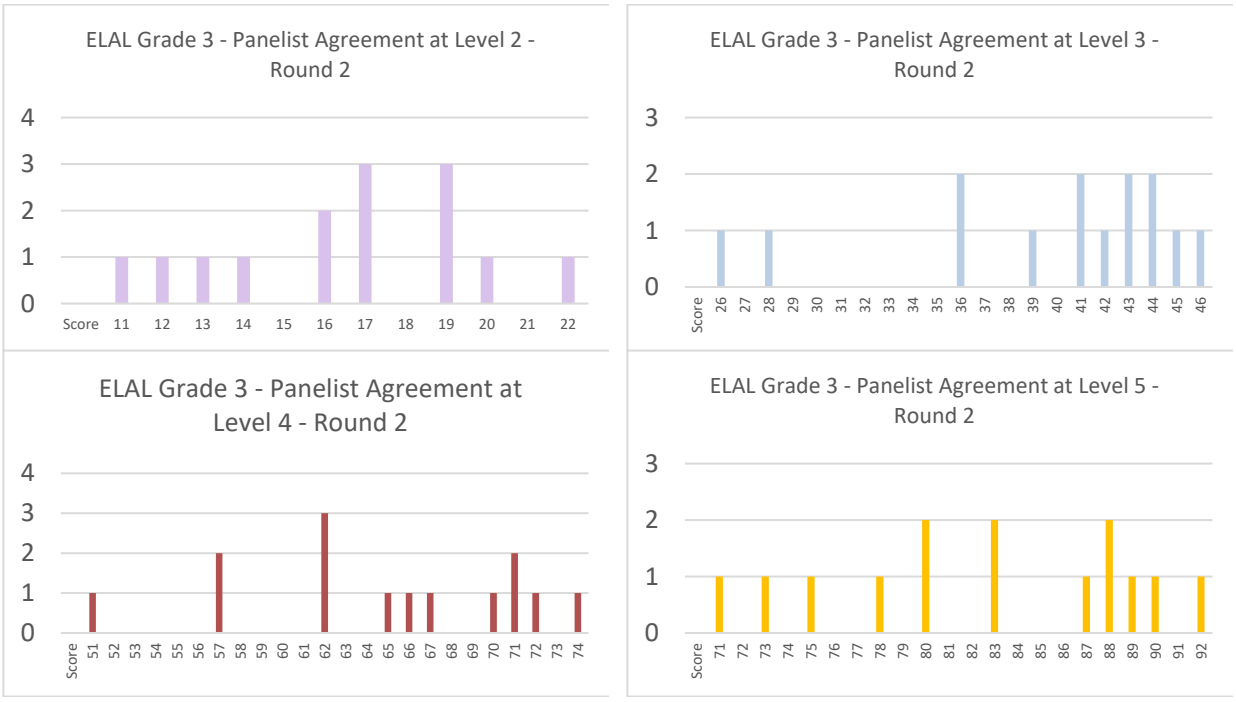



**Figure A10.18:** Grade 3 ELA/L - Round 2

## Mathematics

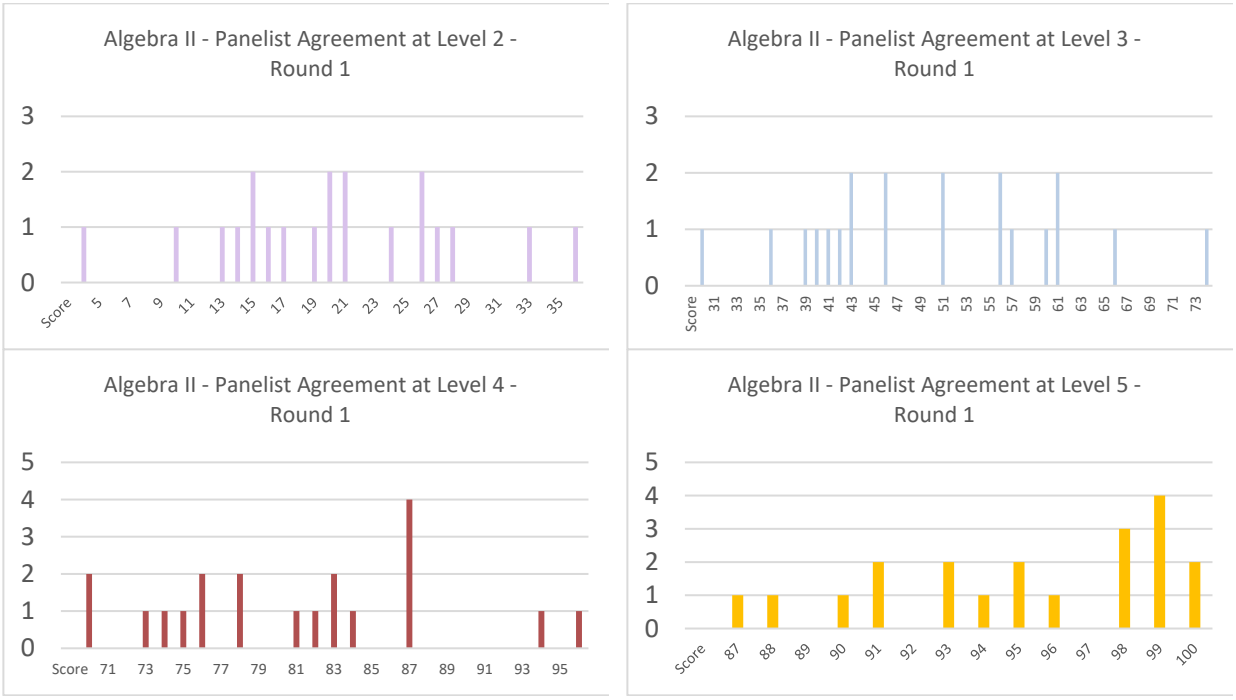



**Figure A10.19:** Algebra II - Round 1

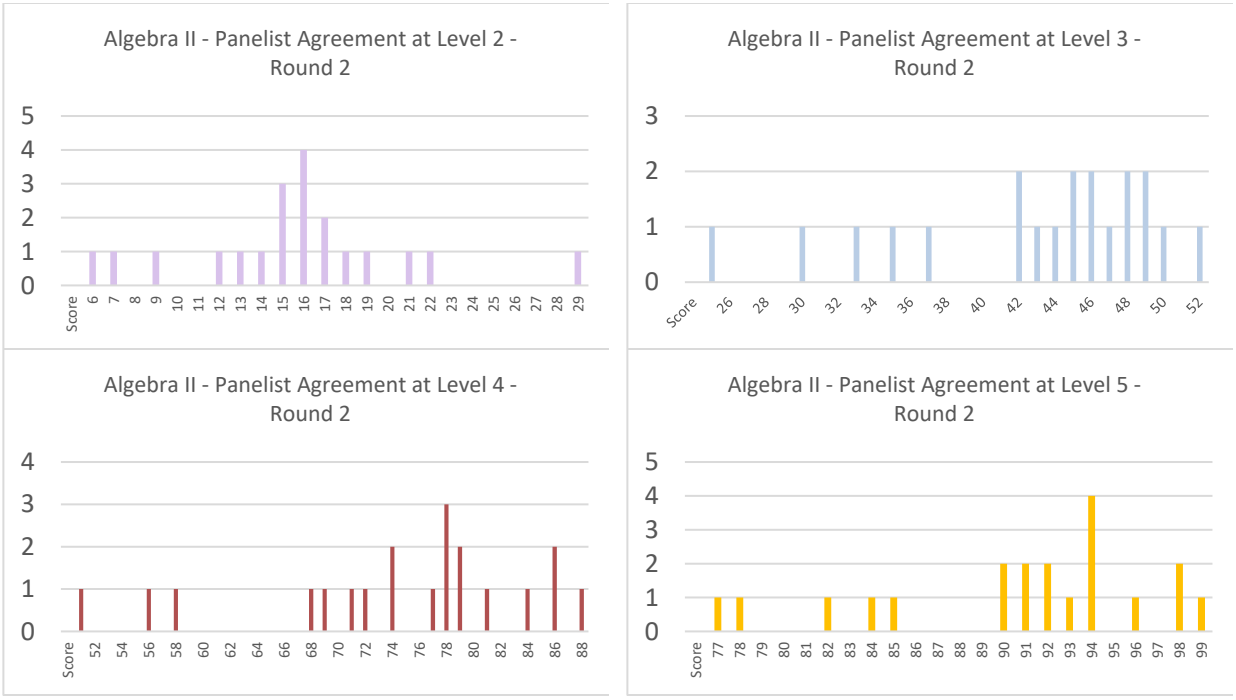



**Figure A10.20:** Algebra II - Round 2

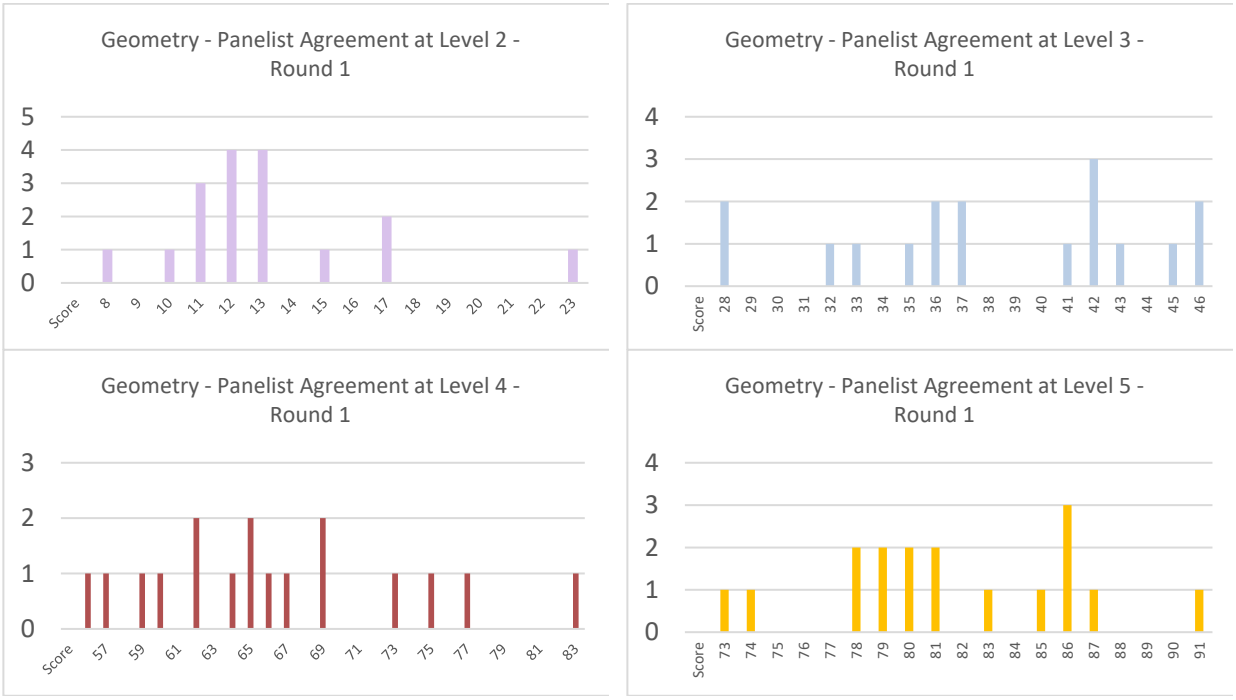



**Figure A10.21:** Geometry - Round 1

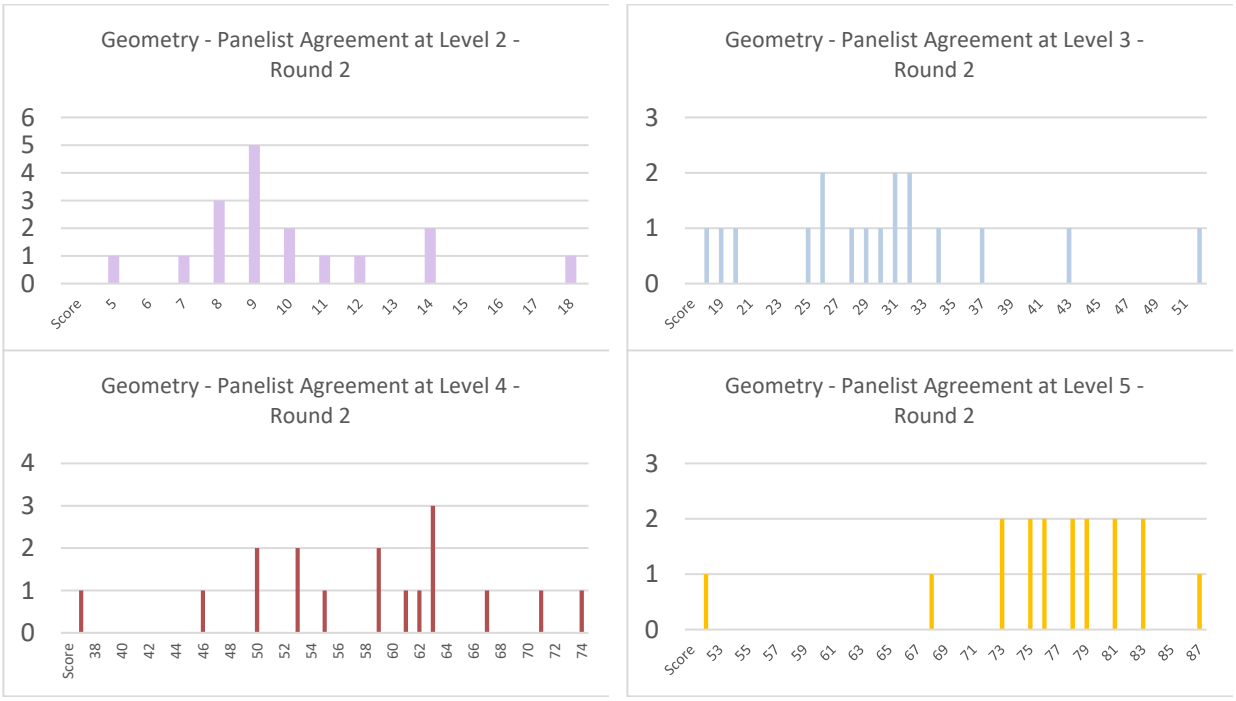



**Figure A10.22:** Geometry - Round 2

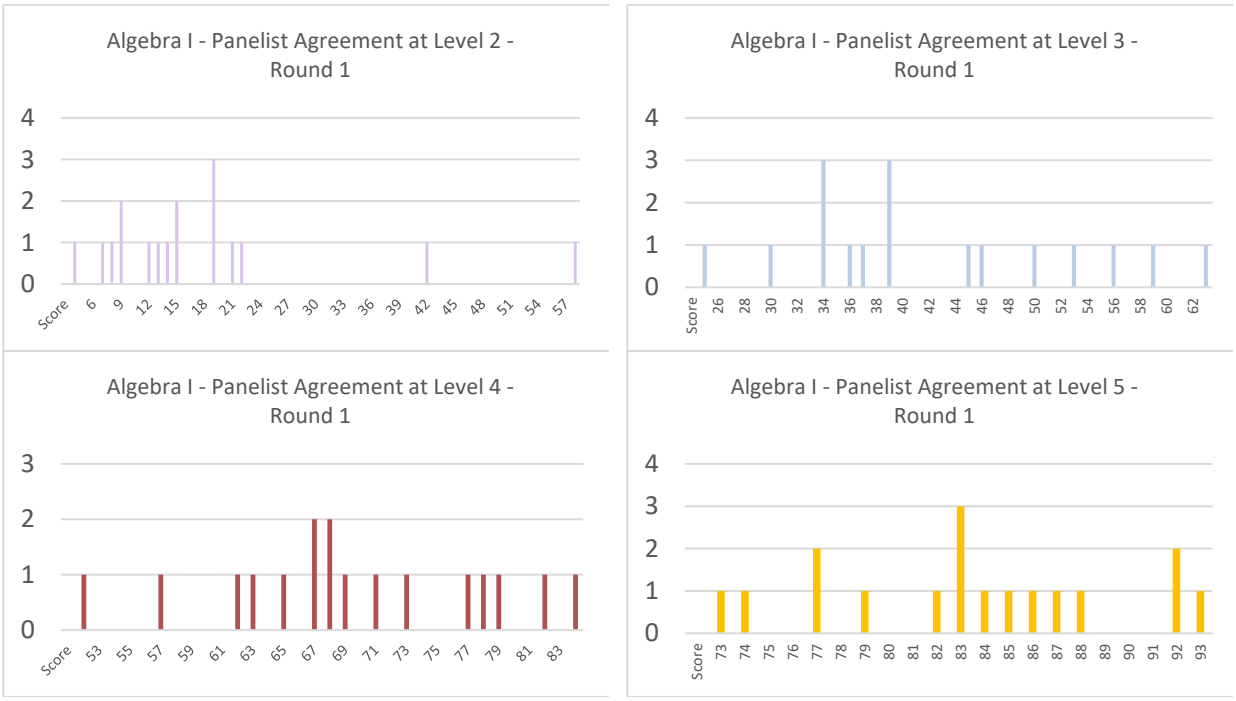



**Figure A10.23:** Algebra I - Round 1

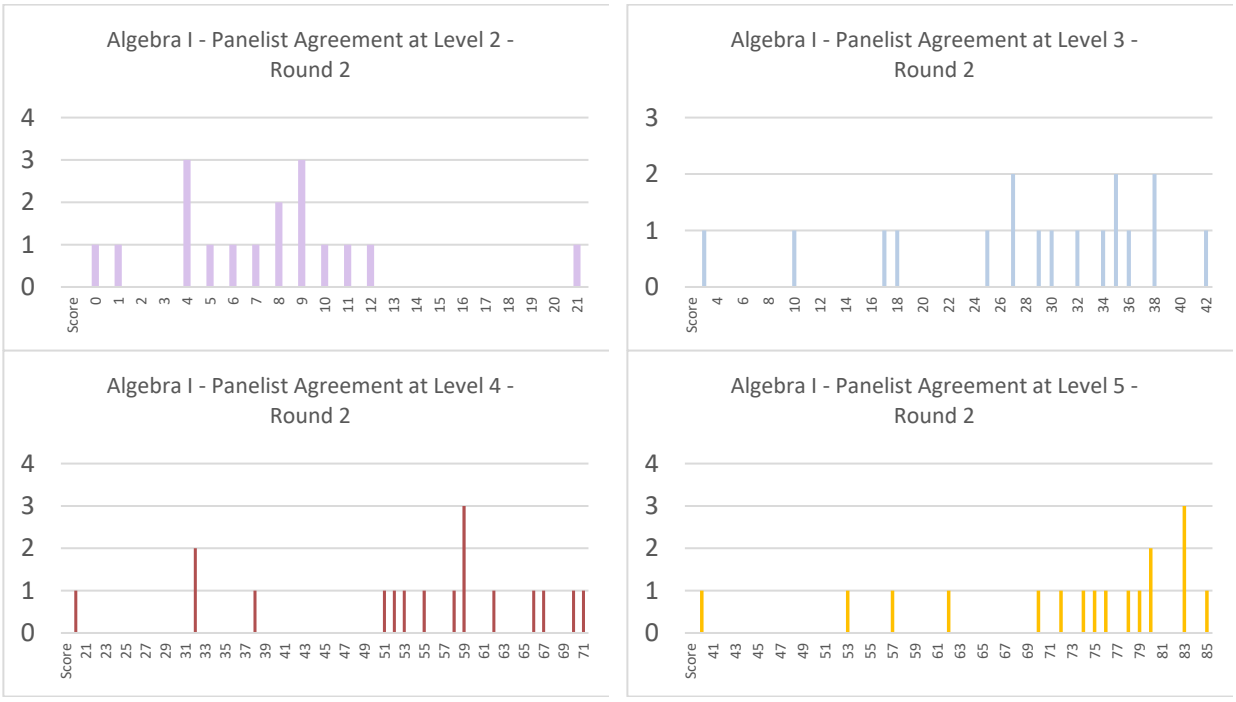



**Figure A10.24:** Algebra I - Round 2

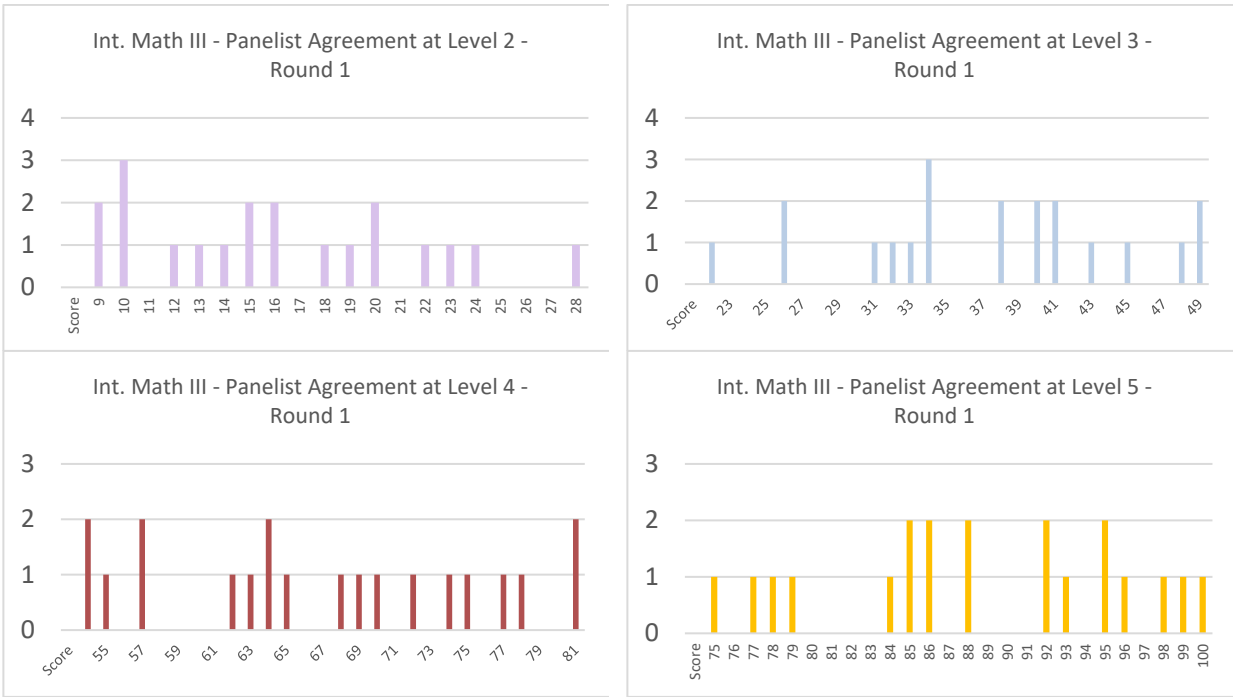



**Figure A10.25:** Integrated Math III - Round 1

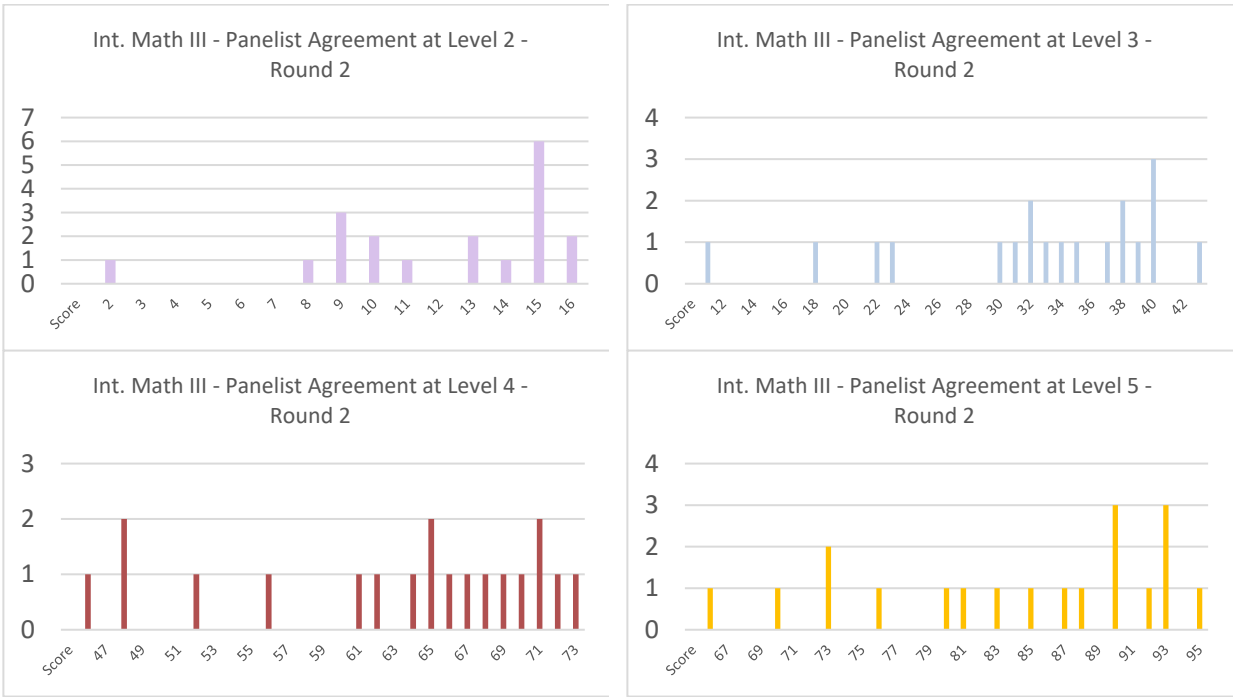



**Figure A10.26:** Integrated Math III - Round 2

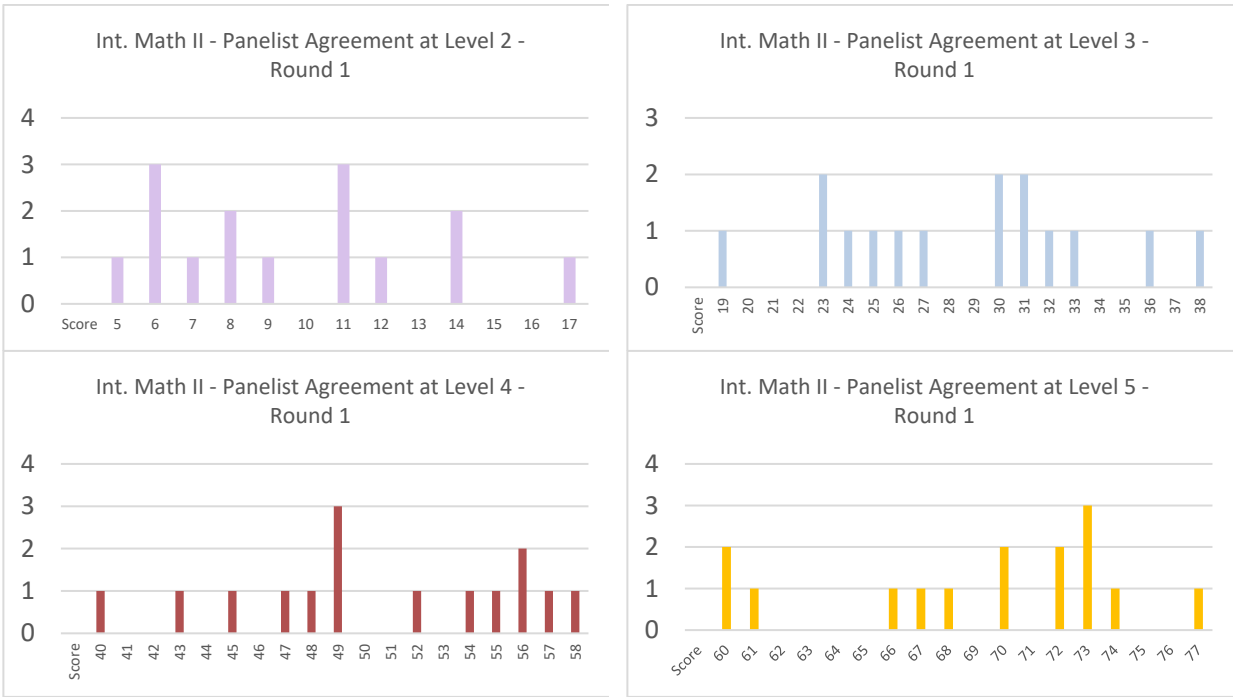



**Figure A10.27:** Integrated Math II - Round 1

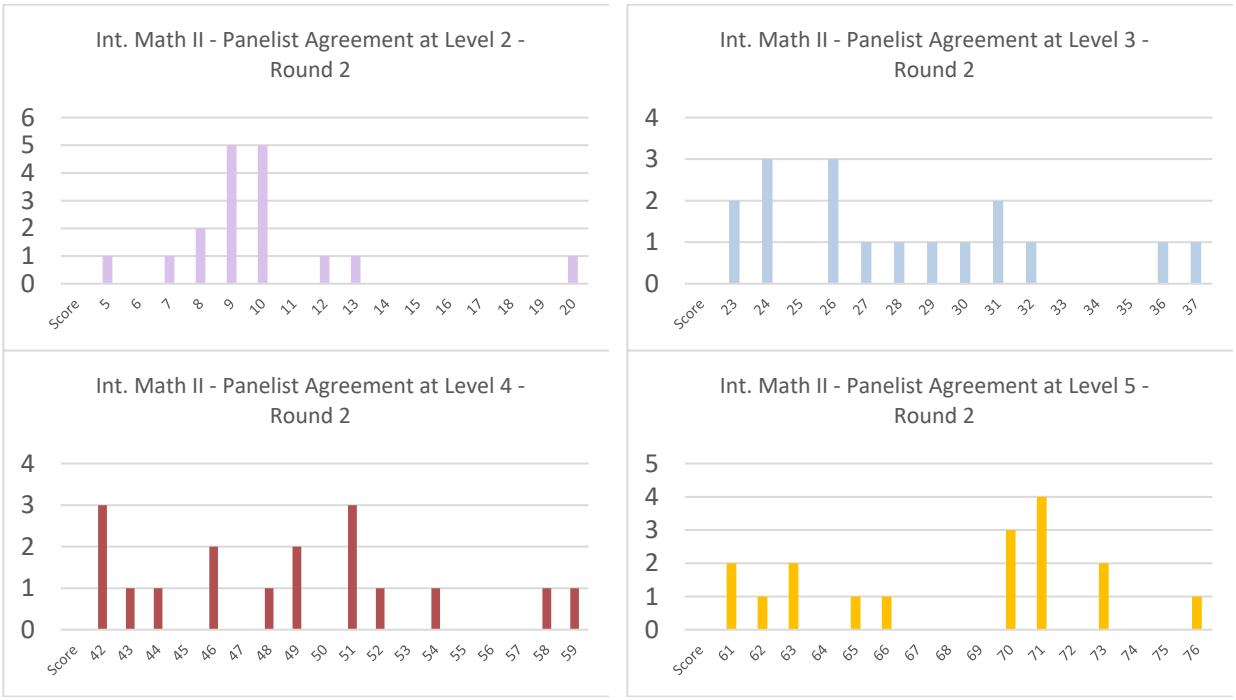



**Figure A10.28:** Integrated Math II - Round 2

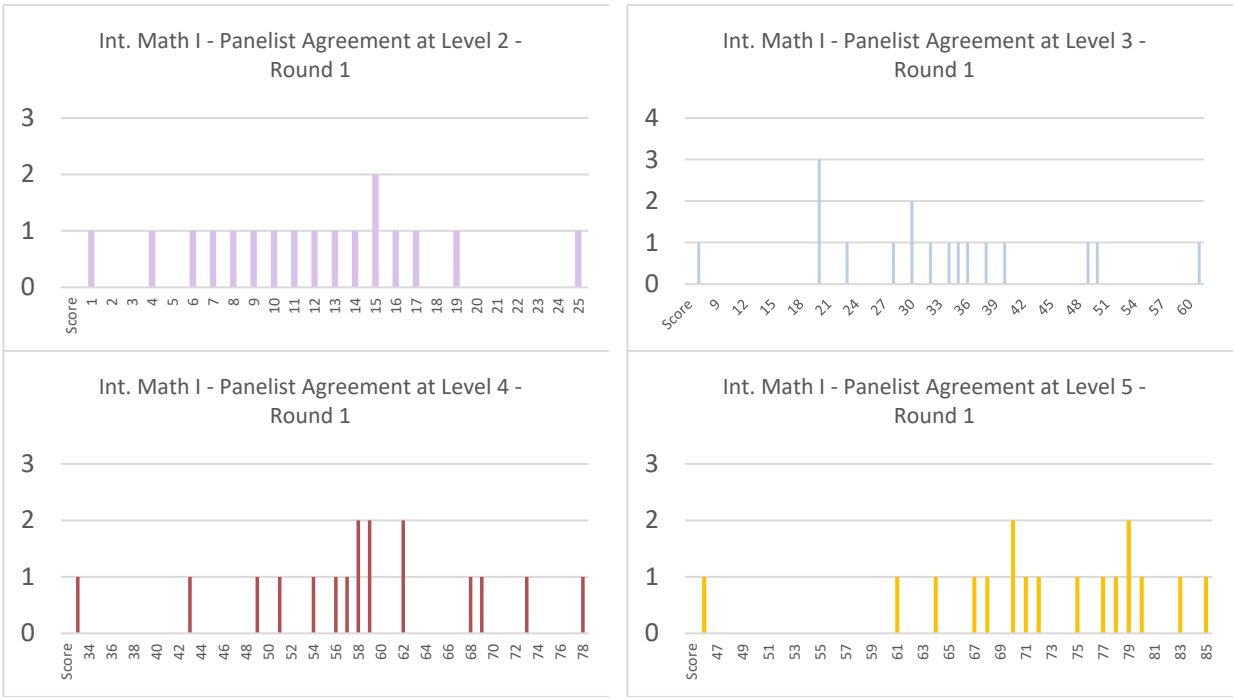



**Figure A10.29:** Integrated Math I - Round 1

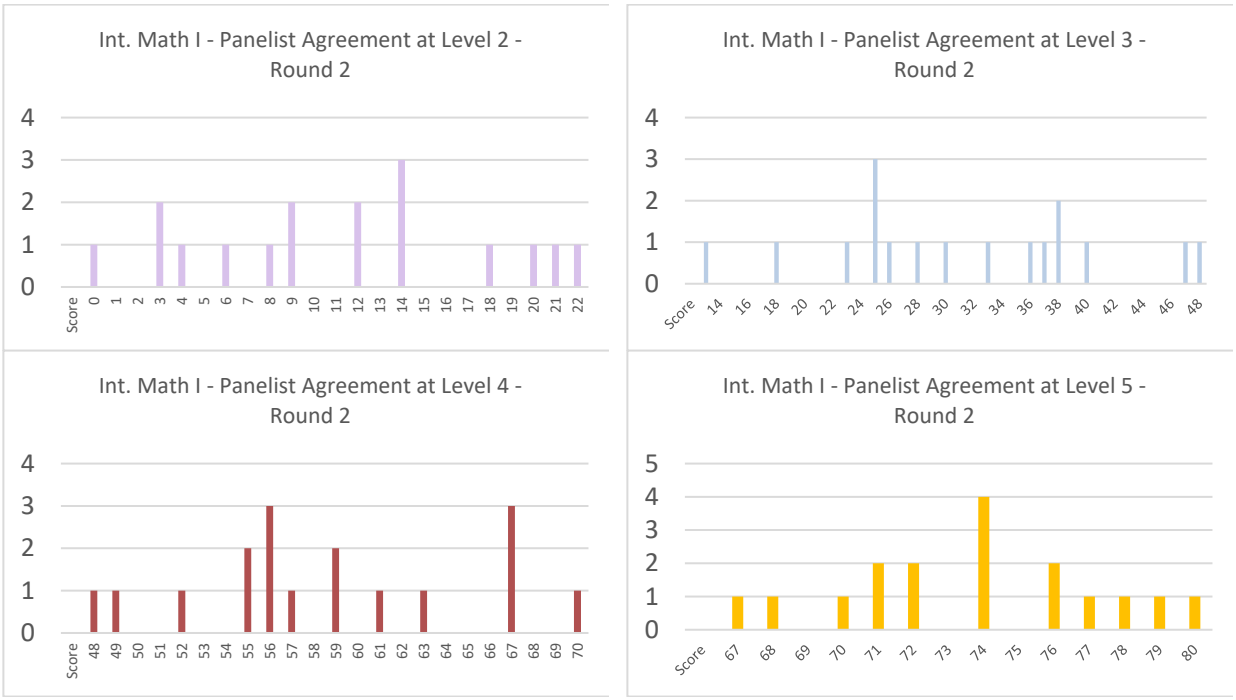



**Figure A10.30:** Integrated Math I - Round 2

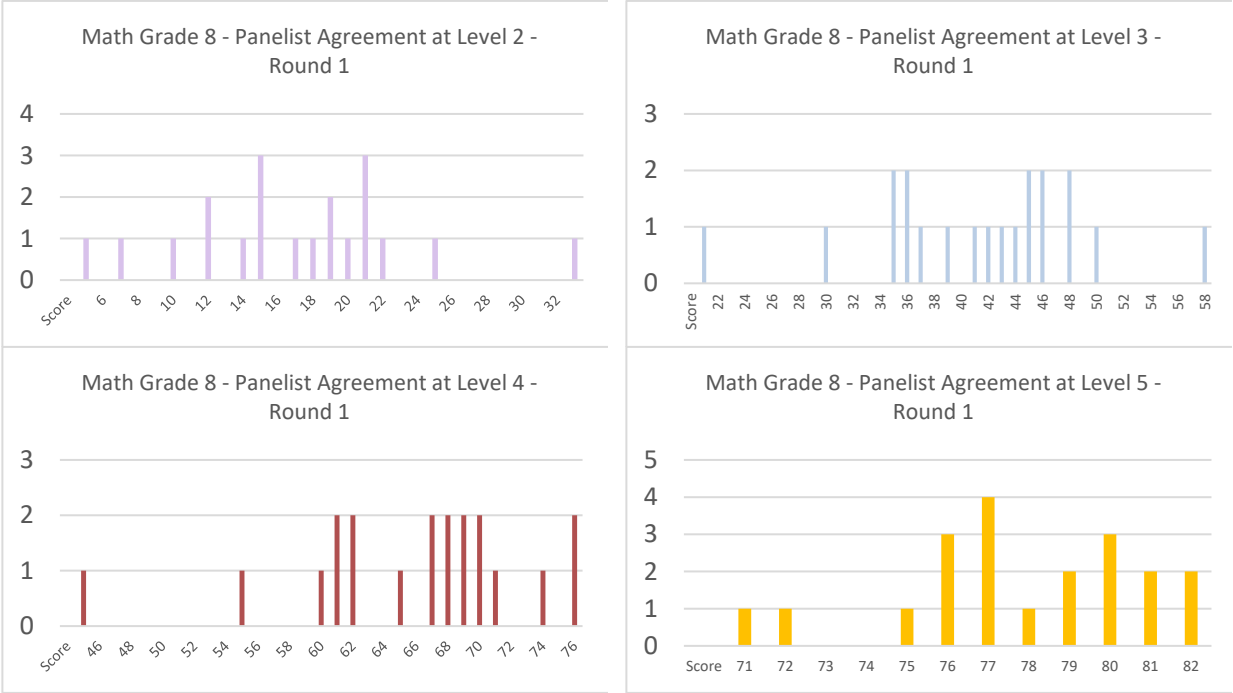



**Figure A10.31:** Grade 8 Mathematics - Round 1

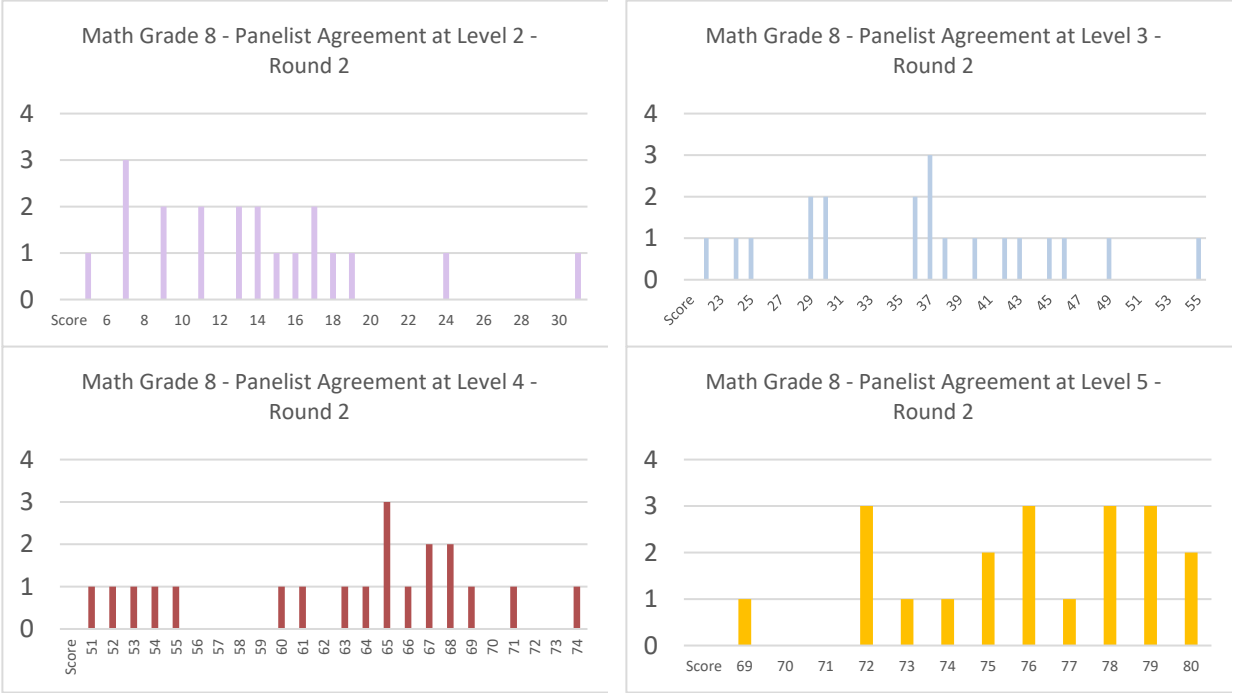



**Figure A10.32:** Grade 8 Mathematics - Round 2

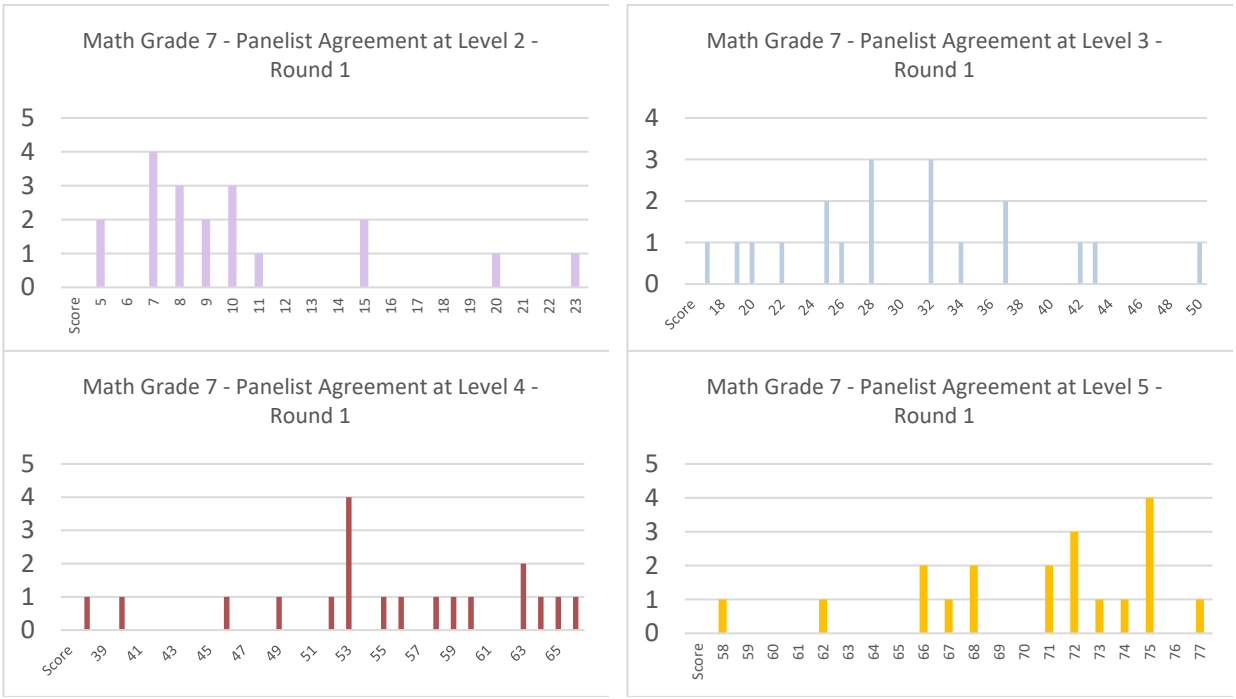



**Figure A10.33:** Grade 7 Mathematics - Round 1

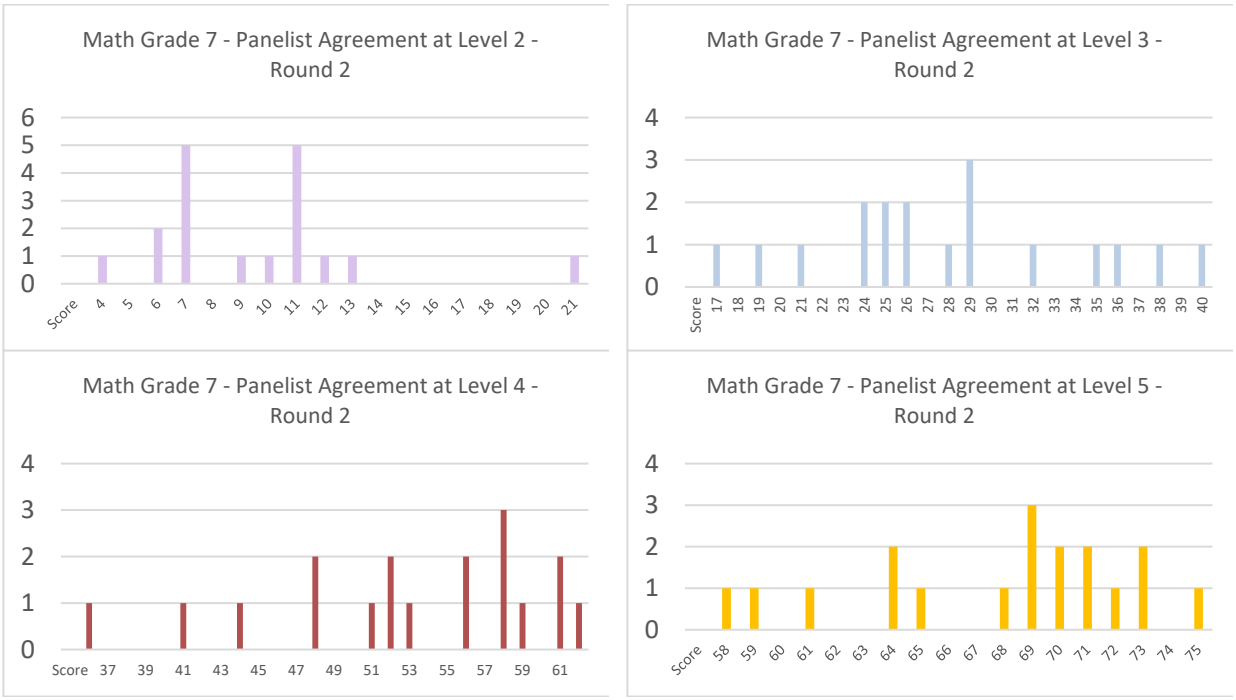



**Figure A10.34:** Grade 7 Mathematics - Round 2

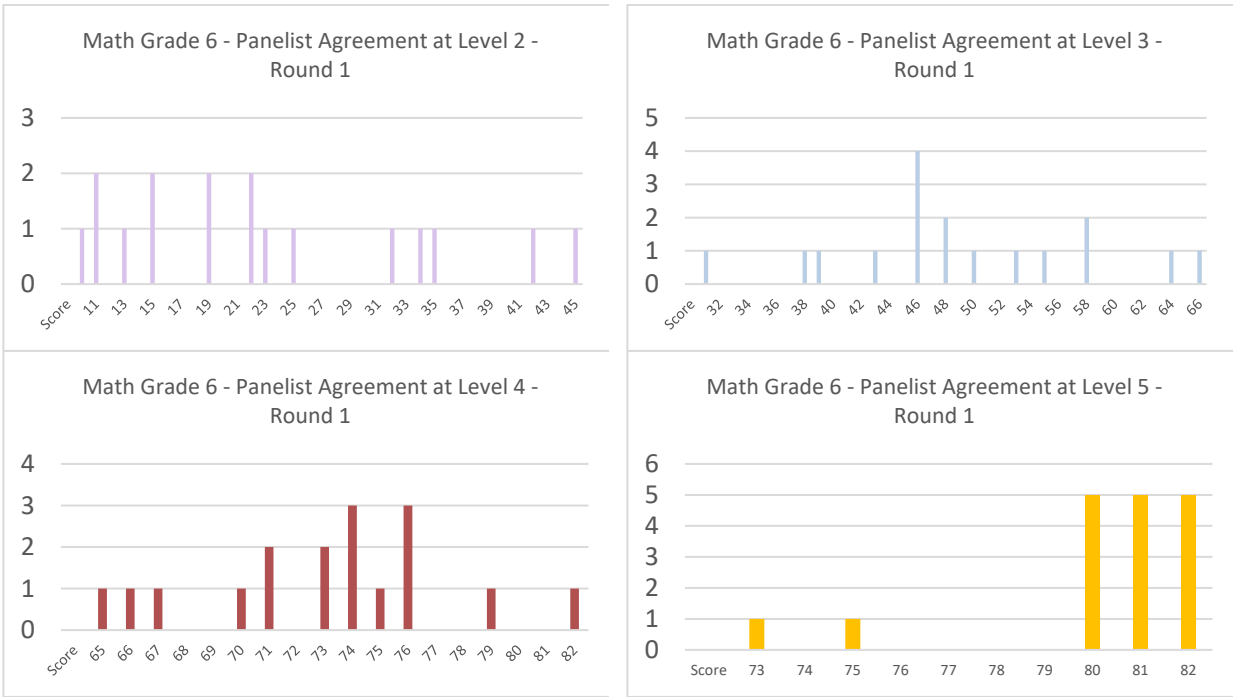



**Figure A10.35:** Grade 6 Mathematics - Round 1

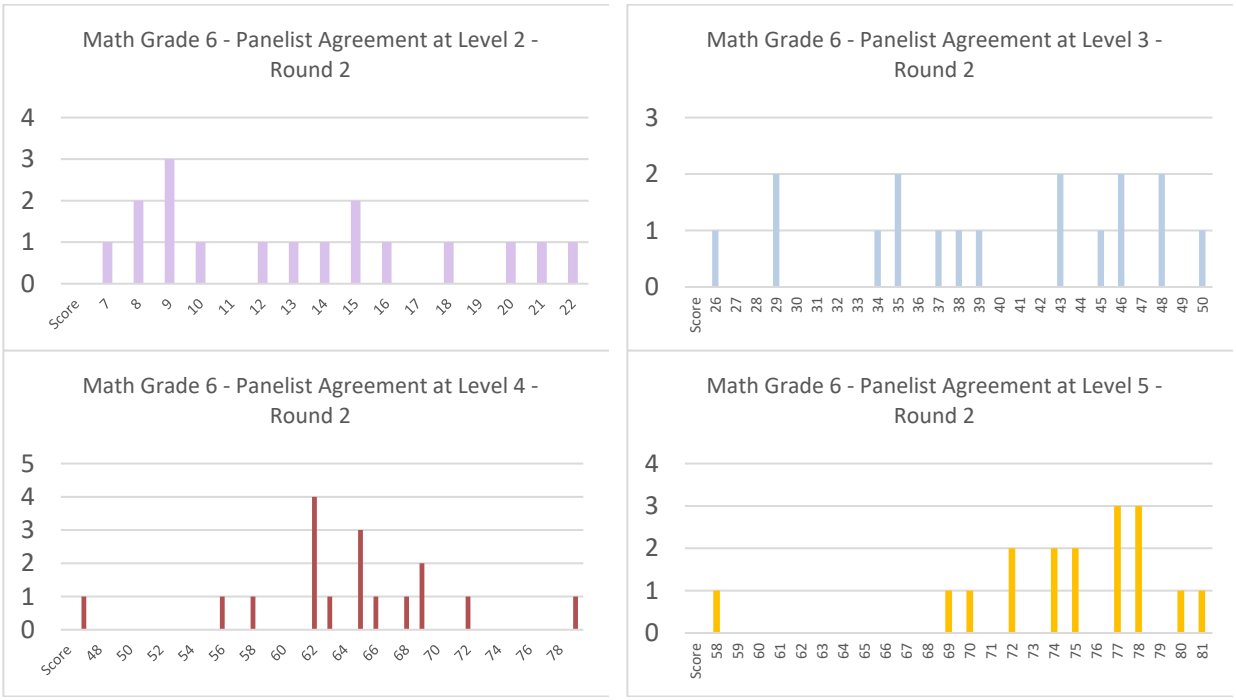



**Figure A10.36:** Grade 6 Mathematics - Round 2

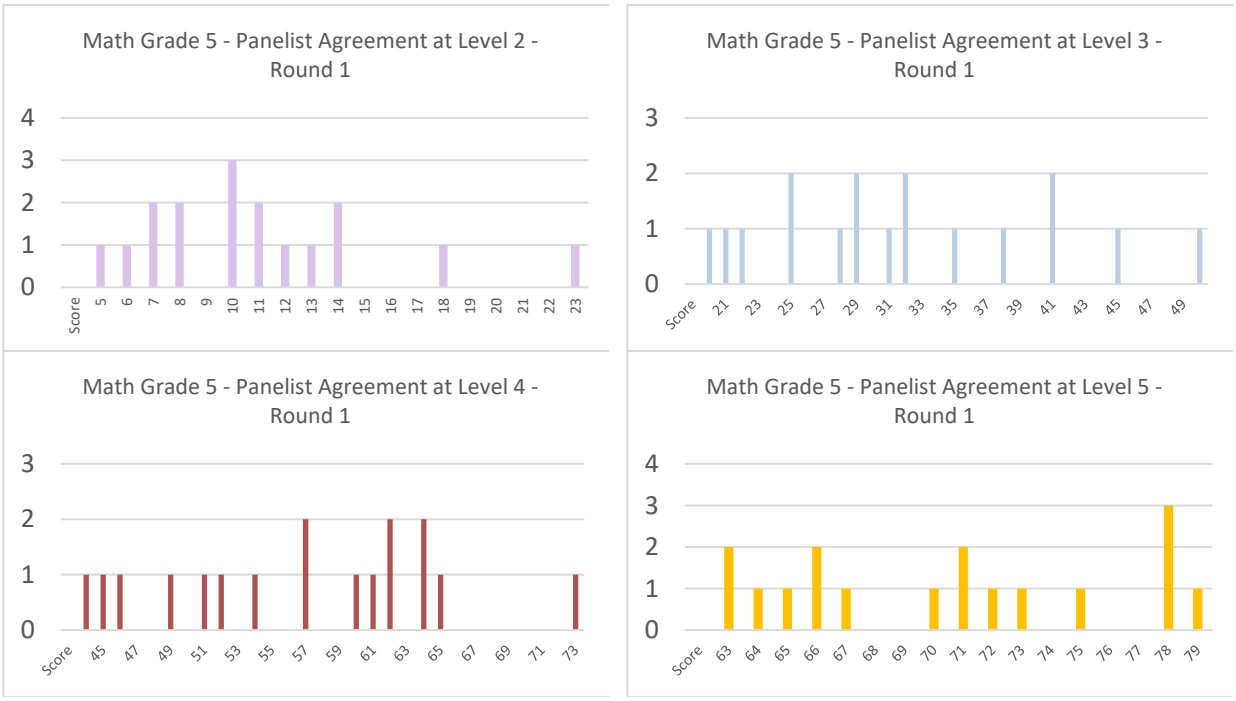



**Figure A10.37:** Grade 5 Mathematics - Round 1

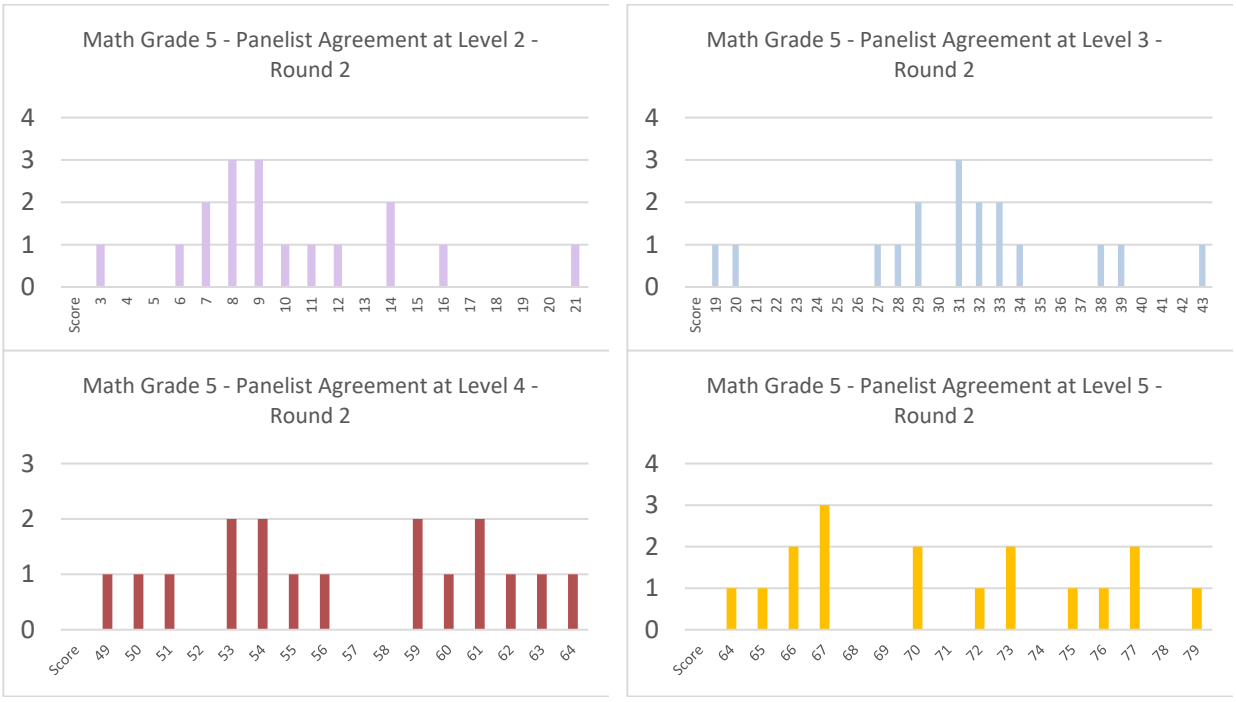



**Figure A10.38:** Grade 5 Mathematics - Round 2

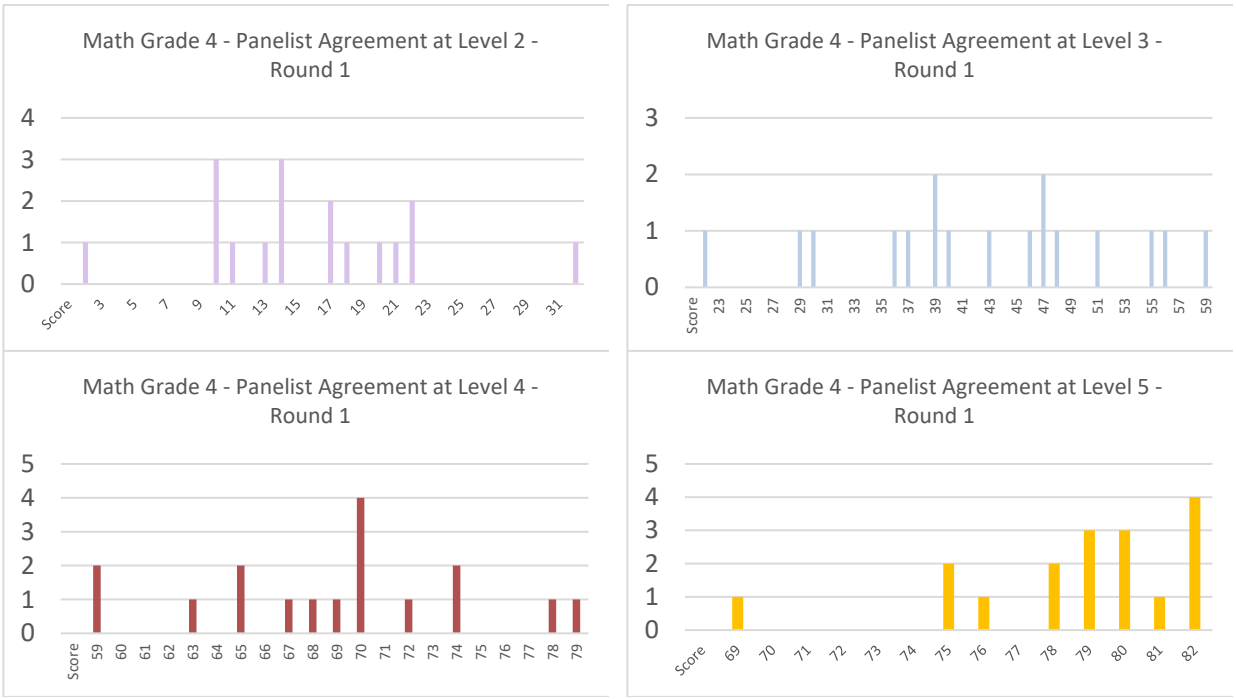



**Figure A10.39:** Grade 4 Mathematics - Round 1

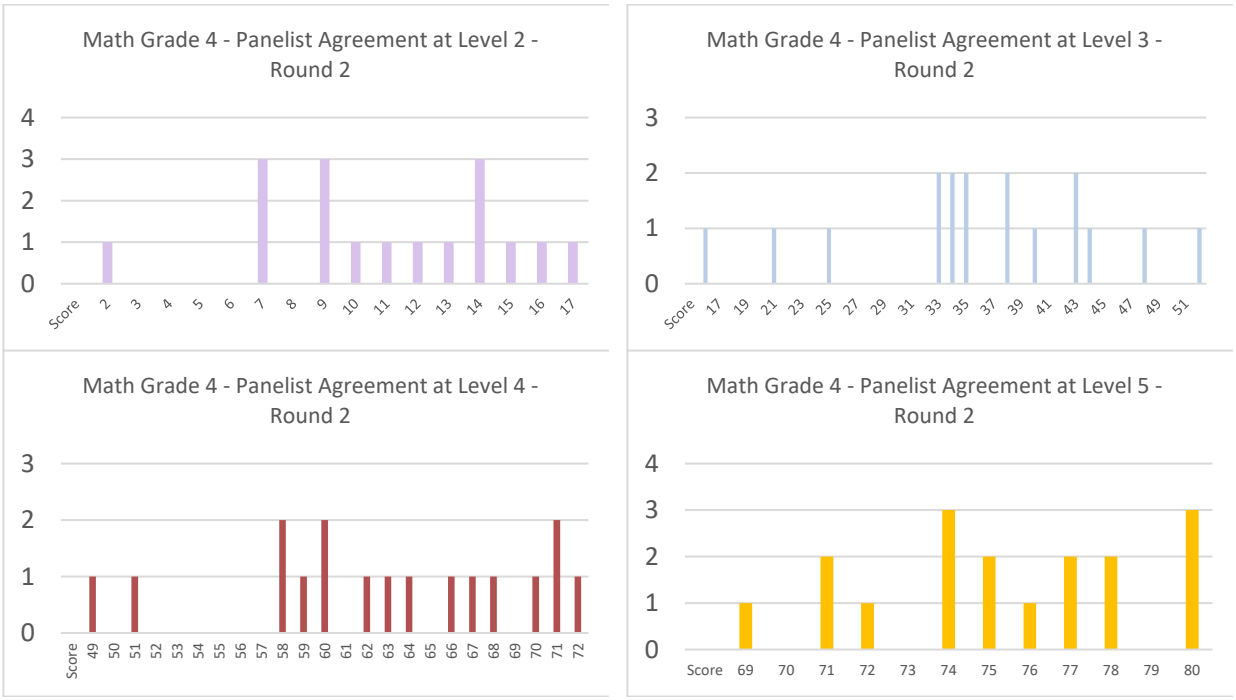



**Figure A10.40:** Grade 4 Mathematics - Round 2

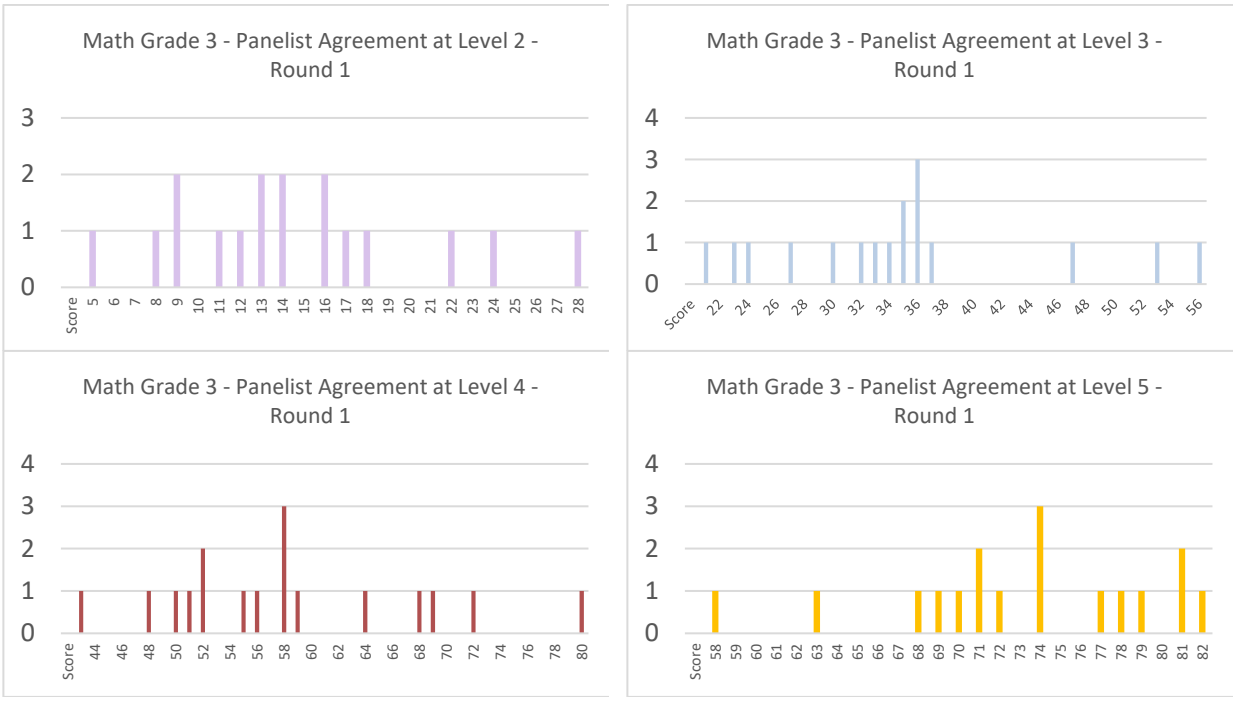



**Figure A10.41:** Grade 3 Mathematics - Round 1

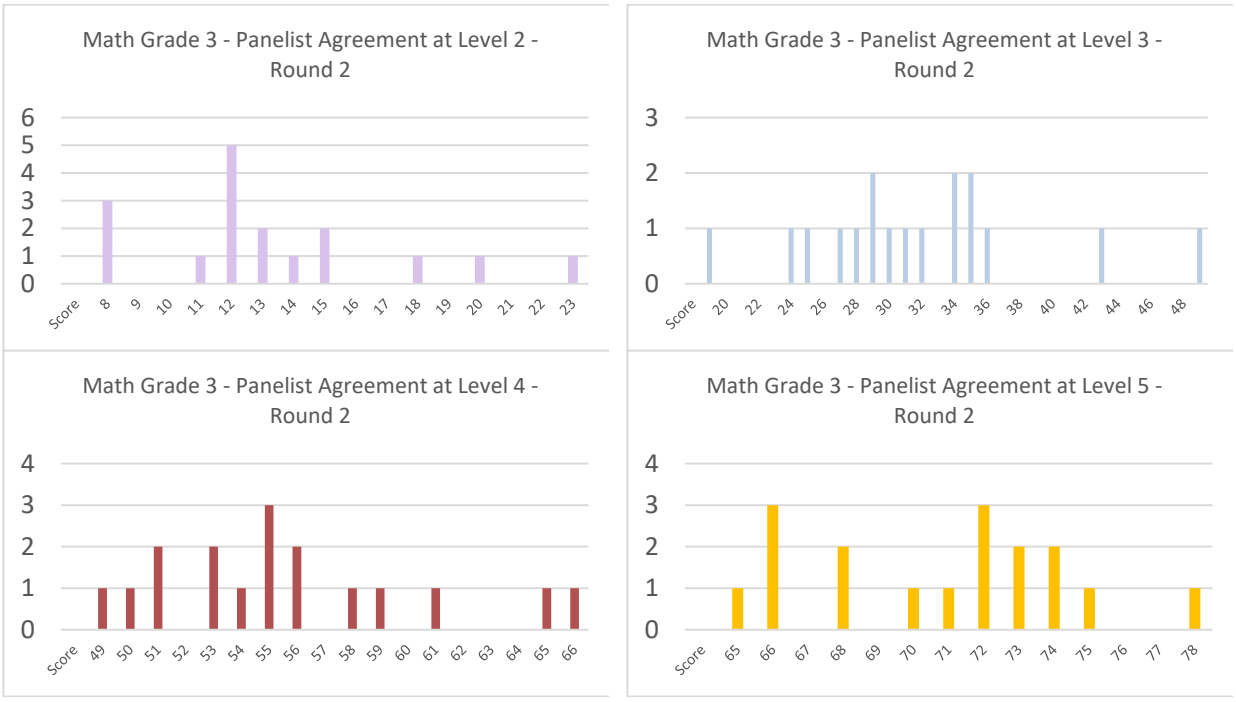



**Figure A10.42:** Grade 3 Mathematics - Round 2

# Appendix 12: Estimated Impact Data

This appendix presents the estimated impact data, or percent of students in each performance level, based on the threshold scores at four different times during the performance level setting process:

- After round 3 judgments of the standard setting committee meetings
- After the vertical articulation of the standard setting committee meetings
- After the reasonableness review of the PARCC Governing Board
- After the shift of the performance levels by the PARCC Governing Board

The impact data presented is based on student performance on the assessments administered during the Spring 2015 administration. The impact data presented after Round 3 judgments and during Vertical Articulation during the PLS meetings include the performance from only those students that took specific online forms of the respective assessment (as discussed in Chapter 6). For the Governing Board and final estimates after the shift of the performance levels, the impact data included all students that took an online form of the respective assessment (as discussed in Chapter 7).

## High School English Language Arts/Literacy (ELA/L)

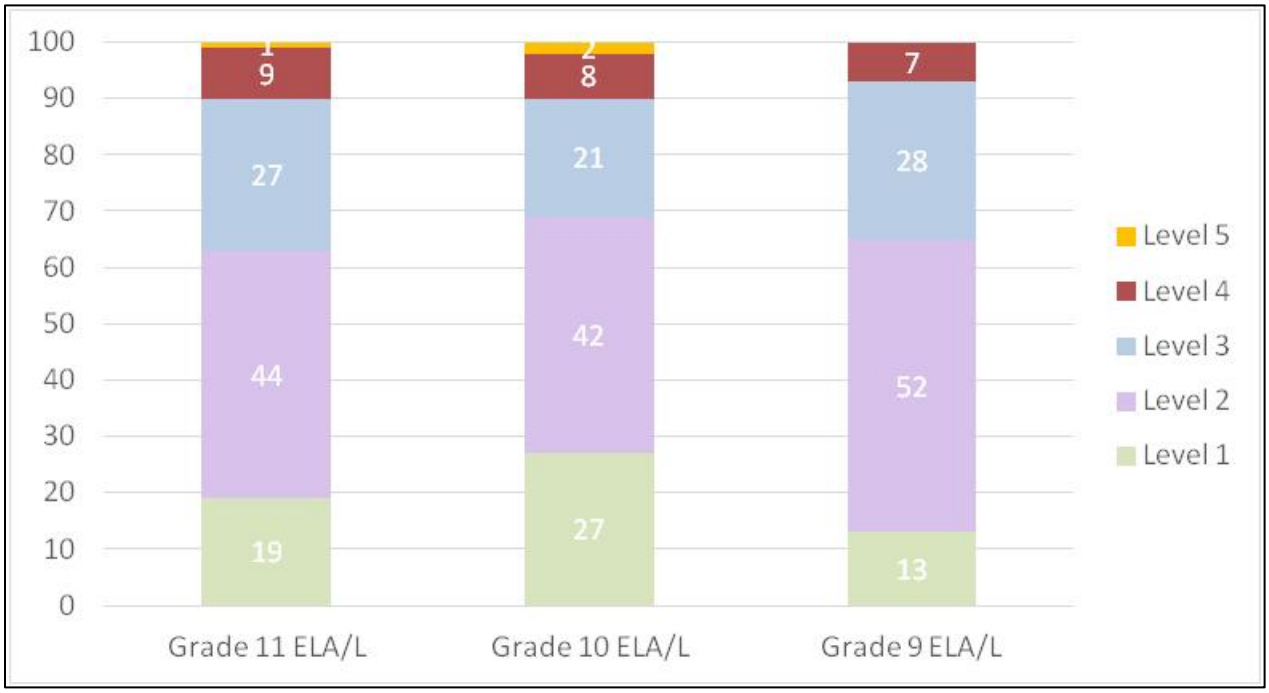



**Figure A11.1:** Estimated Impact for Grades 9-11 ELA/L after Round 3

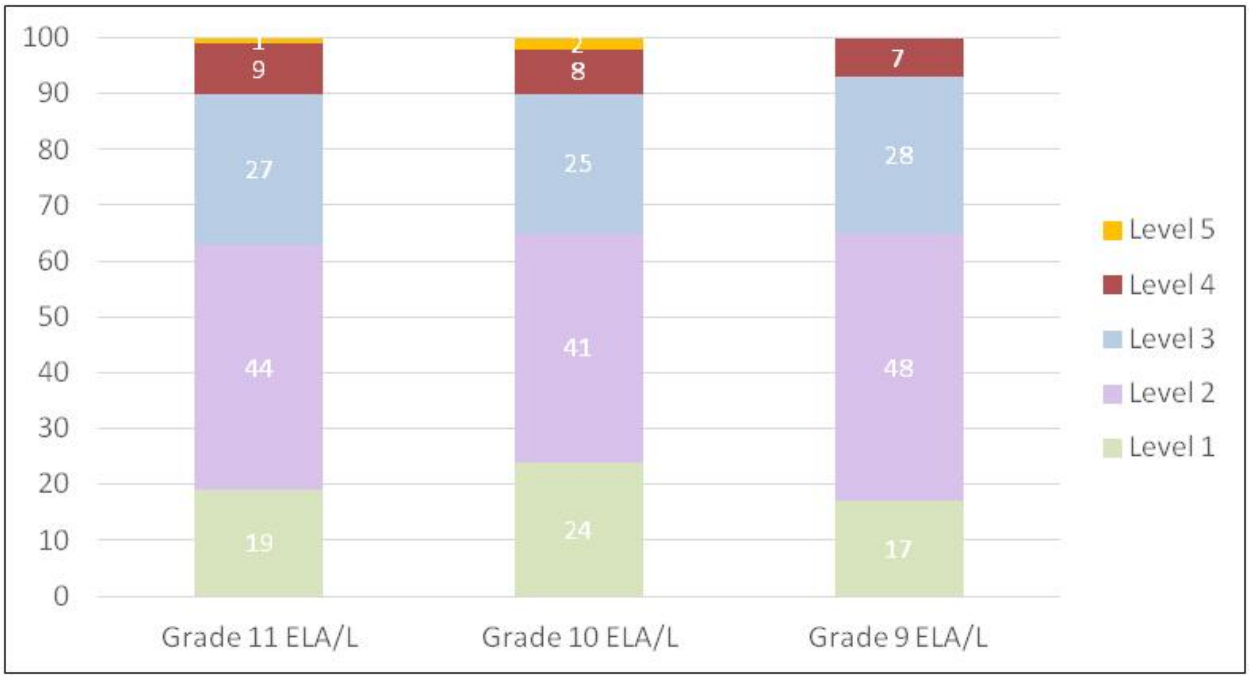



**Figure A11.2:** Estimated Impact for Grades 9-11 ELA/L after Vertical Articulation

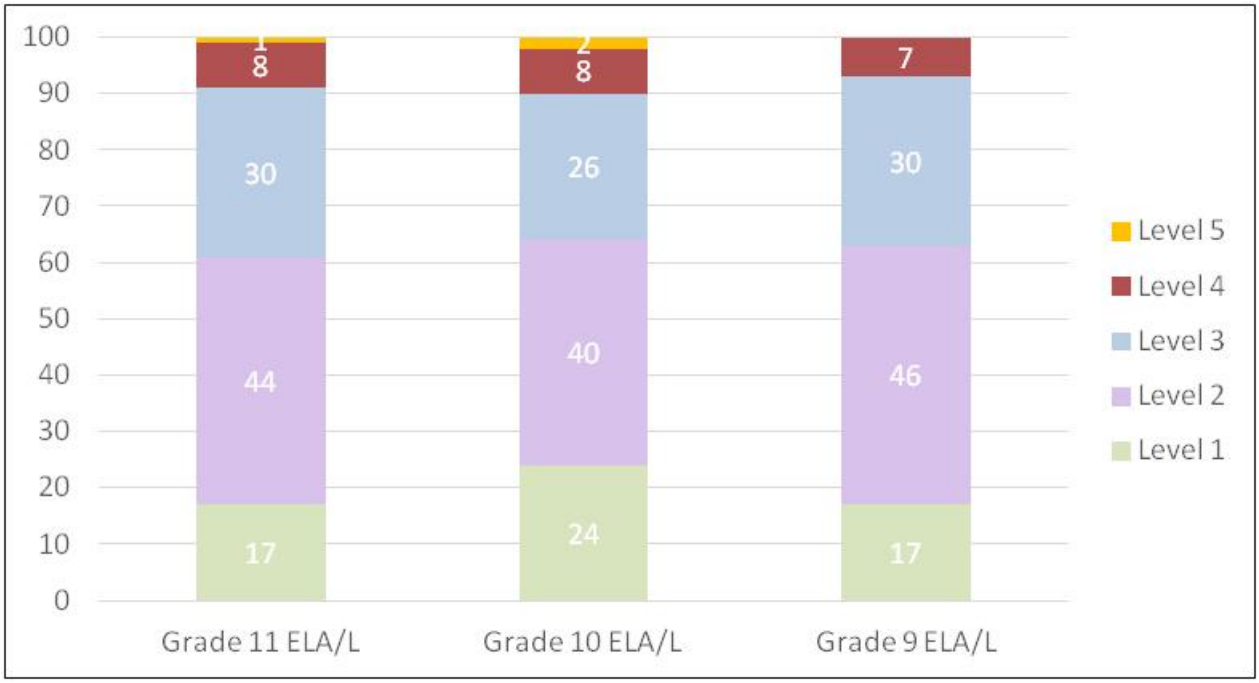



**Figure A11.3:** Estimated Impact for Grades 9-11 ELA/L after Reasonableness Review

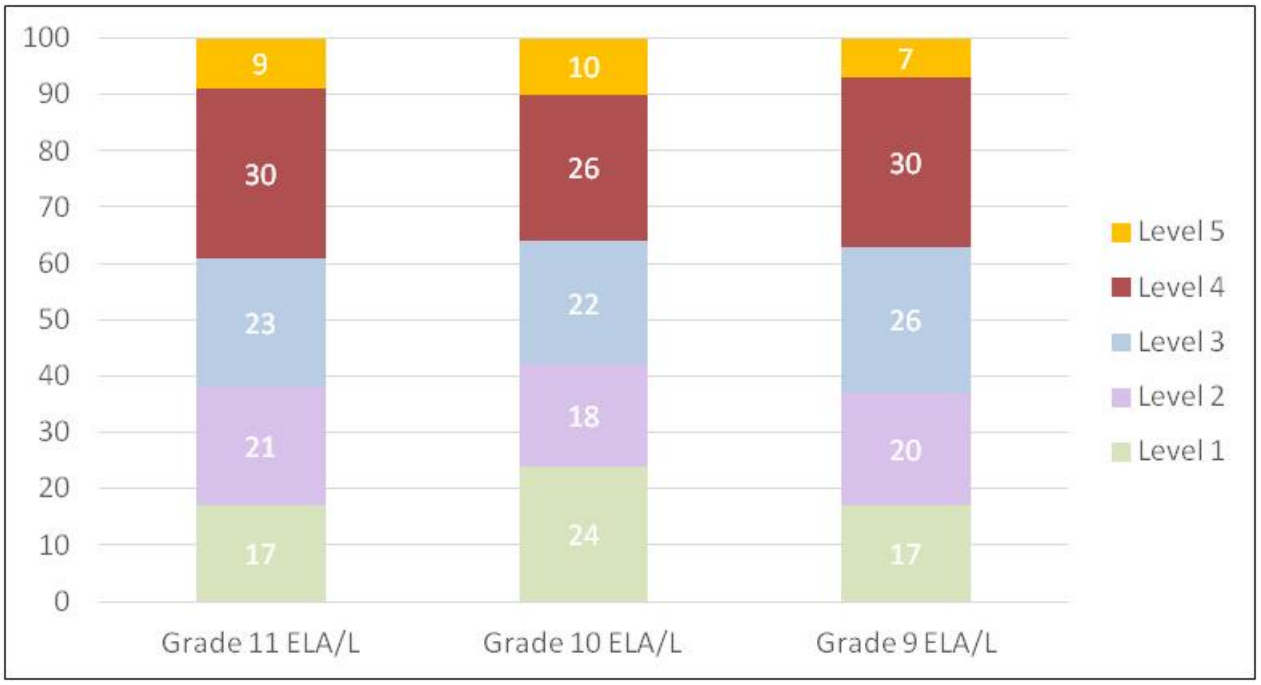



**Figure A11.4:** Estimated Impact for Grades 9-11 ELA/L after performance level shift

## High School Mathematics - Traditional Courses

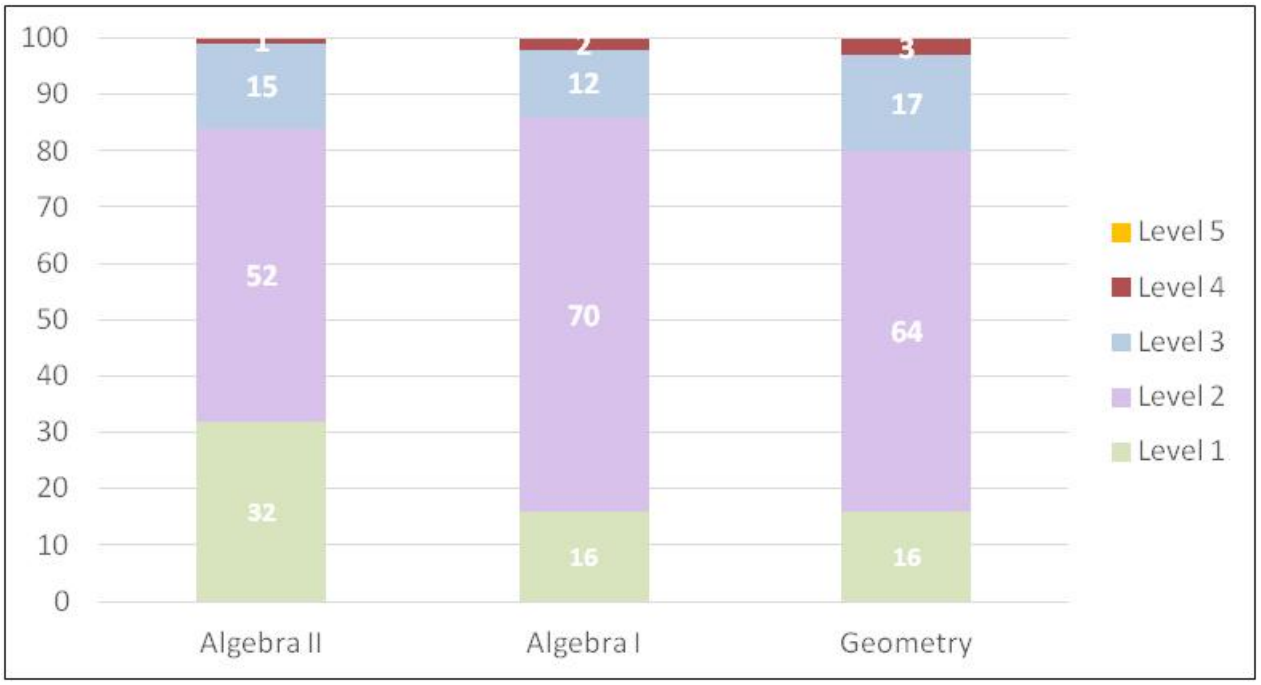



**Figure A11.5:** Estimated Impact for Traditional High School Math Courses after Round 3 Judgments

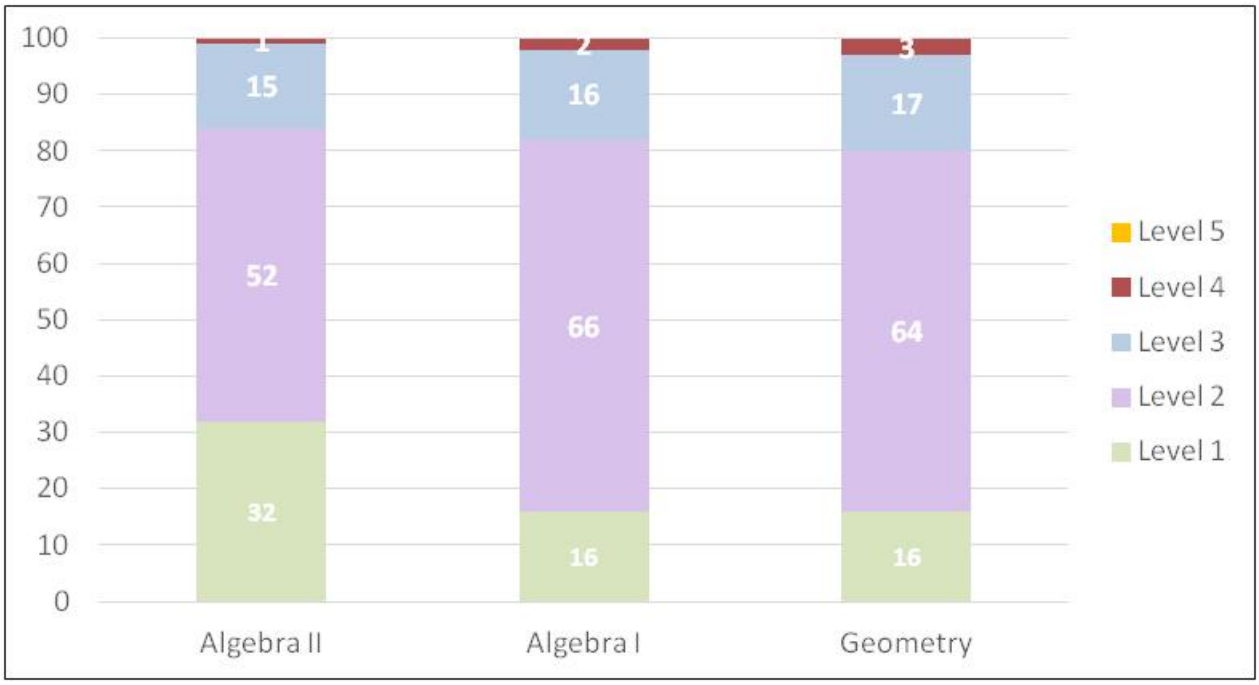



**Figure A11.6:** Estimated Impact for Traditional High School Math Courses after Vertical Articulation

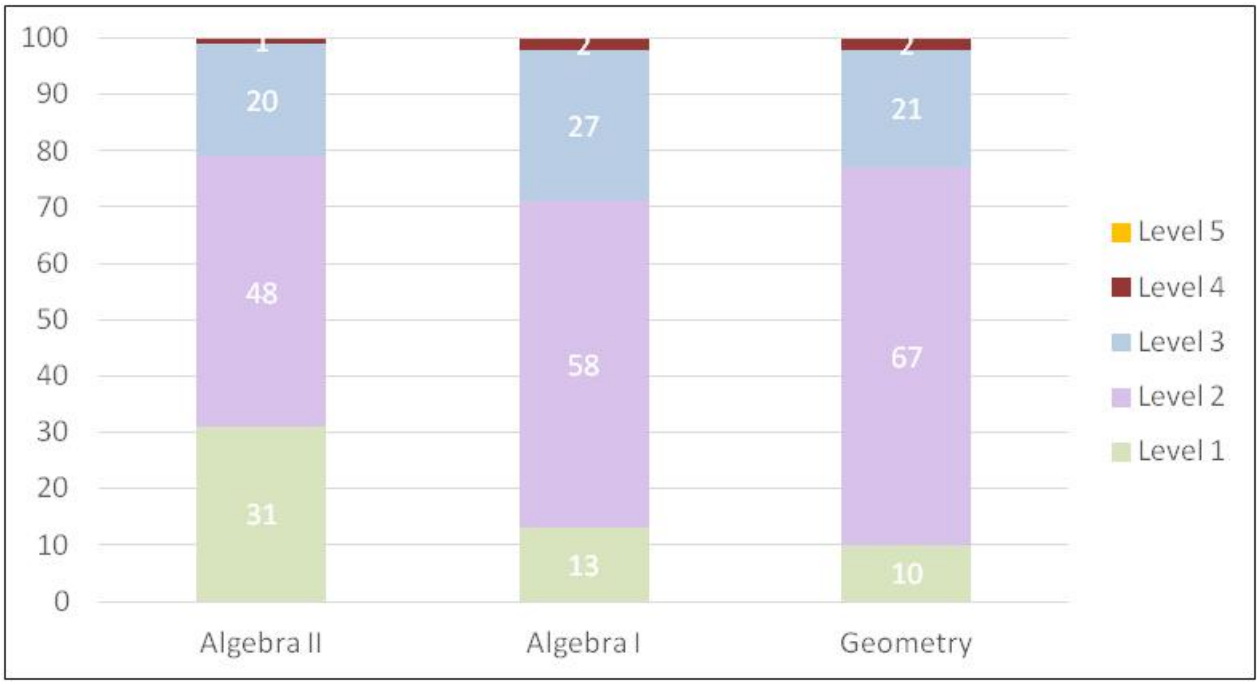



**Figure A11.7:** Estimated Impact for Traditional High School Math Courses after Reasonableness Review

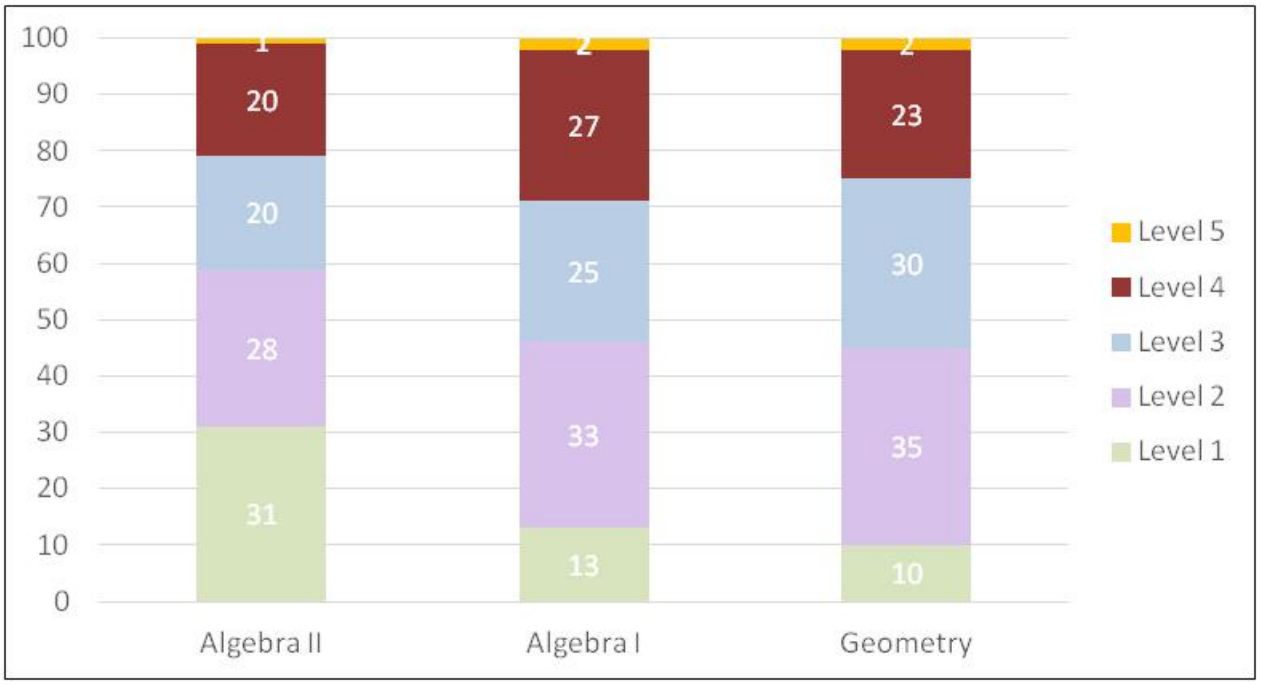



**Figure A11.8:** Estimated Impact for Traditional High School Math Courses after performance level shift

## High School Mathematics - Integrated Courses

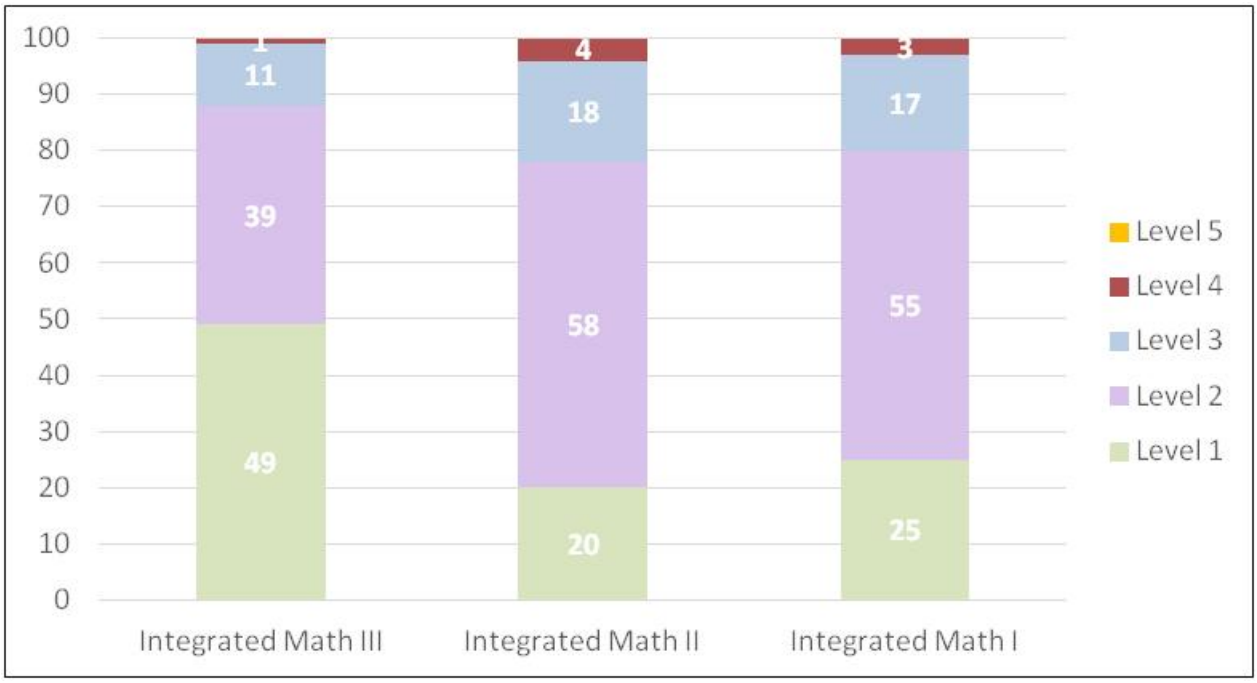



**Figure A11.9:** Estimated Impact for Integrated High School Math Courses after Round 3 Judgments

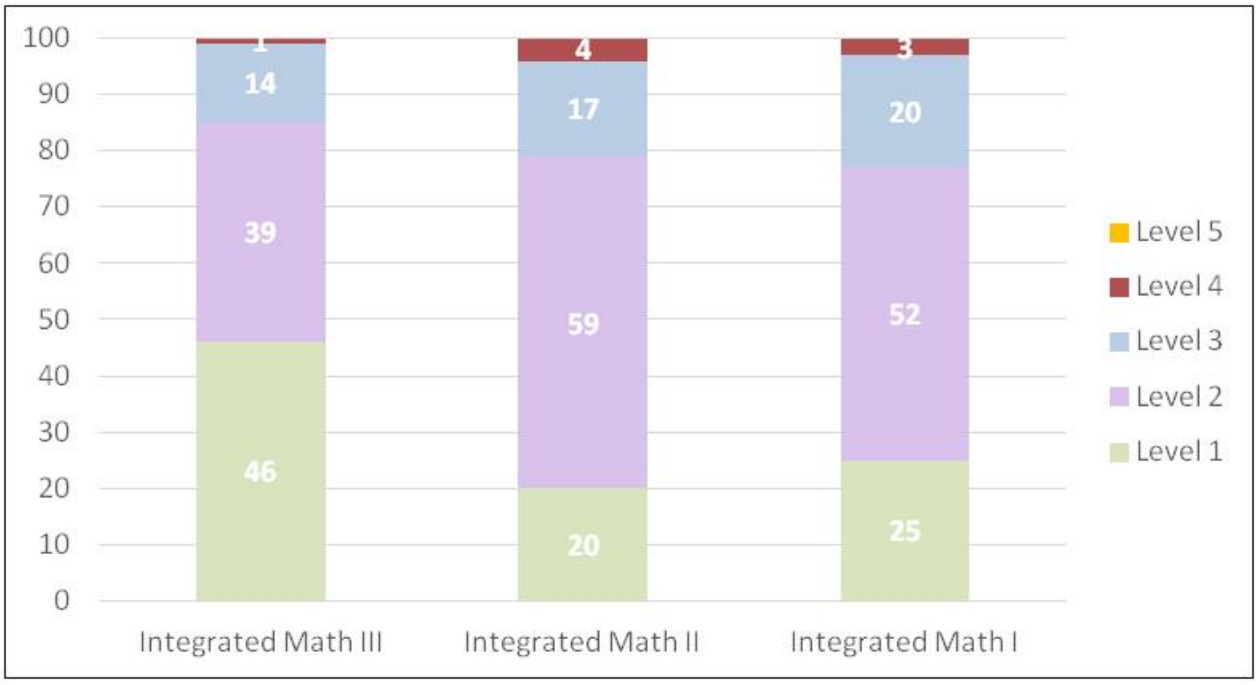



**Figure A11.10:** Estimated Impact for Integrated High School Math Courses after Vertical Articulation

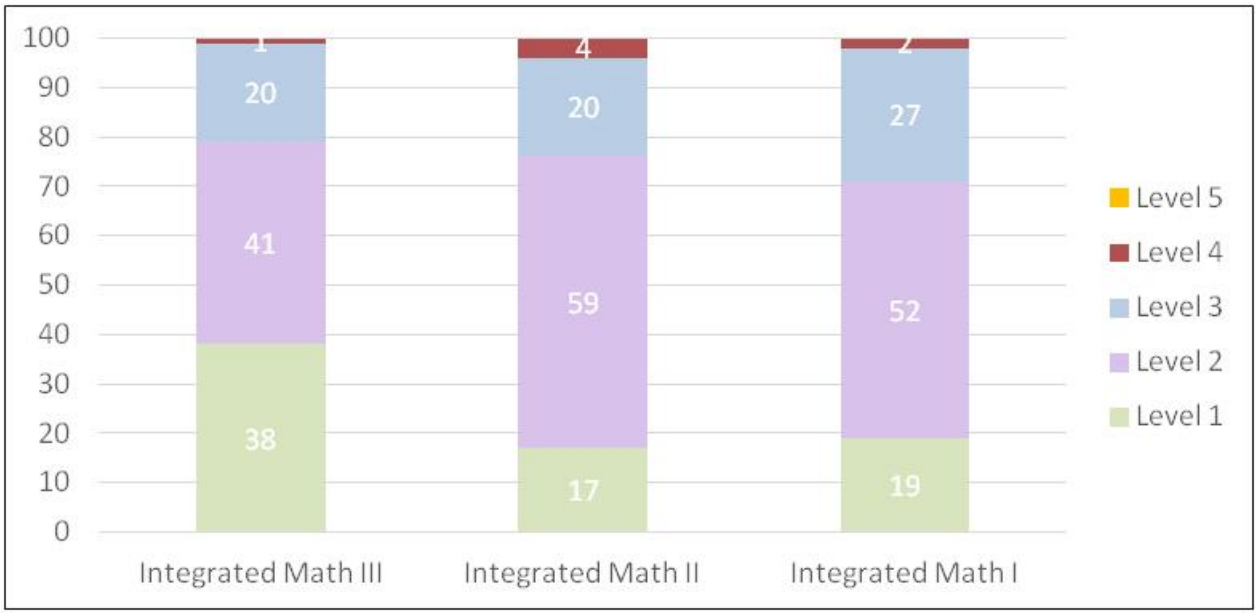



**Figure A11.11:** Estimated Impact for Integrated High School Math Courses after Reasonableness Review

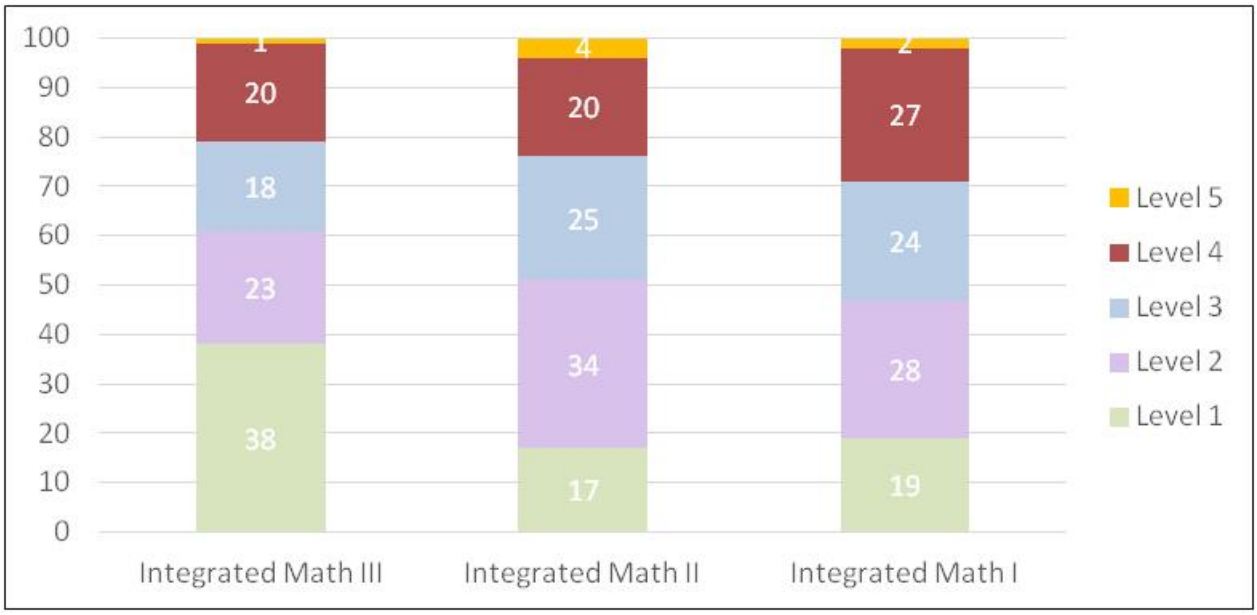



**Figure A11.12**: Estimated Impact for Integrated High School Math Courses after performance level shift

## Grades 3-8 English Language Arts/Literacy (ELA/L)

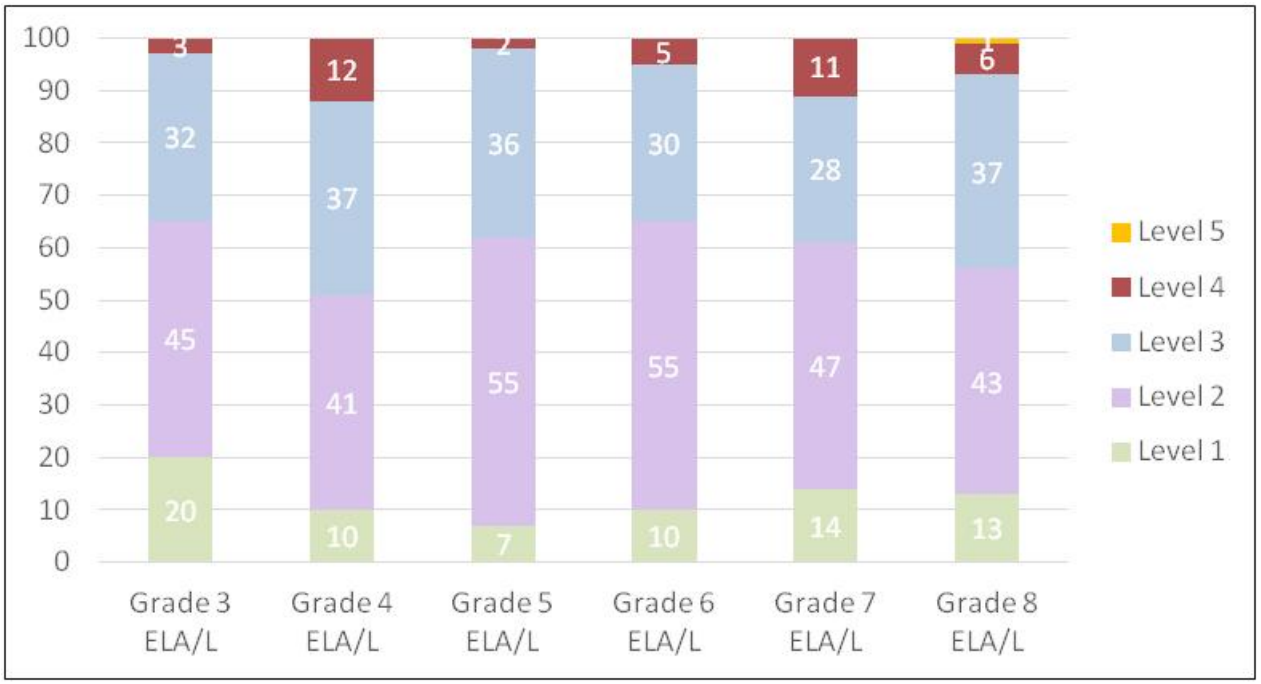



**Figure A11.13:** Estimated Impact for Grades 3-8 ELA/L after Round 3 Judgments

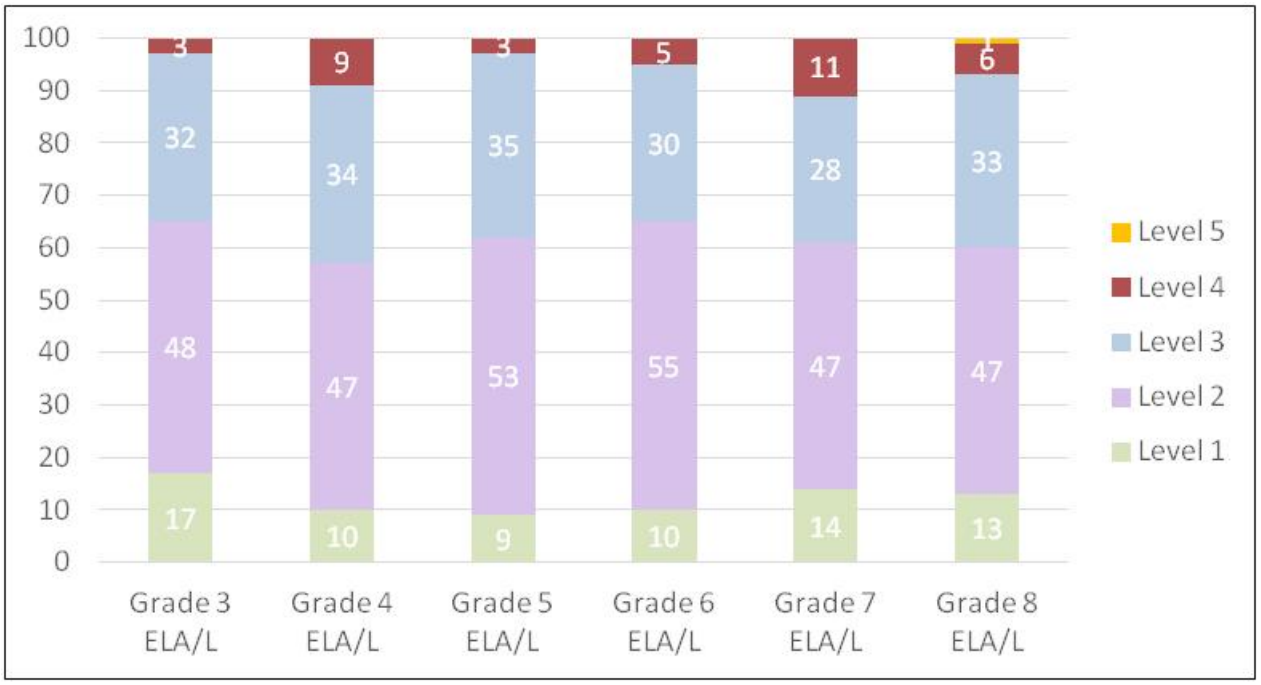



**Figure A11.14:** Estimated Impact for Grades 3-8 ELA/L after Vertical Articulation

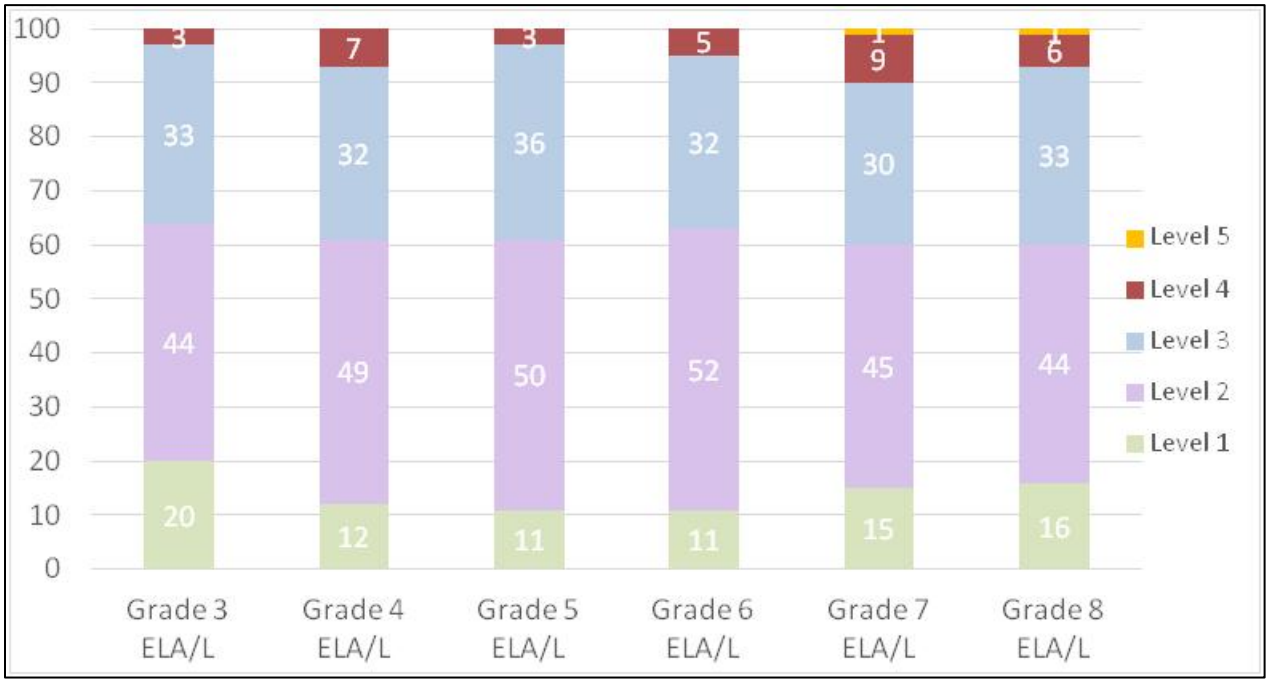



**Figure A11.15:** Estimated Impact for Grades 3-8 ELA/L after Reasonableness Review

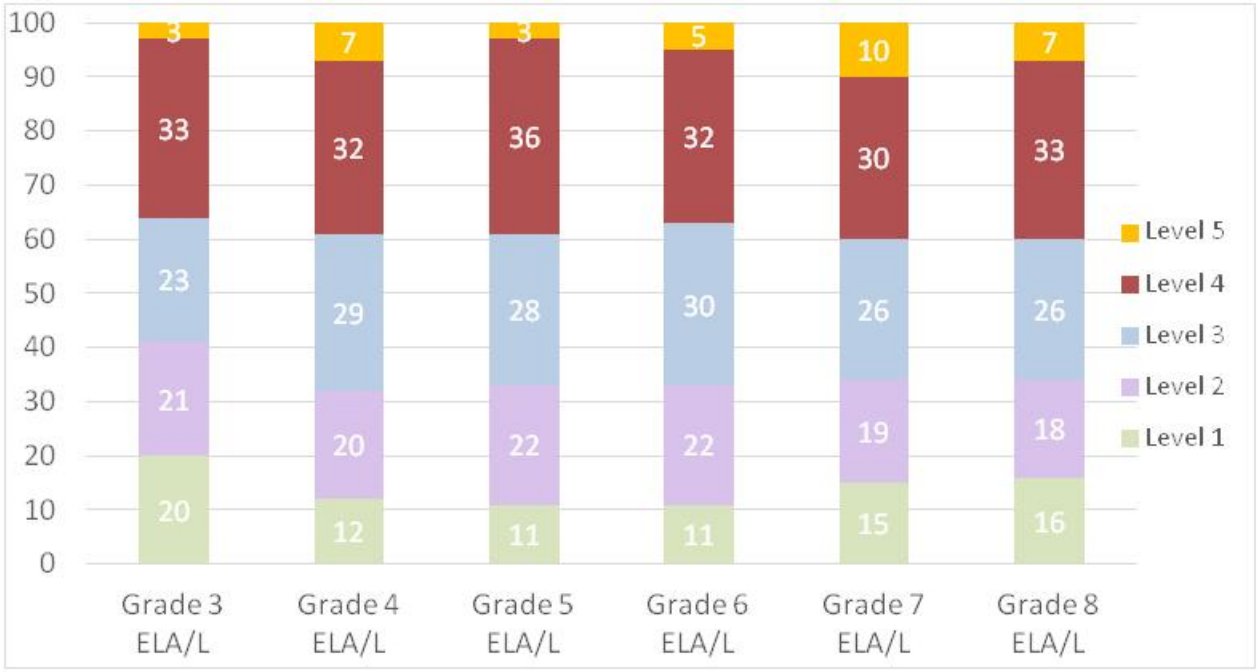



**Figure A11.16:** Estimated Impact for Grades 3-8 ELA/L after performance level shift

## Grades 3-8 Mathematics

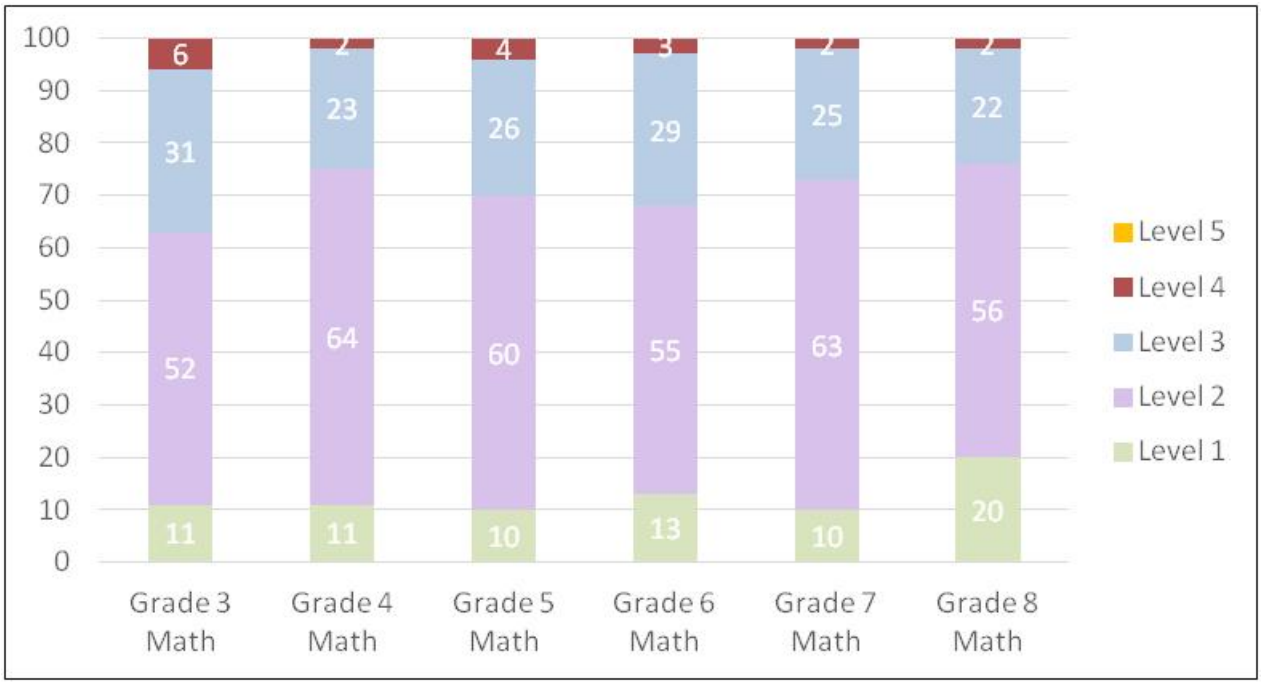



**Figure A11.17:** Estimated Impact for Grades 3-8 Math Courses after Round 3 Judgments

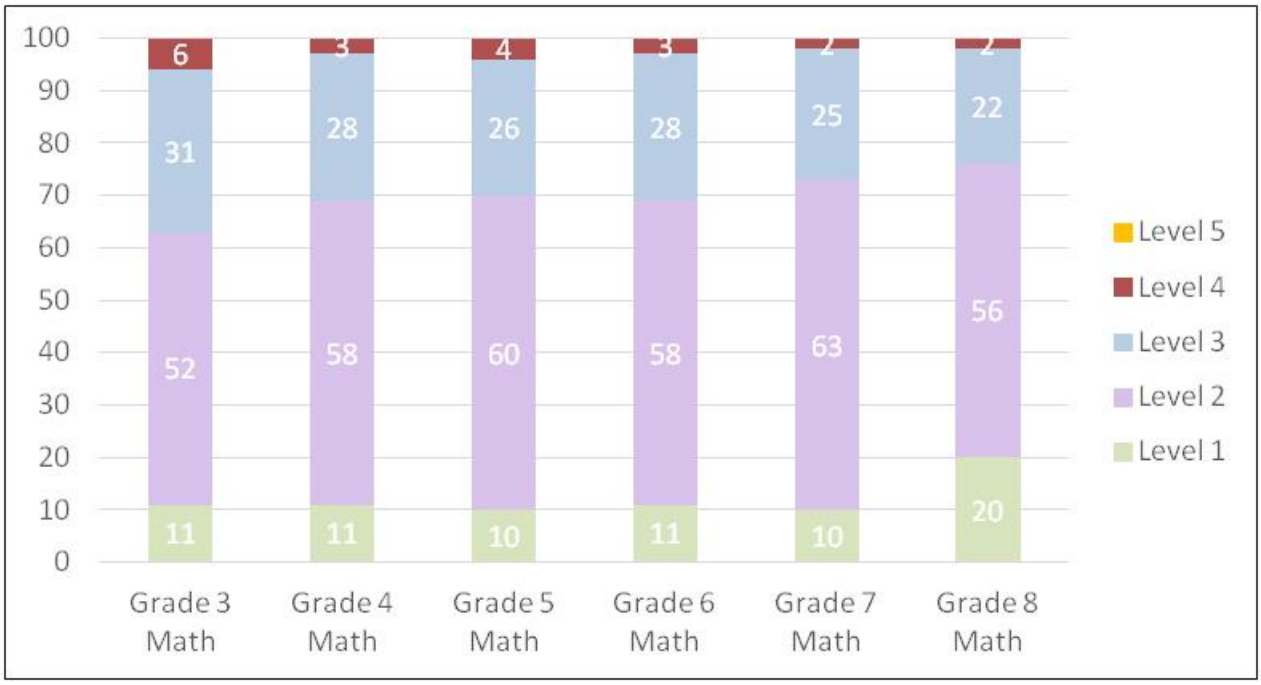



**Figure A11.18:** Estimated Impact for Grades 3-8 Math Courses after Vertical Articulation

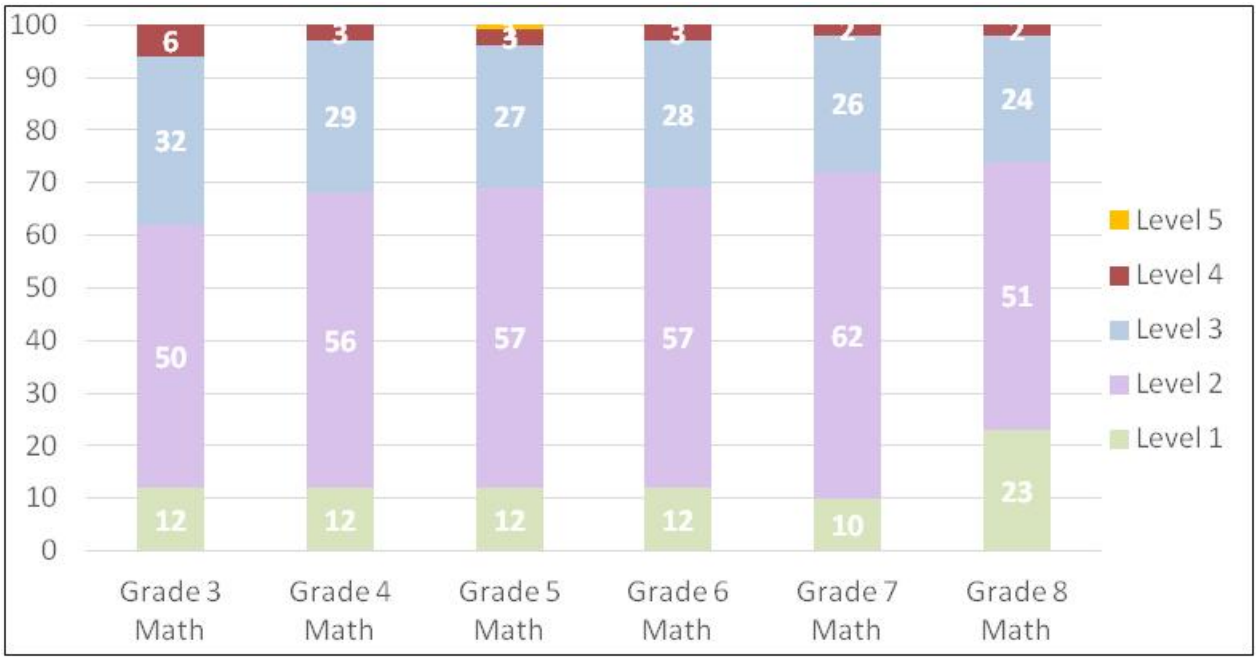



**Figure A11.19:** Estimated Impact for Grades 3-8 Math Courses after Reasonableness Review

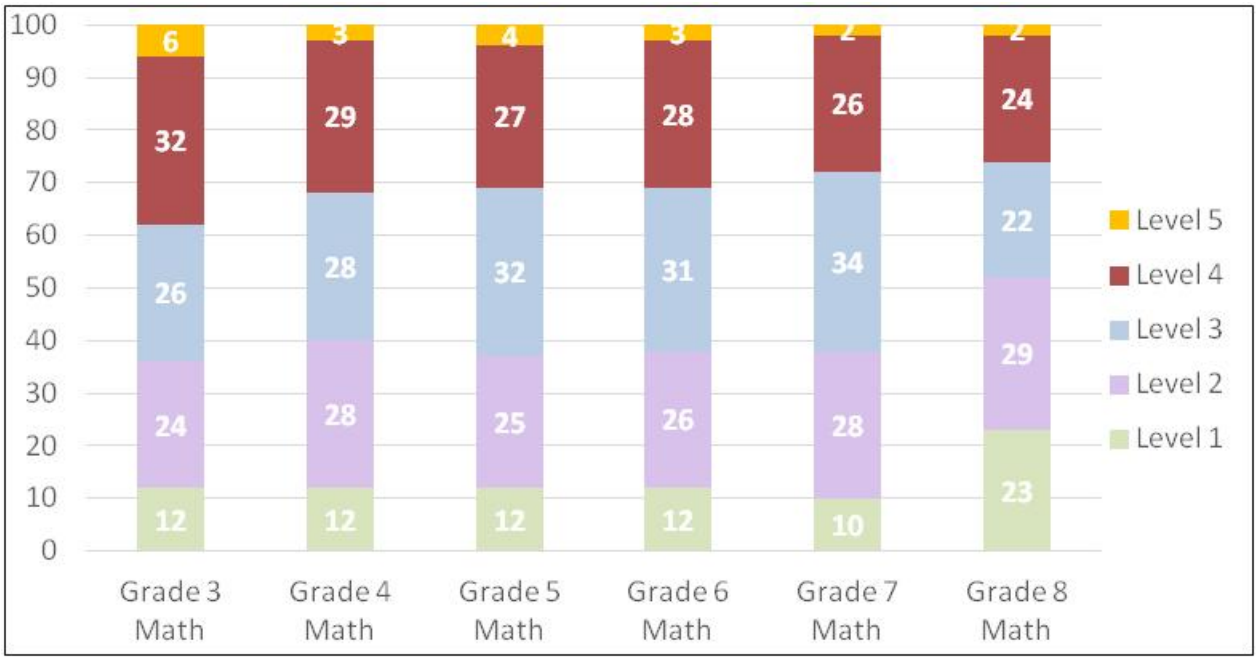



**Figure A11.20:** Estimated Impact for Grades 3-8 Math Courses after performance level shift

## Appendix 13: Reasonableness Review Committee Members

A complete list of all members of the reasonableness review committee, including their names and roles—K-12 PARCC Governing Board (GB)or Higher Education Advisory Committee on College Readiness (ACCR)—at the time of the reasonableness review meeting is provided below.

| Name | State | Role | HS Meeting | 3-8 Meeting |
|---|---|---|---|---|
| Elliott Asp | CO | K-12 GB | Y | -- |
| Joyce Zurkowski* | CO | K-12 GB | -- | Y |
| Rhonda Epper | CO | Higher Ed/ACCR | Y | -- |
| Hanseul Kang | DC | K-12 GB | Y | Y |
| Ronald Mason | DC | Higher Ed/ACCR | Y | -- |
| James Applegate | IL | Higher Ed/ACCR | Y | -- |
| Tony Smith | IL | K-12 GB | Y | Y |
| Mitchell Chester | MA | K-12 GB | Y | Y |
| Susan Lane | MA | Higher Ed/ACCR | Y | -- |
| Jack Smith | MD | K-12 GB & Higher Ed/ACCR | Y | Y |
| Gregg Edwards | NJ | Higher Ed/ACCR | Y | -- |
| David Hespe | NJ | K-12 GB | Y | Y |
| Barbara Darmon | NM | Higher Ed/ACCR | Y | -- |
| Hanna Skandera | NM | K-12 GB | Y | Y |
| Ken Wagner | RI | K-12 GB | Y | Y |
| Jim Purcell | RI | Higher Ed/ACCR | Y | -- |

*Voting as a proxy for Elliott Asp

## References

Beimers, J. N., Way, W. D., McClarty, K. L., & Miles, J. A. (2012, January). Evidence based standard setting: Establishing cut scores by integrating research evidence with expert content judgments. Austin, TX: Pearson. Retrieved from http://researchnetwork.pearson.com/wp-content/uploads/Bulletin21_Evidence_Based_Standard_Setting.pdf

CCSSO. (2015). *Common Core State Standards Initiative*. Retrieved from http://www.corestandards.org/

Moses, T., & von Davier, A. (no date). Using PROC GENMOD for Loglinear Smoothing. Paper SA-065. Educational Testing Service, Princeton, NJ.

Plake, B. S., Ferdous, A. A., Impara, J. C., & Buckendahl, C. W. (2005). Setting Multiple Performance Standards Using the Yes/No Method: An Alternative Item Mapping Method. *Meeting of the National Council on Measurement in Education.* Montreal. Canada.