Texas English Language Proficiency Assessment System (TELPAS) Grades 2–12 Standard-Setting Technical Report

Summer 2018

Table of Contents

Chapter 1 – Overview of the Standard-Setting Process 3
Goals of the Standard-Setting Meeting 3
TELPAS English Language Proficiency Levels 3
The TELPAS Standard-Setting Process 5
Chapter 2 – Pre-meeting Development 6
TELPAS Curriculum Standards 6
TELPAS Language Domains 6
TELPAS Proficiency Level Descriptors 7
Development of Panelist Materials 8
Development of Presentation Materials 8
Facilitator Training 9
Preparation for Data Analysis During the Meetings 9
Chapter 3 – Standard-Setting Meetings 11
Purpose of the Standard-Setting Meetings 11
Standard-Setting Meeting Participants 11
Committee Panelists 12
Standard-Setting Meeting Facilitators and Staff 13
Data Analysts 14
Observers 14
Materials 15
Pearson Standard Setting Website 15
Committee Panelist Folders 16
Computers 16
Procedure 16
Standard-Setting Meeting Proceedings 17
General Session 17
Borderline PLD Sessions 17
Breakout Sessions 19
Introductions and Overview 19
Standard-Setting Training and Judgment Rounds 20
Process Evaluation 25
Recommended TELPAS Cut Scores from Standard-Setting Committees 26
Chapter 4 – Post-Standard Setting 30
Composite Score Meeting 30
Meeting Process 30

Meeting Results 31
Reasonableness Review 32
Reading Standards Review 34
Chapter 5 – Evidence of Procedural Validity of the Standard-Setting Process 37
Committee Representation 37
Committee Training 37
Perceived Validity of the Meeting 38
Process Standardization 40
References 42
Appendix A – Proficiency Level Descriptors 43
Appendix B – Meeting Agenda 47
Appendix C – Participants' Meeting Materials 53
Appendix D – Committee Participant Composition 67
Appendix E – Examples of Feedback Data 82
Individual Test-Level Recommendations 82
Table-level Test-Level Recommendations 82
Overall Test-Level Recommendations 83
Item-Level Judgment Agreement 83
Test-Level Participant Recommendation Agreement 84
Item Score Mean and Score Distribution 85
Impact Data 85
Appendix F – Committee Recommended Cut Scores by Round 87
Appendix G – Recommended Cut Score Summary Statistics 91
Appendix H – Test-Level Participant Judgment Agreement 105
Appendix I – Participant Evaluation Results 147
Appendix J – Composite Score Profiles 190

Chapter 1 – Overview of the Standard-Setting Process

This chapter provides an overview of the standard-setting process used for the Texas English Language Proficiency Assessment System (TELPAS) for grades 2–12, and includes the following sections:

- Goals of setting cut scores
- TELPAS English language proficiency levels
- TELPAS standard-setting process

Goals of the Standard-Setting Meeting

English language proficiency (ELP) assessments are federally required to evaluate the progress that English learners (ELs) make in becoming proficient in the use of academic English. Under the Elementary and Secondary Education Act (ESEA) section 1111(b)(2)(G), all ELs in grades K–12 must participate in an annual ELP assessment that is aligned with ELP standards and covers the domains of listening, speaking, reading, and writing. The TELPAS assessments measure English language proficiency in the four language domains required.

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 proficiency levels, a frame of reference is developed for interpreting student performance. Setting an achievement standard for each language proficiency level is a critical step in developing an assessment program. For a criterion standards-based assessment, such as the TELPAS program, achievement on the assessment is compared to a set of predefined content standards.

The standards define a set of knowledge, skills, and abilities that the students are expected to demonstrate upon completion of each course or grade. A cut score is the minimum score a student must achieve on the assessment in order to be placed in a designated proficiency level. It distinguishes one proficiency level from the next, representing a student's competency level.

TELPAS English Language Proficiency Levels

Student proficiency for each language domain (i.e., listening, speaking, reading, and writing) on the TELPAS assessment is classified into four English language proficiency levels, or stages of increasing English language proficiency. The four levels are beginning, intermediate, advanced, and advanced high.

The original proficiency level standards for TELPAS were established in 2008 when the Texas Assessment of Knowledge and Skills (TAKS) was the academic assessment in Texas. Due to the move from TAKS to the State of Texas Assessments of Academic Readiness (STAAR®)

assessment program in 2012, the TELPAS reading standards were reviewed in 2013 utilizing a standards review process. In spring 2018, there was another change to the TELPAS reading test design in addition to the development of an online test for the listening and speaking domains, which required new cut scores be established for the four TELPAS proficiency levels.

The global definitions and key features for the TELPAS proficiency levels provide general expectations for the classification of student achievement. Global definitions explain what it means for a student to be classified as beginning, intermediate, advanced, or advanced high across language domains. These do not differentiate student proficiency between domains and grade levels and apply to all TELPAS assessments. The global definitions and key features for the proficiency levels are shown in Table 1.

Table 1. *TELPAS Proficiency Level Global Definitions*

Proficiency Level	Global Definition	Key Features
Beginning	Beginning students have little or no ability to understand and use English. They may know a little English but not enough to function meaningfully in social or academic settings.	Little or no English ability
Intermediate	Intermediate students do have some ability to understand and use English. They can function in social and academic settings as long as the tasks require them to understand and use simple language structures and high-frequency vocabulary in routine contexts.	Limited ability, simple language structures, high-frequency vocabulary, routine contexts
Advanced	Advanced students are able to engage in grade-appropriate academic instruction in English, although ongoing second language acquisition support is needed to help them understand and use grade-appropriate language. These students function beyond the level of simple, routinely used English.	Ability to engage in grade-appropriate academic instruction with second language acquisition support
Advanced High	Advanced high students have attained the command of English that enables them, with minimal second language acquisition support, to engage in regular, all-English academic instruction at their grade level.	Ability to engage in grade-appropriate academic instruction with minimal second language acquisition support

The TELPAS Standard-Setting Process

The recommendations by the standard-setting committees represent the level of proficiency students are expected to demonstrate to be classified into each of the proficiency levels. To establish the proficiency levels for each domain, a test-centered, criterion-referenced method was used to guide panelists as they determined their proficiency level cut score recommendations. The method applied was a hybrid of the Angoff method (Angoff, 1971) and Extended Modified (Yes/No) Angoff method (Davis & Moyer, 2015; Plake, Ferdous, Impara, & Buckendahl, 2005). This standard-setting procedure is a systematic method for combining various considerations into the process for recommending cut scores for the different proficiency levels.

The following steps were used for the TELPAS standard-setting process:

- Pre-meeting development — In anticipation of the standard-setting meetings, various tasks were completed, including the development of draft proficiency level descriptors (PLDs) for each grade-band and domain assessed, the development of materials for the panelists, preparation of the Pearson Standard Setting website for panelists and facilitators, presentation materials for the facilitators, and development of data analysis sources and procedures.
- Standard-setting meetings — Committees of panelists referenced the grade-band and domain-specific PLDs to make recommendations for cut scores that define the different proficiency levels for each assessment.
- Composite score review — The rules to establish the composite TELPAS proficiency levels using the domain-level proficiency levels and scores were established.
- Reasonableness review — The Texas Education Agency (TEA) conducted a reasonableness review of the TELPAS cut score recommendations. The cut scores for each domain were reviewed in comparison to the recommendations for the other domains to determine the reasonableness of the system of proficiency standards recommended.
- Standards verification — The recommended cut scores for the reading tests were presented to a committee of educators along with additional external data from validity studies. The committee recommended modifications to the cut scores for proficiency levels associated with the TELPAS reading assessment.

The remaining chapters will describe the specific procedures and activities that occurred during each step of the standard-setting process.

Chapter 2 – Pre-meeting Development

This chapter provides an overview of the work completed prior to the standard-setting meetings for TELPAS grades 2–12, and includes the following sections:

- TELPAS Curriculum Standards
- TELPAS English Language Proficiency Levels
- TELPAS PLDs
- Development of panelist materials
- Development of presentation materials
- Facilitator training
- Preparation for data analysis during the meetings

TELPAS Curriculum Standards

The English language proficiency standards (ELPS) are curriculum standards that support the ability of ELs to acquire academic English, while at the same time allowing them to engage meaningfully in regular, all-English academic instruction at their grade level. Approved by the State Board of Education in 2007–2008, the ELPS are set forth in Title 19, Chapter 74.4 of the Texas Administrative Code (TAC). Districts are required to implement the ELPS as an integral part of each foundation and enrichment subject of the Texas Essential Knowledge and Skills (TEKS) state-required curriculum. The ELPS outline the instruction that ELs must receive to support their ability to develop academic English language proficiency and acquire challenging academic knowledge and skills.

TELPAS Language Domains

The English language proficiency levels classify stages of English language acquisition. Student proficiency for each language domain assessed on TELPAS listening, speaking, reading, and writing will be classified into four levels, or stages of increasing English language proficiency. For assessment purposes, the language domains are defined in Table 2.

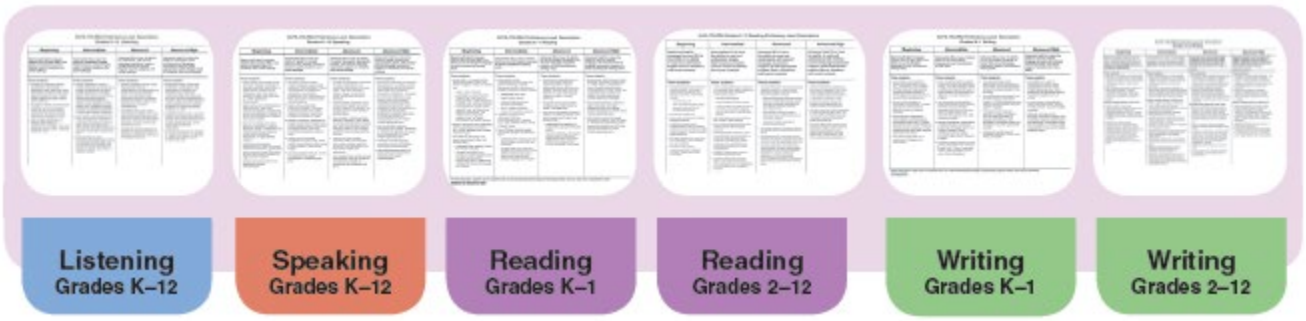
Table 2. *TELPAS Language Domain Definitions*

Domain	Definition
Listening	The ability to understand spoken language, comprehend and extract information, and follow social and instructional discourse through which information is provided
Speaking	The ability to use spoken language appropriately and effectively in learning activities and social interactions
Reading	The ability to comprehend and interpret written text at the grade-appropriate level
Writing	The ability to produce written text with content and format to fulfill grade-appropriate classroom assignments

TELPAS Proficiency Level Descriptors

While the global definitions apply across language domains, the PLDs present the major characteristics of each proficiency level in each language domain. The PLDs are domain-specific and define how well ELs at the four proficiency levels are able to understand and use English in grade-level academic settings. The descriptors show the progression of second language acquisition from one proficiency level to the next. The PLDs are also a critical part of the process used to set the TELPAS standards. They provide a common framework for understanding the language acquisition skills needed to be classified within each proficiency level.

There is one set of PLDs for listening and one set for speaking. For reading and writing, there are two sets. The separate sets of K–12 PLDs address emergent literacy.



The PLDs are presented in a chart format for ease of use and include the proficiency level, a summary statement, and specific descriptors. The PLDs for grades K–12 listening and speaking and grades 2–12 reading can be found in Appendix A.

Development of Panelist Materials

The TELPAS standard setting required a large number of materials for use by panelists during the standard-setting meetings. The Pearson standard-setting team worked with content specialists at TEA to develop the materials used during the meetings and to ensure all materials provided to the panelists communicated correct information. The following materials were developed for use by panelists during the meeting:

- Meeting agenda
- Panelist information survey*
- Non-disclosure agreement
- TELPAS reading, listening, and speaking test forms*
- “Experience the Test” activity response forms
- Test form answer keys
- Open-ended speaking item rubrics and exemplars
- Item comment form
- Practice judgment forms and test maps
- Judgment item maps
- Judgment round surveys*—rounds 1, 2, and 3
- PLDs and PLD comment form
- ELPS
- Process evaluations*

Because the meetings utilized the Pearson Standard Setting website as a tool for facilitation, the website for each subject area needed to be developed by grade. Several of the documents developed, indicated with an asterisk (*), were presented online through the standard-setting website.

The Pearson standard-setting team used approved templates to create documents for each grade-level committee. All documents developed for the website were reviewed and approved by TEA staff before being finalized for publication for the meetings. A sample set of materials for a committee meeting are provided in Appendix C.

Development of Presentation Materials

Presentations with notes were developed to guide facilitators through the presentation of information and materials throughout the standard-setting meetings. The Pearson standard-setting team developed the initial presentations. Staff from TEA reviewed and suggested edits to the presentations, which were resolved by the Pearson standard-setting team. The following presentations were created for the standard-setting meetings:

- Standard-Setting General Session
- Borderline Proficiency Level Meeting
- Standard-Setting Reading Breakout Meeting—Day 1
- Standard-Setting Reading Breakout Meeting—Day 2
- Standard-Setting Listening & Speaking Breakout Meeting—Day 3
- Standard-Setting Listening & Speaking Breakout Meeting—Day 4
- Standard-Setting Listening & Speaking Breakout Meeting—Day 5

- Composite Score Review

The presentations for the breakout meetings, Day 1 through Day 5, were customized to reflect the specific information for the domain and grades for each committee. Additionally, specific information was added to the notes section within each presentation to guide the facilitators through the presentations.

Facilitator Training

The meetings were led by a psychometrician from Pearson, or a process facilitator, with knowledge and experience leading standard-setting meetings. The facilitator was responsible for ensuring appropriate processes were followed throughout all sections of the meeting and that panelists had a solid understanding of the tasks they were being asked to complete.

All facilitators underwent an extensive program of training to prepare them for leading this set of standard-setting meetings. The training included:

- Use of the standard-setting website — Because the standard-setting website was used as a facilitation tool during the meeting, facilitators needed to become familiar with the use of the platform. Specific guidelines for modeling the website and providing access to the panelists were discussed.
- TELPAS — The facilitators were provided an overview of the TELPAS program, including the different item types, scoring rules, proficiency levels, and scaling design.
- Standard-setting process — The facilitators participated in a walk-through of the standard-setting meeting agenda, with a focus on specific issues for these meetings, such as time management, the use of the online platform, and communicating feedback information.
- Training slides and presentation notes — The facilitators were introduced to the standard-setting training slides before the meetings. Notes in the standard-setting training slides provided the facilitators with specific guidance throughout the presentation, including when specific language was to be used during panelist training.

The facilitator training meetings were held for 90 minutes each on May 17, 18, 21, and 24, 2018. Additionally, there was a final training and discussion held on-site on June 3, the day before the standard-setting meetings, to address any final topics. At the end of each day during the standard-setting meetings, a debriefing was held to discuss concerns, positives, and material planned for the next day.

Content experts from Pearson and TEA were also available as observers to assist in answering content and policy questions. A staffing plan was provided to TEA prior to the standard-setting meetings to communicate the psychometric and support staff scheduled to attend.

Preparation for Data Analysis During the Meetings

Creation and testing of analysis programs and the calculation of impact data lookup tables were conducted prior to the standard-setting meeting. To facilitate the independent analysis for each judgment round during the meeting, each analyst independently completed the programming

necessary to conduct all analysis using the SAS statistical software. A trial was run with mock data generated through the standard-setting website to ensure that each independent analysis generated the same results.

Prior to the standard-setting meetings, several sets of analysis were performed using the spring 2018 test results, which provided the data used during the meetings. These analyses included item response theory (IRT) calibration of the items, which were used to determine the association between the recommended raw score cuts and estimated student ability, as well as classical item analysis.

Impact data is the percent of students that fall within a proficiency level based on the recommended cut scores at the given judgment round for a particular domain and grade band. The impact data is provided to panelists during the standard-setting meeting to present the expected results of their recommendations on student proficiency level classifications. The analysis programs used impact data lookup tables, which were created prior to the standard-setting meetings, to produce this output during the meetings.

The impact data lookup tables were created using the data from students taking the online form of each subject and grade assessment during the spring 2018 administration. A frequency distribution of student results on the different test forms was created based on the spring 2018 administration. This frequency distribution was used to determine the expected percent of students classified into each proficiency level based on panelists' judgments.

In addition to the aforementioned programming for determining impact data, Pearson analysts have developed programs to generate all feedback handouts, plots, and tables needed during the standard-setting meeting. For example, following a round of judgment, the analyst produced:

- Individual panelist feedback — the judgments of the panelists to ensure that they were recorded and analyzed accurately (given to all panelists)
- Table-level feedback — a summary of judgments from all panelists at a table, including a frequency distribution of judgments and the mean and median (given to facilitators and table leaders)
- Panel-level feedback — a summary of judgments from all panelists, including a frequency distribution of judgments and the mean and median (given to facilitators and TEA, presented to the panelists using tables and histograms in digital presentations)
- Impact data (*after* judgment rounds 2 and 3) — the percentage of students, not disaggregated by demographic groups, in each proficiency level according to the recommended cut scores at that round (presented to panelists as stacked bar graphs in digital presentations)

Prior to the standard-setting meetings, the Pearson standard-setting team developed templates of each report. Staff from TEA were provided the opportunity to review and suggest final revisions to the reports.

Chapter 3 – Standard-Setting Meetings

This chapter provides details about the standard-setting meeting process. The sections of this chapter include:

- Purpose of standard-setting meetings
- Standard-setting meeting participants
- Materials used during the standard setting
- A description of the standard-setting procedure
- Standard-setting meeting proceedings
- Recommended proficiency level cut scores

Purpose of the Standard-Setting Meetings

Standard setting is based, to a large degree, on the judgment of educators. Committees of educators make expert recommendations about the performance expected for each proficiency 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 cut scores for the proficiency levels. The purpose of the TELPAS standard-setting meetings was to gather expert recommendations from groups of educators from across Texas for the cut scores that define the different proficiency levels on the TELPAS reading, speaking, and listening assessments for grades 2 through 12.

Student performance on each of the TELPAS assessments is classified into one of four proficiency levels. Each committee was asked to recommend three cut scores that would define the boundaries between the different proficiency levels. These recommended cut scores represent the performance on each assessment that a student would need to meet or exceed to be classified into the specific proficiency level.

Standard-Setting Meeting Participants

Standard-setting panelists for the TELPAS meetings included three distinct groups of people:

- Committee panelists
- Meeting facilitators
- Observers and staff

Figure 1 illustrates the room setup for the standard-setting meeting.

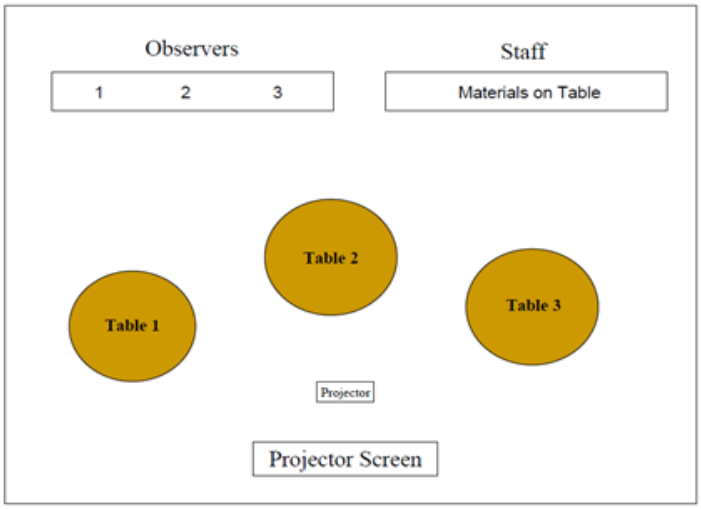


Figure 1. ***Physical Setup for Standard-Setting Meeting***

Committee Panelists

All panelists for the standard-setting committees were selected by TEA to represent educators and key stakeholders from across the state with knowledge and experience with student groups that are administered the TELPAS assessments. Panelists recruited for the meetings attended both the reading meeting and the listening and speaking meetings. The process of selecting committee panelists included selecting a sample of panelists that would be as representative of the state as possible, including demographic variables (gender, race, etc.), geographic representation, and background (educational experience, education, etc.). When selecting panelists, TEA placed an emphasis on those educators who had relevant content knowledge as well as experience with a variety of student groups.

There was a total of 76 panelists at the standard-setting meetings. These panelists were first divided into three committees, one for each domain, to develop general borderline proficiency level descriptors. The committees were then regrouped by grade and formed six committees for reading and then reshuffled based on grade band into four committees for listening and speaking.

The reading meetings consisted of six committees for the following grade or grade bands: grade 2, grade 3, grades 4 and 5, grades 6 and 7, grades 8 and 9, and grades 10 through 12. The listening and speaking meetings consisted of four committees for the following grade bands: grades 2 and 3, grades 4 and 5, grades 6 through 8, and grades 9 through 12. The same panelists participated in both meetings. Each committee focused on providing cut score recommendations for the grade/grade band and domain they were assigned. The panelists were assigned to the committee prior to the meetings based on their teaching experience. Table 3 displays the total number of panelists in each of the standard-setting committees. The tables in Appendix D summarize the characteristics and experience of the panelists in each committee. These tables provide demographic information about the panelists as well as information about the panelists' current positions in education, their experience working with various types of student populations, and the types of districts they represent. Panelists' response to the gender and ethnicity questions was voluntary.

Table 3. ***Number of Panelists in Each Standard-Setting Committee***

Borderline PLDs (June 4)		Reading (June 4–5)		Listening and Speaking (June 6–8)	
Domains	**# of panelists**	**Grade/Grade Band**	**# of panelists**	**Grade Band**	**# of panelists**
Reading	25	2	8	2–3	18
		3	10		
		4–5	14	4–5	14
Listening	22	6–7	13	6–8	23
		8–9	17		
Speaking	24	10–12	14	9–12	21

The panelists in each committee were assigned to table groups. The table groups were selected prior to the meeting to ensure that, to the greatest extent possible, the panelists at each table were representative of the committee. The panelists were placed into table groups to facilitate discussions during the standard-setting meeting and ensure that each panelist had the opportunity to fully engage in the process.

Prior to the standard-setting meeting, an individual from each table group was selected to serve as a table leader based on demonstrated leadership at previous educator meetings (e.g., standard setting, data review, content review). The table leaders assisted the process facilitator during the meeting by facilitating the table discussions, ensuring that all panelists had the opportunity to participate, and ensuring that the discussion remained relevant to the meeting.

Standard-Setting Meeting Facilitators and Staff

Staff members from TEA and Pearson collaborated to conduct the TELPAS standard-setting meeting. These staff members worked in facilitative and observational roles and did not contribute to the cut score recommendations during the meeting.

Facilitators

The process facilitator was a member of the Pearson psychometric staff with experience in facilitating standard-setting meetings and was responsible for leading the panelists through the standard-setting process. This individual was responsible for making sure processes were followed throughout all sections of the meeting and that panelists had a solid understanding of the tasks they were being asked to complete.

Although the facilitators had experience leading standard-setting meetings, all facilitators underwent extensive training to prepare them for this set of standard-setting meetings. The lead

facilitator of the standard-setting meeting was Eric L. Moyer, Ph.D., from Pearson. Table 4 presents the facilitators for each standard-setting committee.

Table 4. ***Process Facilitators for TELPAS Standard-Setting Committees***

Committee		Facilitators
Subject	Grades/Subjects	
PLDs	Reading	Phyllis Echols, Ph.D.
	Listening	Steve Fitzpatrick, Ph.D.
	Speaking	Alvaro Arce, Ph.D.
Reading	2	Melena McBride, Ph.D.
	3	Ha Phan, Ph.D.
	4 & 5	Ou Zhang, Ph.D.
	6 & 7	Alvaro Arce, Ph.D.
	8 & 9	Phyllis Echols, Ph.D.
	10 & 12	Steve Fitzpatrick, Ph.D.
Listening & Speaking	2 & 3	Ha Phan, Ph.D.
	4 & 5	Ou Zhang, Ph.D.
	6 & 8	Phyllis Echols, Ph.D.
	9 &12	Steve Fitzpatrick, Ph.D.

Data Analysts

For the standard-setting meeting, two data analysts performed all of the analyses for all ten committees. The data analysts were Trey Heideman and Andrew Owens. During the meeting, the analysts collected panelist judgment data from the standard-setting website, performed independent analysis to verify analysis results, and prepared panelists' feedback.

Observers

The individuals who attended as observers consisted of TEA staff and Pearson staff. Observers did not participate in the standard-setting process. The purpose of the observers was to allow individuals the opportunity to view the standard-setting process and, in some cases, provide feedback on the process used. Observers, other than vendor staff, were invited to attend the meeting by TEA. The number of observers in a committee meeting were kept to a maximum of ten individuals, so the committee panelists did not feel overwhelmed by the number of

observers. Whenever possible, observers were assigned to a single committee meeting for the duration of the standard-setting meeting.

TEA staff members attended the standard-setting meeting to observe the process, answer assessment and content questions, and address policy questions. TEA staff monitored the cut score recommendations for each proficiency level throughout the standard-setting meetings. TEA staff also monitored the standard-setting meeting, including the content and assessment specialists.

Materials

The following section describes the materials used by the committee members during the standard-setting breakout sessions. Separate materials were developed for each standard-setting meeting.

Pearson Standard Setting Website

The Pearson Standard Setting website was used as the online platform for housing the materials for the standard-setting meeting and collecting panelist judgments throughout the standard-setting process (see example in Figure 2).

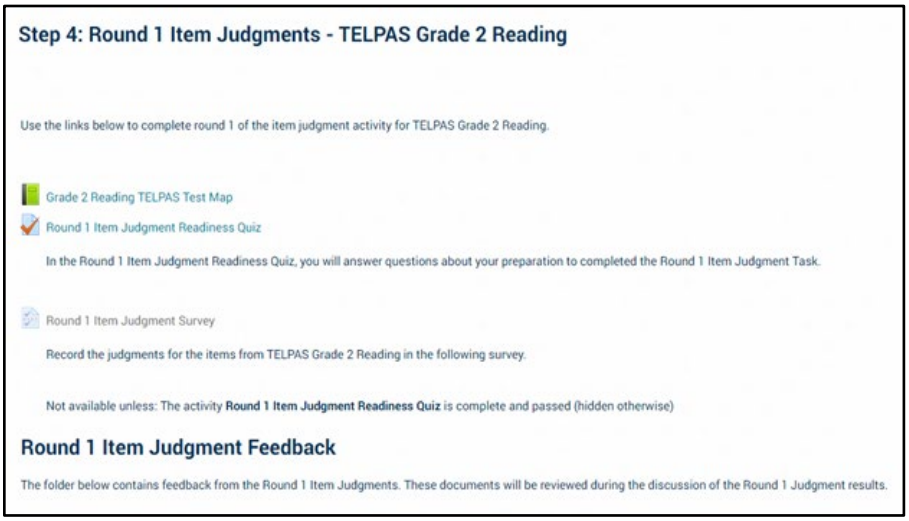


Figure 2. ***Example website interface with links to standard-setting materials***

The website was built using Moodle, an online, open-source collaboration and learning tool. Each panelist was provided a unique user identification and password that provided secure access to the website. Panelist access was restricted to only the sections of the website associated with the standard-setting meeting, as defined by their assigned grade band and content area. Since the TELPAS reading, listening, and speaking assessments are computer-delivered using TestNav 8, the website provided panelists access for viewing the items within the same online platform used by the students during the spring 2018 administration. The website also provided panelists access to online documents that included background

information about the TELPAS assessments. Each standard-setting meeting had a unique site within the Pearson Standard Setting website.

Committee Panelist Folders

In addition to the online resources accessed through the website, panelists were supplied with a meeting folder to organize a variety of hard-copy materials they would need to work with throughout the meeting. These materials included the following:

- Meeting agenda
- Non-disclosure agreement
- “Experience the Test” activity response forms
- Test form answer keys
- Open-ended speaking item rubrics and exemplars
- Item comment form
- Practice judgment forms and test maps
- Judgment item maps
- PLDs and PLD comment form
- ELPS

The folders were prepared in advance. Panelists were required to check-out and check-in their folders at the start and end of each day. Panelists were provided additional materials throughout the meeting, which they were instructed to insert into their folders.

Computers

Each panelist was provided a laptop computer in his or her meeting room to access the online resources through the Moodle site. The laptops were Dell Latitudes with 15.6” screens, standard keyboards with a full-size number pad, and an external mouse. Panelists were not provided with external keyboards or external monitors. Panelists were seated in table groups in pod configuration to provide each panelist with enough space to work with the computer and folder materials. The power supplies were centrally located in the middle of each table. The panelists used Google Chrome to access the Moodle site, which was programmed with a white list of websites to restrict use of the computers to work associated with the standard-setting meeting.

Procedure

A hybrid of the Angoff method (Angoff, 1971) and the Extended Modified (Yes/No) Angoff method (Davis & Moyer, 2015; Plake et. al, 2005) was used during the TELPAS standard-setting meetings to guide committee panelists as they made their cut score recommendations. These were both content- and item-based methods that lead panelists through a standardized process through which they considered student expectations, as defined by PLDs, and the individual items administered to students to recommend cut scores for each proficiency level. The standardized process was used by the committees for each grade and resulted in cut score recommendations.

Standard-Setting Meeting Proceedings

The standard-setting meetings were conducted across five days, June 4–8, 2018, in Austin, Texas. Appendix B includes the complete agenda for the standard-setting meetings. Table 5 presents a high-level agenda for the standard-setting committee meetings.

Table 5. ***Standard-Setting Meeting Agenda Topics***

Meeting Dates				
June 4	June 5	June 6	June 7	June 8
Borderline PLD Meetings	General Session & Reading Standard Setting	Speaking Standard Setting	Listening Standard Setting	Composite Score Review

General Session

The standard-setting meeting began with an orientation of the meeting, during which panelists were presented with an overview of the TELPAS program and the standard-setting process. The purpose of the general session was to welcome the panelists, provide background information about TELPAS, and introduce the standard-setting process. A single general session including all 76 standard-setting panelists was conducted on June 4 at the beginning of the standard-setting meeting.

The overview of the testing program included the following:

- Goals and rationale
- Legislative requirements
- Stakes for the students and teachers
- Uses for state accountability

The facilitator also provided an overview of the standard-setting process. The panelists were introduced to the key concepts and materials that would be used during the Angoff and Extended Modified (Yes/No) Angoff procedures, such as the PLDs, the item map, and the standard-setting website. A clear description of the review process for the cut score recommendation was included as part of the overview process to emphasize that the committees are making cut score recommendations for other groups to review.

Borderline PLD Sessions

After the general session, panelists moved into domain-specific Borderline PLD sessions of the standard-setting meeting. The purpose of the Borderline PLD activity is to have all panelists gain a common understanding of the students at the threshold, or border, of each proficiency level. The PLDs for TELPAS reading, listening, and speaking assessments are the same across grades 2 through 12 and indicate that the general expectations of students within each proficiency level are similar across grades. To reflect this expectation within the standard-setting

process, a single set of borderline PLDs was constructed for each domain and was referenced by each grade and grade band committee. This process reinforced the general assumptions of the constructs, as defined by the PLDs, and built cohesiveness across the committee recommendations.

This activity was held on the morning of the first day of the TELPAS standard-setting meetings. The meeting agenda can be found in Appendix B. The panelists for this activity included the members of the standard-setting meetings. The panelists were divided between three domain-focused committees for reading, listening, and speaking.

Before the panelists were split by domain, they were provided an overview of the standard-setting process that emphasized the role of the borderline PLDs in the process. To provide a single set of training on the development of borderline PLDs, the lead facilitator provided training on the review and development process that was used to create the PLDs for each domain. The panelists were informed that the PLDs provide a snapshot of the typical student's characteristics at each proficiency level, including the breadth and depth of the skills and abilities demonstrated by students within the level. Panelists were told they needed to consider the skills and abilities of a borderline student (i.e., a student that is just barely past the point-of-entry for that level) in order to complete the activity. Because TELPAS has four proficiency levels, the panelists would develop borderline descriptors for three groups of borderline students:

- Intermediate borderline students
- Advanced borderline students
- Advanced high borderline students

Finally, the panelists were introduced to the specific activities that would be used to develop the borderline PLDs, following a three-step jigsaw process.

- Step 1: A representative from each table was assigned a proficiency level. These groups met to discuss the PLDs for their proficiency level and identified general characteristics of a borderline student within that proficiency level.
    - How well did these describe the borderline students at the proficiency level for the domain as we envisioned them?
    - How were these similar or different from the general PLDs for the proficiency level?
- Step 2: Table groups reconvened and discussed what they learned about each borderline student group. Each table worked with the TELPAS PLDs for the respective assessment to create borderline PLDs for each proficiency level.
- Step 3: The facilitator collected the edited borderline PLDs from each group into a single document. The collected borderline PLDs were reviewed with the whole group for consistency in expectations and additional edits or clarifications were made as needed.

Panelists were then divided into breakout committees that focused on the development of borderline PLDs for a single domain. To begin the breakout activity, panelists reviewed the TELPAS PLDs and the ELPS associated with the specific domain. Panelists had already reviewed the PLDs as part of the standard-setting meeting pre-work but spent some time during the activity discussing the PLDs with their table groups and with the whole group. Panelists then were led through the three-step jigsaw process to develop the borderline PLDs.

The result of this final discussion was a single set of borderline PLDs for each domain that was printed and provided for panelists as a reference throughout subsequent activities. The final borderline PLDs were not considered official documents and were not published beyond the meeting. The goal of the borderline PLDs was to help panelists maintain a common understanding of the characteristics possessed by the borderline students.

Breakout Sessions

After the borderline PLD sessions, panelists moved into grade or grade-band and domain-specific breakout sessions for the remainder of the standard-setting meeting. Each committee was responsible for providing recommendations for cut scores for each of the proficiency levels for the grade or grade-band and domain assigned. Table 6 displays an overview of activities conducted during the breakout session for each standard-setting meeting.

Table 6. ***Overview of Activities During Breakout Sessions***

Activity
Introductions and process overview*
Experience the Test activity
Review of domain PLDs
Review of borderline PLDs
Standard-setting training*
Practice judgment task and discussion*
Round 1 recommendations
Discussion of round 1 recommendations and feedback
Round 2 recommendations
Discussion of round 2 recommendations and feedback
Round 3 recommendations
Closing remarks and evaluation*

* These activities were only performed one time and not repeated during each breakout session.

Introductions and Overview

To begin the breakout session for the reading standard setting, the individuals in the room—facilitators, panelists, and observers—introduced themselves, including their names, current teaching experience, and location.

After introductions, the facilitator reviewed the security and non-disclosure expectations for the meeting. The panelists then reviewed the Security and Non-disclosure Agreement on the

standard-setting website and completed and signed the agreement for the TELPAS standard-setting meeting. This acknowledged that they understood the security expectations for the meeting and agreed to follow them as described.

Finally, the facilitator distributed the meeting folders with secure materials. The facilitator reviewed the materials in the folder, the website, and the use of these resources during the standard-setting process. The panelists had the opportunity to ask questions before proceeding.

Experience the Test

Panelists experienced the specific operational test form that the students were administered during the spring 2018 administration. The panelists experienced the test as an online administration administered through the TestNav8 system, just as students did. The online test was accessed through the standard-setting website.

Since the version of the online testing system used during the standard-setting meetings did not store and score panelist responses, panelists recorded their responses on a separate item response form, provided in the panelist folder. During this activity, if the panelists had any concerns with items on the test form, they were asked to record the comments on an Item Comment Form, which was collected at the end of the meeting.

Panelists had less time to complete the Experience the Test activity than students had to complete the test since it is expected that content experts should not need as much time as students. If panelists did not complete the test in the time provided, they still had the opportunity to review items during the item judgment activity.

After the panelists completed the Experience the Test activity, they were provided information about how TELPAS is scored. For all three domains, an online answer key, or test map, was accessed through the standard-setting website. The test map provided information about the item, including a unique item number, reporting category, maximum possible score, the correct response for the item, and any specific scoring rules for the item. For the items on the speaking test, the test map provided a link to the appropriate item rubric and student exemplars so the panelists could hear what was expected, to earn each possible score point.

Borderline Proficiency Level Descriptors

Domain-specific borderline PLDs were developed as part of the first activity of the TELPAS standard setting. Once the participants were in the breakout rooms, they were given the opportunity to make minor changes to the borderline PLDs that were reflective of the grade-level expectations. The changes were made during a whole group discussion and captured by the facilitator. The final working borderline PLDs were printed and provided to the panelists to use during the remaining standard-setting activities.

Standard-Setting Training and Judgment Rounds

The following activities are designed to provide a common and necessary foundation for fully participating in the cut score recommendation process.

Training

The panelists were provided thorough training on how to make their recommendations as part of the standard-setting meetings. For the reading assessments, composed solely of dichotomous

items, the panelists were instructed on using a modified Angoff method (Angoff, 1971). With this method, panelists reviewed each item and answered the following question for each proficiency level:

"*What is the probability that a borderline student at the proficiency level would get this item correct?"*

Significant time was spent describing the thought process the panelists should go through using parts of the question.

- "What is the probability…" — Participants selected an option that would represent a range which contained an expected likelihood.
- "...a borderline student at the proficiency level…" — The panelists referenced the borderline PLDs for the proficiency level to determine how a borderline student would be expected to respond.
- "...would…" — When considering expected student response to an item, the panelists needed to consider how a student would respond rather than how they should respond. Where "should" is an aspirational expectation, "would" is a more realistic expectation of a student response to the item.
- "...get this item correct?" — The panelists reviewed the knowledge, skills, and abilities required to provide a correct response to the item compared to the expected PLDs for the borderline proficiency level student.

Instead of having panelists provide open responses as values between 0 and 100, they selected an option from 0% to 100% in intervals of 10. Table 7 presents the responses with their corresponding ranges.

Table 7. ***Response Probabilities and Ranges***

Option	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Range	0-4%	5-14%	15-24%	25-34%	35-44%	45-54%	55-64%	65-74%	75-84%	85-94%	95-100%

For the speaking domain, which is composed solely of polytomous items, the panelists were instructed on how to use the Extended Modified (Yes/No) Angoff method (Impara & Plake, 1997). The Extended Modified (Yes/No) Angoff method extends the Yes/No method to support judgments with polytomously scored items, where multiple score points are possible, either through rubric-based or partial-credit scoring. Panelists reviewed each item and answered the following question for each proficiency level:

"How many points would a borderline student at the [specific proficiency] level likely earn if he or she answered the question?"

Significant time was spent describing the thought process panelists should go through using parts of the question.

- "How many points…" — Rather than recording "yes" or "no" judgments, panelists recorded the number of points for an item.
- "...would…" — When considering expected student response to an item, the panelists needed to consider how a student would respond rather than how they should respond.

Where “should” is an aspirational expectation, “would” is a more realistic expectation of a student response to the item.

- “...a borderline student at the [specific proficiency] level…” — The panelists referenced the borderline PLDs for the proficiency level to determine how a borderline student would be expected to respond.
- “...likely earn if he or she answered the question?” — In this context, likely is defined as 2 out of 3 times, or 67%.

Because the listening assessments are a mixture of dichotomous and polytomous items, panelists were asked to consider the appropriate question when making item judgments. They were prompted to the question on the paper judgment forms as well as in the surveys on the website.

The training of the standard-setting process included an orientation to the following components and how they would be used during the process.

- Standard-setting website—The website provided panelists access to the items used in the judgment activity, various reference resources, and the item judgment survey where the panelists recorded their individual item judgment recommendations for each proficiency level.
- Online operational test items—A set of items that represented the entire online operational test were presented in the order that the items were administered. The panelists reviewed these items online through the standard-setting website.
- Test map—A summary of the items in the test form. The test map included the following information about the item:
    - Item position from the order of presentation
    - Item scoring key and links to item scoring rubrics, notes, and exemplars
    - Maximum number of possible points
- Item judgment record form—The panelists recorded their item judgments for each round in the standard-setting website and on the paper record form.

Beginning with the intermediate proficiency level, panelists reviewed each item and made a judgment for each borderline proficiency level. Because a student’s response to an item was expected to increase or stay the same as the proficiency level increased, panelists were trained to check their item judgments for expected patterns across proficiency levels. For dichotomously scored items, the selected probability should either increase or stay the same across increasing proficiency levels. For polytomously scored items, the selected number of points should increase or stay the same across proficiency levels. The training included multiple examples of different judgment patterns, which were reviewed with panelists to assist them with understanding the judgment task. The examples included responses that follow and do not follow the expected judgment patterns as well as floor and ceiling patterns in the item judgments.

The panelists also kept a record of their item judgments on the paper Item Judgment Record Form. This document was provided to them as part of the materials in their folder. It included the unique item number, reporting category, and maximum possible points for the item. The panelists were shown how to use the unique item number to ensure that they were referencing the correct item on all documents within the judgment survey and in the online system.

Practice Judgment Round

At the end of the training session for the reading test, panelists made practice judgments prior to beginning the actual judgment rounds. The practice judgment activity only happened as part of the reading standard-setting meeting. The goals of the practice judgment activity were:

- for panelists to become aware of the range of item types and student responses they would encounter during the judgment task,
- to give panelists experience reviewing and making judgments for different item types, and
- to build panelists' confidence so they understood the task they were being asked to complete.

A set of 6–10 practice items was selected for use in this activity. Items were a subset of those the panelists reviewed in the actual judgment rounds and included a range of different item types, item difficulty, and scoring types. For this activity, the panelists worked with a Practice Judgment Item Map and Practice Judgment Record Form.

Following the practice judgments, facilitators showed item-level judgment results interactively through the standard-setting website, including what percentage of panelists selected each point value for each proficiency level. The facilitator walked through the judgment materials for the first two items to ensure panelists knew where to locate key information for making their judgments. The group also had the opportunity to discuss each practice item and to hear perspectives on why panelists selected different point values.

Judgment Rounds and Feedback Discussions

The panelists were led through three independent judgment rounds, with feedback discussions subsequent to each round.

Judgment Rounds

Before making judgments during each round, panelists responded to a survey asking them to state their readiness to participate in the standard-setting activity and to confirm their understanding that judgments should be independent and free from pressure to reach consensus.

- Do you understand your task for the item judgment activity?
- Are you ready to begin the item judgment activity?

As needed, the facilitator answered panelist questions about the upcoming activity.

During each round, panelists individually made judgments for each item starting at the lowest proficiency level (intermediate) based on the borderline descriptions and the skills and abilities required by the item. The panelists then made judgments for the same item for the rest of the proficiency levels, advanced then advanced high, before proceeding to the next item. Judgments were recorded on the website using the Item Judgment Survey for the specific round. Once the panelists recorded judgments for all items, they submitted their judgments for analysis.

After all panelists completed the judgment activity for the round, data analysts from Pearson collected the item judgments from the website, performed the necessary analysis of the data (with appropriate quality control checks), and created feedback reports that were provided to panelists.

Feedback and Discussion

After each judgment round, the panelists were given feedback. The feedback provided was based on their recommendations, the recommendations of others in their group, and relevant data from actual student results on the assessment. Feedback data included the following:

- Information about panelists' cut scores for each proficiency level:
    - Individual cut scores: Item judgments were summed for each proficiency level to obtain a cut score for each level. The panelists were presented with their recommended cut score for each level, along with their item judgment for each level.
    - Table cut score recommendations and statistics: Table-level recommendations were the median cut score across panelists at the table for each proficiency level. The table members were presented with the table-level recommendations and cut score statistics (minimum, maximum, median, mean, and standard deviation) for each level.
    - Committee cut score recommendations and statistics: Committee-level recommendations were the median cut score across all panelists in the committee for each proficiency level. The committee members were presented with the committee-level recommendations and cut score statistics (minimum, maximum, median, mean, and standard deviation) for each level.
    - Panelist agreement data: Bar graphs were displayed showing the frequency of individual recommended cut scores for each proficiency level and across adjacent proficiency levels.
- Item level judgment agreement across panelists: Distributions of individual item judgments for each item and proficiency level were presented.
- Item means (p-values) and score point distributions: The average score earned for each item and the distribution of score points for polytomously scored items were calculated from operational test data.
- Impact data: The proportion of students that would be classified into each proficiency level was displayed based on the current recommended proficiency-level cut scores and reflected the results of students who took the assessment during the spring 2018 administration.

Certain types of information were provided only after specific rounds. Table 8 presents the feedback information that was provided after each judgment round.

Table 8. ***Recommendation Round Feedback for Panelists***

		Round 1	Round 2	Round 3
Item-Level Feedback	Panelist Agreement Data	✓	✓	
	Item Means	✓		
	Score Point Distributions	✓		
Test-Level Feedback	Individual Cut Score Recommendations	✓	✓	
	Table Cut Score Recommendations	✓	✓	
	Committee Cut Score Recommendations	✓	✓	✓
	Panelist Agreement Data	✓	✓	
	Impact Data		✓	✓

Before the discussion, panelists were given guidance regarding the independence of their judgments. That is, they were instructed to listen to other panelists and consider the rationales given for their judgments but not to feel pressured to reach consensus. During the discussion, which was facilitated by table-group leaders, panelists discussed the rationale for their cut score recommendations with other panelists at the same table. After Round 2, the panelists also participated in a whole-group discussion led by the process facilitator.

After the Round 3 judgments for the TELPAS speaking test, the Round 3 results for the reading, listening, and speaking domains were presented to participants for the grade bands for which they provided cut score recommendations. This gave the participants the opportunity to see the cut score recommendations and student impact across the domains and discuss the results as a whole group. During discussion, the facilitator took notes about the participants' perceptions of the recommended cut scores for use during the reasonableness review with TEA.

Process Evaluation

The validity of standard-setting outcomes depends on procedural validity. Evidence of procedural validity was gathered through evaluation surveys administered at the end of the standard-setting meeting for each domain. The evaluations focused on the processes and procedures of the standard-setting meeting, including the panelists' overall views of the standard-setting process, training, materials, meeting facilitation, and ultimately their views on how well they understood the process and how they felt about the final results. The evaluations were anonymous. The results from the evaluations were aggregated and can be found in Appendix I.

Recommended TELPAS Cut Scores from Standard-Setting Committees

Panelists at the standard-setting meeting recommended a range of cut scores for each proficiency level. To determine a committee's single cut score recommendation for a proficiency level, analysts used the median cut score from a set of panelists' recommendations. The Round 3 judgments were considered the committee's recommendation for the standard-setting meeting. The recommended cut scores for each proficiency level based on the Round 3 recommendations for each domain are displayed in Table 9.

Table 9. ***Cut Score Recommendations from Standard-Setting Committees***

Domain	Grade	Maximum Score	Cut Scores		
			Intermediate	Advanced	Advanced High
Reading	2	30	10	19	25
	3	34	9	18	26
	4–5	36	8	19	25
	6–7	37	11	20	28
	8–9	37	9	19	29
	10–12	37	10	22	29
Listening	2–3	27	10	17	22
	4–5	27	12	19	22
	6–8	27	10	17	22
	9–12	27	10	17	23
Speaking	2–3	36	15	24	29
	4–5	36	17	24	29
	6–8	36	13	23	30
	9–12	36	16	22	28

The recommended cut scores for each proficiency level from the three judgment rounds for each standard-setting committee, represented as raw scores, are presented in Appendix G. The summary statistics for the recommended cut scores for each proficiency level from the three judgment rounds for each standard-setting committee are shown in Appendix F. The panelist agreement data for each proficiency level for judgment rounds 1 and 2 for each standard-setting meeting are shown in Appendix H. The estimated impact data after judgment round 3 for each proficiency level for each standard-setting committee are shown in Figures 3, 4, and 5 below.

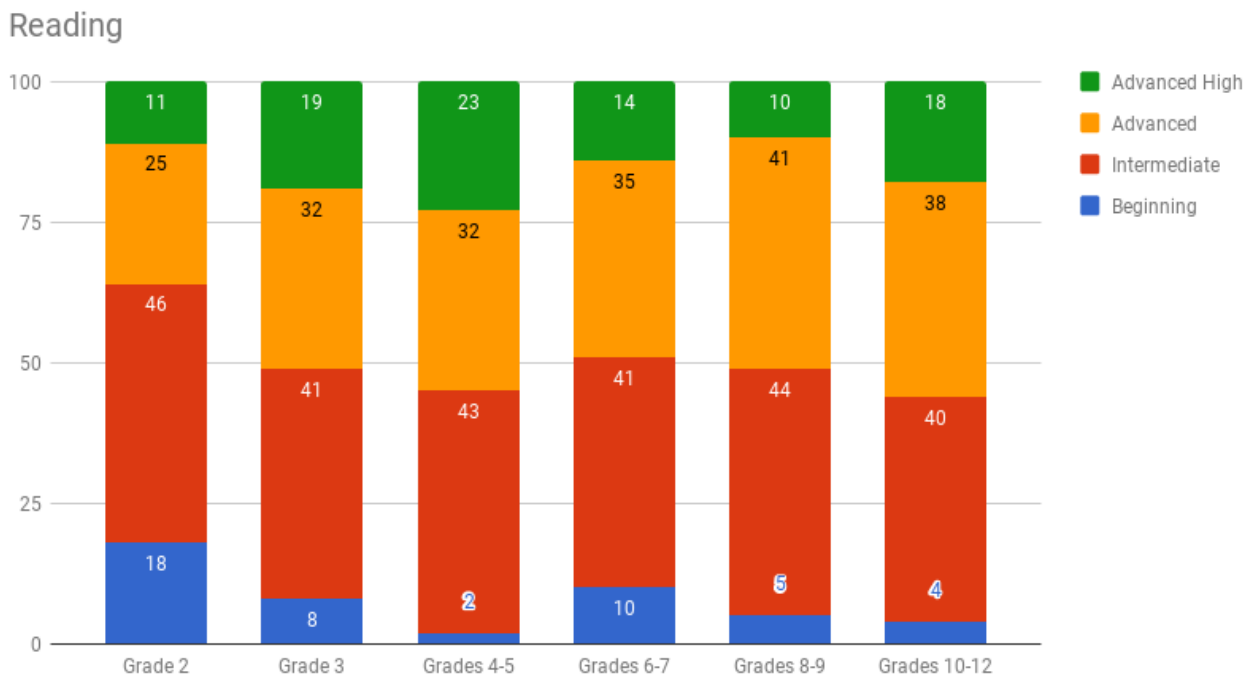


Figure 3. ***Reading impact data from Round 3 recommendations***

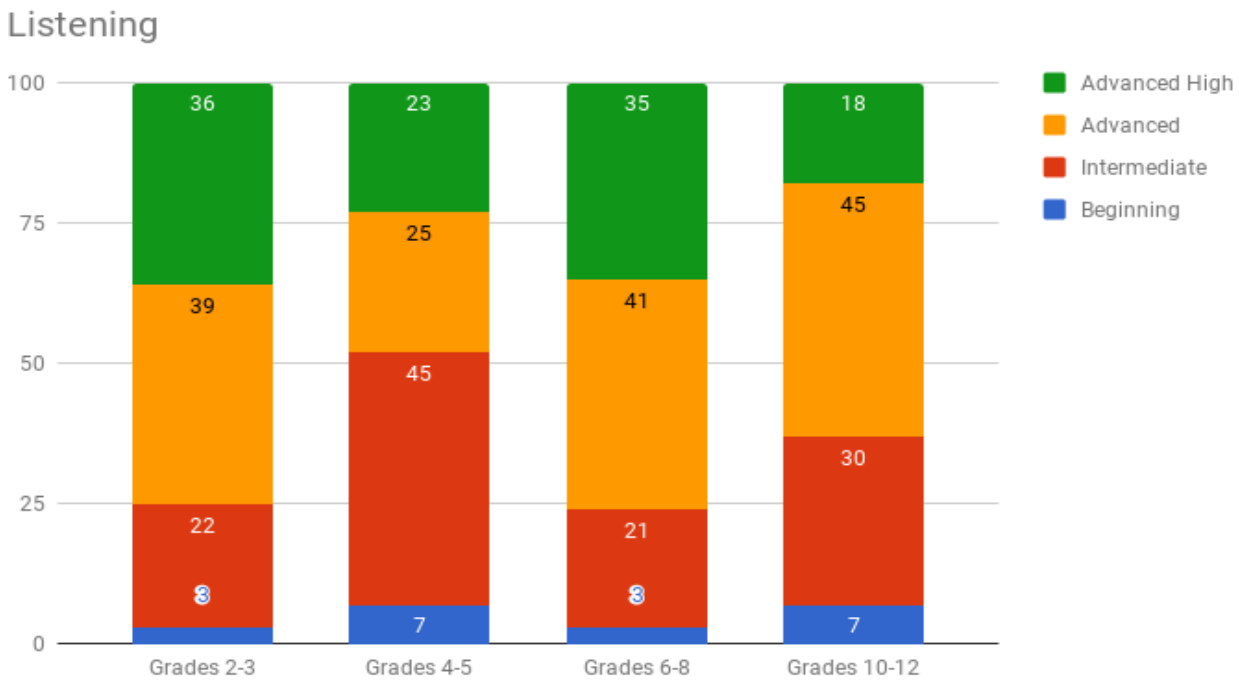


Figure 4. ***Listening impact data from Round 3 recommendations***

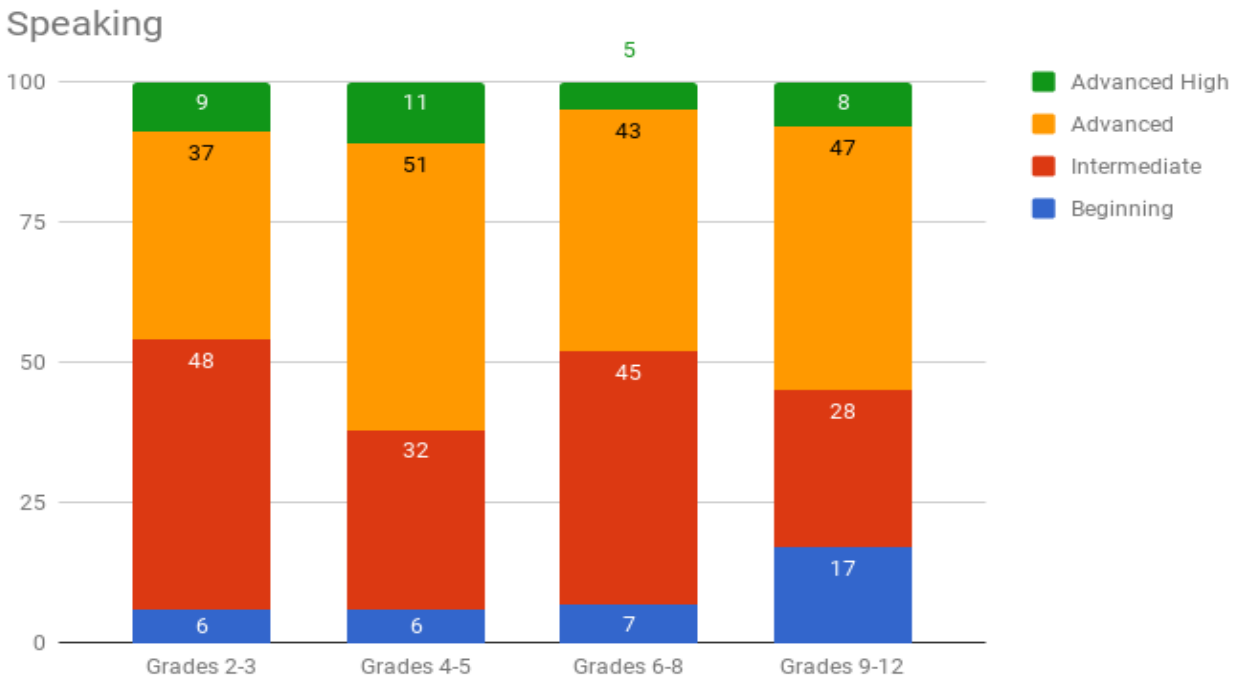


Figure 5. ***Speaking impact data from Round 3 recommendations***

Chapter 4 – Post-Standard Setting

This chapter provides details about the work completed after the standard-setting committee meetings. The sections of this chapter include:

- Composite score meeting
- Reasonableness review
- Reading standards review

Composite Score Meeting

In addition to the individual proficiency level rating of students (beginning, intermediate, advanced, advanced high) for each of the four individual language domains assessed by TELPAS (listening, speaking, reading, writing), students are provided composite scores and composite proficiency ratings. The composite score and composite proficiency rating for students taking the TELPAS assessments provide a single overall level of English language proficiency derived from the proficiency ratings in the four language domains. The individual domain scores on the TELPAS assessment are weighted and averaged to calculate a composite TELPAS score. The domain score for each domain proficiency level is shown in Table 10. If a student does not take the assessment or cannot receive a score for a domain, the domain score received is a 0. The composite score is calculated using an average of the four individual domain scores.

Table 10. *Domain Scores for Domain Proficiency Levels*

	Domain Proficiency Level			
	Beginning	**Intermediate**	**Advanced**	**Advanced High**
Domain Scores	**1**	**2**	**3**	**4**

As part of the standard-setting process, a set of the participants discussed the rules used for assigning composite proficiency ratings for the TELPAS assessments. The process followed is discussed in the next section.

Meeting Process

The composite score process involved the following three steps:

- Review and understand how the composite scores are calculated
- Review the possible composite scores and recommend the minimum composite score for each composite proficiency level
- Review score profiles and define a set of rules

At the beginning of the meeting, the panelists introduced themselves to the rest of the committee. After introductions, the panelists were instructed to the purpose of the composite score meeting and had the opportunity to review the possible composite scores for TELPAS and recommend minimum composite scores for each composite proficiency level.

Prior to the composite score meeting, cut-score neighborhoods were established to delineate appropriate ranges for the cut scores of each composite proficiency level. For example, the cut-score neighborhood for intermediate was from 1 to 2. The rule for establishing the neighborhoods was that the high end of a neighborhood is where the composite score is equal to the composite proficiency rating (e.g., a composite score of a 3 is the highest value the panelists could recommend for the minimum composite score to get a composite proficiency rating of a 3/advanced). The composite score neighborhoods and associated profiles can be found in Appendix J.

As the second step in the composite score process, panelists were asked to examine the scoring profiles associated with each possible composite score for each composite proficiency level. They were then asked to discuss in their table groups which composite score should be the *minimum* composite score a student should earn to be classified into a composite proficiency level based on the neighborhoods. After they discussed a composite proficiency level in their table groups, there was a whole-group discussion where each table shared its minimum cut score for the composite proficiency level and the rationale used by the table group members. This process continued until all composite proficiency levels were discussed.

The third step was to consider the scoring profiles associated with the minimum composite score chosen for each composite proficiency level and determine if there were any profiles panelists would not consider to be classified at the same composite proficiency level.

The final step was to discuss the composite score profiles associated with the recommended minimum scores for each composite proficiency level and create a set of rules. The panelists first discussed the profiles in their table groups. A whole-group discussion then took place and votes were taken if there was not a unanimous consensus.

Meeting Results

The result of the composite score meeting was a set of composite proficiency rating profile descriptors. The descriptors for each composite proficiency level are shown in Table 11.

Table 11. *Composite Proficiency Rating Profile Descriptors*

Composite Proficiency Level	Domain Score Profile Descriptors
Advanced High	• Composite score must be equal to or greater than 3.5. • All domain scores must be either 3 (advanced) or 4 (advanced high).
Advanced	• Composite score must be equal to or greater than 2.5. • All domain scores must be 2 (intermediate), 3 (advanced), or 4 (advanced high) • At least two domains must have domain scores of 3 (advanced) or 4 (advanced high)
Intermediate	• Composite score must be equal to or greater than 1.5. • At least two domains must have domain scores of 2 (intermediate), 3 (advanced), or 4 (advanced high)

To receive a composite proficiency rating, students must have a domain score of 1 or greater on at least two domains.

Reasonableness Review

Recommended cut scores with their corresponding impact data were summarized following the standard-setting meetings and the initial results were presented to members of TEA on June 27, 2018. The presentation included a brief overview of the methodology used to obtain the cut score recommendations, the panelists' recommended cut scores for each assessment and proficiency level, and the impact data associated with the cut scores. Feedback regarding the implementation of the process used during the standard-setting meetings and results was provided.

Additionally, the cut score recommendations from the standard setting for the TELPAS reading assessment were reviewed with regard to various validity studies. The specific validity studies conducted include

- STAAR pass rate comparison
- Average STAAR scale score comparison
- TELPAS and STAAR concordance tables
- TELPAS and STAAR decision consistency and accuracy

The results of these validity studies, along with a review of the cut score in relation to the existing vertical scale, were used to review the reasonableness of the cut score recommendations for the reading assessment.

The goal of this process was to evaluate the reasonableness of the results of the standard-setting meeting, alignment of the results with previous cut scores, expected alignment across grades, and usefulness in communicating results. As part of this review, the alignment of the cut score ranges for each of the proficiency levels across the grades and grade bands was reviewed. The recommended cut scores were translated from the raw score metric to an ability scale, using the respective raw score-to-theta tables for the assessments. The distribution of the recommended cut scores with the ranges was then reviewed on the ability (theta) metric for coherence across grades. Any suggested adjustments to the recommended cut scores were recorded. Changes to the recommended cut scores for each proficiency level were based on the results of the Reasonableness Review process for each course and domain and are displayed in Table 12.

Table 12. ***Changes to the Cut Score Recommendations from the Reasonableness Review***

Domain	Grade	Maximum Score	Cut Score Changes		
			Intermediate	Advanced	Advanced High
Reading	2	30	0	0	0
	3	34	0	0	-2
	4–5	36	+2	0	-1
	6–7	37	0	0	-1
	8–9	37	+1	0	0
	10–12	37	0	0	0
Listening	2–3	27	0	0	0
	4–5	27	0	-2	0
	6–8	27	0	0	0
	9–12	27	0	0	0
Speaking	2–3	36	0	0	0
	4–5	36	0	0	0
	6–8	36	0	0	0
	9–12	36	0	0	0

The final recommended cut scores for each proficiency level, based on the results of the Reasonableness Review process for each course and domain, are presented in Table 13.

Table 13. ***Final Cut Score Recommendations from the Reasonableness Review***

Domain	Grade	Maximum Score	Cut Scores		
			Intermediate	Advanced	Advanced High
Reading	2	30	10	19	25
	3	34	9	18	24
	4–5	36	10	19	24
	6–7	37	11	20	27
	8–9	37	10	19	29
	10–12	37	10	22	29
Listening	2–3	27	10	17	22
	4–5	27	12	17	22
	6–8	27	10	17	22
	9–12	27	10	17	23
Speaking	2–3	36	15	24	29
	4–5	36	17	24	29
	6–8	36	13	23	30
	9–12	36	16	22	28

Reading Standards Review

The TELPAS reading cut scores were not finalized until after the reasonableness review to allow an additional committee to examine the recommended standards from each grade-level committee. On August 9–10, 2018, a standards review committee made recommendations for adjustments to the proficiency level cut scores set by the initial standard-setting committees for TELPAS Reading. The Reading standards review committee included a representative sample of Texas educators who possessed expertise in English language learning and reading. No members of the Reading standards review panel participated in an initial standard-setting committee.

The Reading standards review panelists were separated into three committees, with each group focused on review of the TELPAS Reading tests for specific grades. The grades or grade bands associated with each committee and the number of participants involved are shown in Table 14.

Table 14. ***Reading Standards Review Committees***

Domain	Grade or Grade Band	Participants
Reading	2	5
	3	
	4–5	5
	6–7	
	8–9	6
	10–12	

During the two-day reading standards review, participants evaluated the recommended cut scores from the previous committee, examined the items from the TELPAS reading assessments, and discussed additional data unavailable to the original standard-setting committee. Additional data provided as part of the standards review included test guessing levels, TELPAS reading bridge study results, and TELPAS-STAAR reading validity study results. The process used to facilitate the standards review was an evidence-based bookmark procedure (Lewis, Mitzel, & Green, 1996; Mitzel, Lewis, Patz, & Green, 2001; Schultz & Mitzel 2009), which was different than the method selected for the standard setting. A change in methodology was chosen due to the additional data available to the reading standards review committees and to allow for any subsequent changes in the cut score recommendations that resulted.

The panelists made cut score recommendations for the three proficiency levels (intermediate, advanced, and advanced high) for each reading assessment. The cut score recommendations from the reading standard-setting committees and reading standards review committees are shown for each grade or grade-band in Table 15.

Table 15. ***Cut Score Recommendations from the TELPAS Reading Standard-Setting and Reading Standards Review Committees***

Domain	Grade(s)	Proficiency Level Cut Score Recommendations					
		Intermediate		Advanced		Advanced High	
		Original	Review	Original	Review	Original	Review
Reading	2	10	10	19	18	25	24
	3	9	10	18	17	26	23
	4–5	8	10	19	17	25	23
	6–7	11	11	20	18	28	25
	8–9	9	10	19	19	29	27
	10–12	10	12	22	23	29	30

Note: The cut score recommendations from the committees are based on the raw scores for the TELPAS reading test.

A visual representation of the reading impact data (percentage of students classified into each proficiency level), based on student performance from the spring 2018 administration of the TELPAS assessments and the recommendations from the standards review committee is provided in Figure 6.

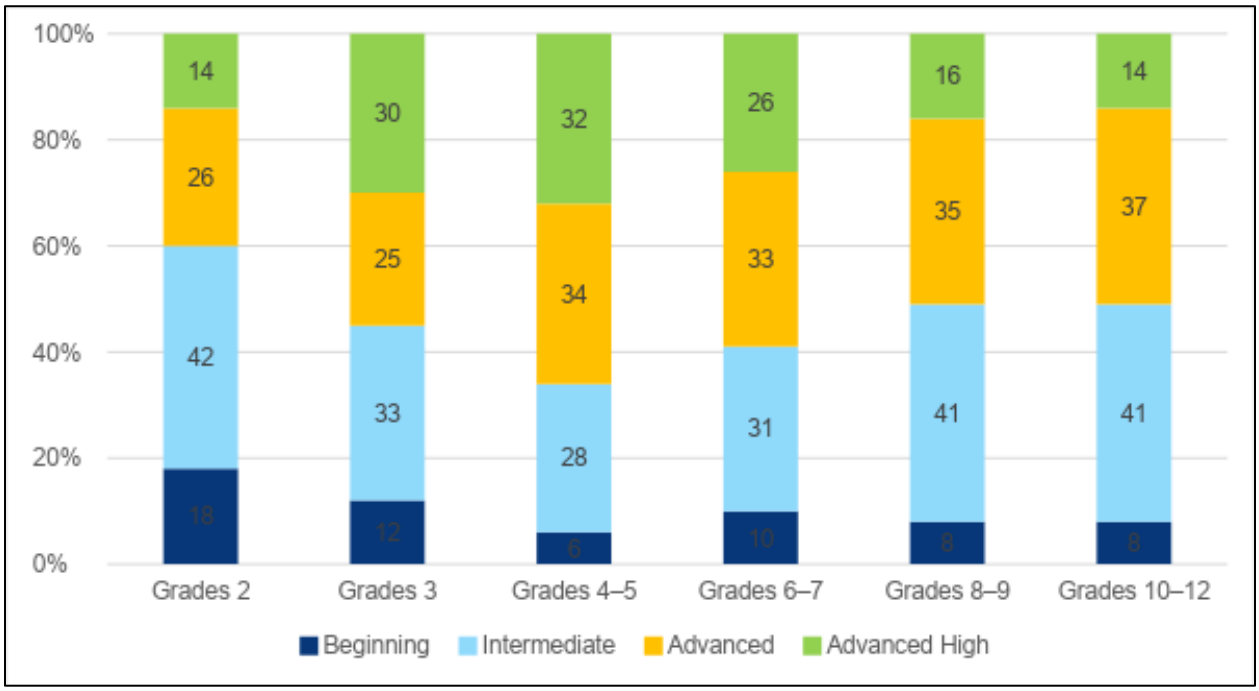


Figure 6. ***Impact data for the TELPAS Reading grades 2–12 final cut scores***

An executive summary was provided to TEA with a brief overview of the methodology and process used to obtain the final cut score recommendations, the final cut score recommendations for each proficiency level on the reporting scale, and the impact data associated with the final recommended cut scores. The Executive Summary was provided to TEA on August 14, 2018 and is available at the beginning of the technical report.

Chapter 5 – Evidence of Procedural Validity of the Standard-Setting Process

This chapter details various evidence for the validity of the processes used during the standard-setting meetings. The sections in this chapter include information regarding the following:

- Committee representation
- Committee training
- Panelists' perceived validity of the meeting
- Process standardization

Committee Representation

As part of the standard-setting evaluation, panelists completed a demographic survey that collected information about their background relevant to educational experience. The results of the self-reported demographic characteristics of the panelists are documented in Appendix D.

As part of the survey, panelists were asked to report their current position (Table D.1), number of years in education (Table D.2), number of years teaching a course related to their standard-setting meeting (Table D.3), and highest level of education (Table D.6). In each of the committees, a majority of panelists had at least a Bachelor's or Master's degree. Additionally, at least 50 percent of the panelists in each committee were K–12 teachers. The panelists in the committees had an extensive range of teaching experience, with at least 50 percent of teachers possessing more than 11 years of experience in education.

A large majority of panelists indicated they had experience teaching in the subject and grade(s) relevant to their committee (presented in Table D.3), which is a relevant factor in relation to the cut score recommendations. In most committees, at least half of the panelists had experience teaching in grades beyond those relevant to their committee. A wide range of teaching experience is an important consideration that ideally increases the cohesiveness of cut score recommendations across grade levels. The teachers in the standard-setting committees also were experienced teaching different student populations, which is displayed in Table D.4. A large majority of panelists in each committee had experience teaching general education, mainstream special education, and English learners.

All panelists were currently working in school districts, as shown in Table D.10. The panelists represented various types of districts across the state, including size, type, and socioeconomic status. The set of panelists for this standard setting was a well-selected sample that represented teachers across the state, which was consistently noticed by the facilitators of the meetings.

Committee Training

During the standard-setting meeting, it was essential that panelists understood how to make judgments as part of the hybrid Modified Angoff method (Angoff, 1971) and Extended Modified

(Yes/No) Angoff standard-setting methodology. The panelists were provided initial training in the standard-setting methodology during the general session and much more extensive preparation in their individual standard-setting committees. The training provided to panelists regarding implementation of the standard-setting process was standardized across committees in the breakout session training presentations.

To provide the panelists an opportunity to implement the standard-setting methodology without consequence, they participated in a practice item judgment round. During the practice round, panelists reviewed a reduced set of items and provided item judgments for three proficiency levels: intermediate, advanced, and advanced high. After the practice judgment round, a whole-group discussion was initiated by the process facilitator to identify and respond to any questions or issues the panelists encountered while implementing the standard-setting process. Additionally, before each round, panelists responded to a readiness survey that asked if they felt prepared to make their item judgments. Panelists were not able to continue to the item judgment survey unless they answered “yes” to both questions on the readiness survey and were encouraged to ask the facilitator questions if they responded “no” to either question.

At various points during the standard-setting meeting, panelists were asked to complete a process evaluation survey to record their impressions of the effectiveness of the materials and methods employed. The results of these process evaluations for each individual committee are presented in Appendix I.

As part of the evaluation survey, panelists were specifically asked about the effectiveness of the training they received on the standard-setting process. One question asked panelists to rate how successful the initial introduction to the standard-setting process during the general session was. More than 50 percent of panelists across all committees responded that it was either *Successful* or *Very Successful*. Another question asked about the overview of the standard-setting process in the breakout session. More than 50 percent of the panelists in each committee responded that the overview was either *Successful* or *Very Successful*. More than 50 percent of panelists in each committee indicated the practice activity for the standard-setting process was either *Successful* or *Very Successful*. In 13 of the 14 committees, greater than 70 percent of panelists indicated the training on the standard-setting process was either *Successful* or *Very Successful*. Taken as a whole, the evaluation survey results seem to indicate that most panelists believed they were prepared to implement the standard-setting procedures and to provide cut score recommendations for each assessment for which they were responsible.

During the composite score meeting, the panelists were provided training on the process and tools used during the meeting. At the end of the meeting, the panelists completed a process evaluation form to record their opinion on the training provided. The results of the composite score process evaluation are presented in Appendix I. For each committee, all panelists indicated the introduction to the composite score process was either *Successful* or *Very Successful.*

Perceived Validity of the Meeting

Panelists communicated their perceived validity of the meeting and the recommended cut scores as part of the standard-setting process evaluation. Evaluations are important evidence for establishing the validity of recommended cut scores for the proficiency levels. Generally, the panelists were satisfied with their recommendations and the standard setting as a whole.

As part of the process evaluation from each committee, the panelists had an opportunity to indicate their confidence that the PLDs were reasonable for each of the proficiency levels. The results for each subject and grade are presented in Appendix I. In the majority of committees, at least 65 percent of the panelists were *Confident* or *Very Confident* the PLDs were reasonable for the proficiency levels and domains. For the intermediate PLDs, the majority of the committees indicated they were *Confident* or *Very Confident* with the PLDs. For the advanced PLDs, more than 75% of the panelists across committees were *Confident* or *Very Confident* with the PLDs. This same pattern was true for the advanced high proficiency level as well. These responses provide evidence that the PLDs, a foundation for the standard-setting process, were perceived by the panelists as reasonable expectations for each proficiency level.

The panelists were also given an opportunity to indicate their confidence in the cut scores recommended by the standard-setting committees. The results of the evaluation survey across subject-level committees regarding their level of confidence in the recommended cut scores are shown in Figure 7. The results for each subject and grade are presented in Appendix I.

For the intermediate proficiency level, a majority of committees had at least 80 percent of panelists indicate they were *Confident* or *Very Confident* in the cut score recommendations. For the advanced proficiency level, over 75 percent of panelists reported they were either *Confident* or *Very Confident* in the cut score recommendations. For the advanced high proficiency level, all committees had at least 85 percent of the panelists respond they were *Confident* or *Very Confident* with the cut score recommendations.

The panelists were also provided the opportunity to provide additional feedback regarding their overall agreement with the process and the cut score recommendations.

Overall, results from the feedback given by standard-setting panelists provides supportive validity evidence for the cut score recommendations for all proficiency levels from each standard-setting committee.

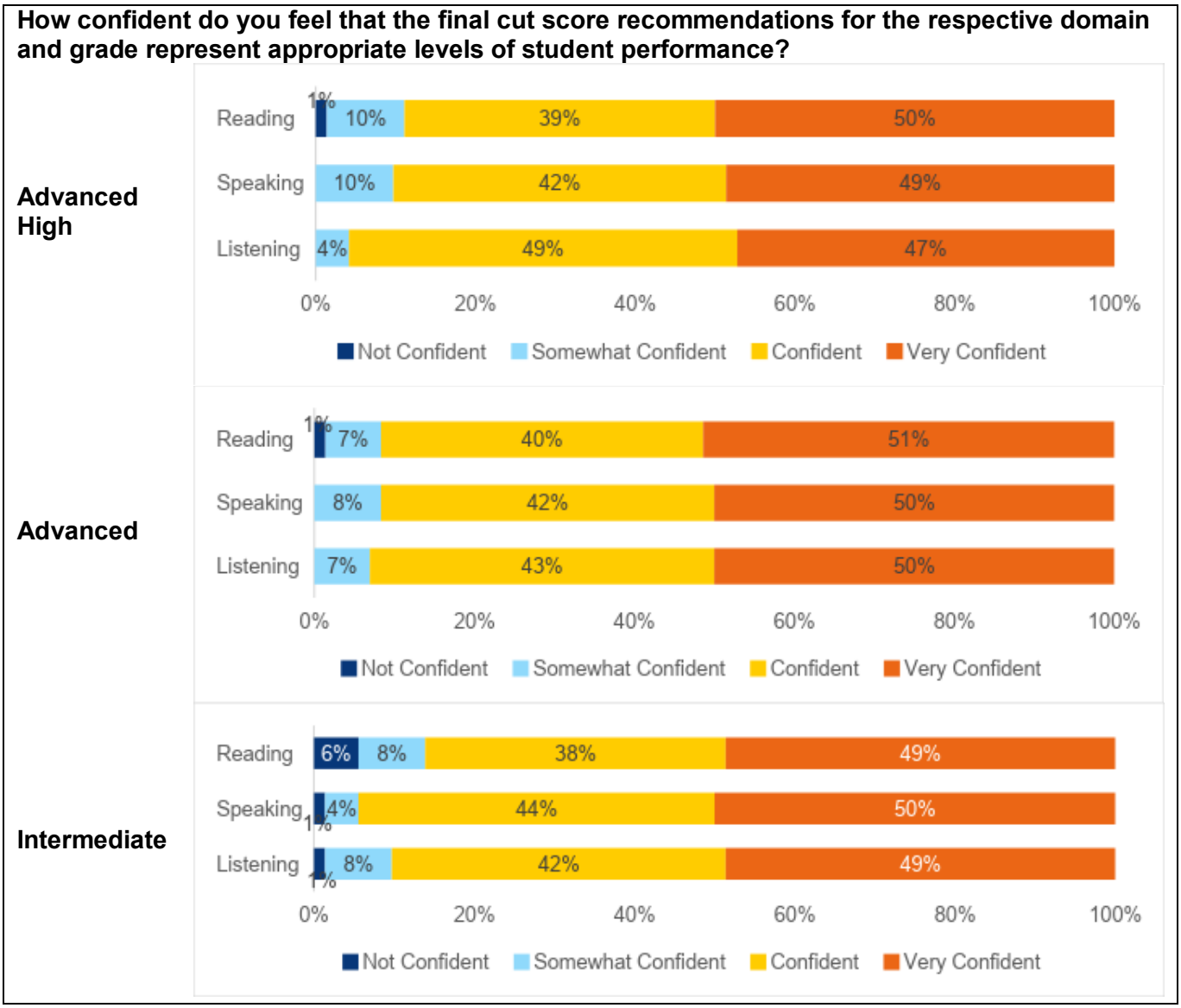


Figure 7. ***Evaluation results on reasonableness of cut scores for each proficiency level***

Process Standardization

An important part of standard-setting meetings is that standardized procedures are implemented by several facilitators working independently across different domain and grade or grade-band panels. During the TELPAS standard-setting meetings, six facilitators worked with 10 panels (6 reading panels and 4 listening and speaking panels) to determine cut scores for four proficiency levels (beginning, intermediate, advanced, and advanced high) for each test.

The organizers of the meeting paid careful attention to the selection and training of facilitators and the preparation of standard-setting meeting materials to ensure standardization of key aspects of the process. Although it is understood that some variation will occur in a dynamic

process that involves independent facilitators working for multiple days with panels of educators, the ultimate goal is to achieve an appropriate balance between standardization and flexibility. An appropriate balance of standard protocol and adaptability allows for individual differences in facilitators and panelists while also ensuring critical steps in the process that might impact panelists' ratings are implemented consistently across panels.

The training of facilitators provided consistent instruction of the process and procedures used throughout the standard setting. Facilitator training was held for 90 minutes each on May 17, 18, 21, and 24, 2018. Additionally, a final preparation meeting was convened one day prior to the standard setting and at the conclusion of each day. The training was focused on consistent use of the materials provided for facilitating the meeting.

Materials were used to facilitate each of the meetings and were prepared in advance to ensure consistency of the presentation and recording of the information. The materials included presentation slides that facilitators presented to panelists as a guide through the training process. Additionally, a script was included to remind facilitators at various points in the presentation of critical steps in the training process. The Pearson Standard Setting website was also an important resource used to distribute materials and collect panelist judgments.

The utilization of standardized materials and procedures ensured that critical steps in the process were implemented consistently across the different meetings. There were no reports of any deviations from the procedures that might have impacted the panelist ratings.

References

19 Tex. Admin. Code § 74.4 (2007).

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 508–600). Washington, DC: American Council in Education.

Davis, L. L. & Moyer, E. L. (2015). *PARCC performance level setting technical report.* Available from Partnership for Assessment of Readiness for College and Careers (PARCC), Washington, D.C.

Impara, J. C., & Plake, B. S. (1997). Standard setting: An alternative approach. *Journal of Educational Measurement*, *34*, 353–366.

Lewis, D. M., Mitzel, H. C., & Green, D. R. (1996, June). Standard setting: A Bookmark approach. In D. R. Green (Chair), *IRT-based standard setting procedures utilizing behavioral anchoring.* Symposium conducted at the Council of Chief State School Officers National Conference on Large-Scale Assessment, Phoenix, AZ.

Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The Bookmark procedure: Psychological perspectives. In G. Cizek (Ed.), *Setting performance standards: Concepts, methods and perspectives* (pp. 249–281). Mahwah, NJ: Erlbaum.

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.

Schulz, E. M., & Mitzel, H. C. (2009). A Mapmark method of standard setting as implemented for the national assessment governing board. In E. V. Smith Jr. & G. E. Stone (Eds.), *Criterion referenced testing: Practice analysis to score reporting using Rasch measurement models* (pp. 194–235). Maple Grove, MN: JAM Press.

United States (1965). *Elementary and secondary education act of 1965: H. R. 2362, 89th Cong., 1st sess., Public law 89-10. Reports, bills, debate and act.* Washington: U.S. Govt. Print. Off.

Appendix A – Proficiency Level Descriptors

ELPS–TELPAS Proficiency Level Descriptors, 19 TAC 74.4(d)(4)
Grades 2–12 Reading

BEGINNING	INTERMEDIATE	ADVANCED	ADVANCED HIGH
(A) Beginning ELLs have little or no ability to read and understand English used in academic and social contexts.	**(B) Intermediate ELLs have the ability to read and understand simple, high-frequency English used in routine academic and social contexts.**	**(C) Advanced ELLs have the ability to read and understand, with second language acquisition support, grade-appropriate English used in academic and social contexts.**	**(D) Advanced high ELLs have the ability to read and understand, with minimal second language acquisition support, grade-appropriate English used in academic and social contexts.**
These students: **(i)** read and understand the very limited recently practiced, memorized, or highly familiar English they have learned; vocabulary predominantly includes **(I)** environmental print **(II)** some very high-frequency words **(III)** concrete words that can be represented by pictures **(ii)** read slowly, word by word **(iii)** have a very limited sense of English language structures **(iv)** comprehend predominantly isolated familiar words and phrases; comprehend some sentences in highly routine contexts or recently practiced, highly familiar text **(v)** are highly dependent on visuals and prior knowledge to derive meaning from text in English **(vi)** are able to apply reading comprehension skills in English only when reading texts written for this level	**These students:** **(i)** read and understand English vocabulary on a somewhat wider range of topics and with increased depth; vocabulary predominantly includes **(I)** everyday oral language **(II)** literal meanings of common words **(III)** routine academic language and terms **(IV)** commonly used abstract language such as terms used to describe basic feelings **(ii)** often read slowly and in short phrases; may re-read to clarify meaning **(iii)** have a growing understanding of basic, routinely used English language structures **(iv)** understand simple sentences in short, connected texts, but are dependent on visual cues, topic familiarity, prior knowledge, pretaught topic-related vocabulary, story predictability, and teacher/peer assistance to sustain comprehension **(v)** struggle to independently read and understand grade-level texts **(vi)** are able to apply basic and some higher-order comprehension skills when reading texts that are linguistically accommodated and/or simplified for this level	**These students:** **(i)** read and understand, with second language acquisition support, a variety of grade-appropriate English vocabulary used in social and academic contexts **(I)** with second language acquisition support, read and understand grade-appropriate concrete and abstract vocabulary, but have difficulty with less commonly encountered words **(II)** demonstrate an emerging ability to understand words and phrases beyond their literal meaning **(III)** understand multiple meanings of commonly used words **(ii)** read longer phrases and simple sentences from familiar text with appropriate rate and speed **(iii)** are developing skill in using their growing familiarity with English language structures to construct meaning of grade-appropriate text **(iv)** are able to apply basic and higher-order comprehension skills when reading grade-appropriate text, but are still occasionally dependent on visuals, teacher/peer assistance, and other linguistically accommodated text features to determine or clarify meaning, particularly with unfamiliar topics	**These students:** **(i)** read and understand vocabulary at a level nearly comparable to that of their native English-speaking peers, with some exceptions when low-frequency or specialized vocabulary is used **(ii)** generally read grade-appropriate, familiar text with appropriate rate, speed, intonation, and expression **(iii)** are able to, at a level nearly comparable to native English-speaking peers, use their familiarity with English language structures to construct meaning of grade-appropriate text **(iv)** are able to apply, with minimal second language acquisition support and at a level nearly comparable to native English-speaking peers, basic and higher-order comprehension skills when reading grade-appropriate text

ELPS–TELPAS Proficiency Level Descriptors, 19 TAC 74.4(d)(1)
Grades K–12 Listening

BEGINNING	INTERMEDIATE	ADVANCED	ADVANCED HIGH
(A) Beginning ELLs have little or no ability to understand spoken English used in academic and social settings.	**(B) Intermediate ELLs have the ability to understand simple, high-frequency spoken English used in routine academic and social settings.**	**(C) Advanced ELLs have the ability to understand, with second language acquisition support, grade-appropriate spoken English used in academic and social settings.**	**(D) Advanced high ELLs have the ability to understand, with minimal second language acquisition support, grade-appropriate spoken English used in academic and social settings.**
These students: (i) struggle to understand simple conversations and simple discussions even when the topics are familiar and the speaker uses linguistic supports such as visuals, slower speech and other verbal cues, and gestures (ii) struggle to identify and distinguish individual words and phrases during social and instructional interactions that have not been intentionally modified for ELLs (iii) may not seek clarification in English when failing to comprehend the English they hear; frequently remain silent, watching others for cues	**These students:** (i) usually understand simple or routine directions, as well as short, simple conversations and short, simple discussions on familiar topics; when topics are unfamiliar, require extensive linguistic supports and adaptations such as visuals, slower speech and other verbal cues, simplified language, gestures, and preteaching to preview or build topic-related vocabulary (ii) often identify and distinguish key words and phrases necessary to understand the general meaning during social and basic instructional interactions that have not been intentionally modified for ELLs (iii) have the ability to seek clarification in English when failing to comprehend the English they hear by requiring/requesting the speaker to repeat, slow down, or rephrase speech	**These students:** (i) usually understand longer, more elaborated directions, conversations, and discussions on familiar and some unfamiliar topics, but sometimes need processing time and sometimes depend on visuals, verbal cues, and gestures to support understanding (ii) understand most main points, most important details, and some implicit information during social and basic instructional interactions that have not been intentionally modified for ELLs (iii) occasionally require/request the speaker to repeat, slow down, or rephrase to clarify the meaning of the English they hear	**These students:** (i) understand longer, elaborated directions, conversations, and discussions on familiar and unfamiliar topics with occasional need for processing time and with little dependence on visuals, verbal cues, and gestures; some exceptions when complex academic or highly specialized language is used (ii) understand main points, important details, and implicit information at a level nearly comparable to native English-speaking peers during social and instructional interactions (iii) rarely require/request the speaker to repeat, slow down, or rephrase to clarify the meaning of the English they hear

ELPS–TELPAS Proficiency Level Descriptors, 19 TAC 74.4(d)(2)
Grades 2–12 Speaking

BEGINNING	INTERMEDIATE	ADVANCED	ADVANCED HIGH
(A) Beginning ELLs have little or no ability to speak English in academic and social settings.	**(B) Intermediate ELLs have the ability to speak in a simple manner using English commonly heard in routine academic and social settings.**	**(C) Advanced ELLs have the ability to speak using grade-appropriate English, with second language acquisition support, in academic and social settings.**	**(D) Advanced high ELLs have the ability to speak using grade-appropriate English, with minimal second language acquisition support, in academic and social settings.**
These students: **(i)** mainly speak using single words and short phrases consisting of recently practiced, memorized, or highly familiar material to get immediate needs met; may be hesitant to speak and often give up in their attempts to communicate **(ii)** speak using a very limited bank of high-frequency, high-need, concrete vocabulary, including key words and expressions needed for basic communication in academic and social contexts **(iii)** lack the knowledge of English grammar necessary to connect ideas and speak in sentences; can sometimes produce sentences using recently practiced, memorized, or highly familiar material **(iv)** exhibit second language acquisition errors that may hinder overall communication, particularly when trying to convey information beyond memorized, practiced, or highly familiar material **(v)** typically use pronunciation that significantly inhibits communication	**These students:** **(i)** are able to express simple, original messages, speak using sentences, and participate in short conversations and classroom interactions; may hesitate frequently and for long periods to think about how to communicate desired meaning **(ii)** speak simply using basic vocabulary needed in everyday social interactions and routine academic contexts; rarely have vocabulary to speak in detail **(iii)** exhibit an emerging awareness of English grammar and speak using mostly simple sentence structures and simple tenses; are most comfortable speaking in present tense **(iv)** exhibit second language acquisition errors that may hinder overall communication when trying to use complex or less familiar English **(v)** use pronunciation that can usually be understood by people accustomed to interacting with ELLs	These students: (i) are able to participate comfortably in most conversations and academic discussions on familiar topics, with some pauses to restate, repeat, or search for words and phrases to clarify meaning (ii) discuss familiar academic topics using content-based terms and common abstract vocabulary; can usually speak in some detail on familiar topics (iii) have a grasp of basic grammar features, including a basic ability to narrate and describe in present, past, and future tenses; have an emerging ability to use complex sentences and complex grammar features (iv) make errors that interfere somewhat with communication when using complex grammar structures, long sentences, and less familiar words and expressions (v) may mispronounce words, but use pronunciation that can usually be understood by people not accustomed to interacting with ELLs	These students: (i) are able to participate in extended discussions on a variety of social and grade-appropriate academic topics with only occasional disruptions, hesitations, or pauses (ii) communicate effectively using abstract and content-based vocabulary during classroom instructional tasks, with some exceptions when low-frequency or academically demanding vocabulary is needed; use many of the same idioms and colloquialisms as their native English-speaking peers (iii) can use English grammar structures and complex sentences to narrate and describe at a level nearly comparable to native English-speaking peers (iv) make few second language acquisition errors that interfere with overall communication (v) may mispronounce **words, but rar**ely use pronunciation that interfer**es with overall** communicat**ion**

Appendix B – Meeting Agenda

TELPAS Standard-Setting Meeting

Agenda

Day 1

7:30–8:00 a.m.	*Breakfast*
	General Session
8:00–8:30 a.m.	Welcome and orientation
8:30–9:30 a.m.	Standard-Setting Overview
	Borderline PLD Development (in domain groups)
9:30–9:45 a.m.	Welcome and Overview
9:45–10:15 a.m.	Review and Discuss Proficiency Level Descriptors
10:15–10:30 a.m.	Borderline PLD Training
10:30–11:00 a.m.	Table Discussion—Borderline proficiency level discussion (in domain groups)
11:00–11:45 a.m.	Table Discussion—Borderline PLD creation (in domain groups)
11:45 a.m.–12:30 p.m.	*Lunch*
12:30–1:30 p.m.	Whole-Group Discussion (in domain groups)
	Breakout Session—Reading Standard Setting
1:30–2:00 p.m.	Welcome, Introductions, and Orientation
2:00–2:45 p.m.	Experience the Test—TELPAS Reading
2:45–3:00 p.m.	**Break**
3:00–3:15 p.m.	Scoring the TELPAS Reading Assessment
3:15–3:45 p.m.	Review and Discuss Proficiency Level Descriptors—TELPAS Reading

3:45–4:30 p.m. Borderline Proficiency Level Descriptors Review—Reading

Day 2

7:30–8:00 a.m.	*Breakfast*
	Breakout Session—Reading Standard Setting
8:00–8:30 a.m.	Standard-Setting Training
8:30–9:15 a.m.	Practice Judgment Activity and Discussion
9:15–10:15 a.m.	Round 1 Judgments Round 1 Readiness Form Panelists work independently to make Round 1 judgments
10:15–10:45 a.m.	**Break**
10:45–11:15 a.m.	Round 1 Judgment Feedback Item Level - Item level panelist agreement Test Level - Cut score recommendations; panelist agreement
11:15 a.m.– 12:00 p.m.	Table Discussion—Round 1 Feedback Panelists discuss feedback data at their tables
12:00–12:45 p.m.	*Lunch*
12:45–1:30 p.m.	Round 2 Judgments Round 2 Readiness form Panelists work independently to make Round 2 judgments
1:30–2:00 p.m.	**Break**
2:00–2:15 p.m.	Round 2 Judgment Feedback Item Level - Item level panelist agreement Test Level - Cut score recommendations, panelist agreement, impact data
2:15–2:45 p.m.	Table Discussion—Round 2 Feedback
2:45–3:15 p.m.	Whole-Group Discussion—Round 2 Feedback
3:15–4:00 p.m.	Round 3 Judgments Round 3 Readiness form Panelists work independently to make Round 3 judgments
4:00–4:30 p.m.	Wrap-up for Individuals Leaving after Reading.

Day 3

7:30–8:30 a.m.	*Breakfast*
	Breakout Session—Listening Standard Setting
8:30–9:00 a.m.	Welcome and Introductions
9:00–9:45 a.m.	Experience the Test—TELPAS Listening and Speaking
9:45–10:00 a.m.	**Break**
10:00–10:30 a.m.	Scoring the TELPAS Listening and Speaking Assessment
10:30–11:00 a.m.	Review and Discuss Proficiency Level Descriptors—TELPAS Listening
11:00–11:30 a.m.	Borderline Proficiency Level Descriptors Review—Listening
11:30 a.m.– 12:30 p.m.	*Lunch*
12:30–1:00 p.m.	Standard-Setting Judgment Review
1:00–2:00 p.m.	Round 1 Judgments Round 1 Readiness Form Panelists work independently to make Round 1 judgments
2:00– 2:30 p.m.	**Break**
2:30–3:00 p.m.	Round 1 Judgment Feedback Item Level - Item level panelist agreement Test Level - Cut score recommendations, panelist agreement
3:00–3:30 p.m.	Table Discussion - Round 1 Feedback Panelists discuss feedback data at their tables
3:30–4:30 p.m.	Round 2 Judgments Round 2 Readiness form Panelists work independently to make Round 2 judgments

Day 4

7:30–8:30 a.m.	*Breakfast*
	Breakout Session—Listening Standard Setting
8:30–9:00 a.m.	Round 2 Judgment Feedback Item Level - Item level panelist agreement Test Level - Cut score recommendations, panelist agreement, Impact data
9:00–9:30 a.m.	Table Discussion—Round 2 Feedback
9:30–10:00 a.m.	Whole-Group Discussion—Round 2 Feedback
10:00–10:15 a.m.	**Break**
10:15–10:30 a.m.	Standard-Setting Judgment Review
10:30–11:30 a.m.	Round 3 Judgments Round 3 Readiness form Panelists work independently to make Round 3 judgments
11:30 a.m.– 12:30 p.m.	*Lunch*
	Breakout Session—Speaking Standard Setting
12:30–1:00 p.m.	Review and Discuss Proficiency Level Descriptors—TELPAS Speaking
1:00–1:30 p.m.	Borderline Proficiency Level Descriptors Review—Speaking
1:30–1:45 p.m.	Standard-Setting Judgment Review
1:45–2:45 p.m.	Round 1 Judgments Round 1 Readiness form Panelists work independently to make Round 1 judgments
2:45–3:15 p.m.	**Break (Judgment Analysis)**
3:15–3:45 p.m.	Round 1 Judgment Feedback Item Level - Item means and distributions Test Level - Threshold score recommendations, panelist agreement
3:45–4:30 p.m.	Table Discussion—Round 1 Feedback Panelists discuss feedback data at their tables

Day 5

7:30–8:30 a.m.	*Breakfast*
	Breakout Session—Speaking Standard Setting
8:30–9:30 a.m.	Round 2 Judgments Round 2 Readiness form Panelists work independently to make Round 2 judgments
9:30–10:00 a.m.	**Break**
10:00–10:30 a.m.	Round 2 Judgment Feedback Item Level - Item level panelist agreement Test Level - Cut score recommendations, panelist agreement, impact data
10:30–11:00 a.m.	Table Discussion—Round 2 Feedback
11:00–11:30 a.m.	Whole-Group Discussion—Round 2 Feedback
11:30 a.m.– 12:30 p.m.	*Lunch*
12:30–1:30 p.m.	Round 3 Judgments Round 3 Readiness form Panelists work independently to make Round 3 judgments
1:30–2:00 p.m.	**Break**
2:00–2:30 p.m.	Round 3 Judgment Feedback—Across Domains Committee level cut score recommendations, panelist agreement, impact data
2:30–3:00 p.m.	Whole-Group Discussion
3:00–3:30 p.m.	Close-out and Evaluations

Appendix C – Participants' Meeting Materials

The materials developed for the grade 4 and 5 reading standard-setting committee are provided as an example of what was developed and provided to the participants. Since the materials provided to participants contained secure information, any place where secure information would be provided was removed. The following materials will not be provided within the appendix:

- Test form — This was presented to participants through TestNav8, the online testing platform used during the spring 2018 administration.
- Open-ended item rubrics — These documents presented the scoring rubrics, notes, and student-produced response examples for each open-ended item presented to participants.
- Practice item judgment set — This was presented to participants through TestNav 8, the online testing platform used during the spring 2018 administration.

TELPAS Standard-Setting Meeting
June 2018

Agenda

Day 1

General Session

Welcome and Orientation

Standard-Setting Overview

Borderline PLD Development Groups (Reading, Speaking, or Listening)

Proficiency Level Descriptors (PLD)

Borderline PLD Development Training

Borderline PLD Development

Reading Standard Setting (Grade 2, Grade 3, Grades 4–5, Grades 6–7, Grades 8–9, or Grades 10–12)

Introductions and Orientation

Experience the Test—Reading

Scoring the Reading Assessment

Review Reading PLDs and Borderline PLDs

Day 2

Reading Standard Setting (Grade 2, Grade 3, Grades 4–5, Grades 6–7, or Grades 8–9, Grades 10–12)

Standard-Setting Training

Practice Judgment Activity—Reading

Round 1 Judgments—Reading

Round 1 Judgment Feedback and Discussion

Round 2 Judgments—Reading

Round 2 Judgment Feedback and Discussion

Round 3 Judgments—Reading

Day 3

Listening & Speaking Standard Setting (Grades 2–3, Grades 4–5, Grades 6–8, Grades 9–12)

Experience the Test—Listening & Speaking

Scoring the Listening & Speaking Assessment

Review Speaking PLDs and Borderline PLDs

Standard-Setting Judgment Review

Round 1 Judgments—Speaking

Round 1 Judgment Feedback and Discussion

Round 2 Judgments—Speaking

Day 4

Listening & Speaking Standard Setting (Grades 2–3, Grades 4–5, Grades 6–8, Grades 9–12)

Round 2 Judgment Feedback and Discussion

Round 3 Judgments

Review Listening PLDs and Borderline PLDs

Standard-Setting Judgment Review

Round 1 Judgments—Listening

Round 1 Judgment Feedback and Discussion

Round 2 Judgments—Listening

Day 5

Listening & Speaking Standard Setting (Grades 2–3, Grades 4–5, Grades 6–8, Grades 9–12)

Round 2 Judgment Feedback and Discussion

Round 3 Judgments—Listening

Round 3 Judgment Feedback and Discussion (Reading, Speaking, and Listening)

Close-out and Evaluations

TELPAS Composite Score (Table Leaders)

TELPAS Reading

Standards Review Meeting
August 2018

Agenda

Day 1

General Session

Welcome and Orientation

Standard Review Overview

Proficiency Level Descriptors (PLDs)

Borderline PLDs

Lunch

Reading (Grades 2 and 3, Grades 4–5 and 6–7, Grades 8–9 and 10–12)

Introductions and Orientation

Experience the Assessment—1st test

Scoring the Reading Assessment

Review of the Standards Review Process

Individual Judgment Activity—1st test

Review of Judgment Results—1st test

Day 2

Reading (Grades 2 and 3, Grades 4–5 and 6–7, Grades 8–9 and 10–12)

Experience the Assessment—2nd test

Scoring the Reading Assessment

Review of the Standards Review Process

Individual Judgment Activity—2nd test

Review of Judgment Results—2nd test

Close-out and Evaluation

State of Texas

County of _____

Texas Education Agency

Texas Student Assessment Program

PERSONAL OATH OF SECURITY AND CONFIDENTIALITY

I, _____, do solemnly swear, or affirm, that I will
(Print Full Name)
faithfully execute the duty imposed upon me by Sections 39.030 and 39.0303 of the Texas Education Code (TEC) to insure the security of the assessment instruments and achievement tests, and by my oath or affirmation do agree to safeguard the confidentiality of all assessment instruments, assessment instrument items, or achievement tests.

This oath or affirmation is intended by me to extend to any meeting or portion of meetings held pursuant to TEC Section 39.030 or other applicable law, in which assessment instruments or assessment instrument items are discussed. I acknowledge that failure to abide by this, my oath or affirmation, will make me subject to the maximum criminal and professional penalties that can be imposed by law. Penalties involved include:

- a permanent reprimand affixed to the face of all Texas Teacher Certificates and other educator credentials,
- a one-year suspension of all Texas Teacher Certificates and other education credentials,
- a permanent cancellation of all Texas Teacher Certificates and other education credentials, and
- a Class C misdemeanor.

As a testament to this oath, I affix my signature below:

Executed this _____ day of _____, 20____.

_____
(School Name/Organization Affiliation)

_____
(Signature)

_____
(Work Address)

_____
(Home Address)

_____
(City and Zip Code)

_____
(City and Zip Code)

_____
(Telephone Number)

_____
(Telephone Number)

Texas English Language Proficiency Assessment (TELPAS)
Standard Setting Meeting

Participant Information Survey

Grades 4-5 Reading

Professional Experience

What is your current position?

- Teacher (K-12 Education)
- Teacher (Higher Education)
- Administrator (School)
- Administrator (District)
- Other Position:

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

- None
- 1 to 5 years
- 6 to 10 years
- 11 to 15 years
- 16 to 20 years
- More than 20 years

How many years of professional experience do you have teaching reading grades 4 and 5?

- None
- 1 to 5 years
- 6 to 10 years
- 11 to 15 years
- 16 to 20 years
- More than 20 years

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

(Check all that apply.)

- [ ] Students receiving mainstream special education services
- [ ] Students receiving self-contained special education services
- [ ] Students who are English language learners
- [ ] Students who are receiving general education instruction
- [ ] Students who are receiving vocational technical instruction

In addition to English, for which of the following languages do you have educational experience?

(Check all that apply.)

- [ ] Spanish
- [ ] Vietnamese
- [ ] Chinese (Cantonese, Mandarin or other Chinese language)
- [ ] Tagalog
- [ ] German
- [ ] French
- [ ] Hindi
- [ ] Urdu
- [ ] Korean
- [ ] Arabic
- [ ] Other

What is the highest degree you have completed?

- ○ High School Diploma
- ○ Associates degree (A.A., A.S.)
- ○ Bachelors degree (B.A., B.S.)
- ○ Masters degree (M.A., M.S.)
- ○ Doctoral degree (Ph.D., Ed.D.)

Demographic Information

What is your gender?

- ○ Male ○ Female ◉ No answer

What is your ethnicity?

- ○ Hispanic or Latino ○ Not Hispanic or Latino ◉ No answer

What is your race?

- ○ American Indian or Alaskan Native
- ○ Asian
- ○ Black or African American
- ○ Native Hawaiian or Pacific Islander
- ○ White
- ◉ No answer

Do you currently work in a school district?

- ○ Yes
- ○ No

School District Information

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

- ○ Small
- ○ Medium
- ○ Large

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

- ○ Rural
- ○ Metropolitan/Urban
- ○ Suburban

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

- ○ Low
- ○ Moderate
- ○ High

TELPAS Writing Training

Have you attended the TELPAS Writing Scoring Training in the last 2 years?

- ○ Yes
- ○ No

TELPAS Standard Setting Meeting
June 2018

Experience the Test Response Form

Grades 4-5 Reading

Sequence	Item	Passage	Reporting Category	Maximum Points	Response/Notes
1	04R201C9DBC15010		1	1	
2	04R201F9EMC16012		1	1	
3	04R201C9DBC16007		1	1	
4	04R301F9EMC16017		1	1	
5	04R401F9ETC15026		1	1	
6	04R301F9EMC16018		1	1	
7	04R301F9ETC16015		1	1	
8	04R403J2EST16034		3	1	
9	04R403J3EST16038		3	1	
10	04R203J3ESZ16083		3	1	
11	04R302I3ESZ16081		2	1	
12	04R203K9ESZ16084		3	1	
13	04R303J2ESZ16082		3	1	
14	04R202I1ESZ16079		2	1	
15	04R401F9ESZ16240		1	1	
16	04R402I1ESZ16244		2	1	
17	04R402I1ESZ16241		2	1	

Note: Only the first page of this document is presented as an example.

TELPAS Standard Setting Meeting
June 2018

TELPAS
Texas English Language Proficiency Assessment System

Item Comment Form

Grades 4-5 Reading

Name _____

Directions: If you have any comments or suggestions about specific items, please record them here.

Item ID	Comment

TELPAS Standard Setting Meeting
June 2018

Judgment Round Record Sheet

Grades 4-5 Reading

"What is the probability that a borderline student of the proficiency level would answer the question correctly?"

Sequence	Item	Passage	Reporting Category	Maximum Points	Judgment Round								
					1			2			3		
					I	A	AH	I	A	AH	I	A	AH
1	04R201C9DBC15010		1	1									
2	04R201F9EMC16012		1	1									
3	04R201C9DBC16007		1	1									
4	04R301F9EMC16017		1	1									
5	04R401F9ETC15026		1	1									
6	04R301F9EMC16018		1	1									
7	04R301F9ETC16015		1	1									
8	04R403J2EST16034		3	1									
9	04R403J3EST16038		3	1									
10	04R203J3ESZ16083		3	1									
11	04R302I3ESZ16081		2	1									
12	04R203K9ESZ16084		3	1									
13	04R303J2ESZ16082		3	1									

Note: I=Intermediate; A=Advanced; AH=Advanced High

Round 1 Judgment Survey

Print Blank

Texas English Language Proficiency Assessment System (TELPAS)
Standard Setting Meeting
Grades 4-5 Reading

Round 1 Item Judgment Survey

You are now ready to begin!

For each item in the Judgment item set, do the following for each proficiency level:

- Review the item in the online system.
- Review the information provided about the item in the test map and answer key.
- Review the borderline proficiency level descriptors (PLDs) for the proficiency level.
- Answer the following question:

"What is the probability that a borderline student of the proficiency level would answer the question correctly?"

To answer the question, you will select the option for the probability range that would best answer the question.

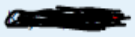
Option	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Range	0-4%	5-14%	15-24%	25-34%	35-44%	45-54%	55-64%	65-74%	75-84%	85-94%	95-100%

- Record your response to the question for the proficiency level for the specific item on the judgment record sheet and in the online survey.

Continue reviewing the items until you have provided judgments for each proficiency level for all of the items.

You will now start the Item Judgment Process for the items in the item judgment set.

Item: 04R201C9DBC15010



	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Intermediate	◉	○	○	○	○	○	○	○	○	○	○
Advanced	◉	○	○	○	○	○	○	○	○	○	○
Advanced High	◉	○	○	○	○	○	○	○	○	○	○

Process Evaluation #2

Print Blank

Texas English Language Proficiency Assessment System (TELPAS)
Standard Setting Meeting

Process Evaluation Survey #2

Grades 4–5 Reading

The purpose of this evaluation is to collect information about your experience with the activities of the standard setting meeting to this point. Your opinions are an important part of our evaluation of this meeting.

* Select the option that best reflects your opinion about the level of success of the various components of the Reading meeting in which you participated. The activities were designed to help you both understand the process and be supportive of the recommendations made by the committee.

	Not Successful	Partially Successful	Successful	Very Successful
Overview of the TELPAS assessments	○	○	○	○
Experiencing the actual assessment	○	○	○	○
Discussion of the scoring of items on the assessment	○	○	○	○
Discussion of proficiency level descriptors (PLDs)	○	○	○	○
Discussion and revision of the borderline proficiency level descriptors	○	○	○	○
Overview of the standard-setting procedure	○	○	○	○
Practice exercise for the standard-setting procedure	○	○	○	○
Judgment rounds	○	○	○	○
Judgment round feedback - table-level statistics	○	○	○	○
Judgment round feedback - committee-level statistics	○	○	○	○
Judgment round feedback - panelist agreement data	○	○	○	○
Judgment round feedback - impact data	○	○	○	○
Discussions after each round	○	○	○	○

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

	Very Useful	Useful	Somewhat Useful	Not Useful
Proficiency Level Descriptors (PLDs)				
Borderline proficiency level descriptors				
Table-level statistics after Rounds 1 and 2				
Committee-level statistics after Round 2				
Panelist agreement data provided after Round 1				
Panelist agreement data provided after Round 2				
Impact data after Round 2				
Discussion after each judgment round				

How adequate were the following elements of the session?

	Not Adequate	Somewhat Adequate	Adequate	More Than Adequate
Training provided on the standard-setting process				
Amount of time spent training				
Total amount of time to review and discuss borderline proficiency level descriptors				
Total amount of time to discuss the practice judgments				
Amount of time to make judgments				
Visual presentation of the feedback provided				
Number of judgment rounds				

In applying the standard-setting method, you were asked to recommend cut scores (separating four proficiency levels) for student performance on TELPAS assessments.

How confident do you feel that the Proficiency Level Descriptors (PLDs) for grades 4–5 Reading are reasonable for each student proficiency level?

	Not Confident	Somewhat Confident	Confident	Very Confident
Intermediate				
Advanced				
High Advanced				

Which Borderline Proficiency Level group did you participate in?

- **Reading**
- **Speaking**
- **Listening**

Appendix D – Committee Participant Composition

Table D.1. ***Participant Position***

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
Teacher (K–12)	5	7	12	8	12	8
Teacher (Higher Ed.)	0	0	0	0	0	0
Administrator (School)	1	1	1	0	2	0
Administrator (District)	1	1	0	3	1	1
Other	1	1	1	2	2	5
Total	**8**	**10**	**14**	**13**	**17**	**14**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
Teacher (K–12)	12	12	15	13
Teacher (Higher Ed.)	0	0	0	0
Administrator (School)	2	1	1	1
Administrator (District)	2	0	4	1
Other	2	1	3	6
Total	**18**	**14**	**23**	**21**

Table D.2. ***Years of Teaching Experience***

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
None	0	0	0	0	0	0
1 to 5 years	1	1	0	0	0	0
6 to 10 years	3	2	5	1	3	1
11 to 15 years	1	1	2	2	3	5
16 to 20 years	2	4	5	5	7	3
More than 20 years	1	2	2	5	4	5
Total	**8**	**10**	**14**	**13**	**17**	**14**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
None	0	0	0	0
1 to 5 years	2	0	0	0
6 to 10 years	5	5	2	3
11 to 15 years	2	2	4	6
16 to 20 years	6	5	10	5
More than 20 years	3	2	7	7
Total	**18**	**14**	**23**	**21**

Table D.3. ***Years of Teaching Experience in Subject Within Grades***

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
None	1	0	1	0	3	1
1 to 5 years	4	5	0	1	4	3
6 to 10 years	2	2	8	6	2	2
11 to 15 years	0	2	3	4	6	4
16 to 20 years	1	0	2	2	1	1
More than 20 years	0	1	0	0	1	3
Total	**8**	**10**	**14**	**13**	**17**	**14**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
None	1	1	1	3
1 to 5 years	9	0	3	5
6 to 10 years	4	8	7	3
11 to 15 years	2	3	8	6
16 to 20 years	1	2	3	1
More than 20 years	1	0	1	3
Total	**18**	**14**	**23**	**21**

Table D.4. ***Experience Teaching Student Populations***

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
Mainstream special education	5	6	7	6	14	9
Self-contained special education	4	4	4	2	4	2
English learners (ELs)	7	10	14	12	17	14
General education	6	7	10	12	17	10
Vocational technical education	0	1	1	0	2	0
Total	**22**	**28**	**36**	**32**	**54**	**35**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
Mainstream special education	11	7	15	14
Self-contained special education	8	4	4	4
English learners (EL)	17	14	22	21
General education	13	10	22	17
Vocational technical education	1	1	1	1
Total	**50**	**36**	**64**	**57**

Table D.5. ***Experience Teaching Languages***

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
Spanish	6	10	13	13	16	13
Vietnamese	2	0	1	0	2	5
Chinese	2	1	0	1	2	3
Tagalog	1	0	0	0	2	3
German	0	0	1	0	0	3
French	0	1	1	1	1	4
Hindi	1	0	0	0	2	3
Urdu	0	0	0	0	1	2
Korean	1	0	0	0	1	5
Arabic	2	1	0	1	2	4
Other	2	0	2	0	2	6
Total	**17**	**13**	**18**	**16**	**31**	**51**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
Spanish	16	13	23	19
Vietnamese	2	1	1	6
Chinese	3	0	2	4
Tagalog	1	0	1	4
German	0	1	0	3
French	1	1	2	4
Hindi	1	0	1	4
Urdu	0	0	1	2
Korean	1	0	1	5
Arabic	3	0	2	5
Other	2	2	0	8
Total	**30**	**18**	**34**	**64**

Table D.6. ***Highest Education Degree***

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
High School Diploma	0	0	0	0	0	0
Associate's degree	0	0	0	0	0	0
Bachelor's degree	3	3	8	3	7	7
Master's degree	5	7	6	10	10	7
Doctoral degree	0	0	0	0	0	0
Total	**8**	**10**	**14**	**13**	**17**	**14**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
High School Diploma	0	0	0	0
Associate's degree	0	0	0	0
Bachelor's degree	6	8	7	10
Master's degree	12	6	16	11
Doctoral degree	0	0	0	0
Total	**18**	**14**	**23**	**21**

Table D.7. ***Demographic: Gender***

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
Male	0	1	1	2	2	3
Female	8	9	12	11	14	11
No answer	0	0	1	0	1	0
Total	**8**	**10**	**14**	**13**	**17**	**14**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
Male	1	1	3	4
Female	17	12	19	17
No answer	0	1	1	0
Total	**18**	**14**	**23**	**21**

Table D.8. ***Demographic: Ethnicity***

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
Hispanic or Latino	5	7	13	8	7	7
Not Hispanic or Latino	3	3	1	4	8	7
No answer	0	0	0	1	2	0
Total	**8**	**10**	**14**	**13**	**17**	**14**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
Hispanic or Latino	12	13	12	10
Not Hispanic or Latino	6	1	9	10
No answer	0	0	2	1
Total	**18**	**14**	**23**	**21**

Table D.9. *Demographic: Race*

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
American Indian or Alaskan Native	1	0	0	0	0	1
Asian	0	0	0	0	0	0
Black or African American	0	1	0	1	2	0
Native Hawaiian or Pacific Islander	0	0	0	0	0	0
White	6	6	13	11	15	13
No answer	1	3	1	1	0	0
Total	**8**	**10**	**14**	**13**	**17**	**14**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
American Indian or Alaskan Native	1	0	0	1
Asian	0	0	0	0
Black or African American	1	0	1	2
Native Hawaiian or Pacific Islander	0	0	0	0
White	12	13	21	18
No answer	4	1	1	0
Total	**18**	**14**	**23**	**21**

Table D.10. ***Currently Work in a School District***

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
Yes	8	10	14	13	17	14
No	0	0	0	0	0	0
Total	**8**	**10**	**14**	**13**	**17**	**14**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
Yes	18	14	23	21
No	0	0	0	0
Total	**18**	**14**	**23**	**21**

Table D.11. ***Size of School District***

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
Small	2	4	1	1	6	1
Medium	5	2	3	4	5	5
Large	1	4	10	8	6	8
Total	**8**	**10**	**14**	**13**	**17**	**14**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
Small	6	1	7	1
Medium	7	3	5	9
Large	5	10	11	11
Total	**18**	**15**	**23**	**21**

Table D.12. ***Type of School District***

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
Rural	2	2	3	2	5	5
Metropolitan/Urban	2	5	6	7	7	4
Suburban	4	3	5	4	5	5
Total	**10**	**10**	**14**	**13**	**17**	**14**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
Rural	4	3	6	6
Metropolitan/Urban	7	6	12	6
Suburban	7	5	5	9
Total	**18**	**14**	**23**	**21**

Table D.13. *Socioeconomic Status of School District*

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
Low	5	8	12	6	8	5
Moderate	3	2	2	7	9	9
High	0	0	0	0	0	0
Total	**8**	**10**	**14**	**13**	**17**	**14**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
Low	13	12	10	9
Moderate	5	2	13	12
High	0	0	0	0
Total	**18**	**14**	**23**	**21**

Table D.14. ***Attendance of the TELPAS Writing Scoring Training in the last 2 years***

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
Yes	2	6	6	4	7	8
No	6	4	8	9	10	6
Total	**8**	**10**	**14**	**13**	**17**	**14**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
Yes	8	6	8	11
No	10	8	15	10
Total	**18**	**14**	**23**	**21**

Table D.15. ***Attendance Year of the TELPAS Writing Scoring Training***

	Reading					
	Grade 2	Grade 3	Grades 4–5	Grades 6–7	Grades 8–9	Grades 10–12
2016–2017	1	0	1	1	3	2
2017–2018	1	6	4	3	4	6
Total	**2**	**6**	**5**	**4**	**7**	**8**

	Listening & Speaking			
	Grades 2–3	Grades 4–5	Grades 6–8	Grades 9–12
2016–2017	1	1	3	3
2017–2018	7	4	5	8
Total	**8**	**5**	**8**	**11**

Appendix E – Examples of Feedback Data

Feedback data was provided to participants after each judgment round. The following are examples of feedback data provided to participants.

Individual Test-Level Recommendations

This provided the participant with the recommendations for test-level cut scores based on their item judgments for the intermediate, advanced, and advanced high proficiency levels.

Reading Grades 4-5 - Individual Score Points - Round 1

Table=1 Name =

I Raw Score	I Roundup Raw Score	A Raw Score	A Roundup Raw Score	AH Raw Score	AH Roundup Raw Score
13	13	27.4	28	32.7	33

Table-level Test-Level Recommendations

This provided the participant with the aggregate test-level recommendation, based on the individual participants at the table, including the number of participants, the mean recommendation, the median recommendation, the minimum and maximum recommendation, and the first and third quartiles for each achievement level.

Reading Grades 4-5 Round 1 Summary Statistics - Table 1

	N	Mean	Median	Roundup Median	Min	Max	Q1	Q3
I Raw Score	5	12.40	13.00	13	7.50	15.40	11.90	14.20
A Raw Score	5	21.46	20.50	21	15.00	27.40	18.60	25.80
AH Raw Score	5	27.80	27.00	27	22.70	32.70	25.40	31.20

Overall Test-Level Recommendations

This provided the participant with the aggregate test-level recommendation, based on the individual participants in the committee, including the number of participants, the mean recommendation, the median recommendation, the minimum and maximum recommendation, and the first and third quartiles for each achievement level.

Reading Grades 4-5 Round 1 Summary Statistics - Overall

	N	Mean	Median	Roundup Median	Min	Max	Q1	Q3
I Raw Score	14	10.79	10.25	11	4.60	20.10	7.50	14.20
A Raw Score	14	21.62	21.00	21	15.00	27.90	18.60	25.80
AH Raw Score	14	28.44	28.20	29	20.60	34.30	25.40	32.10

Item-Level Judgment Agreement

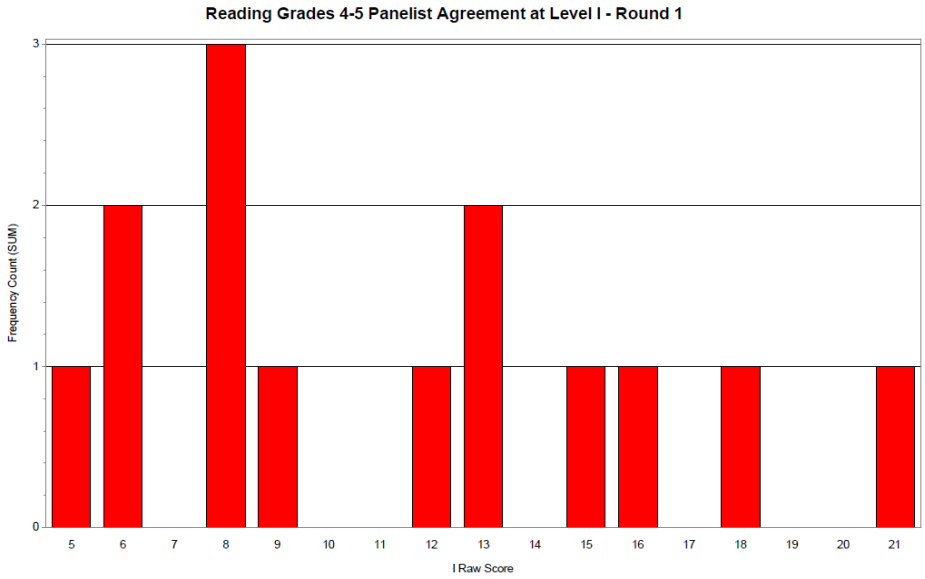
This provided the participants with item-level judgment distributions for the committee for each item. Additionally, for each achievement level, the items with the greatest level of judgment disagreement were identified.

Reading Grades 4-5 Round 1 Level I

UIN	Max Points	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
04R201C9DBC15010	1	0%	0%	0%	7%	21%	36%	29%	7%	0%	0%	0%
04R201F9EMC16012	1	0%	7%	29%	7%	43%	7%	7%	0%	0%	0%	0%
04R201C9DBC16007	1	0%	14%	7%	14%	14%	29%	14%	7%	0%	0%	0%
04R301F9EMC16017	1	0%	21%	21%	7%	50%	0%	0%	0%	0%	0%	0%
04R401F9ETC15026	1	0%	21%	21%	29%	14%	7%	7%	0%	0%	0%	0%
04R301F9EMC16018	1	0%	29%	7%	21%	14%	21%	7%	0%	0%	0%	0%

Test-Level Participant Recommendation Agreement

This feedback was presented to participants by the facilitator. It presented bar graphs displaying the distribution of participant recommendations for the cut score, by raw score, for each proficiency level: beginning, intermediate, advanced, and advanced high proficiency levels. Graphs displaying consecutive achievement levels (beginning, intermediate, advanced, and advanced high) on the scale graph were also presented.



Item Score Mean and Score Distribution

This provided, for each item, the mean score and the distribution of scores received by students during the spring 2018 administration. The results presented were based on the sample of data used to create the impact data.

Item Score Mean and Distribution

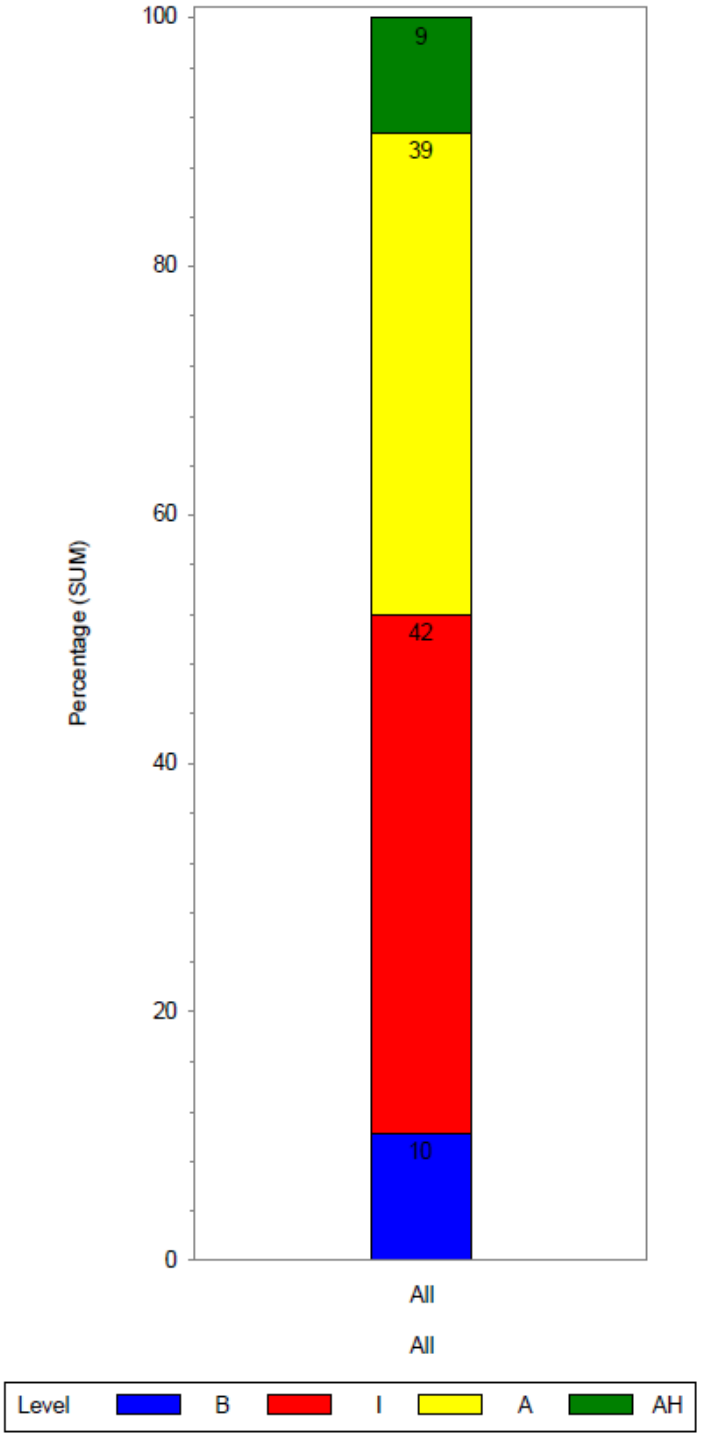
Grades 4-5 Listening

Sequence	Item	Maximum Points	Score Mean	Score Distribution				
				0 pts	1 pt	2 pts	3 pts	4 pts
1	04L101C9DBI16007	1	0.93	7%	93%			
2	04L101C9DBI16014	1	0.93	7%	93%			
3	04L302G9ESZ16204-1602I	1	0.49	51%	49%			
4	04L303H9ESZ16206-1602I	1	0.64	36%	64%			
5	04L302G9ESZ16205-1602I	1	0.70	30%	70%			
6	04L302G9ESZ16207-1602I	1	0.84	16%	84%			
7	04L202I2ESP16018	1	0.80	20%	80%			
8	04L302G9DBZ16209-1603O	1	0.59	41%	59%			
9	04L303H9DBZ16211-1603O	1	0.48	52%	48%			
10	04L302G9DBZ16210-1603O	1	0.66	34%	66%			

Impact Data

This provided the percentage of students expected to be classified into each proficiency level, beginning, intermediate, advanced, and advanced high, based on the committee test-level cut score recommendations for that round. These results were based on the sample of student data from the spring 2018 administration.

Listening Grades 4-5 Round1 Impact Data - Whole Group



Appendix F – Committee Recommended Cut Scores by Round

Table F.1. ***Reading Grade 2***

Proficiency Level	Grade	Maximum Score	Rounds			Final
			1	2	3	
Intermediate	2	30	10	8	10	10
Advanced			21	19	19	19
Advanced High			27	25	25	25

Table F.2. ***Reading Grade 3***

Proficiency Level	Grade	Maximum Score	Rounds			Final
			1	2	3	
Intermediate	3	34	8	8	9	9
Advanced			19	17	18	18
Advanced High			28	26	26	26

Table F.3. ***Reading Grades 4–5***

Proficiency Level	Grade	Maximum Score	Rounds			Final
			1	2	3	
Intermediate	4–5	36	11	11	8	8
Advanced			21	23	19	19
Advanced High			29	30	25	25

Table F.4. ***Reading Grades 6–7***

Proficiency Level	Grade	Maximum Score	Rounds			Final
			1	2	3	
Intermediate	6–7	37	11	11	11	11
Advanced			21	21	20	20
Advanced High			31	30	28	28

Table F.5. ***Reading Grades 8–9***

Proficiency Level	Grade	Maximum Score	Rounds			Final
			1	2	3	
Intermediate	8–9	37	9	9	9	9
Advanced			20	19	19	19
Advanced High			30	28	29	29

Table F.6. ***Reading Grades 10–12***

Proficiency Level	Grade	Maximum Score	Rounds			Final
			1	2	3	
Intermediate	10–12	37	10	10	10	10
Advanced			23	23	22	22
Advanced High			32	32	29	29

Table F.7. ***Listening Grades 2–3***

Proficiency Level	Grade	Maximum Score	Rounds			Final
			1	2	3	
Intermediate	2–3	27	10	10	10	10
Advanced			17	17	17	17
Advanced High			23	23	22	22

Table F.8. ***Listening Grades 4–5***

Proficiency Level	Grade	Maximum Score	Rounds			Final
			1	2	3	
Intermediate	4–5	27	13	12	12	12
Advanced			19	19	19	19
Advanced High			24	23	22	22

Table F.9. ***Listening Grades 6–8***

Proficiency Level	Grade	Maximum Score	Rounds			Final
			1	2	3	
Intermediate	6–8	27	8	10	10	10
Advanced			16	17	17	17
Advanced High			23	22	22	22

Table F.10. ***Listening Grades 9–12***

Proficiency Level	Grade	Maximum Score	Rounds			Final
			1	2	3	
Intermediate	9–12	27	10	10	10	10
Advanced			18	18	17	17
Advanced High			23	23	23	23

Table F.11. ***Speaking Grades 2–3***

Proficiency Level	Grade	Maximum Score	Rounds			Final
			1	2	3	
Intermediate	2–3	36	16	16	15	15
Advanced			25	24	24	24
Advanced High			32	29	29	29

Table F.12. ***Speaking Grades 4–5***

Proficiency Level	Grade	Maximum Score	Rounds			Final
			1	2	3	
Intermediate	4–5	36	12	12	17	17
Advanced			22	22	24	24
Advanced High			29	27	29	29

Table F.13. ***Speaking Grades 6–8***

Proficiency Level	Grade	Maximum Score	Rounds			Final
			1	2	3	
Intermediate	6–8	36	13	13	13	13
Advanced			23	23	23	23
Advanced High			31	31	30	30

Table F.14. ***Speaking Grades 9–12***

Proficiency Level	Grade	Maximum Score	Rounds			Final
			1	2	3	
Intermediate	9–12	36	15	15	16	16
Advanced			23	22	22	22
Advanced High			29	28	28	28

Appendix G – Recommended Cut Score Summary Statistics

Reading Grade 2

Round	Statistic	Intermediate	Advanced	Advanced High
1	Mean	10.12	20.32	26.48
	Minimum	6	17	24
	Q1	7	18	26
	Median	**10**	**21**	**27**
	Q3	12	21	28
	Maximum	17	26	29
2	Mean	8.70	17.61	23.88
	Minimum	6	14	20
	Q1	7	16	23
	Median	**8**	**19**	**25**
	Q3	10	20	25
	Maximum	13	20	26
3	Mean	8.70	17.61	23.88
	Minimum	6	14	20
	Q1	7	16	23
	Median	**8**	**19**	**25**
	Q3	10	20	25
	Maximum	13	20	26

Reading Grade 3

Round	Statistic	Intermediate	Advanced	Advanced High
1	Mean	7.57	18.7	27.41
	Minimum	4	13	23
	Q1	4	16	25
	Median	**8**	**19**	**28**
	Q3	9	21	30
	Maximum	13	25	31
2	Mean	7.39	17.07	24.55
	Minimum	3	12	17
	Q1	6	15	23
	Median	**8**	**17**	**26**
	Q3	10	20	28
	Maximum	10	24	30
3	Mean	7.39	17.07	24.55
	Minimum	3	12	17
	Q1	6	15	23
	Median	**8**	**17**	**26**
	Q3	10	20	28
	Maximum	10	24	30

Reading Grades 4–5

Round	Statistic	Intermediate	Advanced	Advanced High
1	Mean	10.79	21.62	28.44
	Minimum	5	15	21
	Q1	8	19	25
	Median	**11**	**21**	**29**
	Q3	14	26	32
	Maximum	20	28	34
2	Mean	11.00	22.21	28.56
	Minimum	5	16	21
	Q1	7	19	24
	Median	**11**	**23**	**30**
	Q3	14	26	32
	Maximum	21	28	35
3	Mean	11.00	22.21	28.56
	Minimum	5	16	21
	Q1	7	19	24
	Median	**11**	**23**	**30**
	Q3	14	26	32
	Maximum	21	28	35

Reading Grades 6–7

Round	Statistic	Intermediate	Advanced	Advanced High
1	Mean	10.38	20.42	30.02
	Minimum	4	12	23
	Q1	8	18	30
	Median	**11**	**21**	**31**
	Q3	13	23	32
	Maximum	16	29	34
2	Mean	10.39	20.11	29.46
	Minimum	5	12	23
	Q1	7	18	29
	Median	**11**	**21**	**30**
	Q3	14	22	30
	Maximum	16	28	33
3	Mean	10.39	20.11	29.46
	Minimum	5	12	23
	Q1	7	18	29
	Median	**11**	**21**	**30**
	Q3	14	22	30
	Maximum	16	28	33

Reading Grades 8–9

Round	Statistic	Intermediate	Advanced	Advanced High
1	Mean	8.81	19.1	28.54
	Minimum	3	15	20
	Q1	6	16	26
	Median	**9**	**20**	**30**
	Q3	11	22	31
	Maximum	16	27	35
2	Mean	8.95	18.71	27.72
	Minimum	6	16	21
	Q1	7	16	27
	Median	**9**	**19**	**28**
	Q3	10	20	30
	Maximum	16	24	33
3	Mean	8.95	18.71	27.72
	Minimum	6	16	21
	Q1	7	16	27
	Median	**9**	**19**	**28**
	Q3	10	20	30
	Maximum	16	24	33

Reading Grades 10–12

Round	Statistic	Intermediate	Advanced	Advanced High
1	Mean	8.56	22.54	31.29
	Minimum	3	18	26
	Q1	5	19	30
	Median	**10**	**23**	**32**
	Q3	12	24	33
	Maximum	15	27	35
2	Mean	8.36	22.61	31.37
	Minimum	3	20	26
	Q1	6	21	30
	Median	**10**	**23**	**32**
	Q3	10	23	33
	Maximum	12	28	35
3	Mean	8.36	22.61	31.37
	Minimum	3	20	26
	Q1	6	21	30
	Median	**10**	**23**	**32**
	Q3	10	23	33
	Maximum	12	28	35

Listening Grades 2–3

Round	Statistic	Intermediate	Advanced	Advanced High
1	Mean	9.23	16.91	22.33
	Minimum	5	12	19
	Q1	8	15	20
	Median	**10**	**17**	**23**
	Q3	11	19	24
	Maximum	13	22	25
2	Mean	9.54	16.79	21.97
	Minimum	6	12	19
	Q1	9	16	21
	Median	**10**	**17**	**23**
	Q3	10	18	23
	Maximum	13	21	25
3	Mean	9.54	16.79	21.97
	Minimum	6	12	19
	Q1	9	16	21
	Median	**10**	**17**	**23**
	Q3	10	18	23
	Maximum	13	21	25

Listening Grades 4–5

Round	Statistic	Intermediate	Advanced	Advanced High
1	Mean	12.16	19.28	22.69
	Minimum	6	15	20
	Q1	11	18	21
	Median	**13**	**19**	**24**
	Q3	14	21	24
	Maximum	18	23	25
2	Mean	11.83	18.44	21.78
	Minimum	7	15	20
	Q1	11	18	21
	Median	**12**	**19**	**23**
	Q3	13	20	22
	Maximum	17	21	23
3	Mean	11.83	18.44	21.78
	Minimum	7	15	20
	Q1	11	18	21
	Median	**12**	**19**	**23**
	Q3	13	20	22
	Maximum	17	21	23

Listening Grades 6–8

Round	Statistic	Intermediate	Advanced	Advanced High
1	Mean	8.13	15.59	21.53
	Minimum	4	10	18
	Q1	6	14	20
	Median	**8**	**16**	**23**
	Q3	11	17	23
	Maximum	15	20	25
2	Mean	8.64	15.59	20.82
	Minimum	5	10	13
	Q1	7	14	20
	Median	**10**	**17**	**22**
	Q3	10	17	22
	Maximum	13	20	23
3	Mean	8.64	15.59	20.82
	Minimum	5	10	13
	Q1	7	14	20
	Median	**10**	**17**	**22**
	Q3	10	17	22
	Maximum	13	20	23

Listening Grades 9–12

Round	Statistic	Intermediate	Advanced	Advanced High
1	Mean	9.52	16.8	21.95
	Minimum	3	11	17
	Q1	7	15	20
	Median	**10**	**18**	**23**
	Q3	12	18	23
	Maximum	18	22	25
2	Mean	9.76	17.00	21.9
	Minimum	5	11	17
	Q1	8	16	21
	Median	**10**	**18**	**23**
	Q3	11	18	23
	Maximum	14	19	24
3	Mean	9.76	17.00	21.9
	Minimum	5	11	17
	Q1	8	16	21
	Median	**10**	**18**	**23**
	Q3	11	18	23
	Maximum	14	19	24

Speaking Grades 2–3

Round	Statistic	Intermediate	Advanced	Advanced High
1	Mean	16.00	25.39	32.11
	Minimum	10	20	29
	Q1	14	24	31
	Median	**16**	**25**	**32**
	Q3	19	27	34
	Maximum	20	28	36
2	Mean	15.94	24.06	29.17
	Minimum	13	21	24
	Q1	15	24	28
	Median	**16**	**24**	**29**
	Q3	18	25	31
	Maximum	19	27	32
3	Mean	15.94	24.06	29.17
	Minimum	13	21	24
	Q1	15	24	28
	Median	**16**	**24**	**29**
	Q3	18	25	31
	Maximum	19	27	32

Speaking Grades 4–5

Round	Statistic	Intermediate	Advanced	Advanced High
1	Mean	12.14	22.36	28.93
	Minimum	10	20	25
	Q1	10	21	27
	Median	**12**	**22**	**29**
	Q3	12	24	32
	Maximum	18	26	33
2	Mean	12.21	21.43	27.36
	Minimum	10	18	24
	Q1	11	20	26
	Median	**12**	**22**	**27**
	Q3	13	23	28
	Maximum	15	24	32
3	Mean	12.21	21.43	27.36
	Minimum	10	18	24
	Q1	11	20	26
	Median	**12**	**22**	**27**
	Q3	13	23	28
	Maximum	15	24	32

Speaking Grades 6–8

Round	Statistic	Intermediate	Advanced	Advanced High
1	Mean	13.09	23.52	30.78
	Minimum	10	20	25
	Q1	12	22	30
	Median	**13**	**23**	**31**
	Q3	14	25	32
	Maximum	20	30	35
2	Mean	13.61	23.83	31.09
	Minimum	11	22	29
	Q1	13	23	31
	Median	**13**	**23**	**31**
	Q3	13	24	31
	Maximum	18	29	34
3	Mean	13.61	23.83	31.09
	Minimum	11	22	29
	Q1	13	23	31
	Median	**13**	**23**	**31**
	Q3	13	24	31
	Maximum	18	29	34

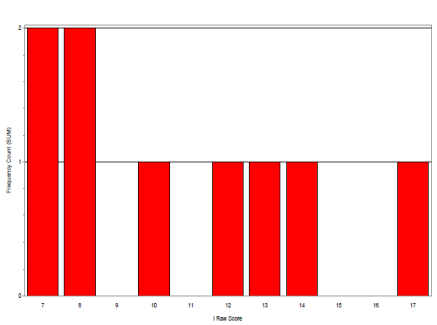
Speaking Grades 9–12

Round	Statistic	Intermediate	Advanced	Advanced High
1	Mean	15.37	23.05	29.47
	Minimum	10	18	26
	Q1	14	21	28
	Median	**15**	**23**	**29**
	Q3	17	25	30
	Maximum	22	27	34
2	Mean	15.19	22.29	28
	Minimum	10	18	25
	Q1	14	21	27
	Median	**15**	**22**	**28**
	Q3	17	24	28
	Maximum	20	26	33
3	Mean	15.19	22.29	28
	Minimum	10	18	25
	Q1	14	21	27
	Median	**15**	**22**	**28**
	Q3	17	24	28
	Maximum	20	26	33

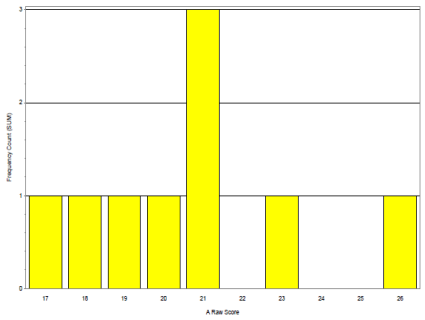
Appendix H – Test-Level Participant Judgment Agreement

Reading Grade 2

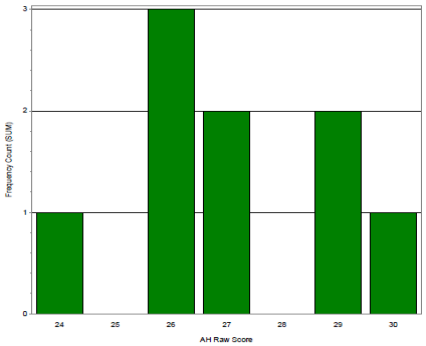
Round 1:



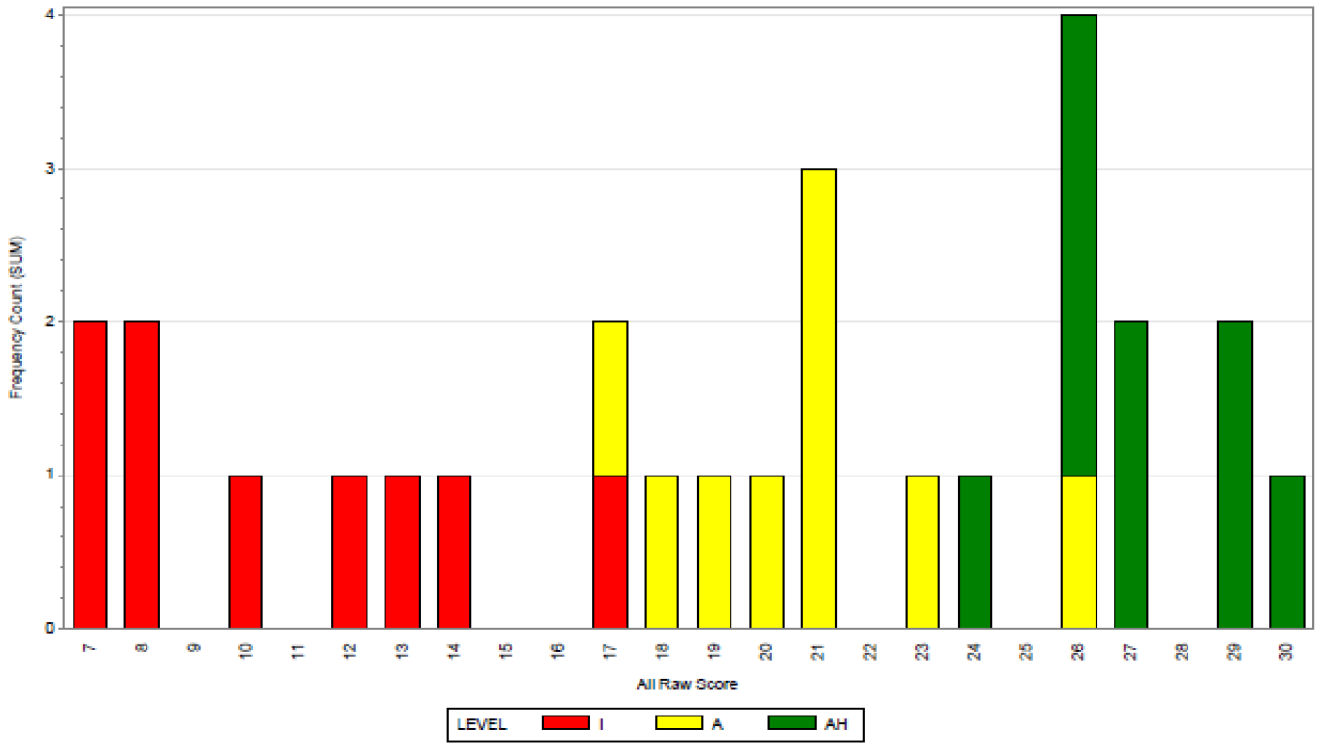
Intermediate



Advanced



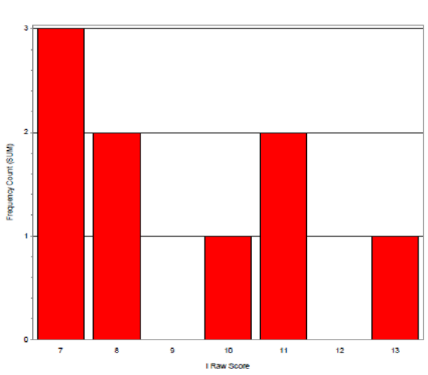
Advanced High



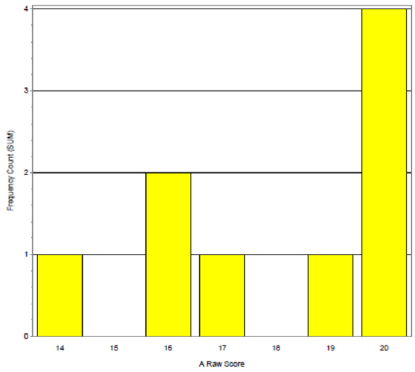
All Three Proficiency Levels Concurrently

Reading Grade 2

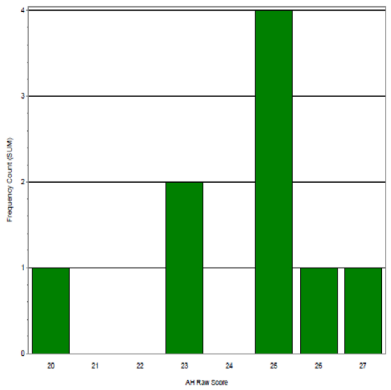
Round 2:



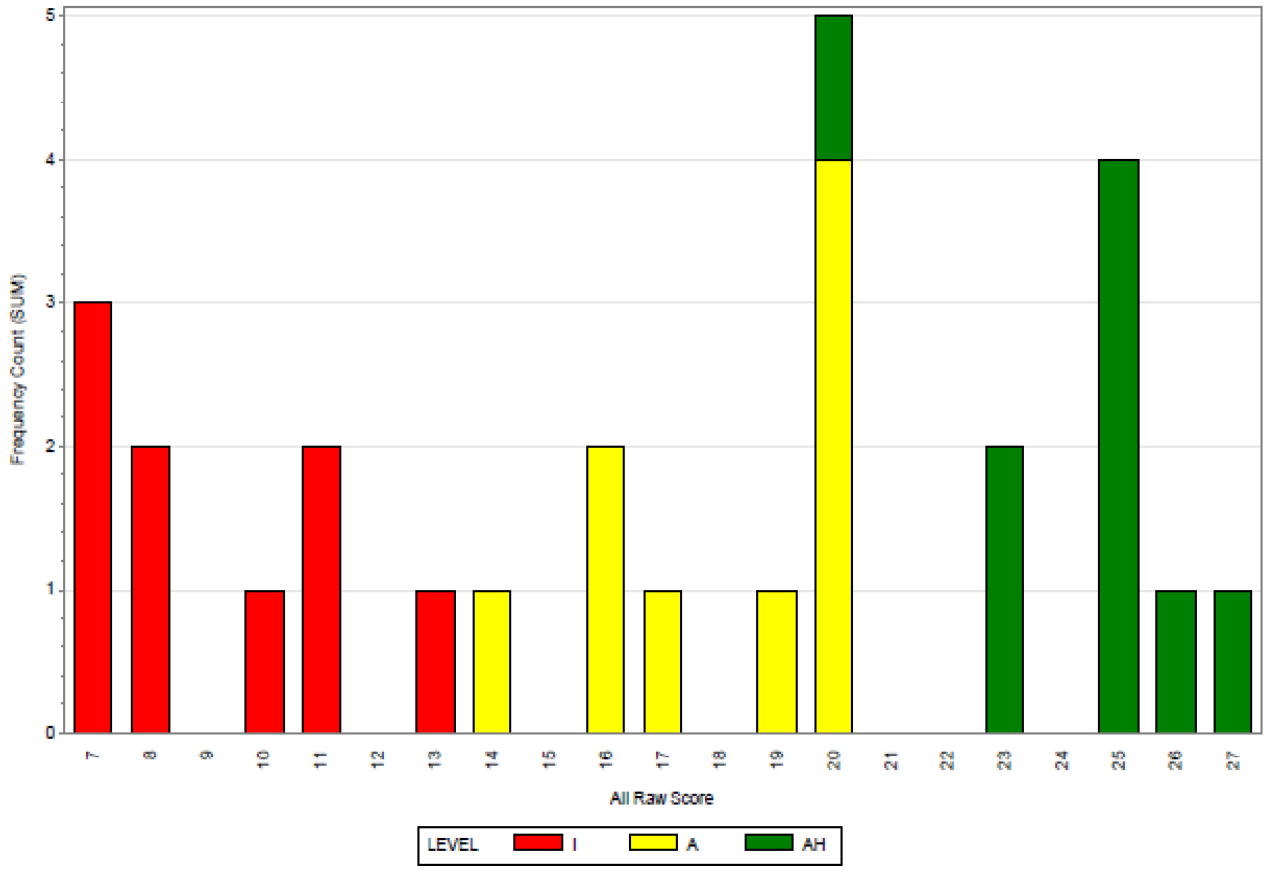
Intermediate



Advanced



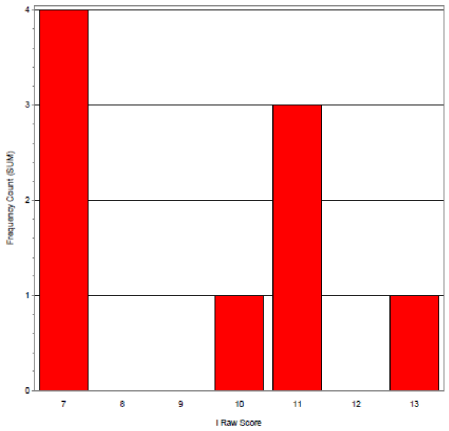
Advanced High



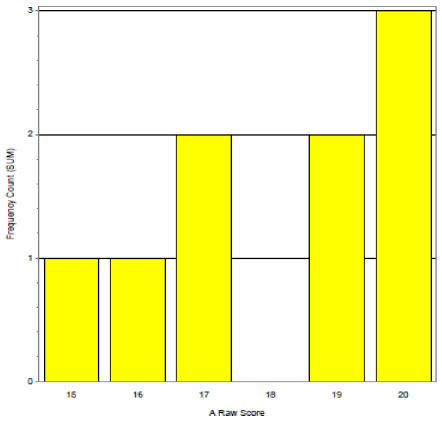
All Three Proficiency Levels Concurrently

Reading Grade 2

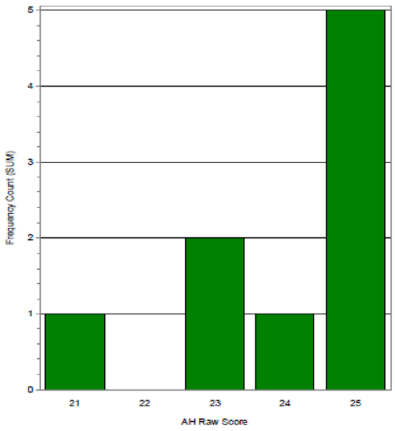
Round 3:



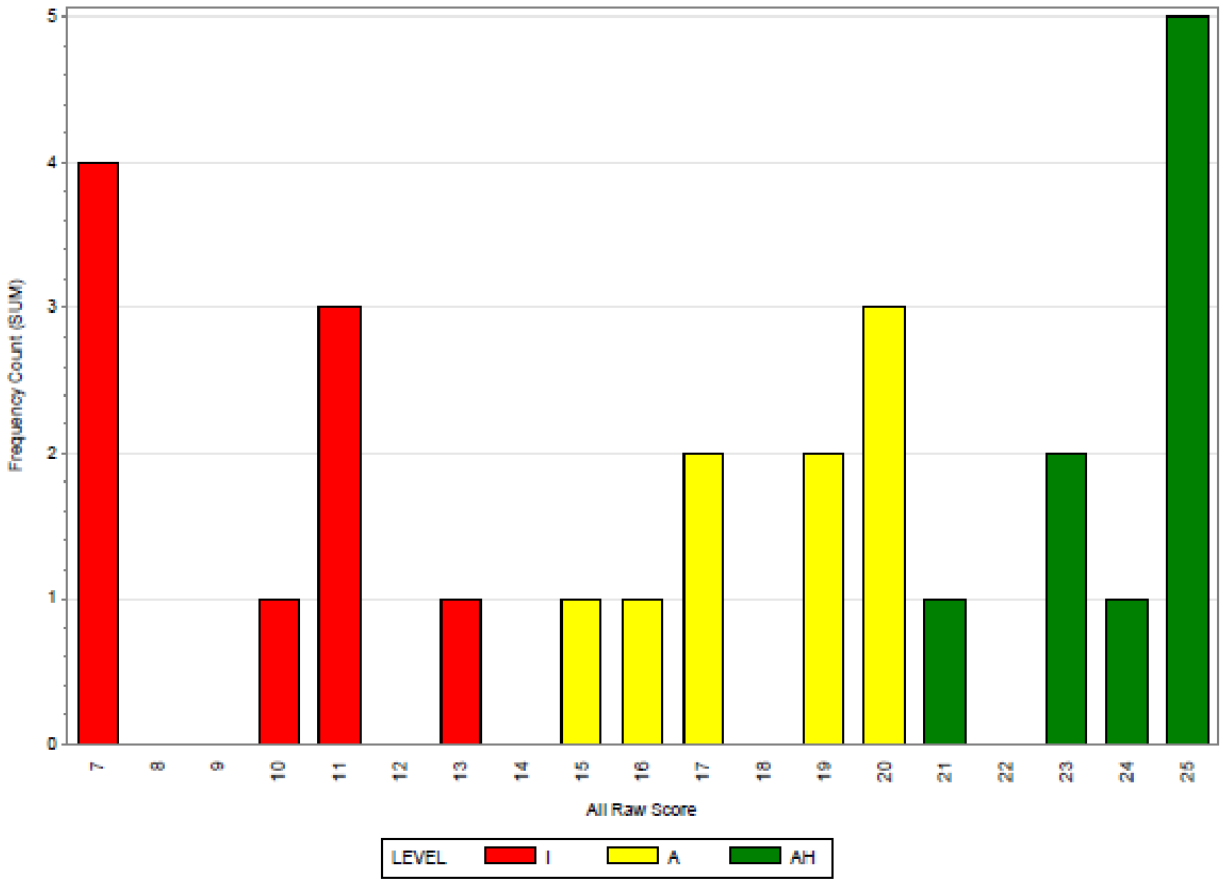
Intermediate



Advanced



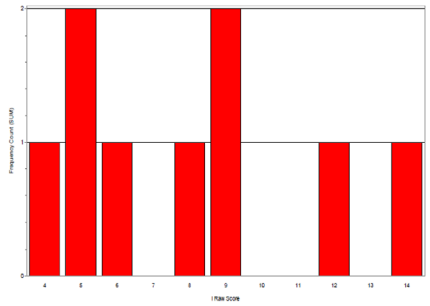
Advanced High



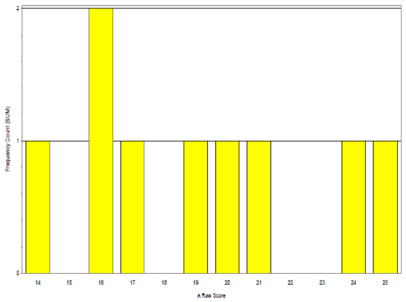
All Three Proficiency Levels Concurrently

Reading Grade 3

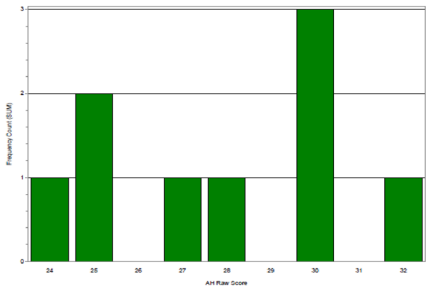
Round 1:



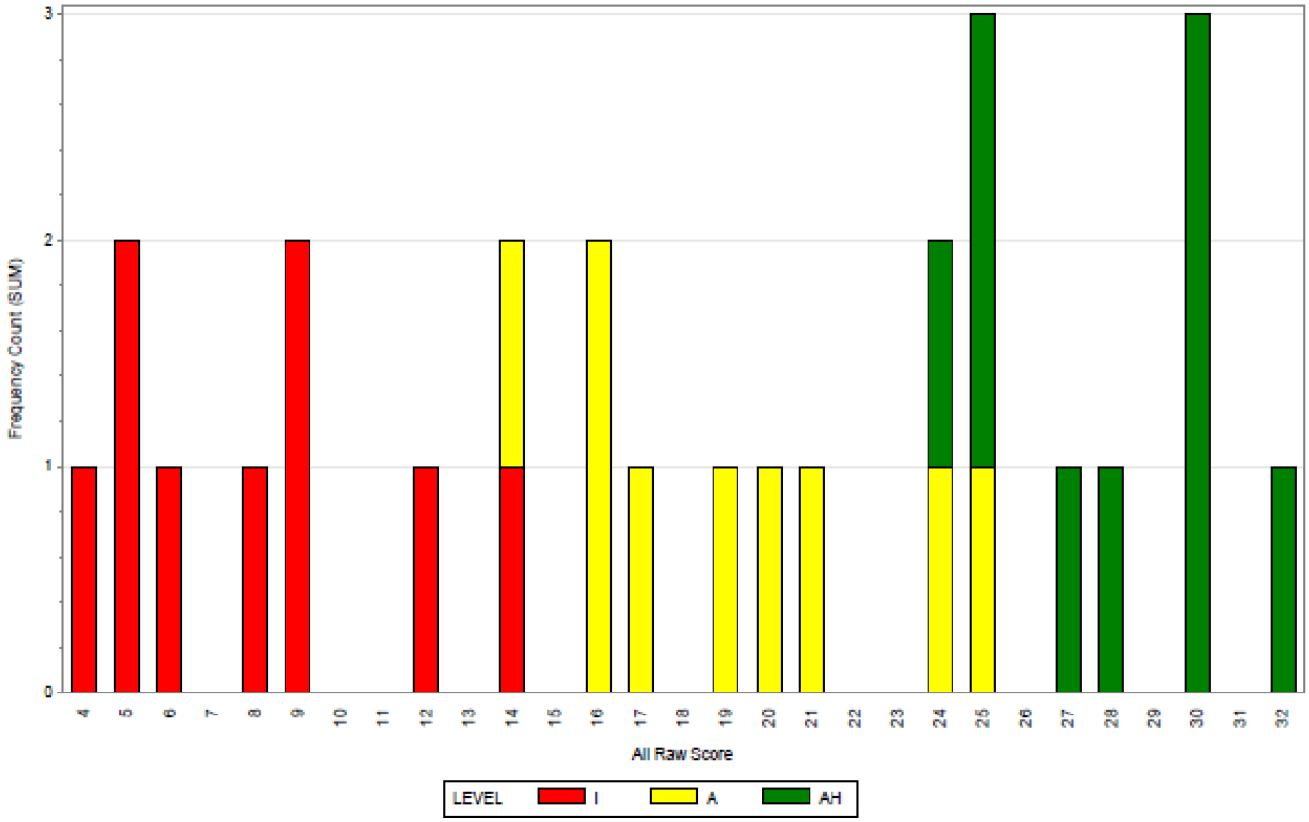
Intermediate



Advanced



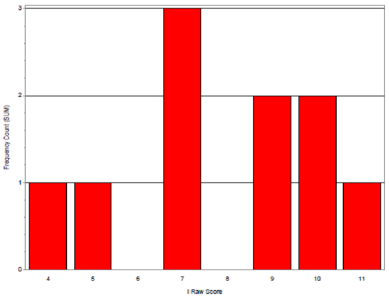
Advanced High



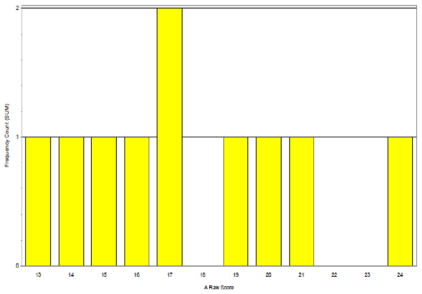
All Three Proficiency Levels Concurrently

Reading Grade 3

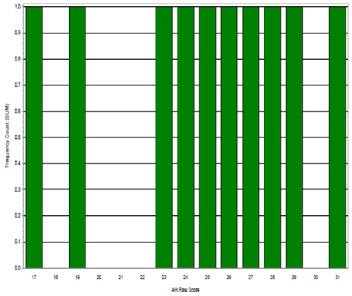
Round 2:



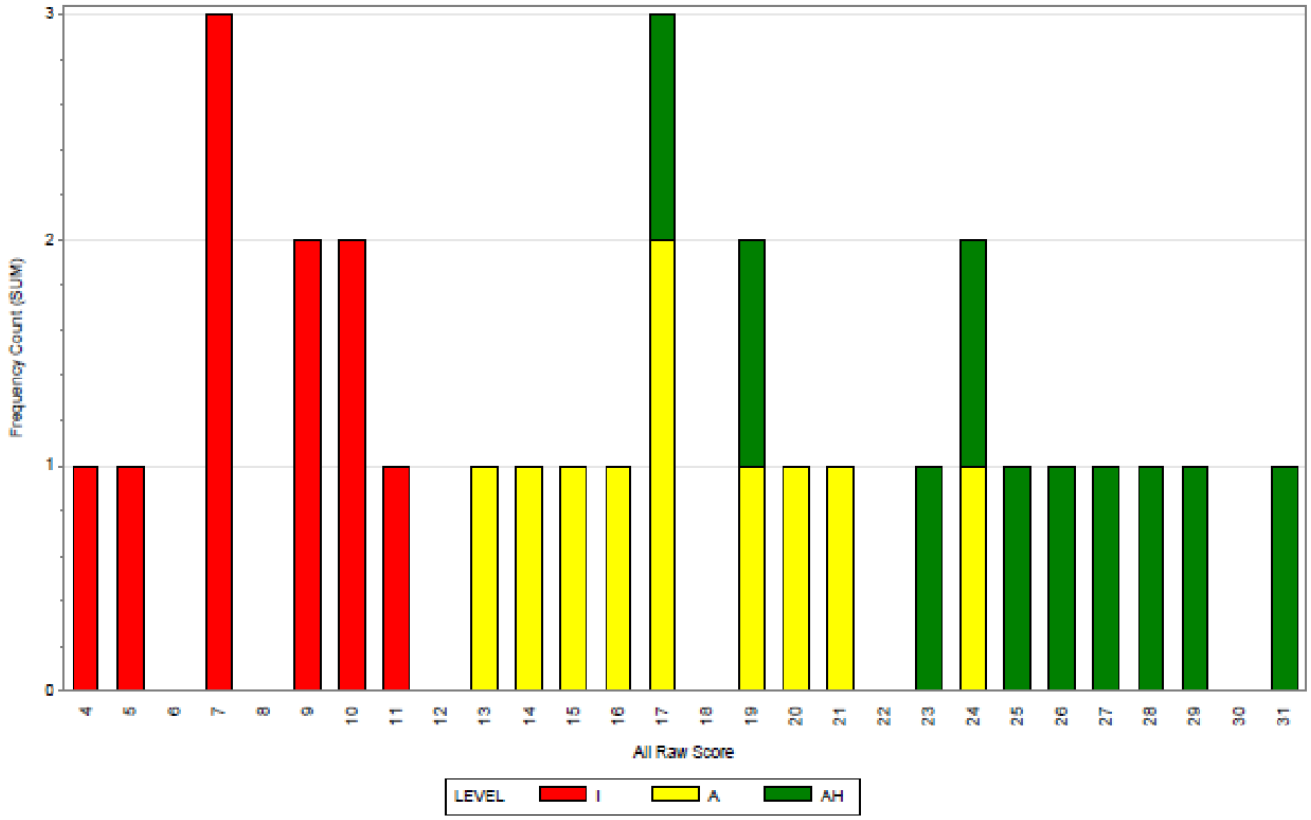
Intermediate



Advanced



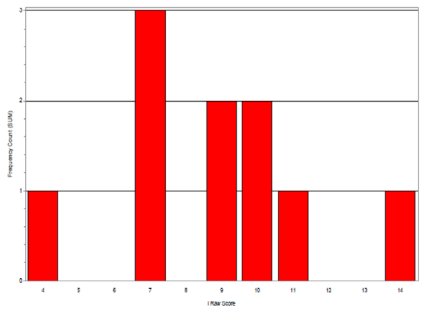
Advanced High



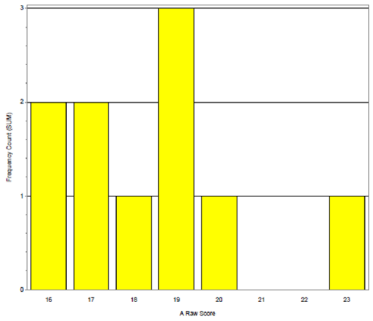
All Three Proficiency Levels Concurrently

Reading Grade 3

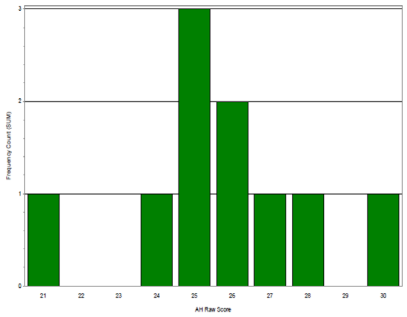
Round 3:



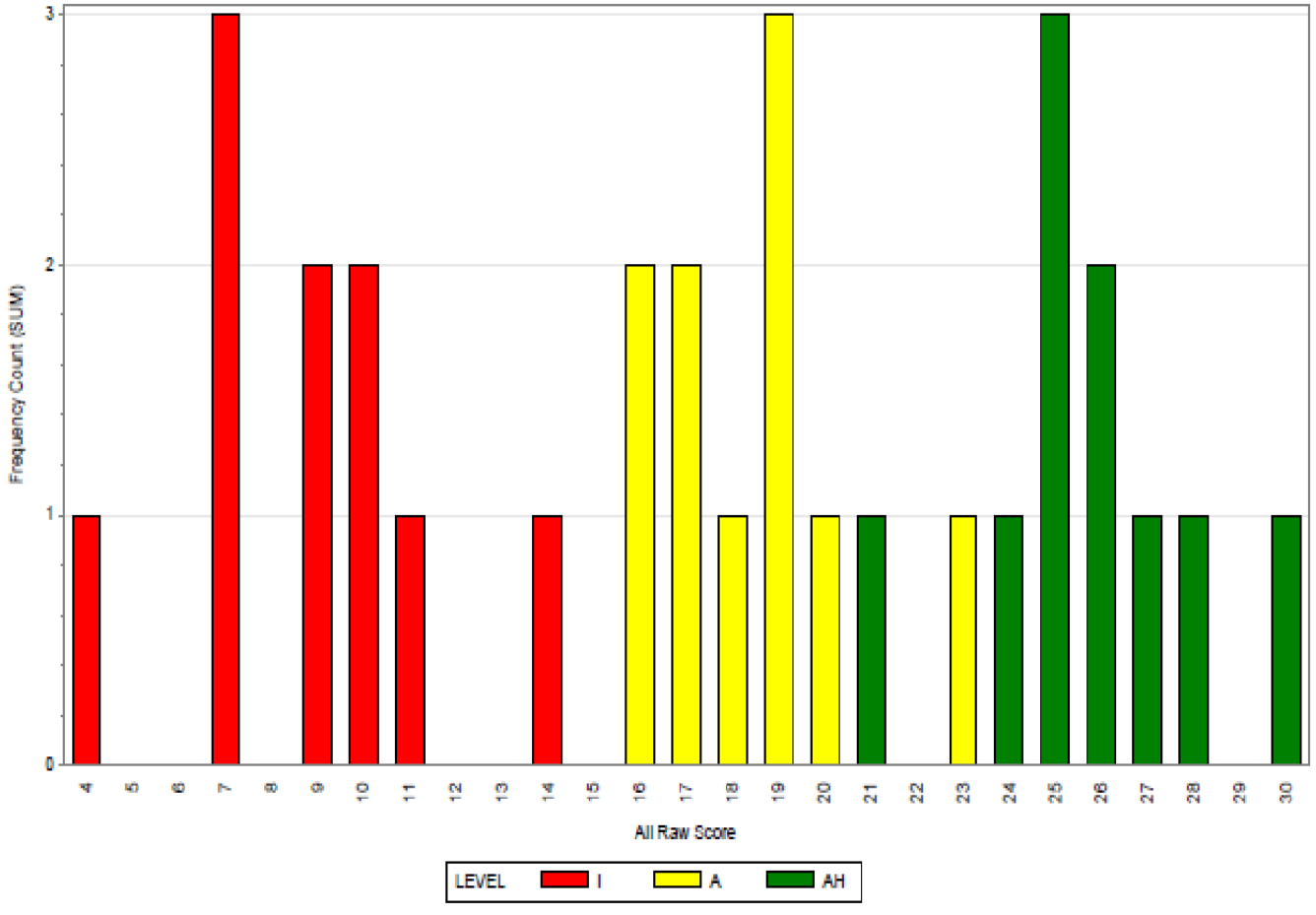
Intermediate



Advanced



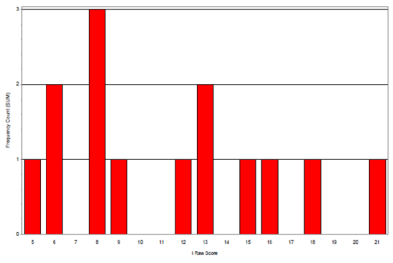
Advanced High



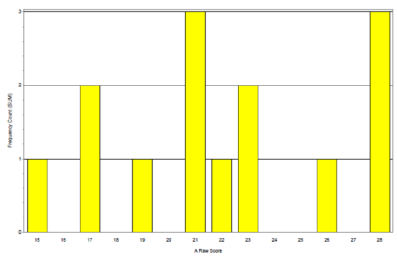
All Three Proficiency Levels Concurrently

Reading Grades 4–5

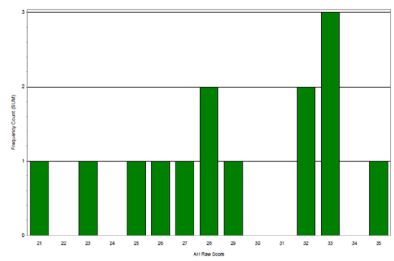
Round 1:



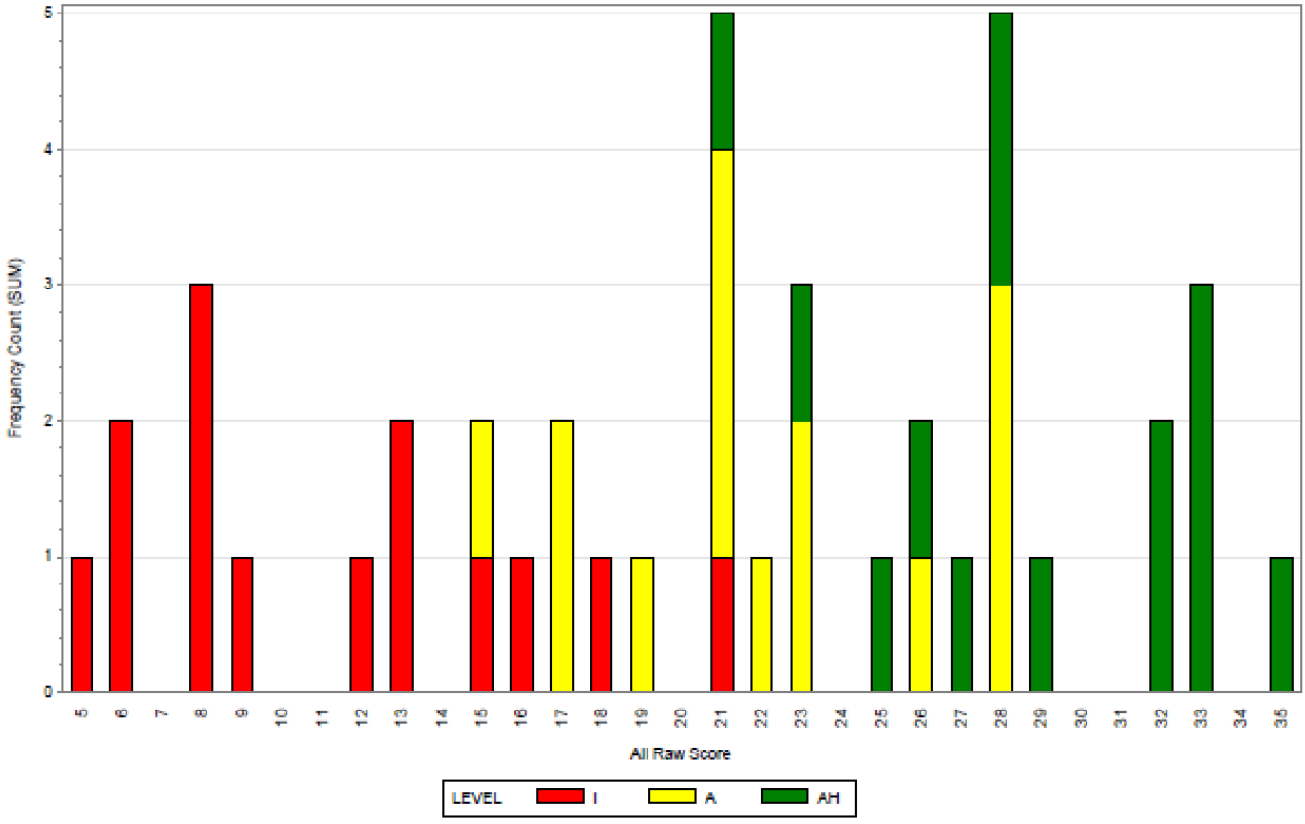
Intermediate



Advanced



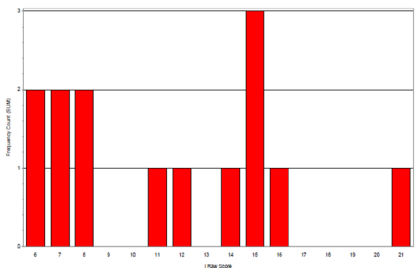
Advanced High



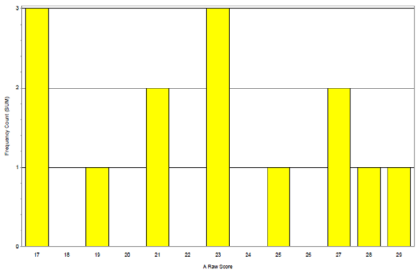
All Three Proficiency Levels Concurrently

Reading Grades 4–5

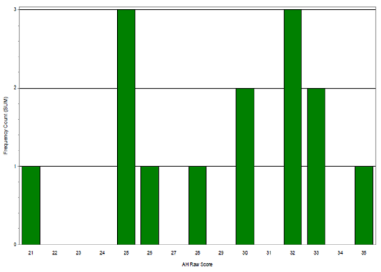
Round 2:



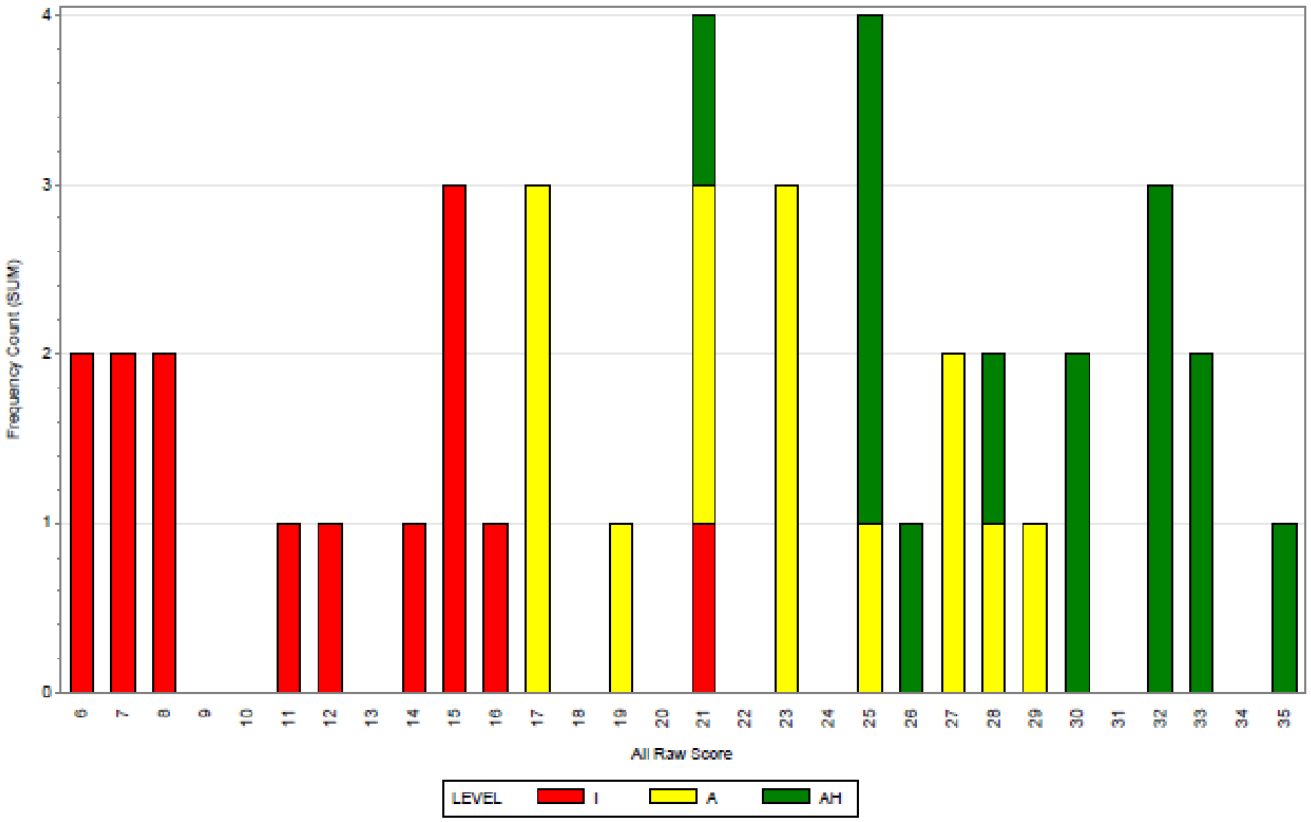
Intermediate



Advanced



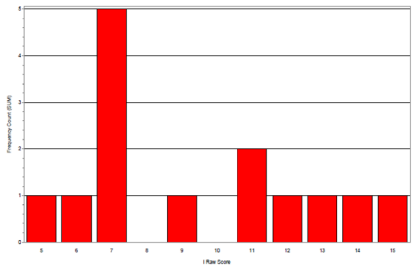
Advanced High



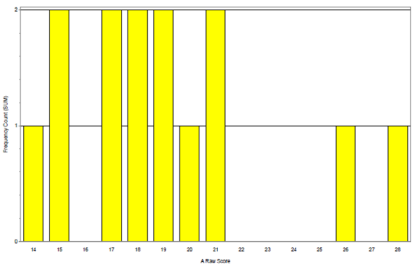
All Three Proficiency Levels Concurrently

Reading Grades 4–5

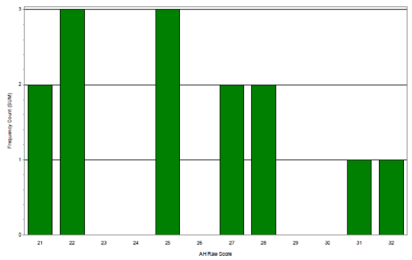
Round 3:



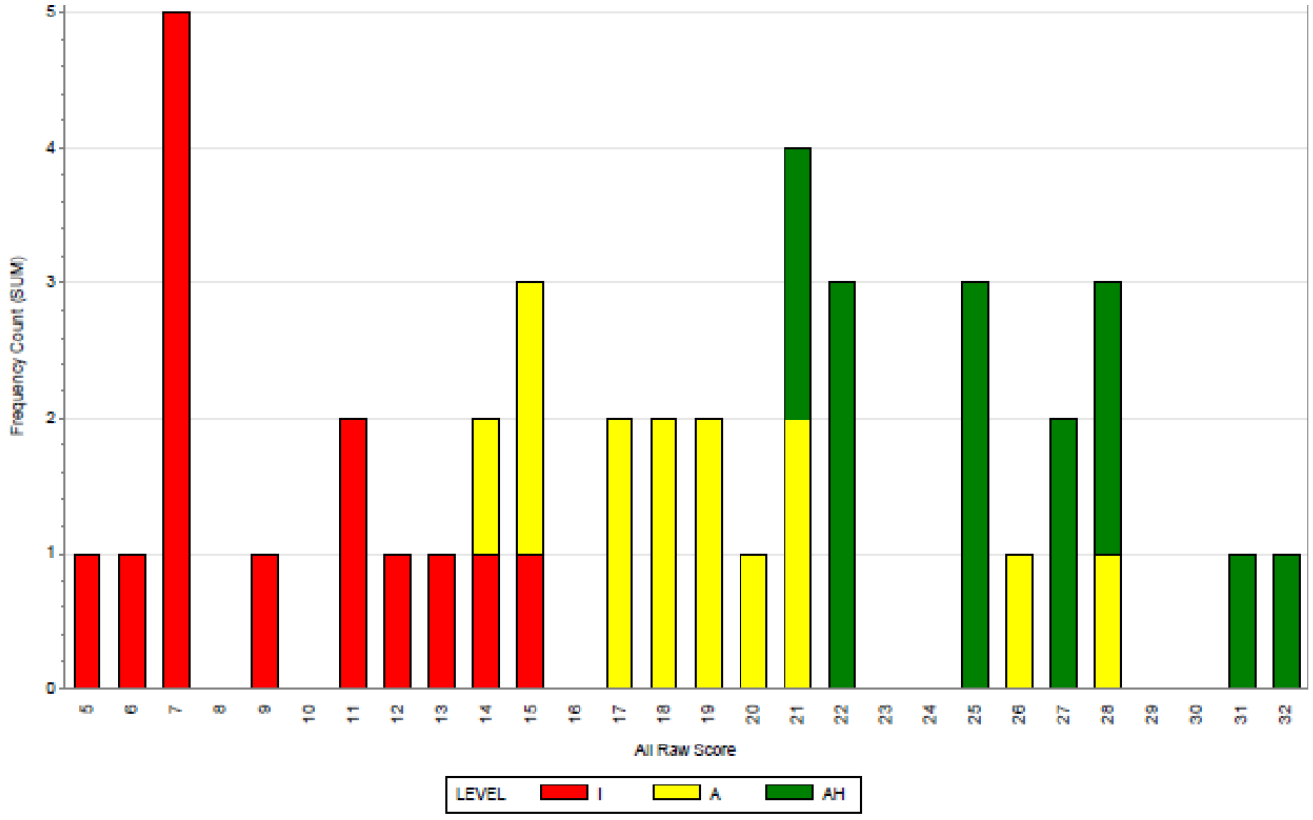
Intermediate



Advanced



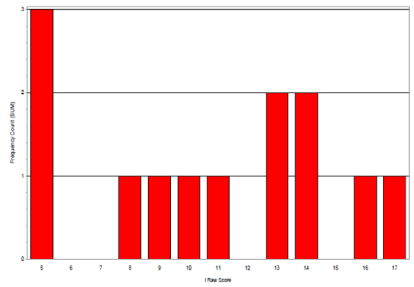
Advanced High



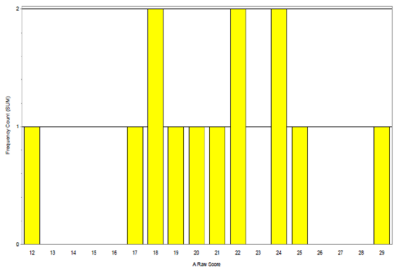
All Three Proficiency Levels Concurrently

Reading Grades 6–7

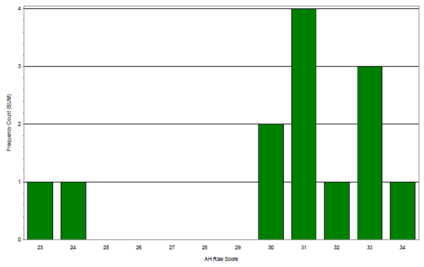
Round 1:



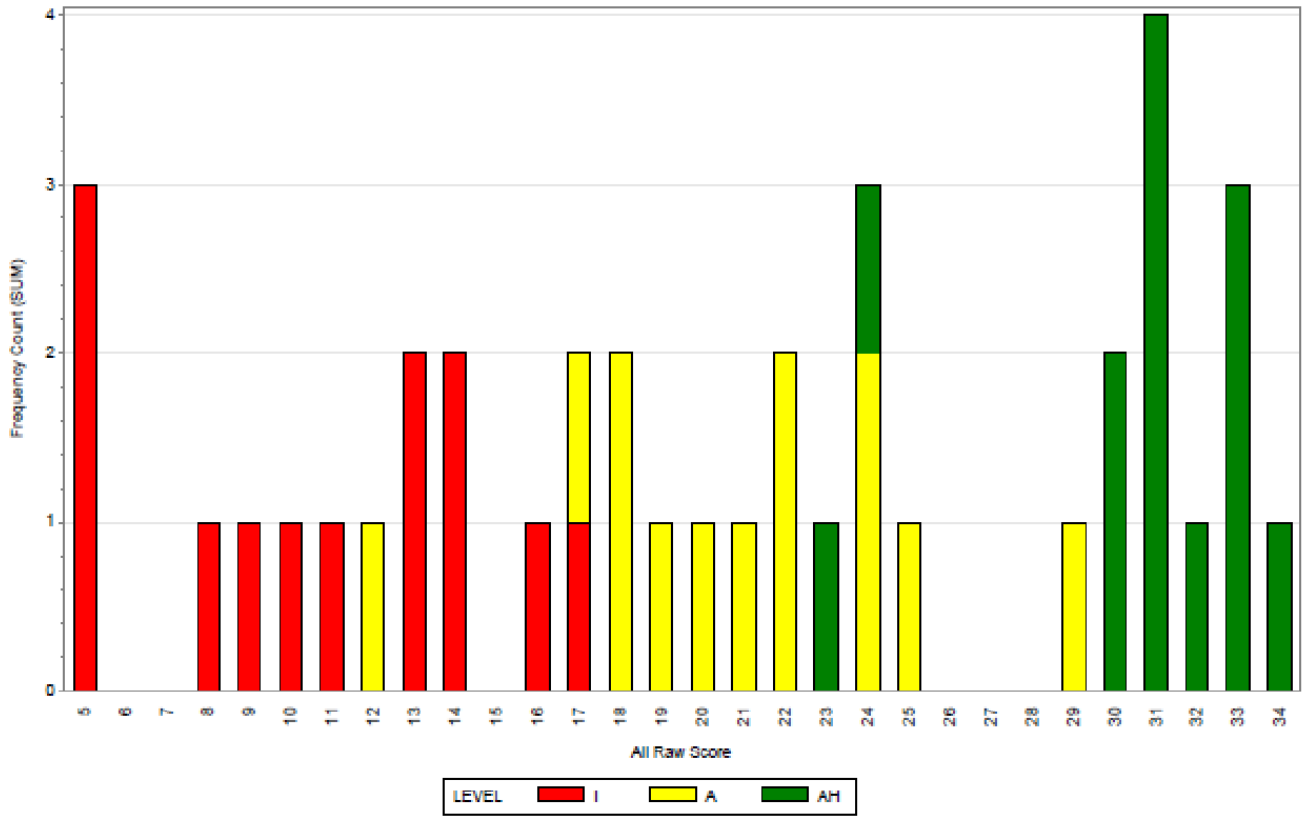
Intermediate



Advanced



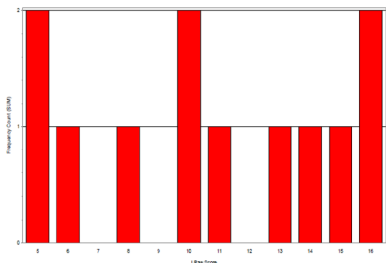
Advanced High



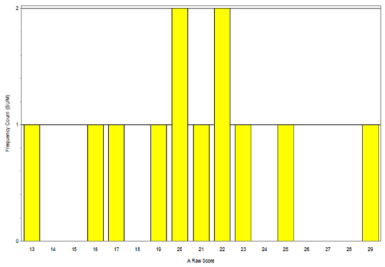
All Three Proficiency Levels Concurrently

Reading Grades 6–7

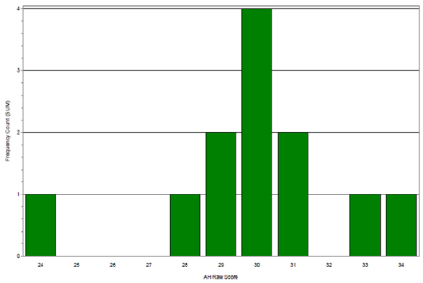
Round 2:



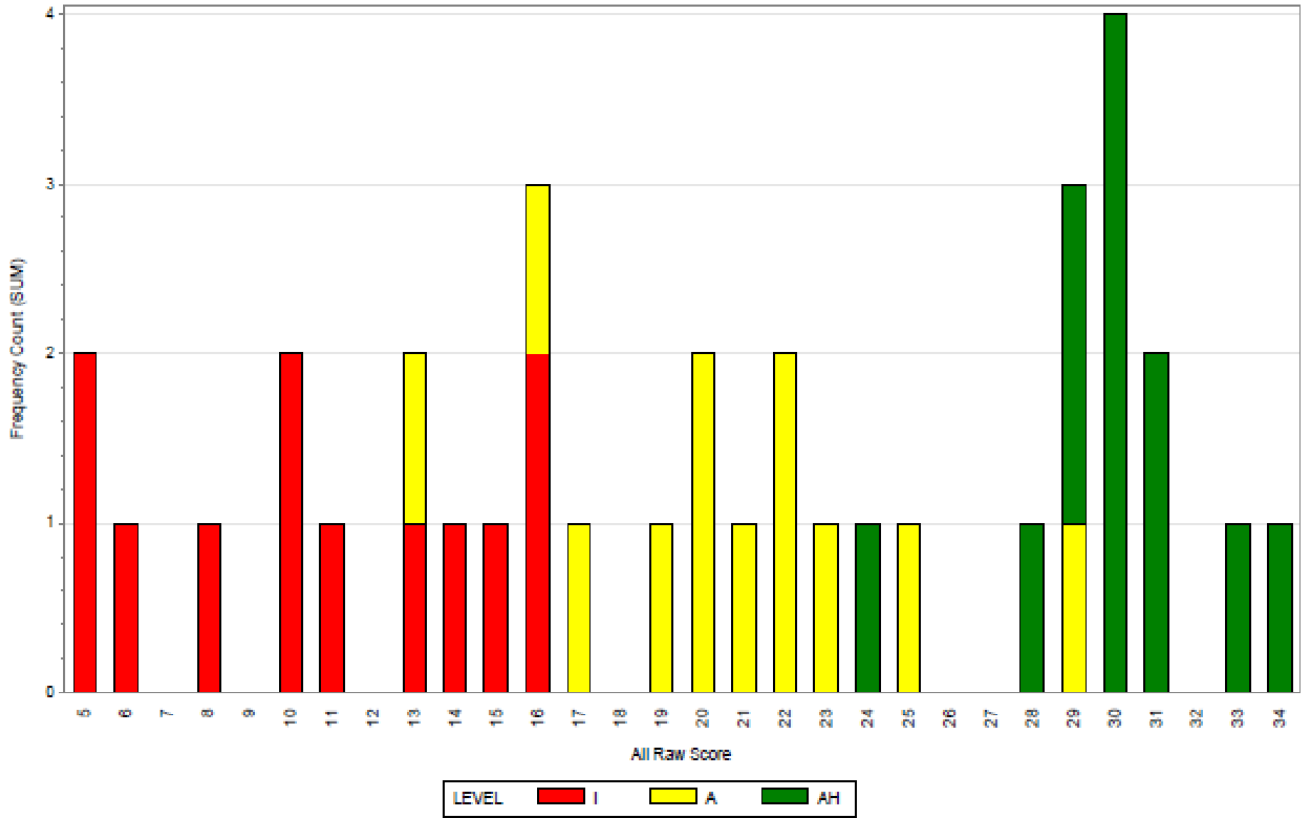
Intermediate



Advanced



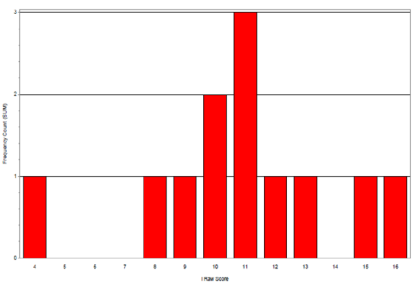
Advanced High



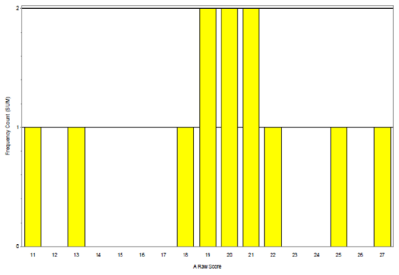
All Three Proficiency Levels Concurrently

Reading Grades 6–7

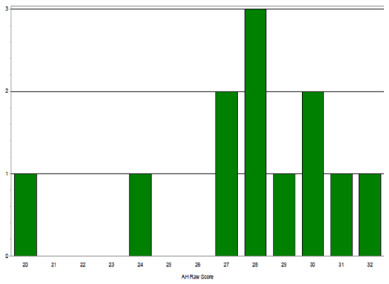
Round 3:



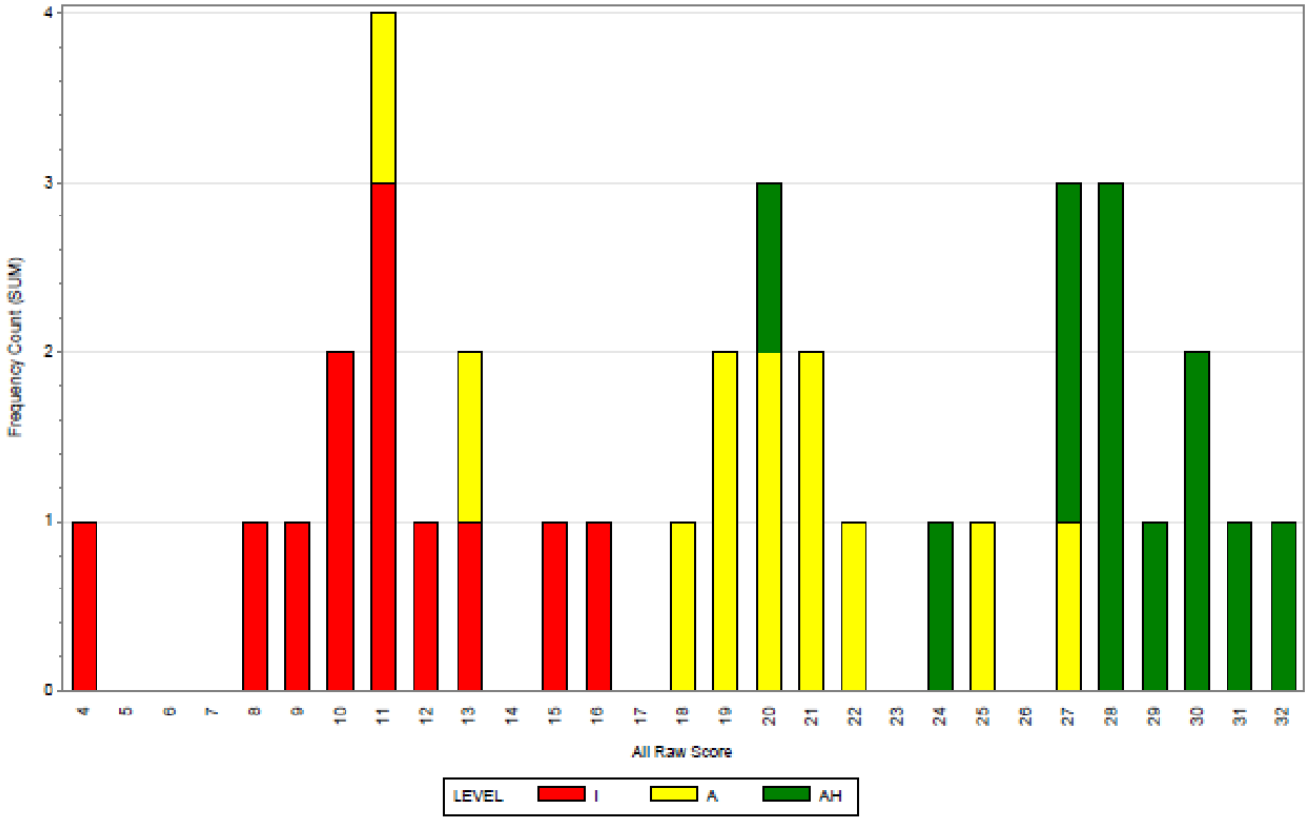
Intermediate



Advanced



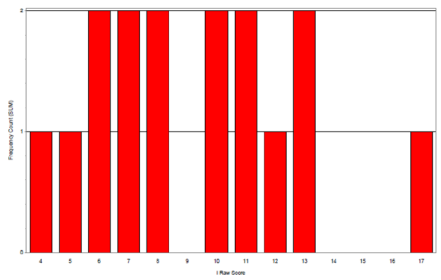
Advanced High



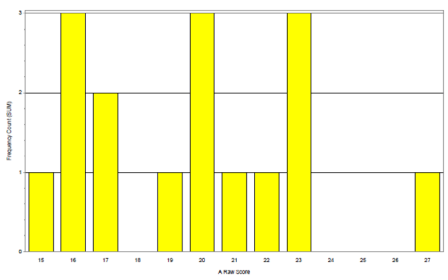
All Three Proficiency Levels Concurrently

Reading Grades 8–9

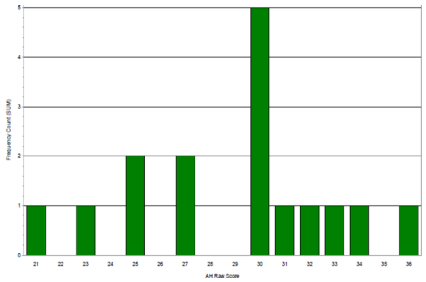
Round 1:



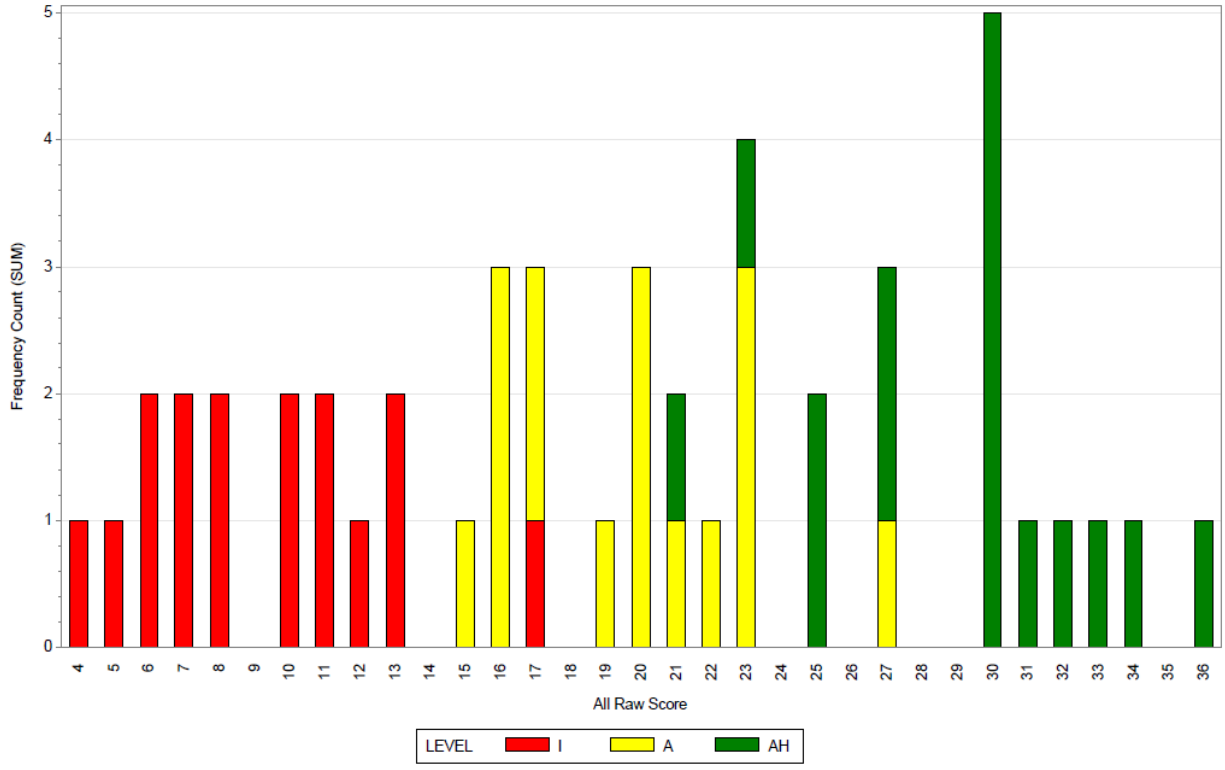
Intermediate



Advanced



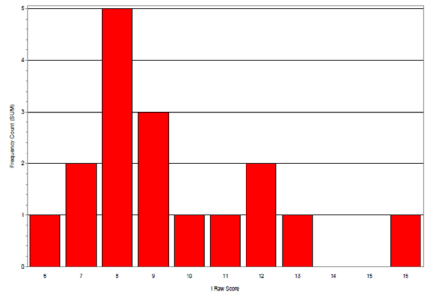
Advanced High



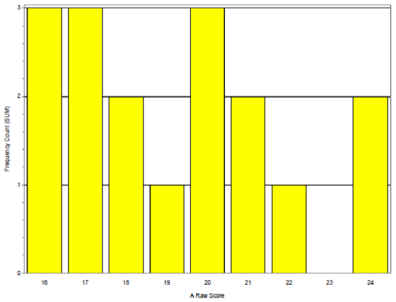
All Three Proficiency Levels Concurrently

Reading Grades 8–9

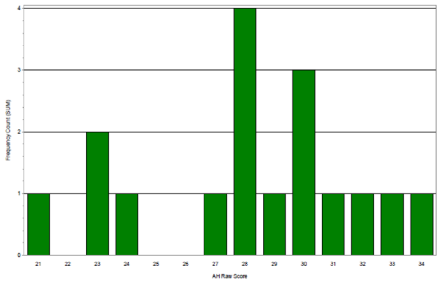
Round 2:



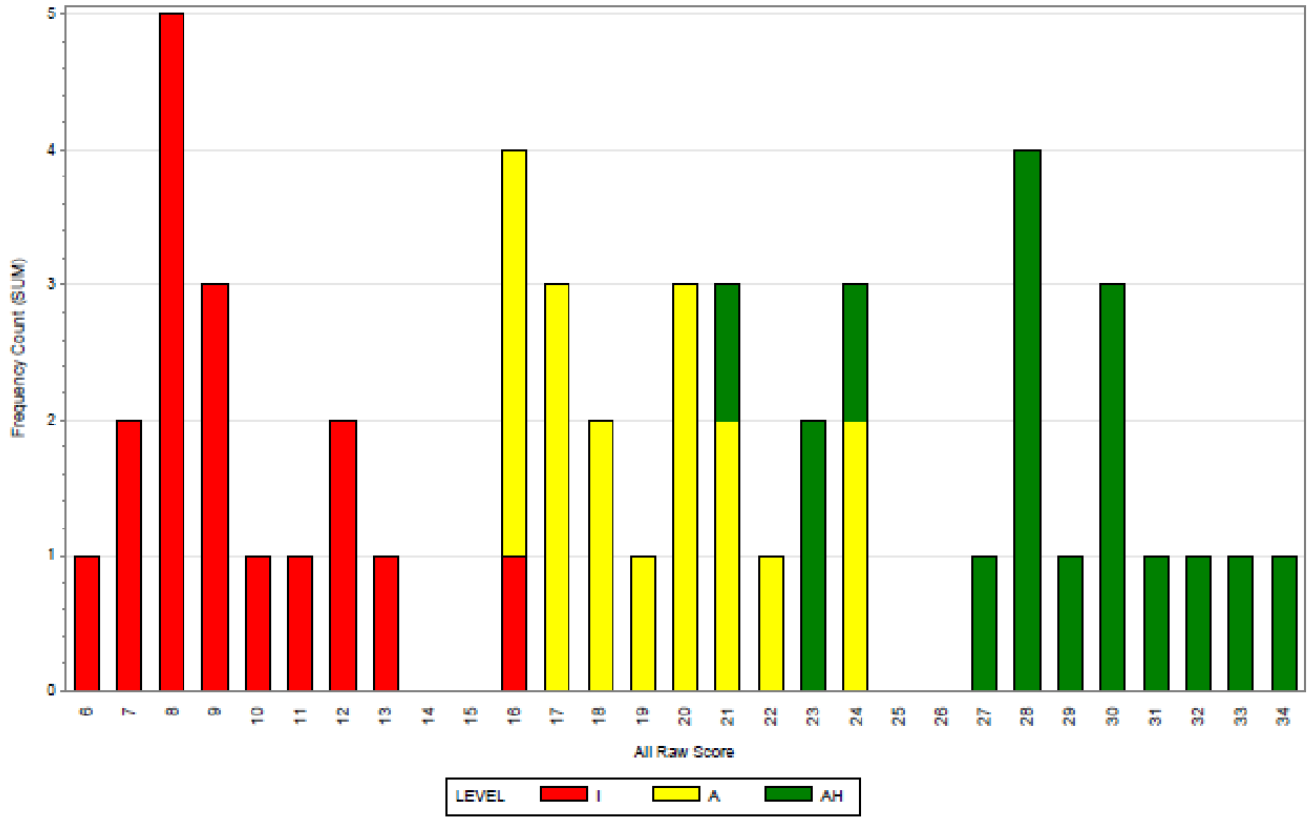
Intermediate



Advanced



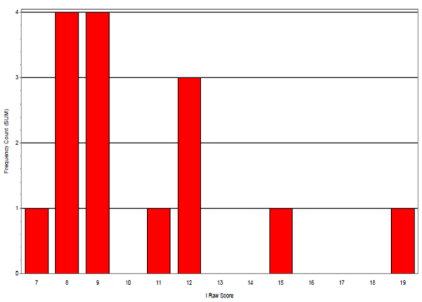
Advanced High



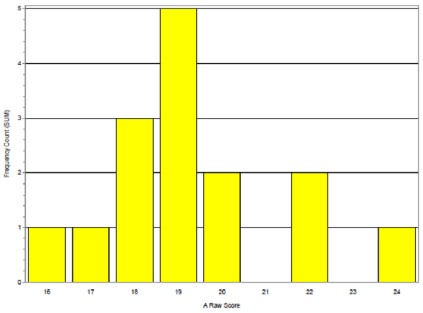
All Three Proficiency Levels Concurrently

Reading Grades 8–9

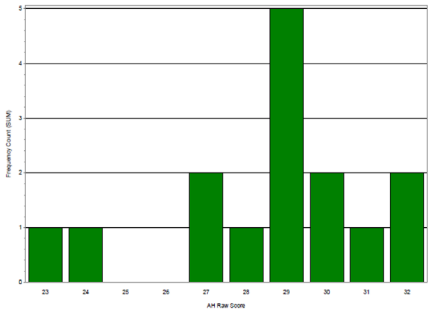
Round 3:



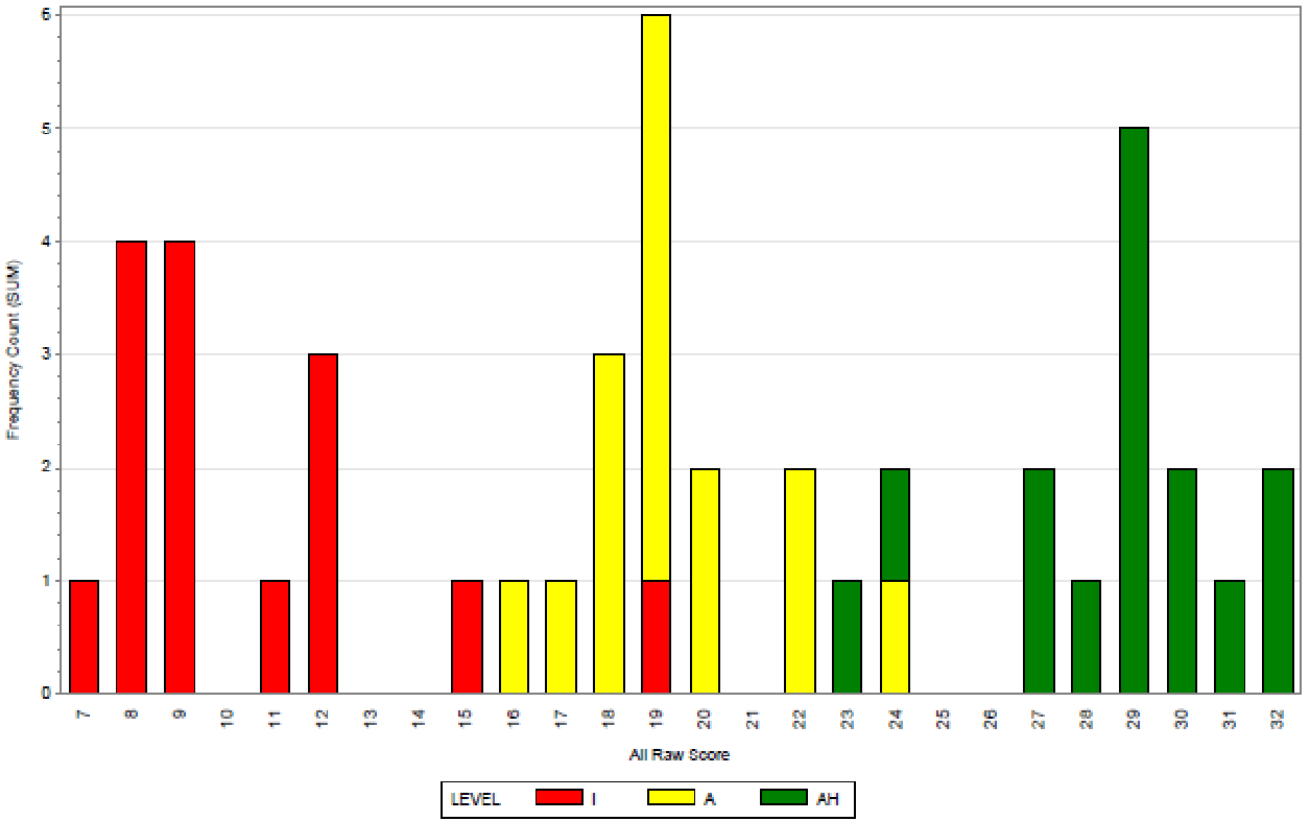
Intermediate



Advanced



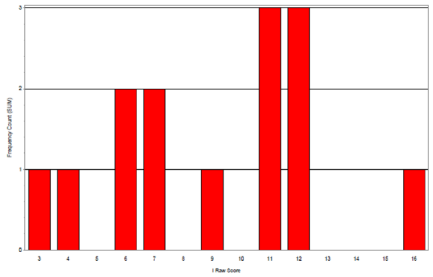
Advanced High



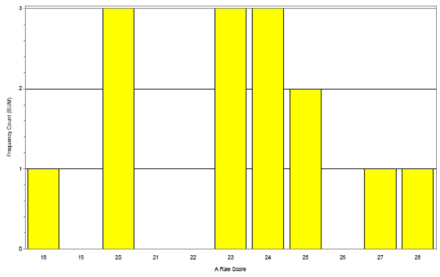
All Three Proficiency Levels Concurrently

Reading Grades 10–12

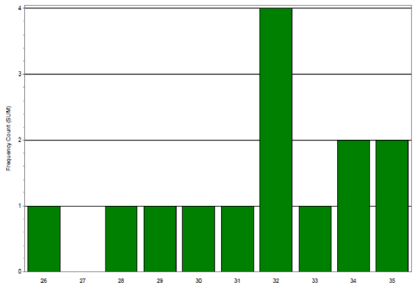
Round 1:



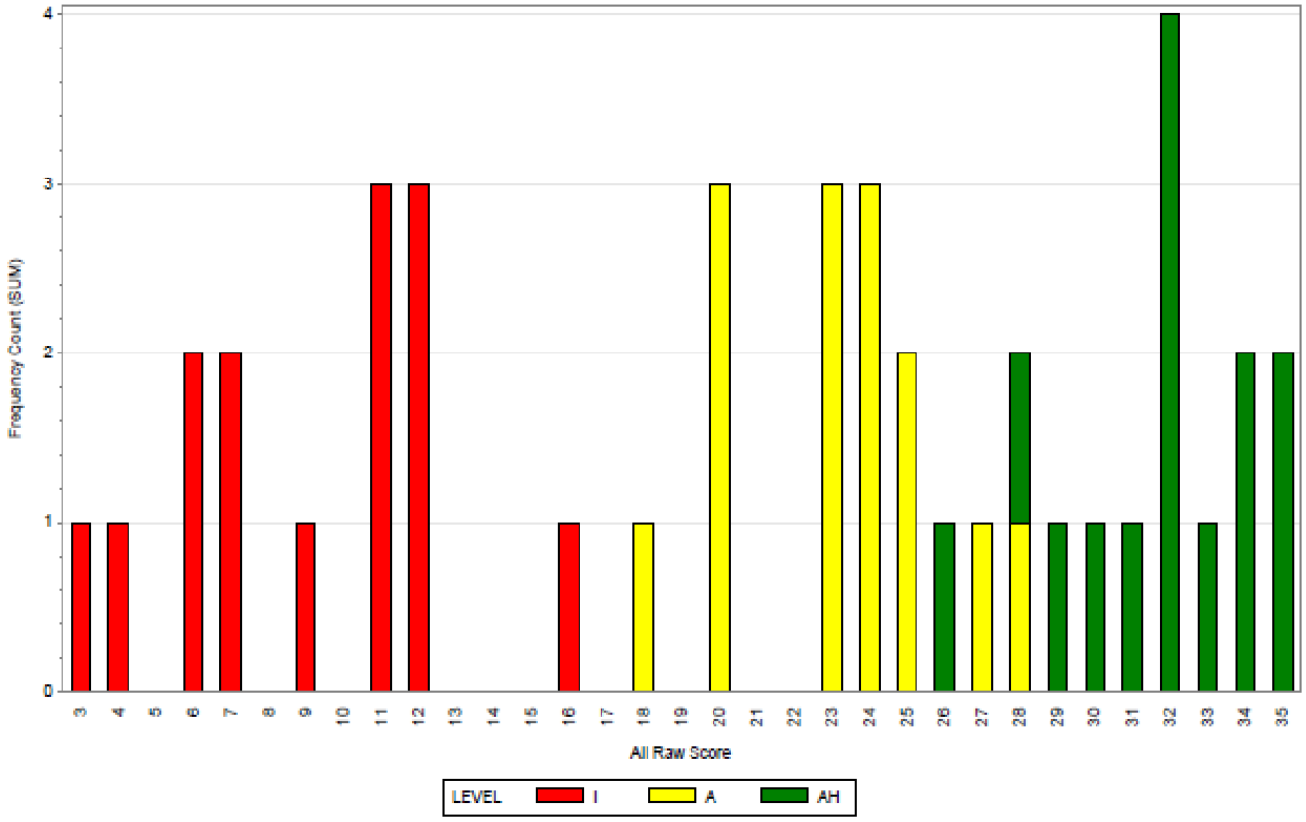
Intermediate



Advanced



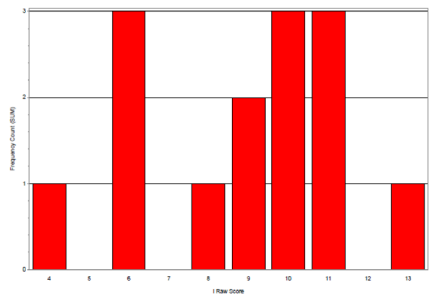
Advanced High



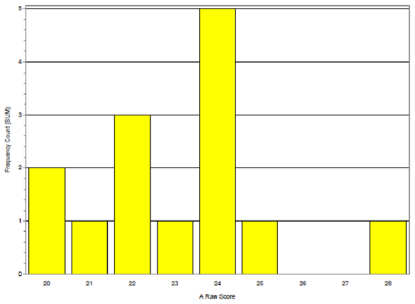
All Three Proficiency Levels Concurrently

Reading Grades 10–12

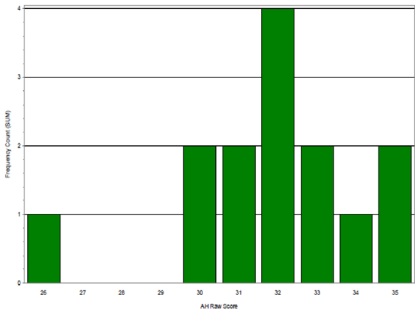
Round 2:



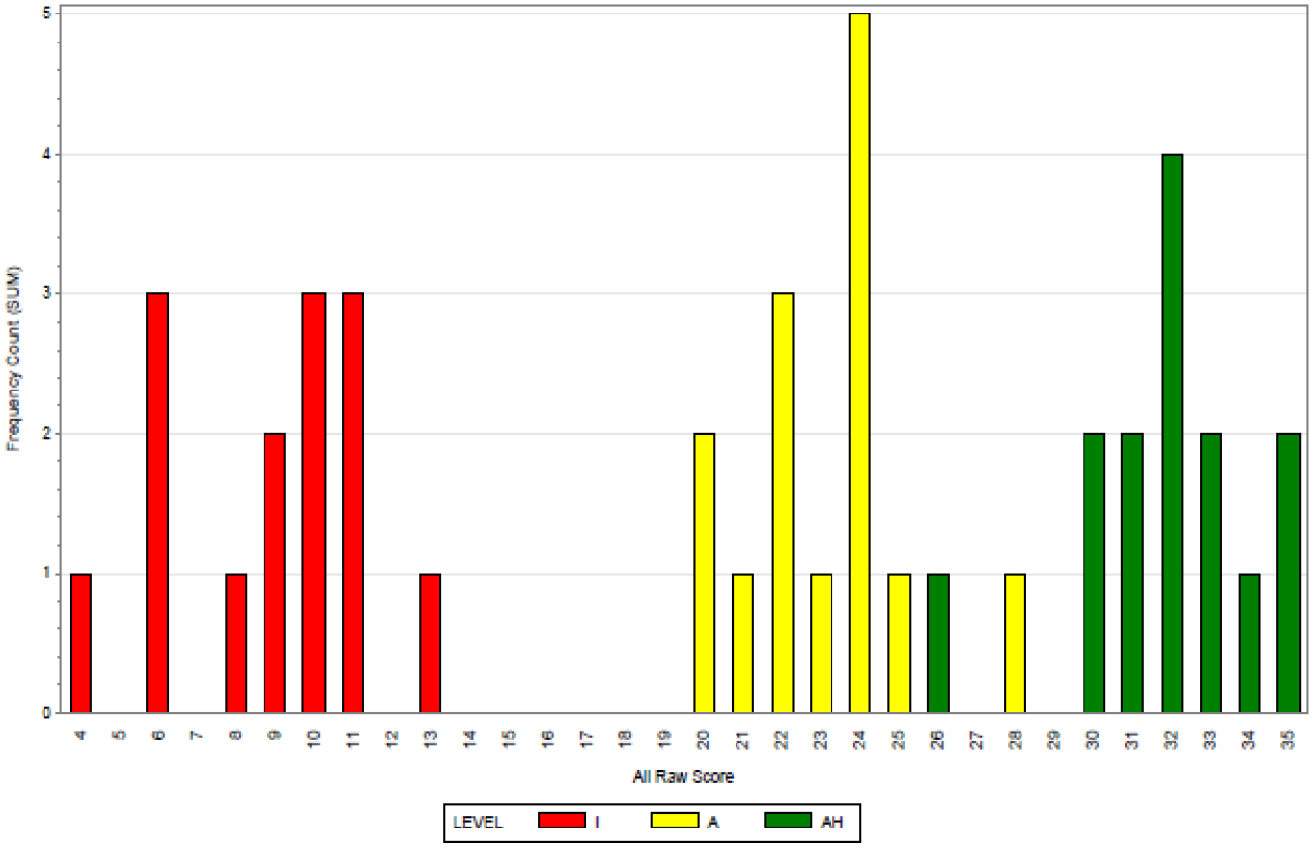
Intermediate



Advanced



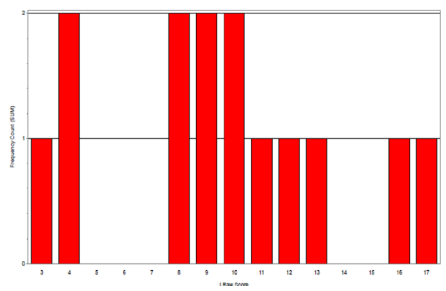
Advanced High



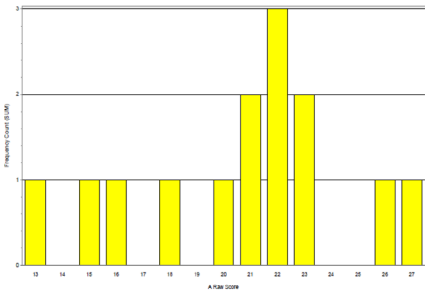
All Three Proficiency Levels Concurrently

Reading Grades 10–12

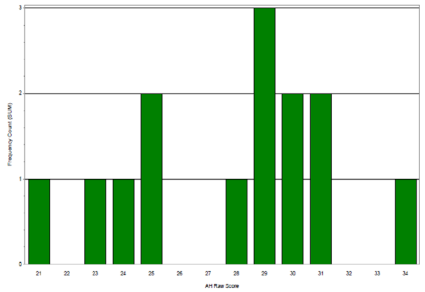
Round 3:



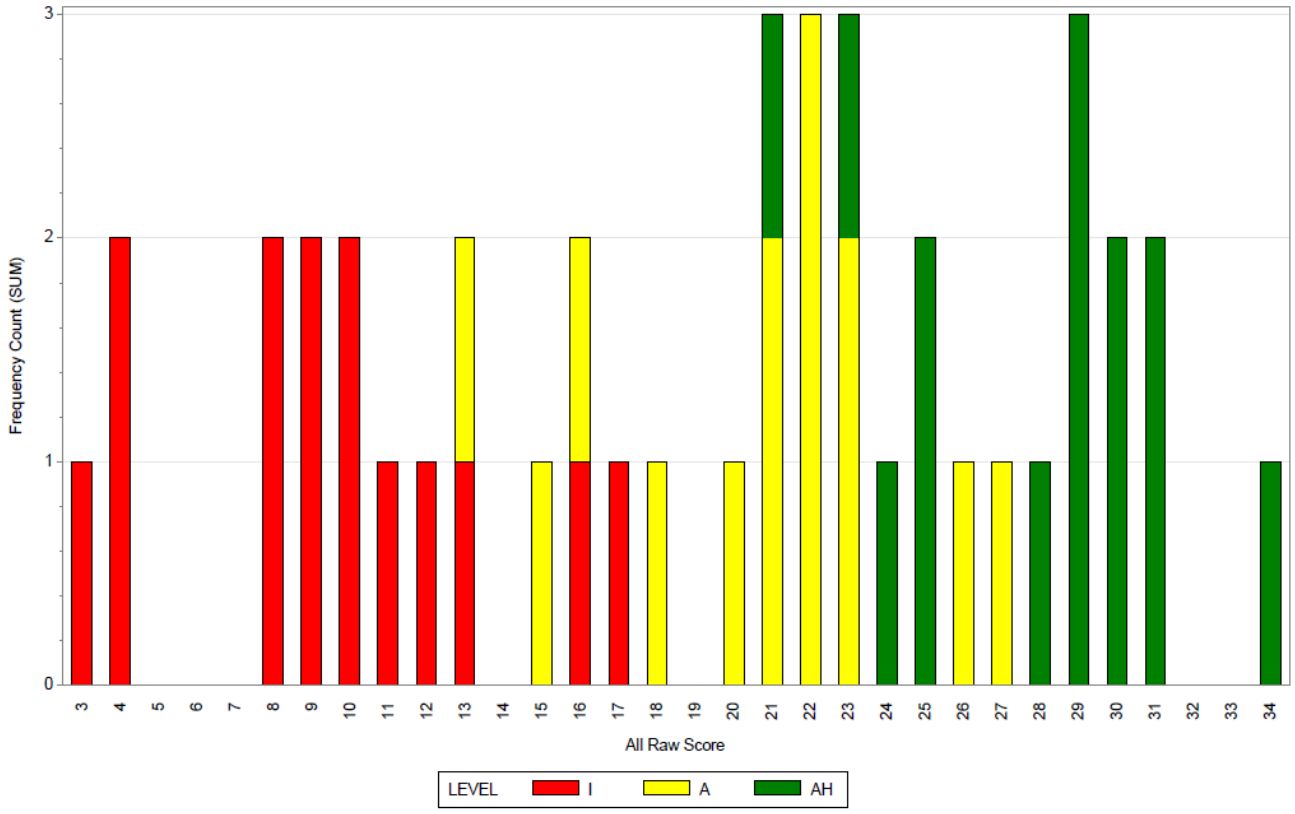
Intermediate



Advanced



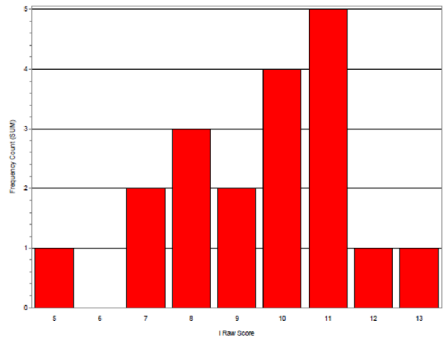
Advanced High



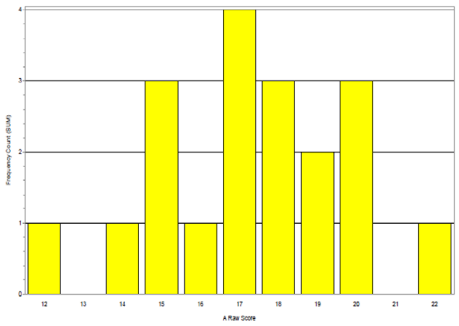
All Three Proficiency Levels Concurrently

Listening Grades 2–3

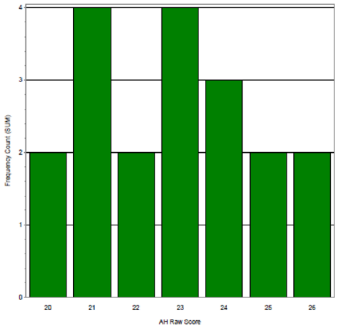
Round 1:



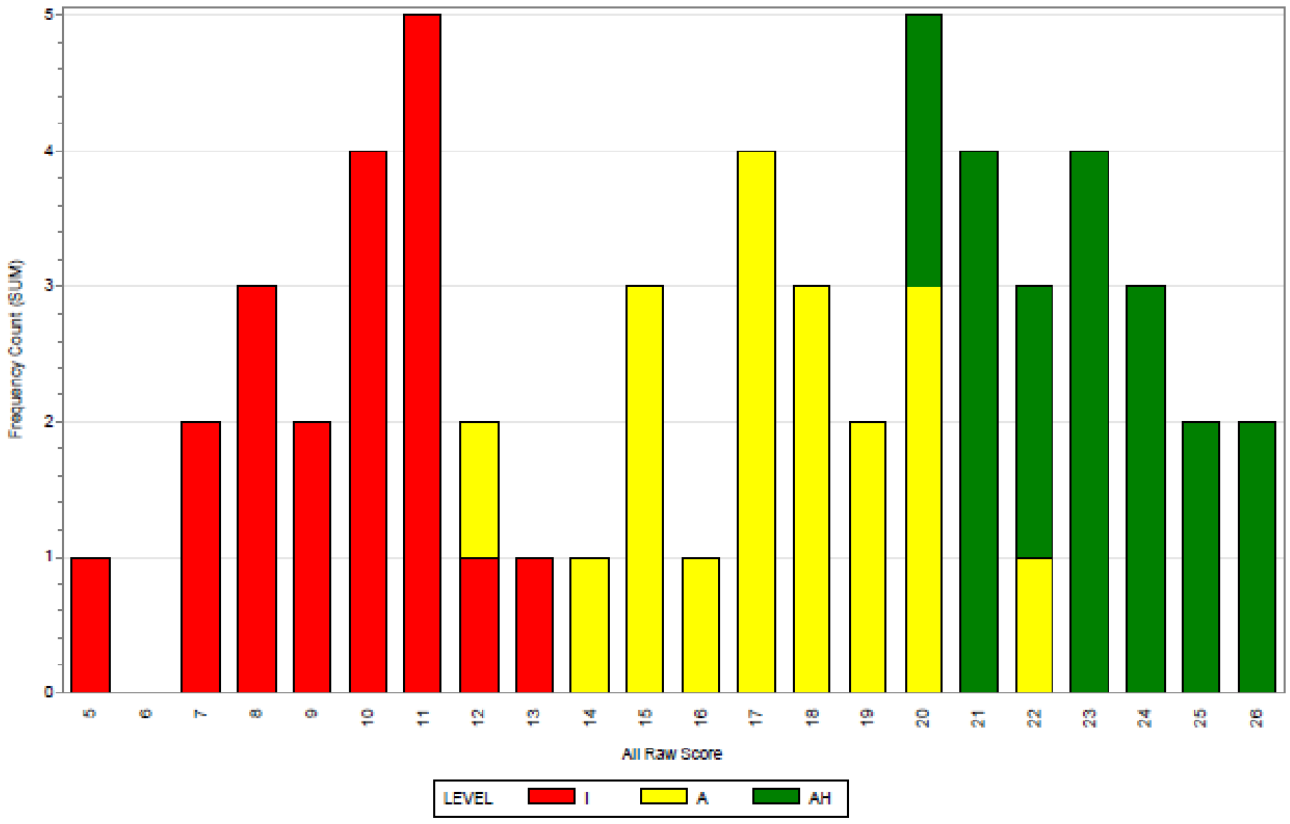
Intermediate



Advanced



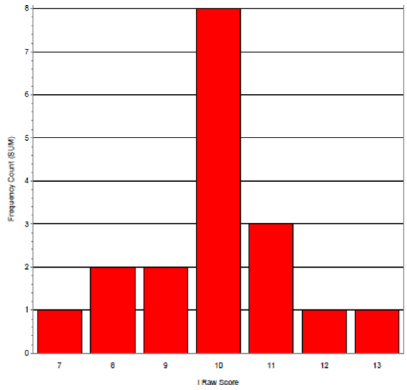
Advanced High



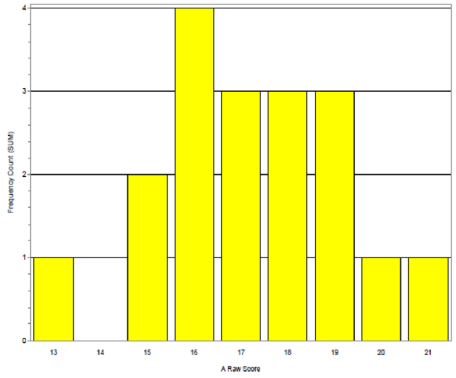
All Three Proficiency Levels Concurrently

Listening Grades 2–3

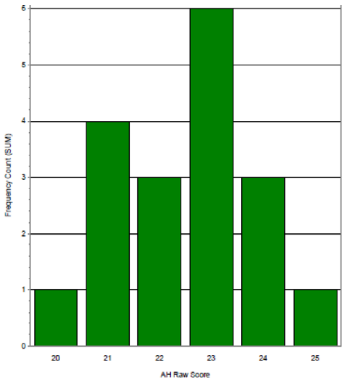
Round 2:



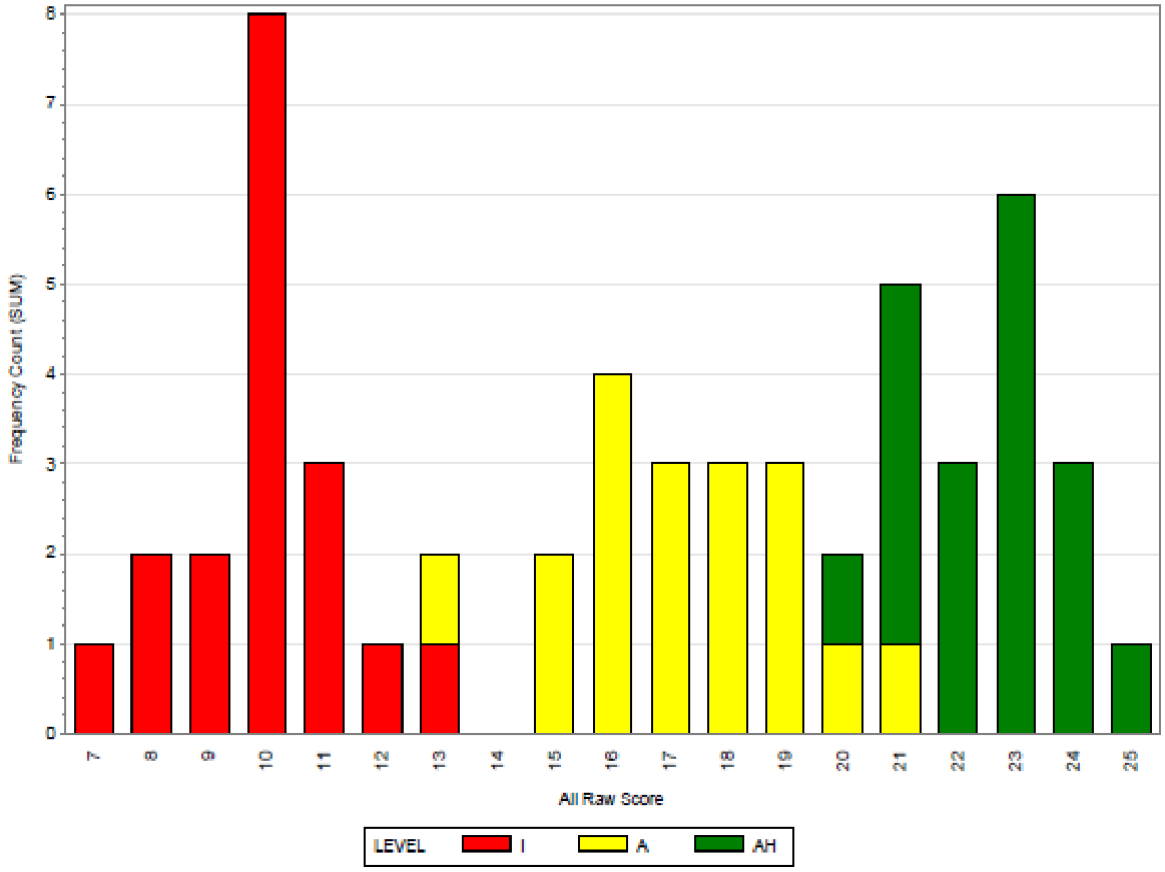
Intermediate



Advanced



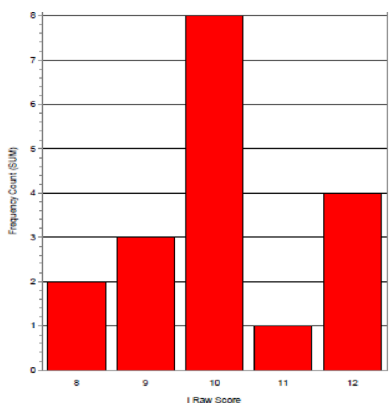
Advanced High



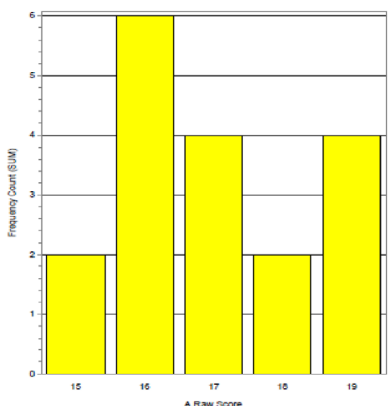
All Three Proficiency Levels Concurrently

Listening Grades 2–3

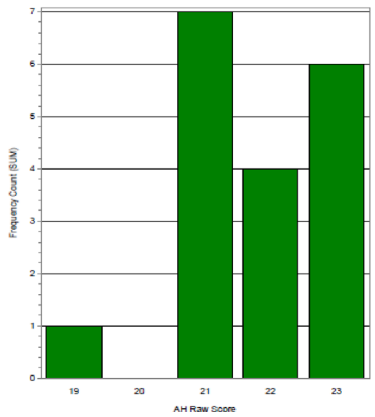
Round 3:



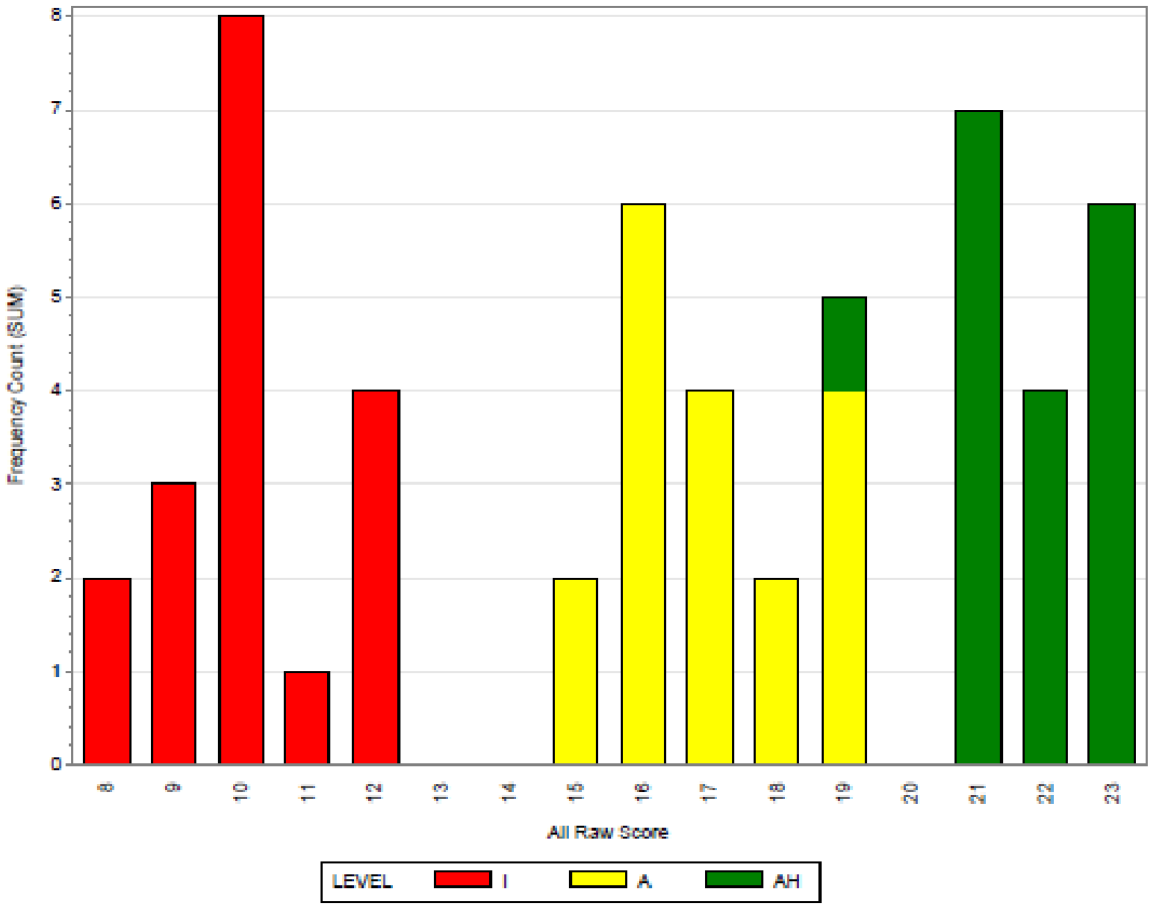
Intermediate



Advanced



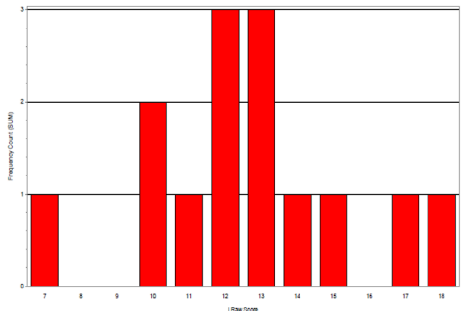
Advanced High



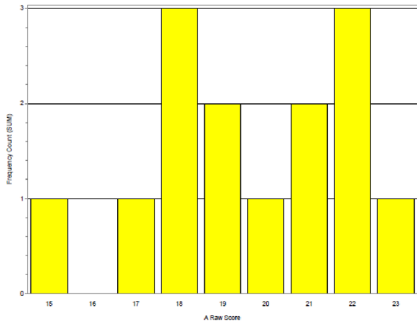
All Three Proficiency Levels Concurrently

Listening Grades 4–5

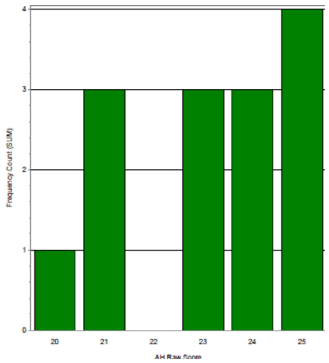
Round 1:



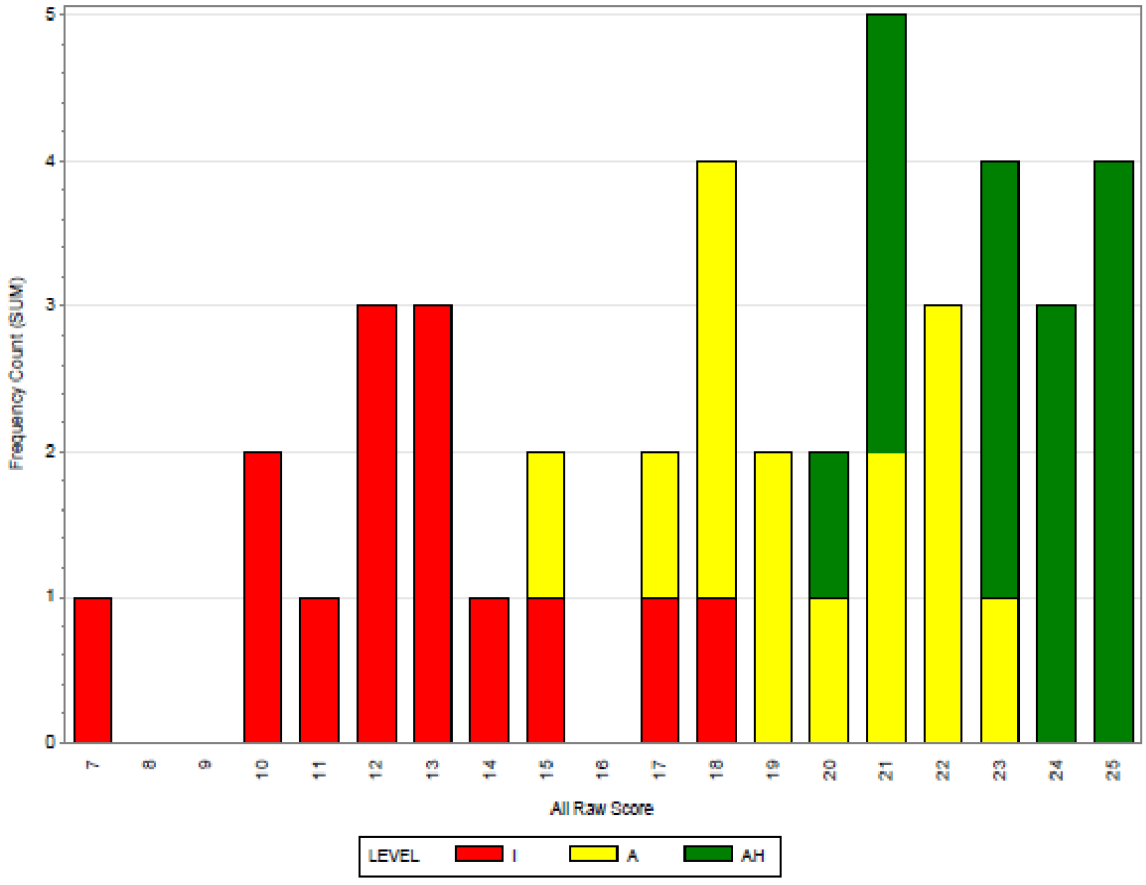
Intermediate



Advanced



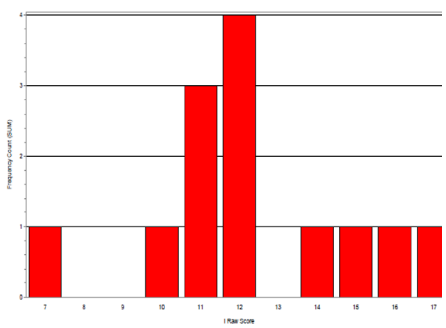
Advanced High



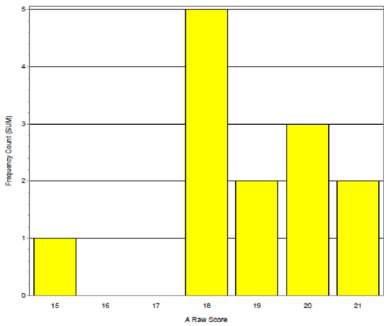
All Three Proficiency Levels Concurrently

Listening Grades 4–5

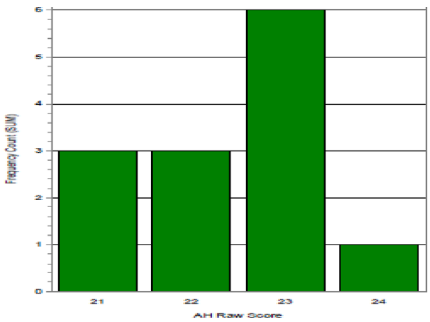
Round 2:



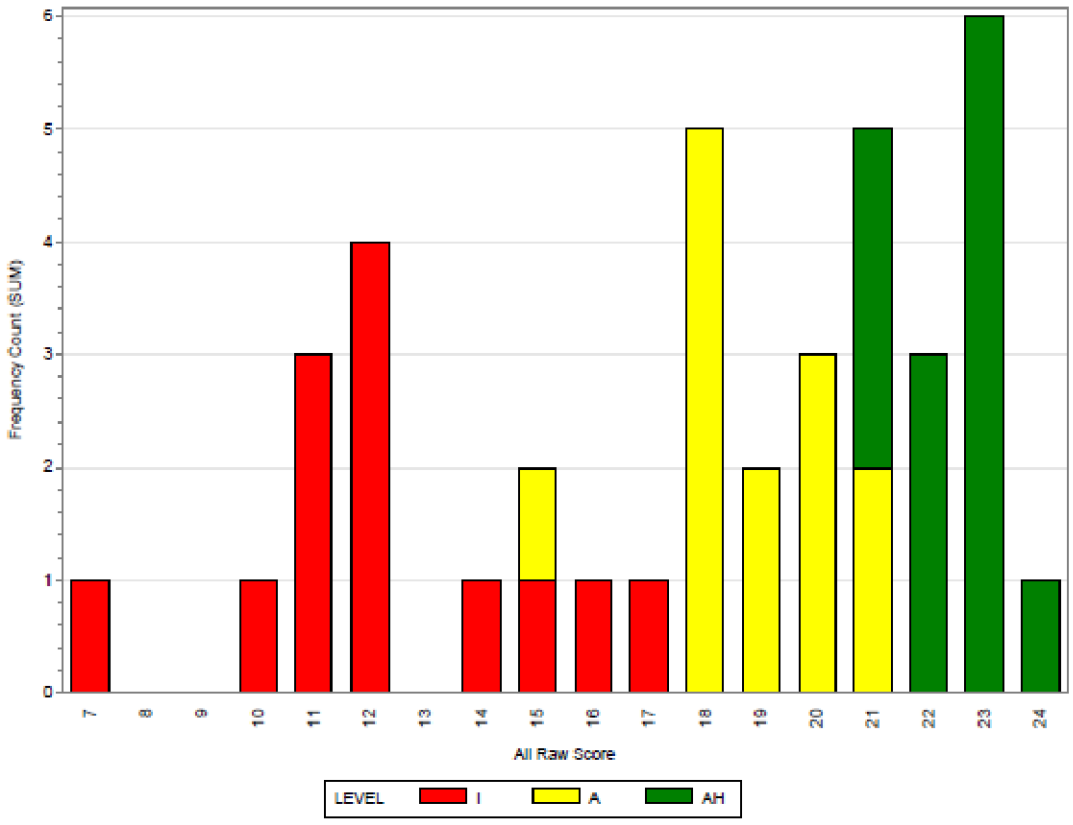
Intermediate



Advanced



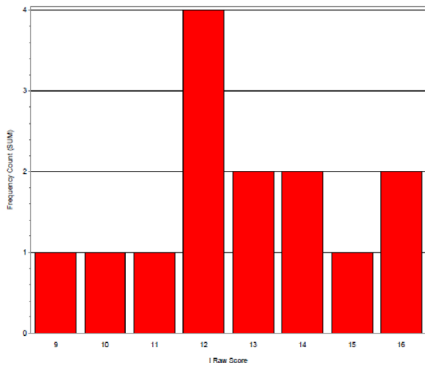
Advanced High



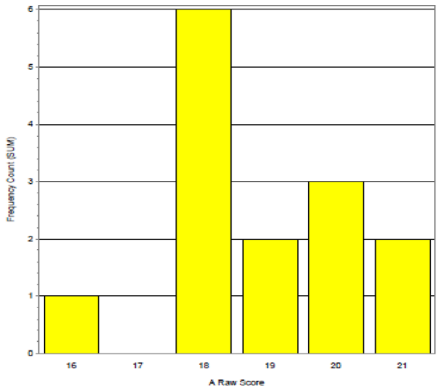
All Three Proficiency Levels Concurrently

Listening Grades 4–5

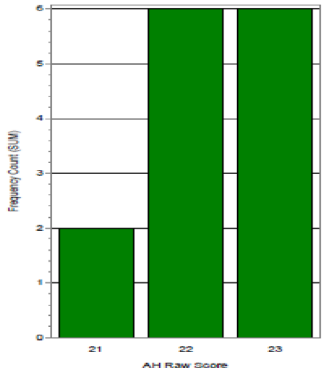
Round 3:



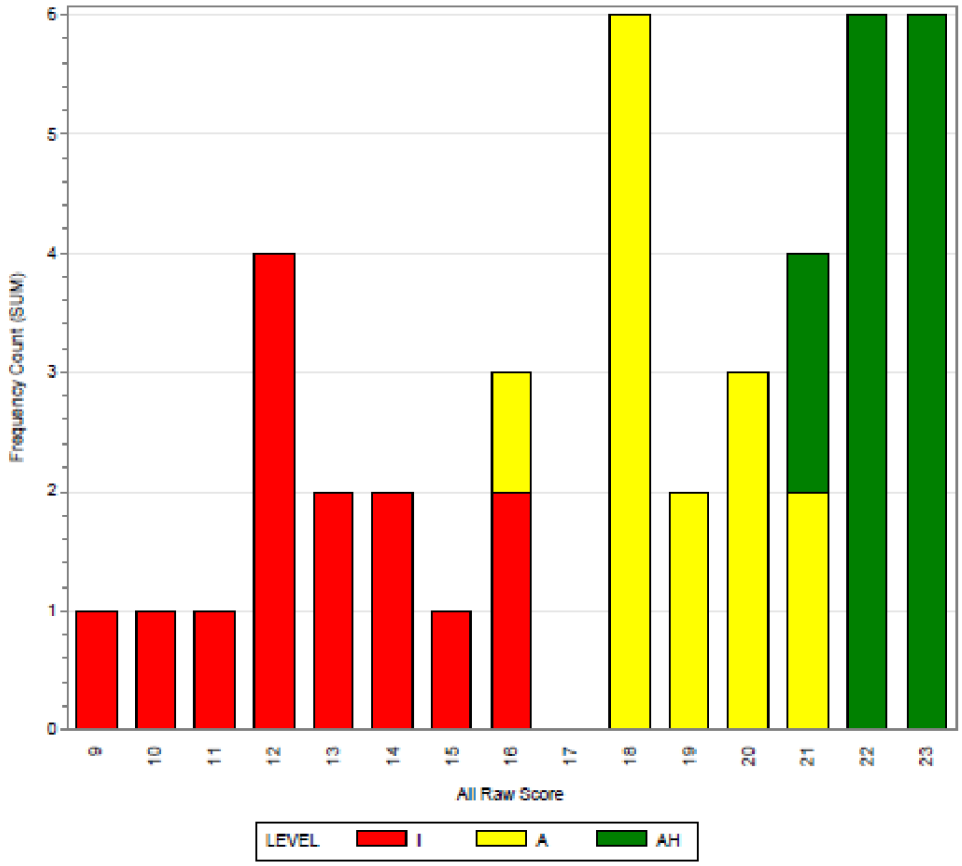
Intermediate



Advanced



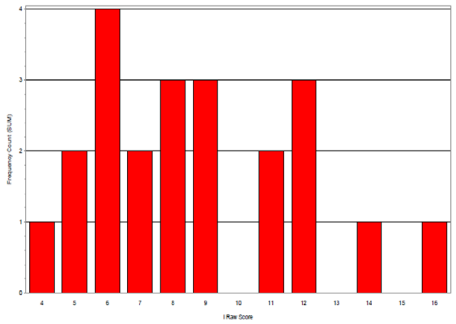
Advanced High



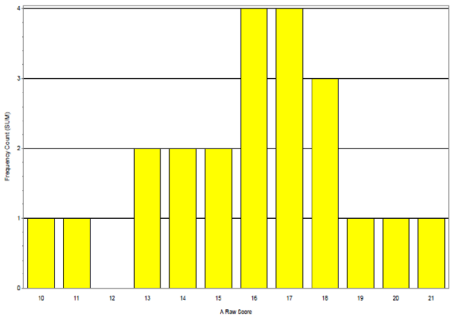
All Three Proficiency Levels Concurrently

Listening Grades 6–8

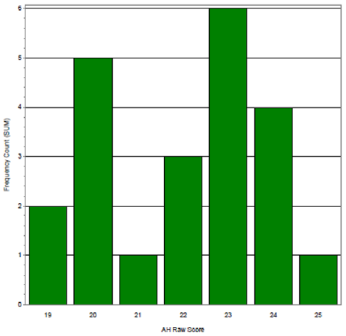
Round 1:



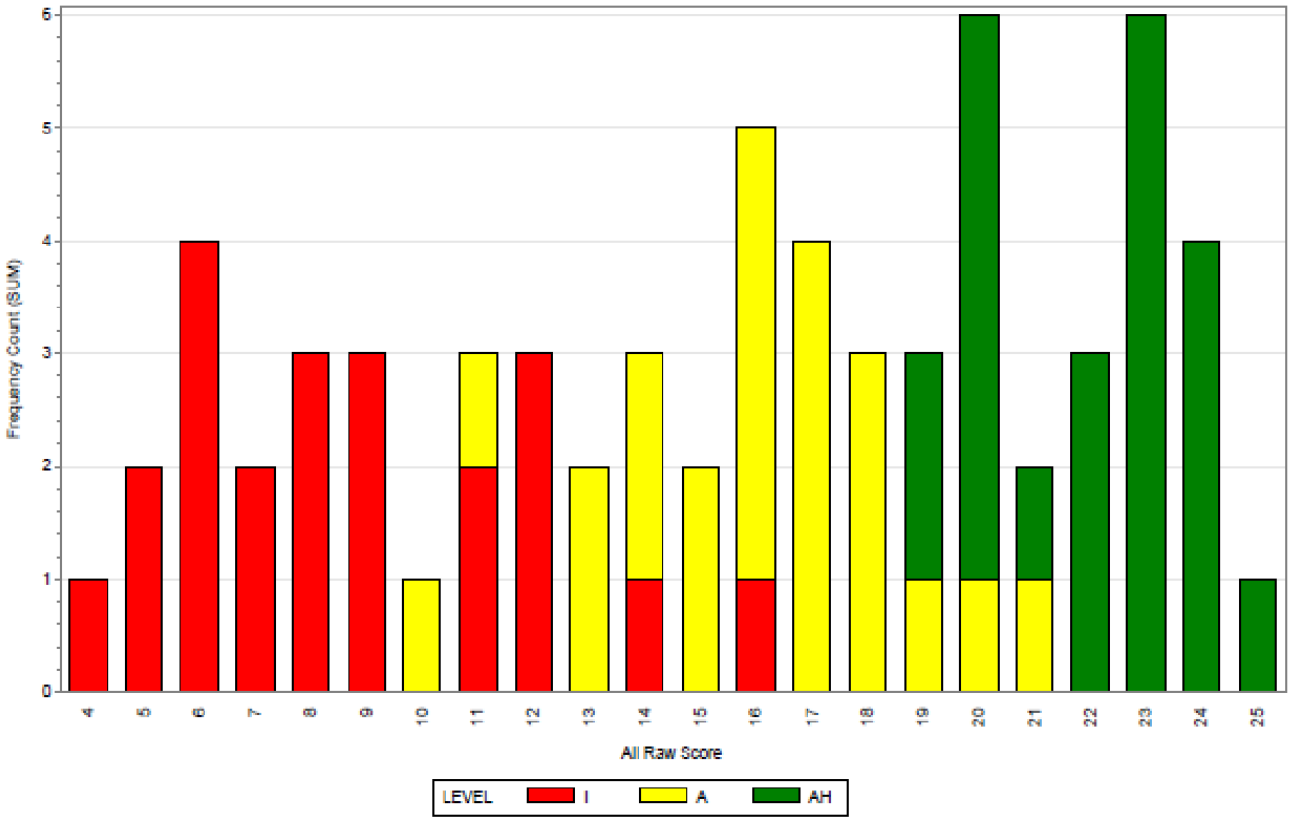
Intermediate



Advanced



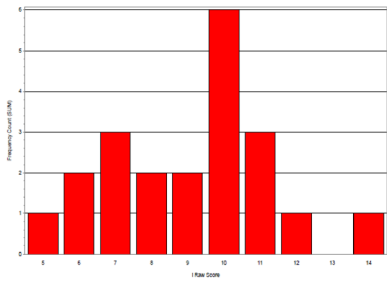
Advanced High



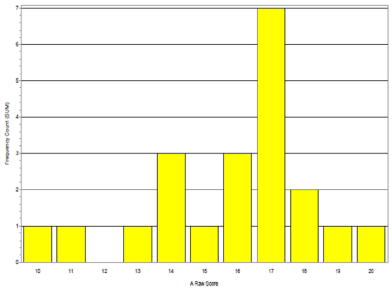
All Three Proficiency Levels Concurrently

Listening Grades 6–8

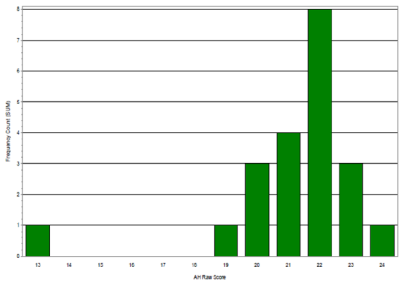
Round 2:



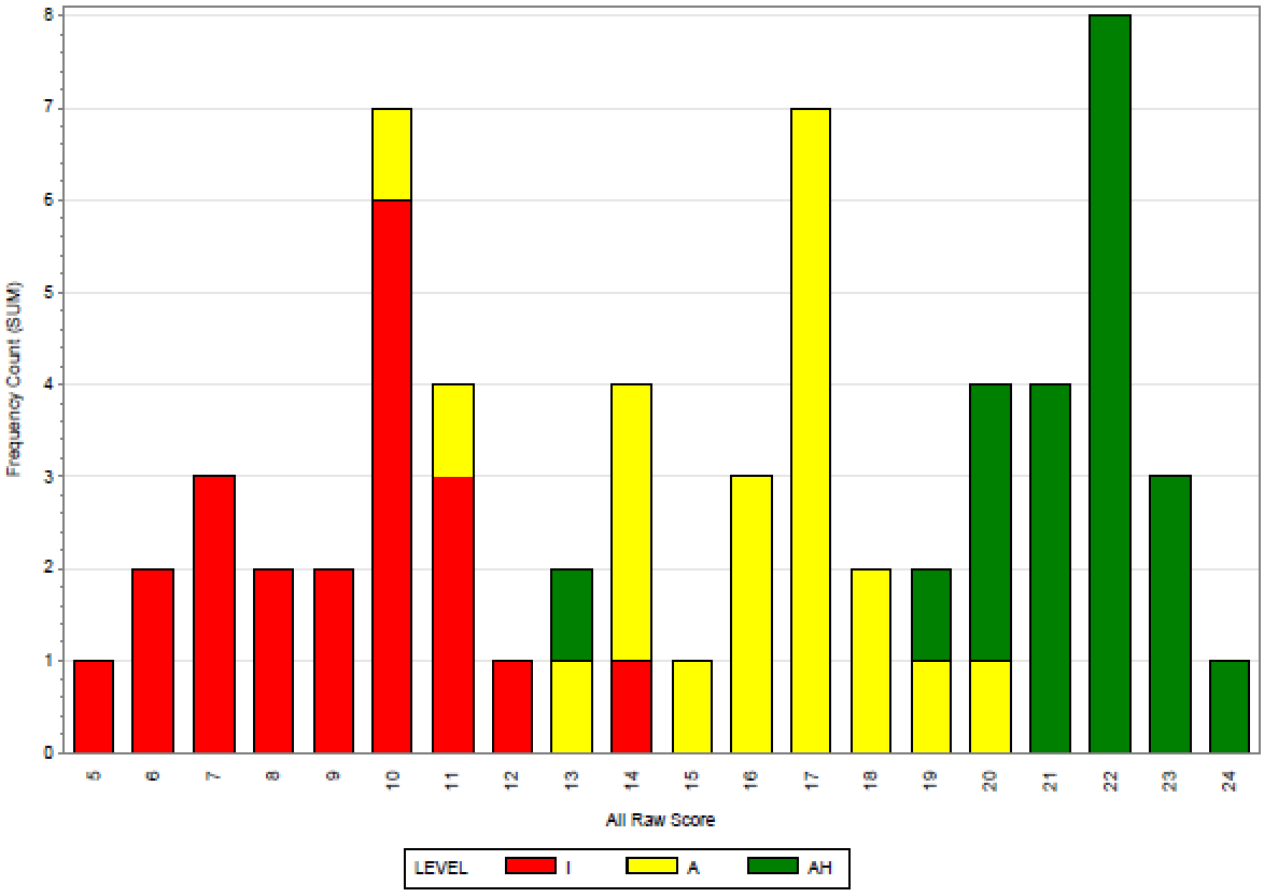
Intermediate



Advanced



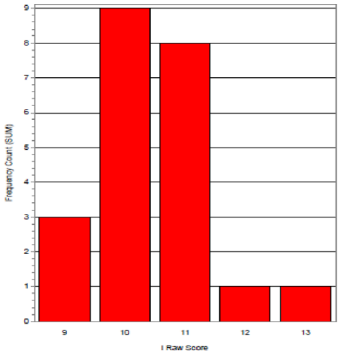
Advanced High



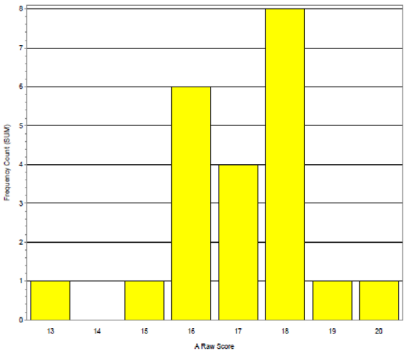
All Three Proficiency Levels Concurrently

Listening Grades 6–8

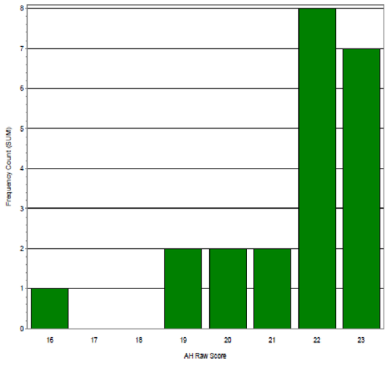
Round 3:



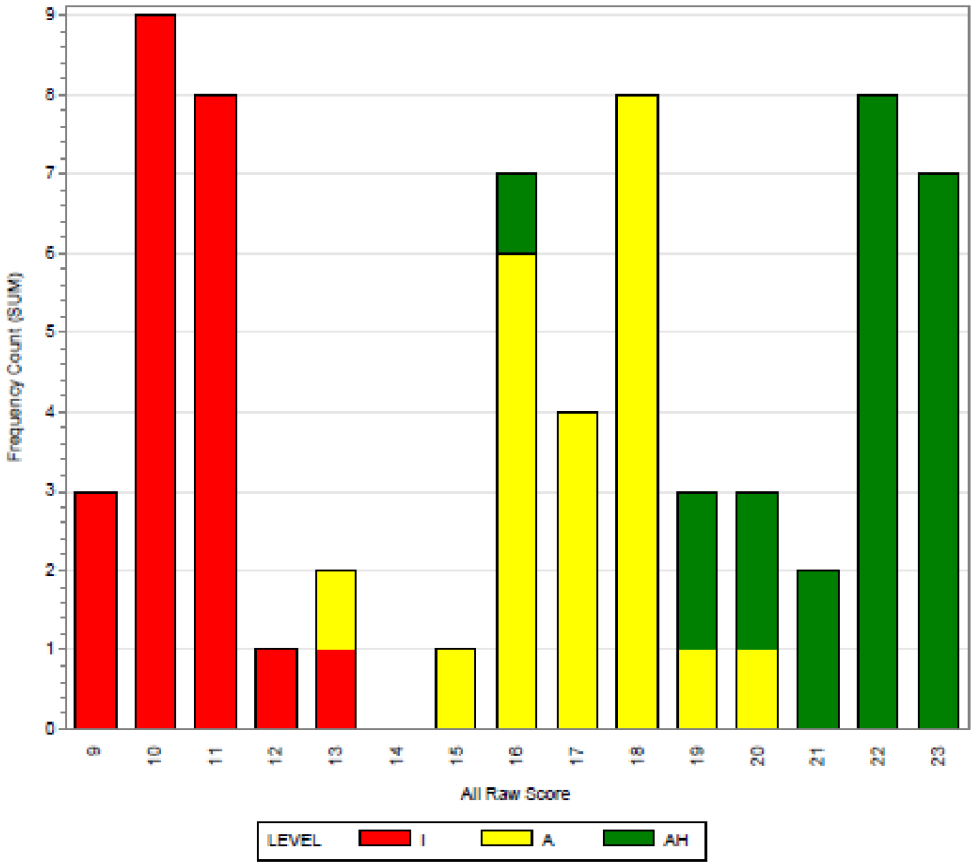
Intermediate



Advanced



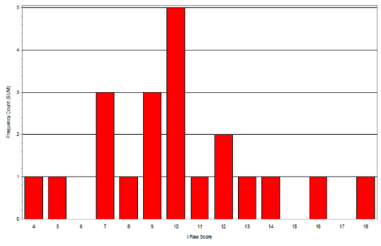
Advanced High



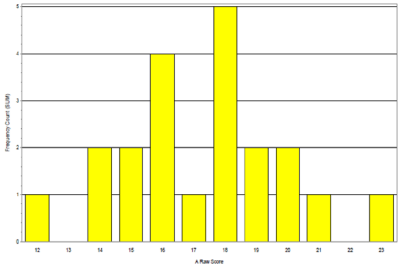
All Three Proficiency Levels Concurrently

Listening Grades 9–12

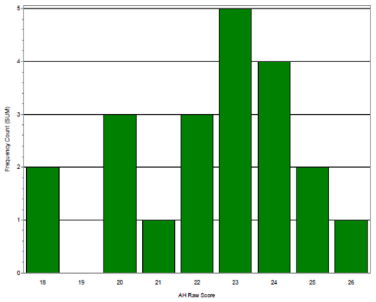
Round 1:



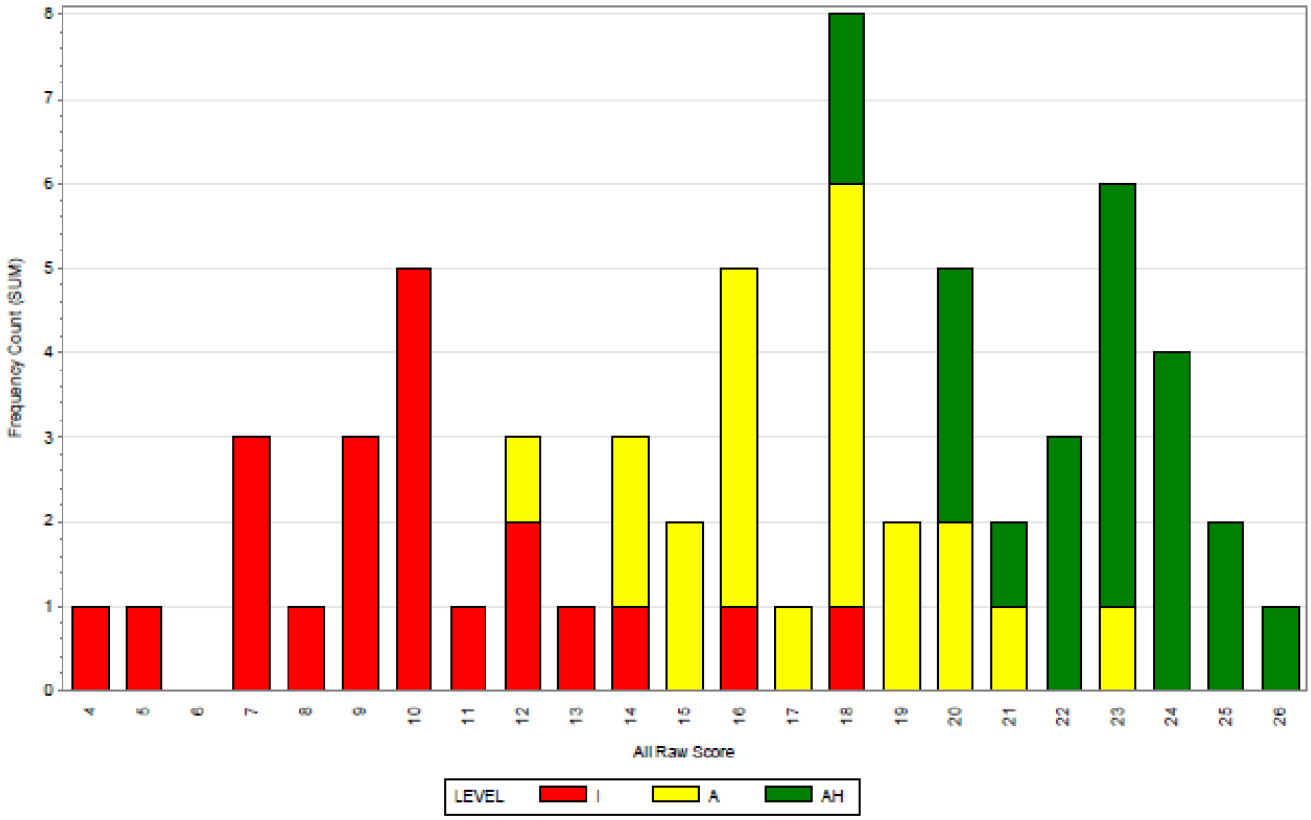
Intermediate



Advanced



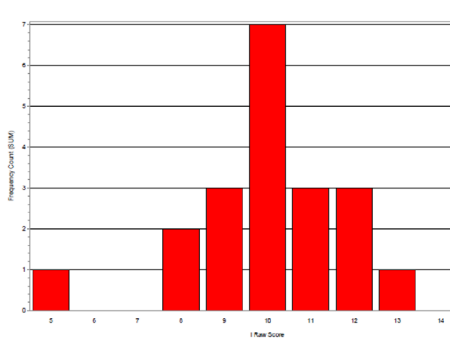
Advanced High



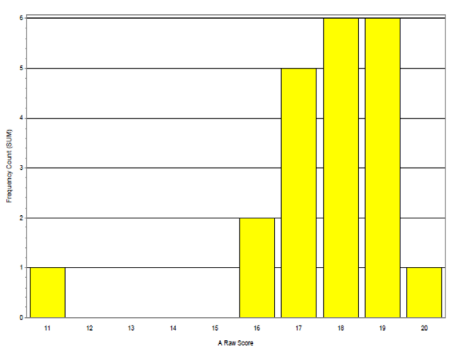
All Three Proficiency Levels Concurrently

Listening Grades 9–12

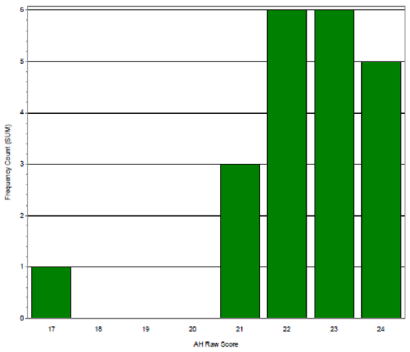
Round 2:



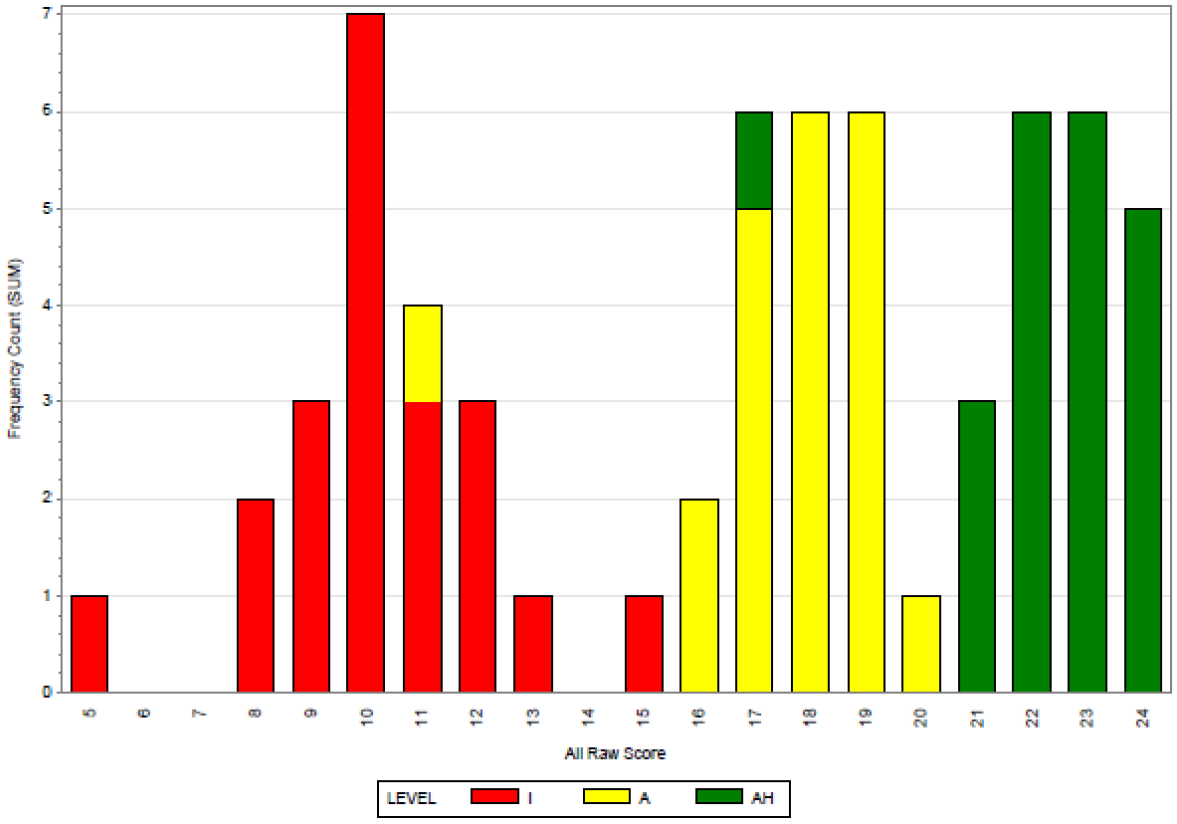
Intermediate



Advanced



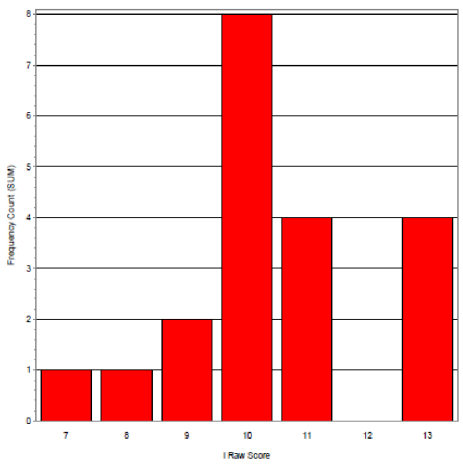
Advanced High



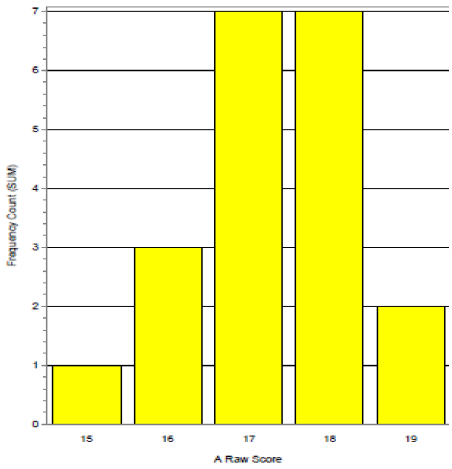
All Three Proficiency Levels Concurrently

Listening Grades 9–12

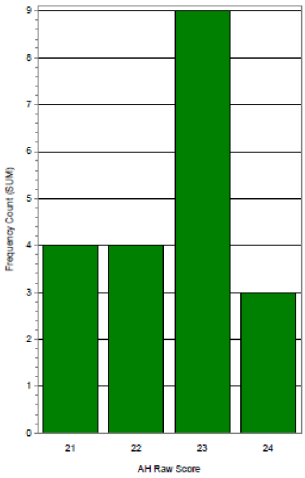
Round 3:



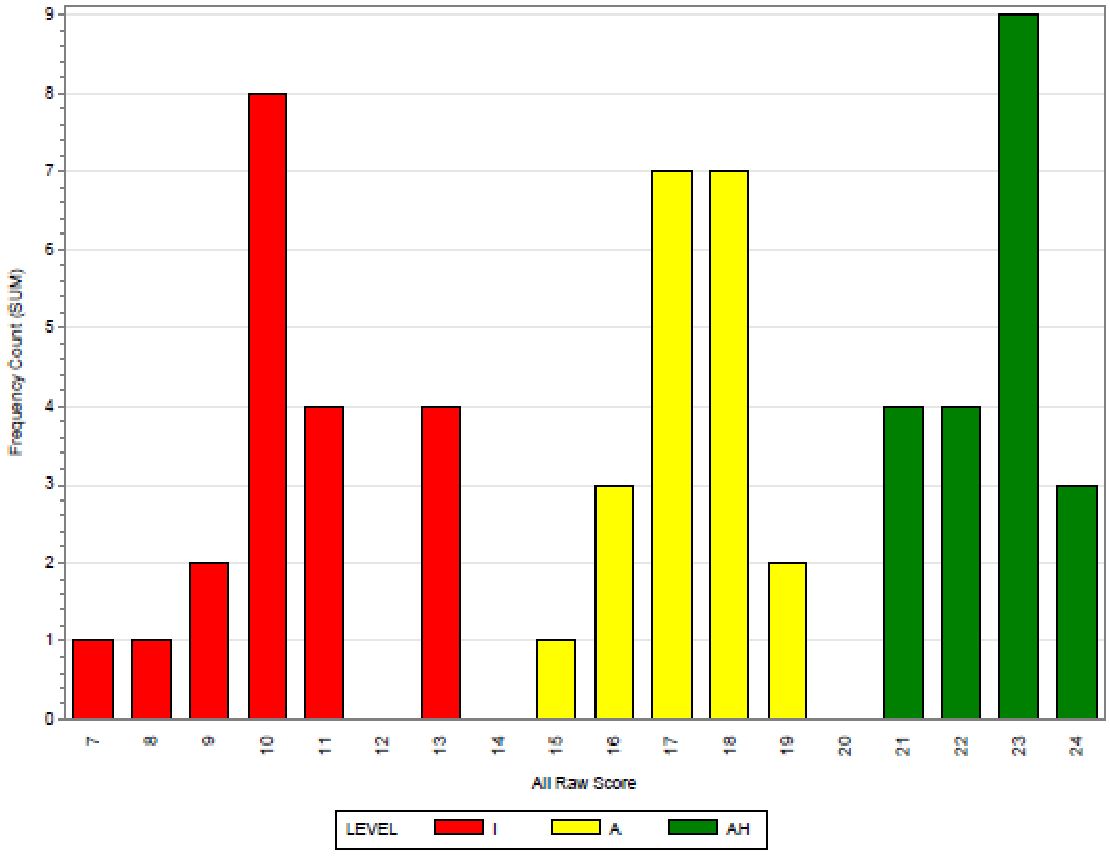
Intermediate



Advanced



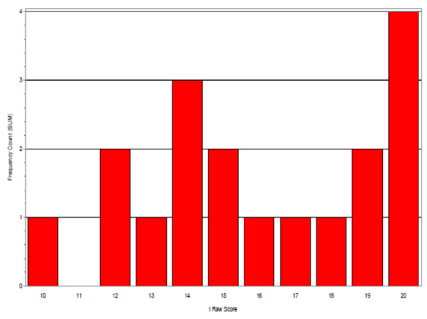
Advanced High



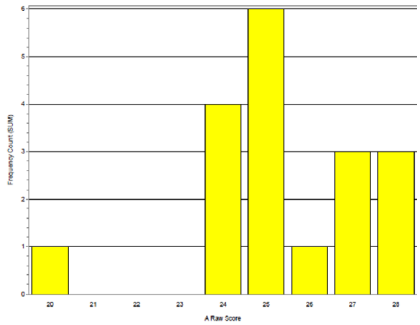
All Three Proficiency Levels Concurrently

Speaking Grades 2–3

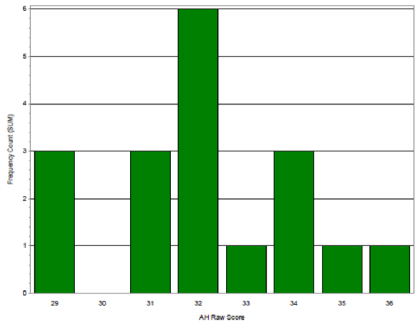
Round 1:



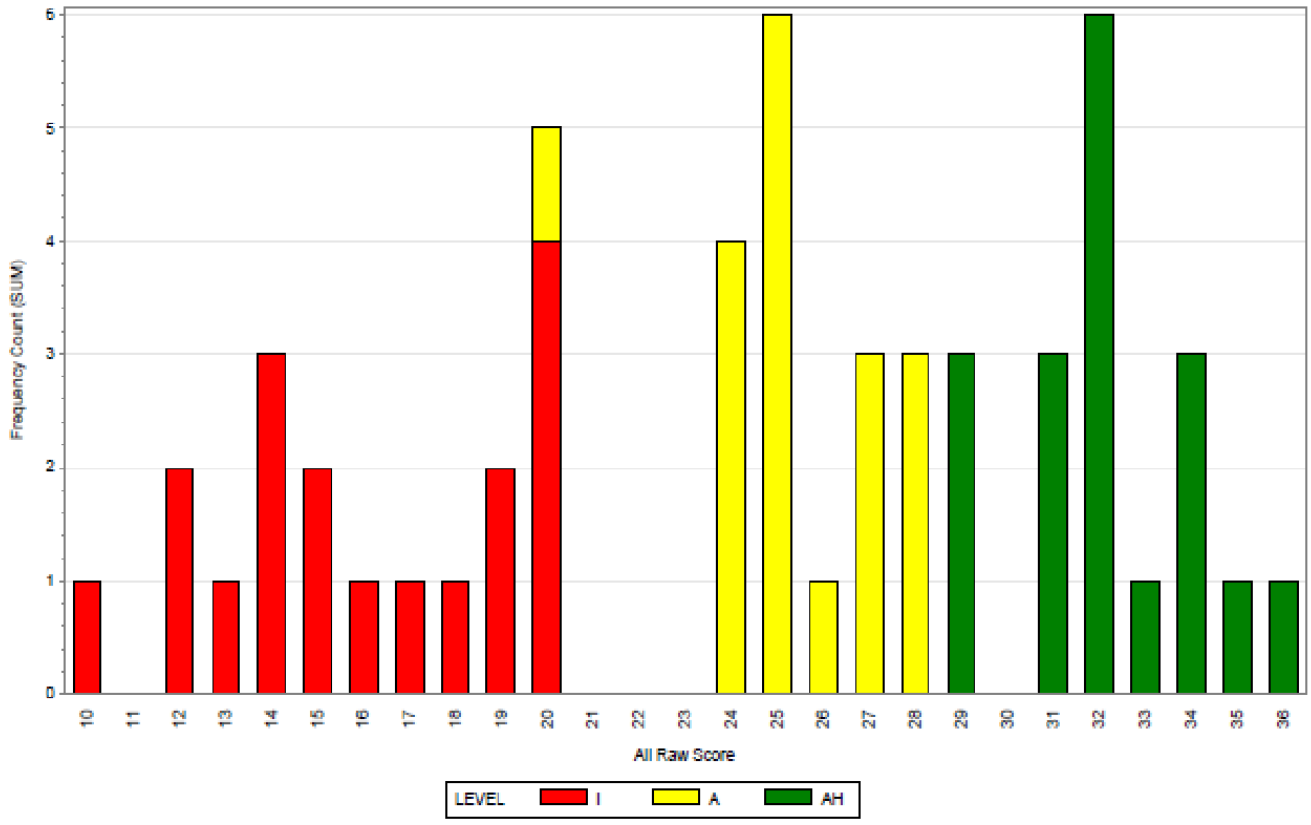
Intermediate



Advanced



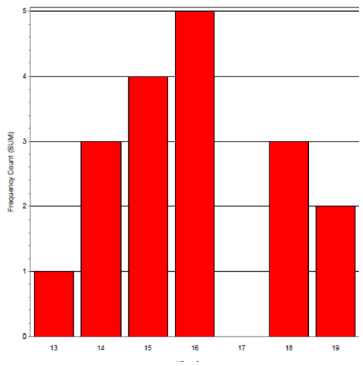
Advanced High



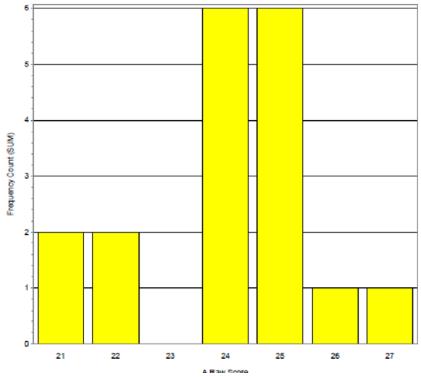
All Three Proficiency Levels Concurrently

Speaking Grades 2–3

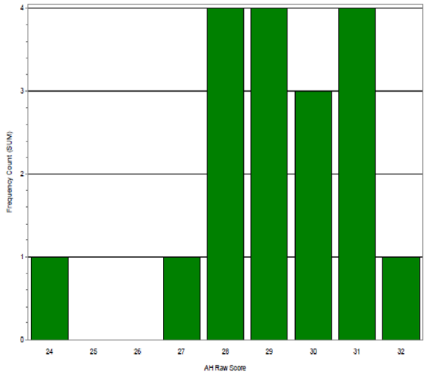
Round 2:



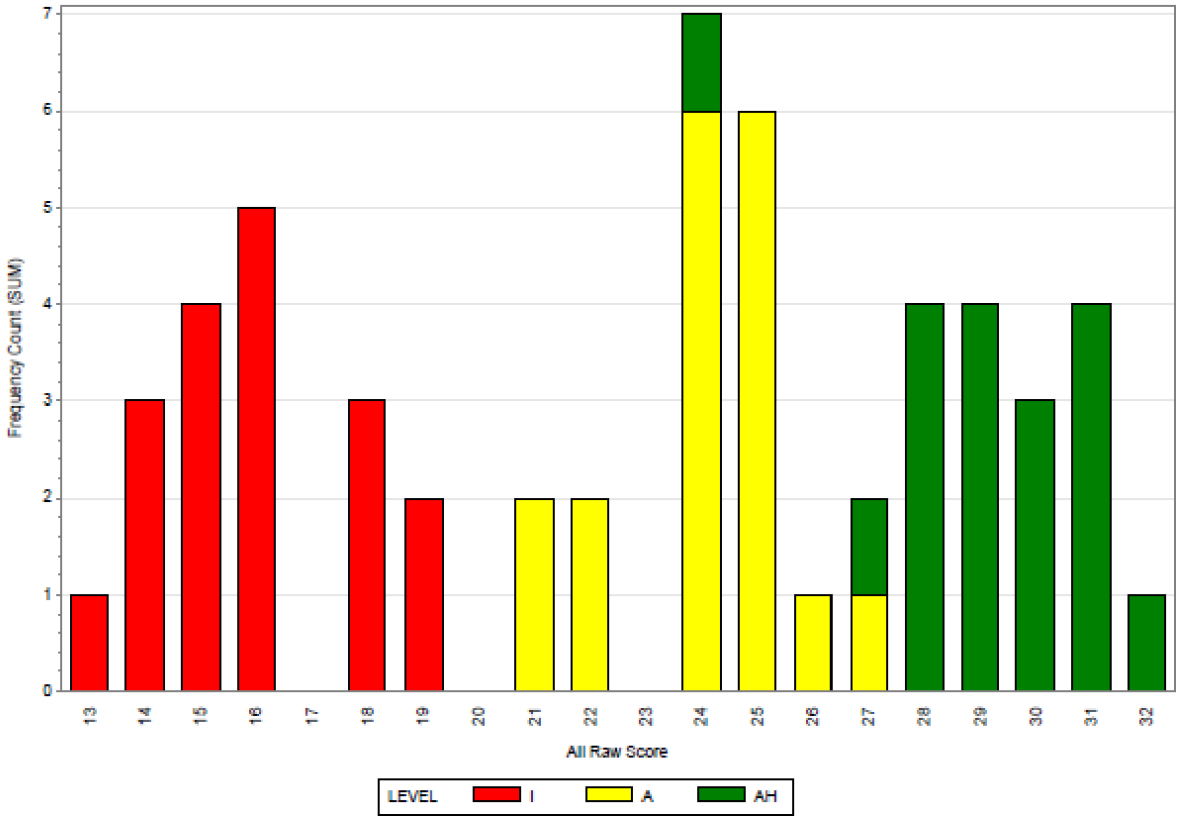
Intermediate



Advanced



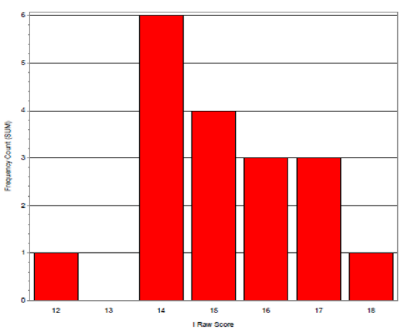
Advanced High



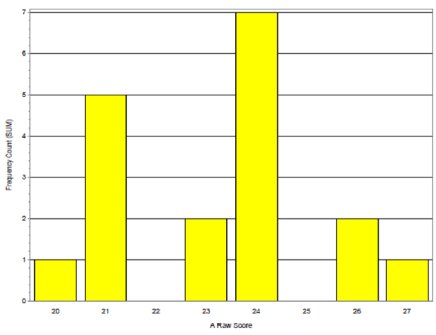
All Three Proficiency Levels Concurrently

Speaking Grades 2–3

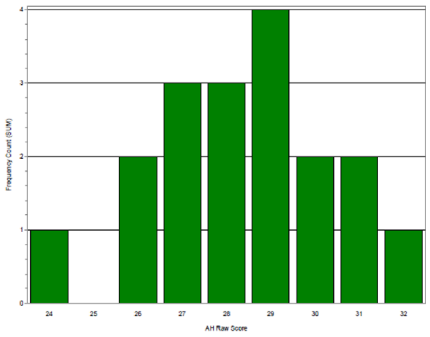
Round 3:



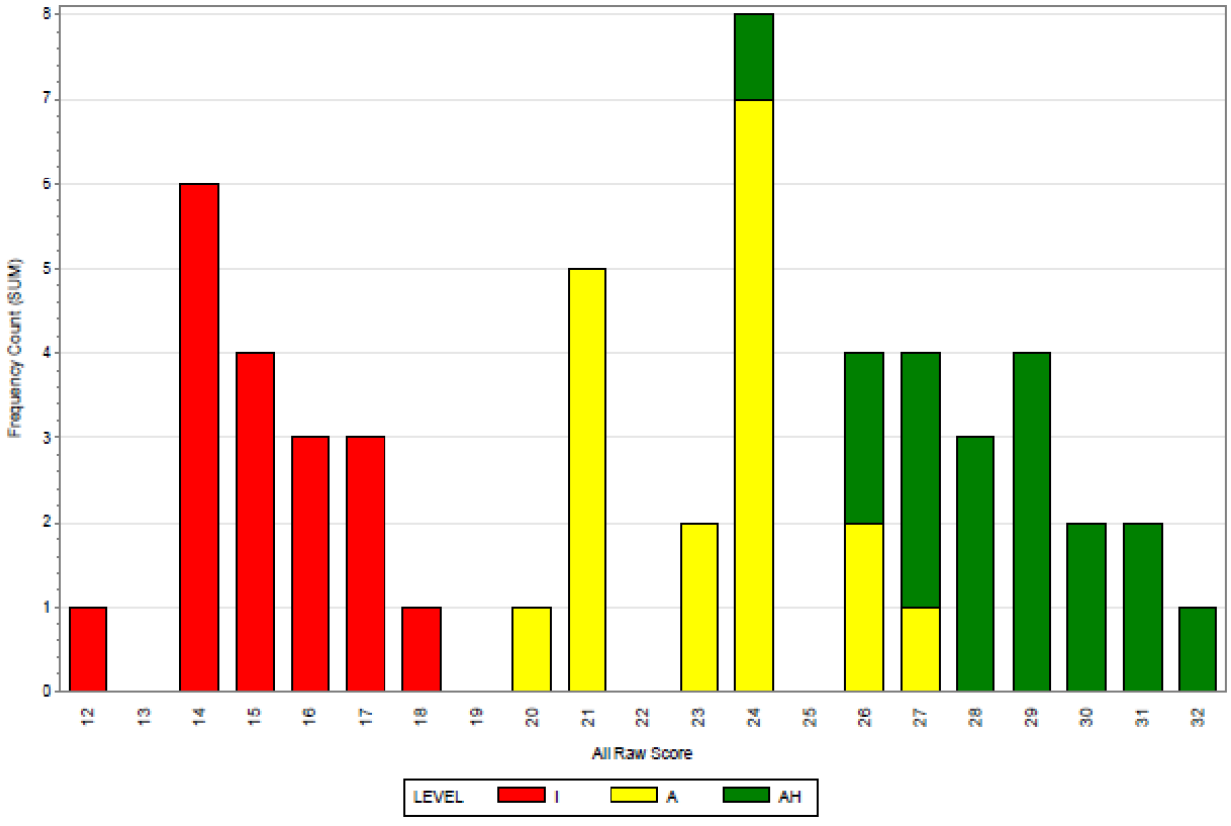
Intermediate



Advanced



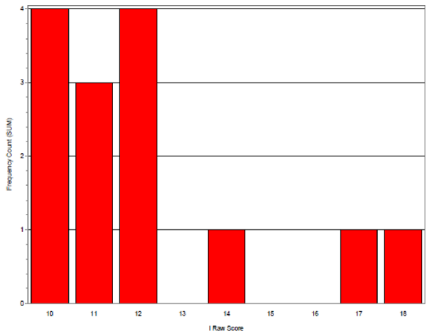
Advanced High



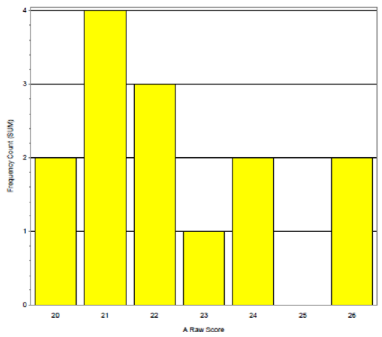
All Three Proficiency Levels Concurrently

Speaking Grades 4–5

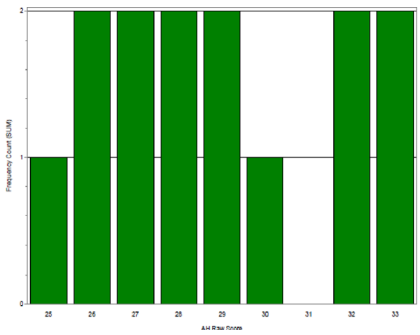
Round 1:



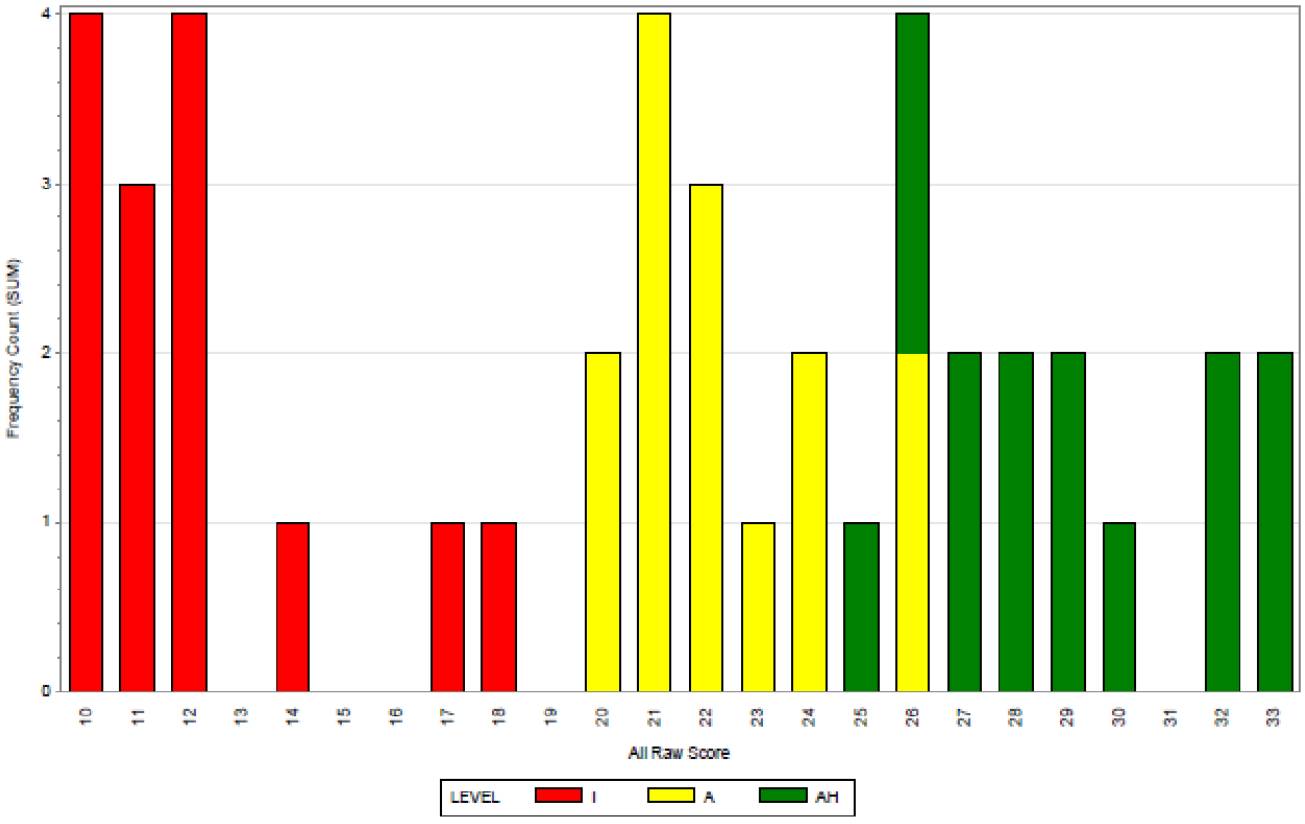
Intermediate



Advanced



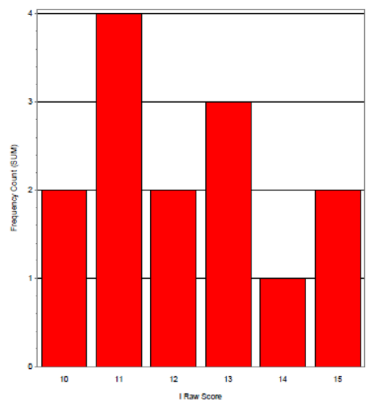
Advanced High



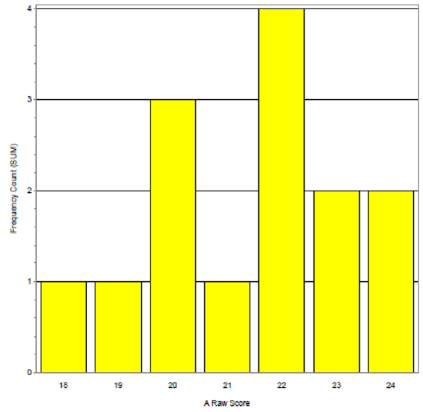
All Three Proficiency Levels Concurrently

Speaking Grades 4–5

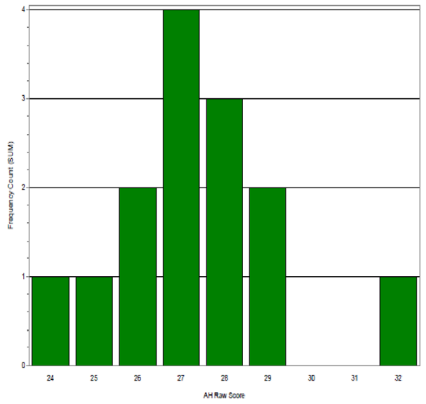
Round 2:



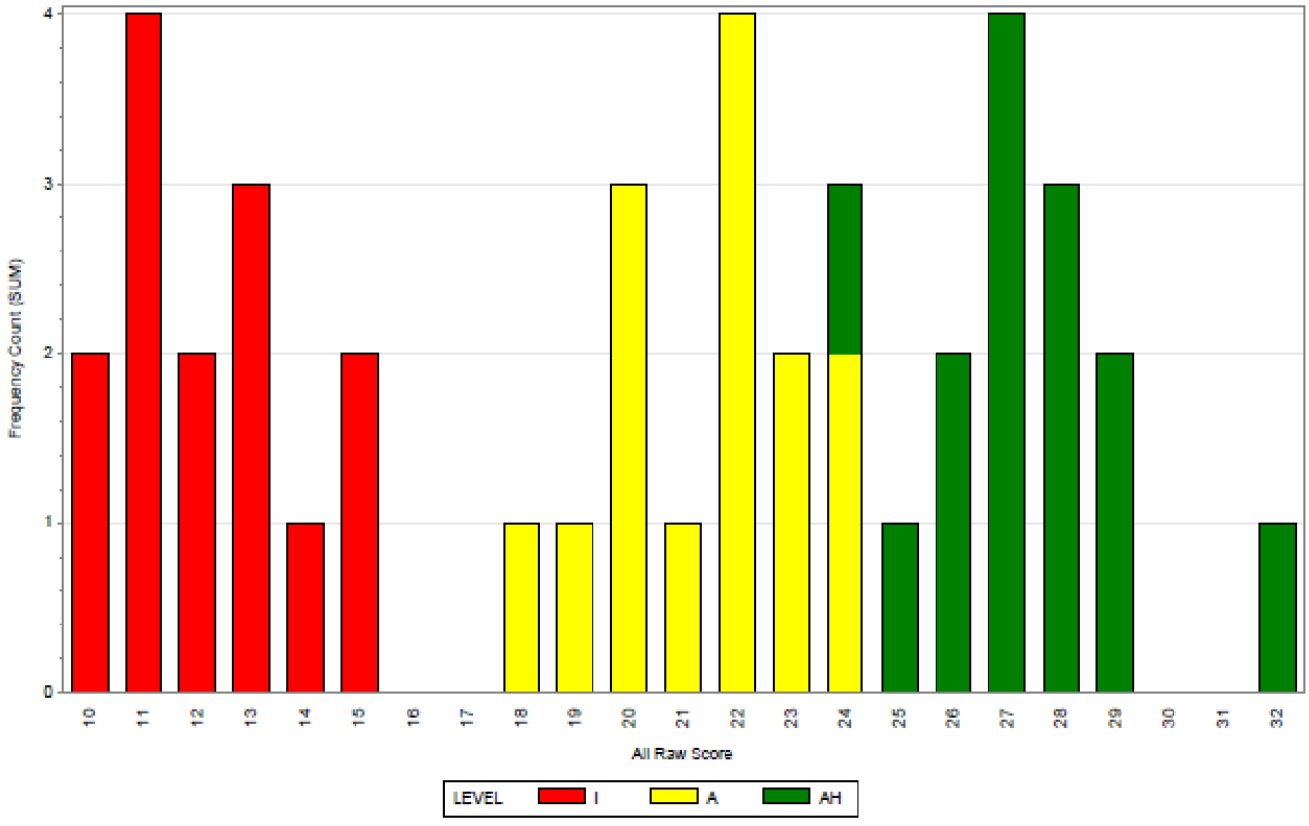
Intermediate



Advanced



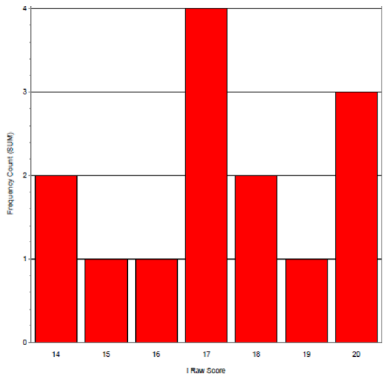
Advanced High



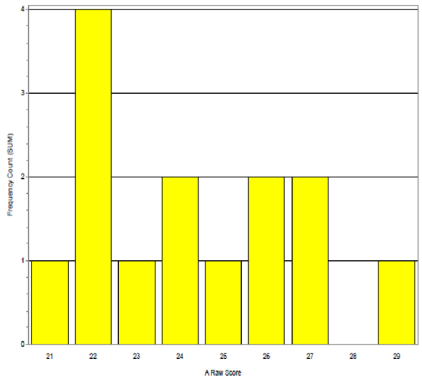
All Three Proficiency Levels Concurrently

Speaking Grades 4–5

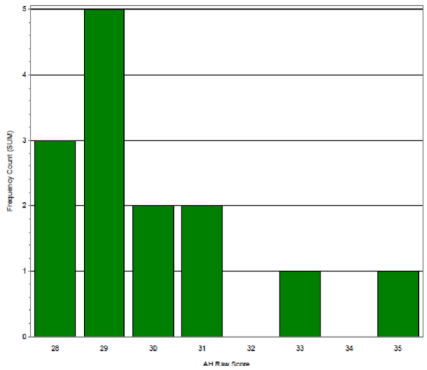
Round 3:



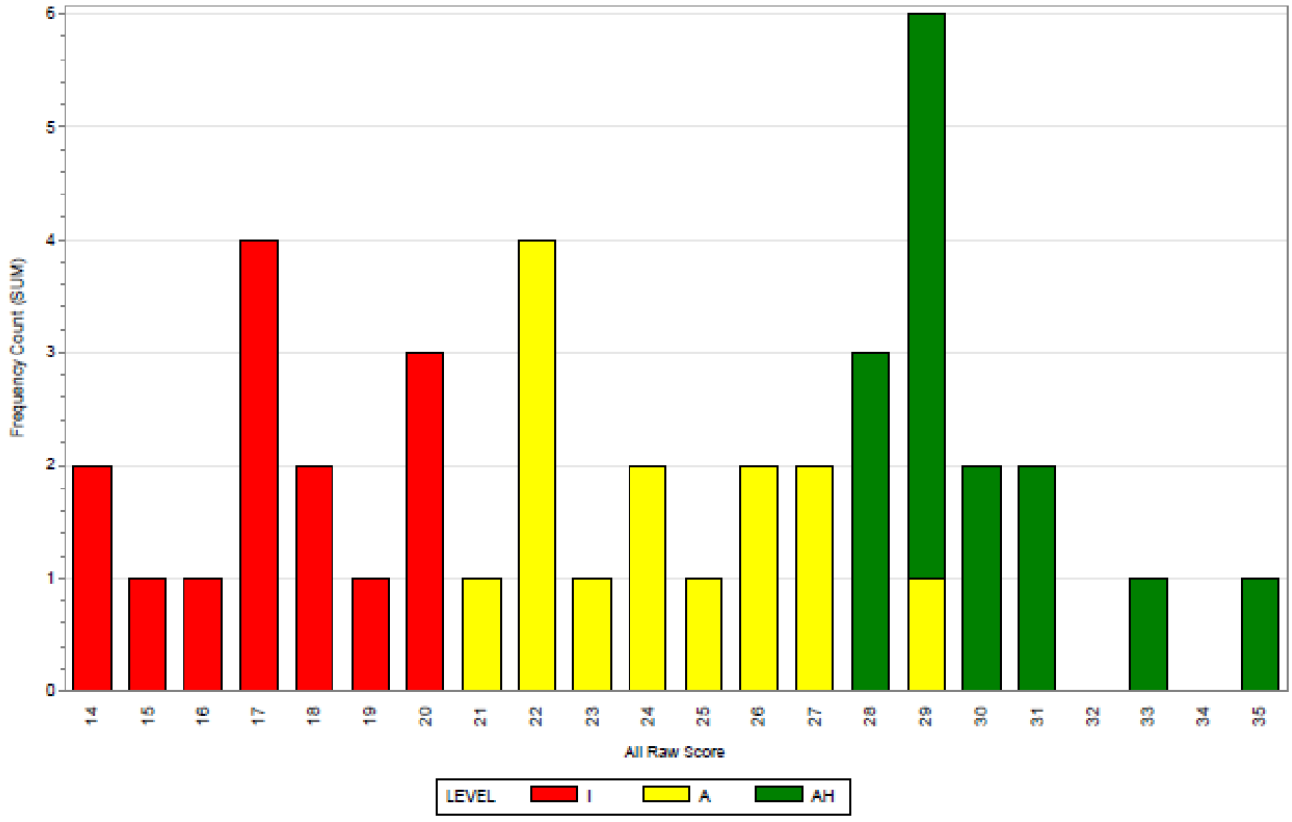
Intermediate



Advanced



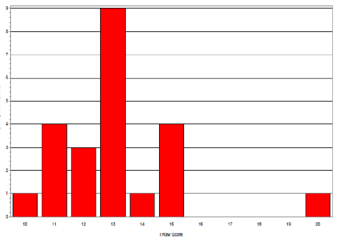
Advanced High



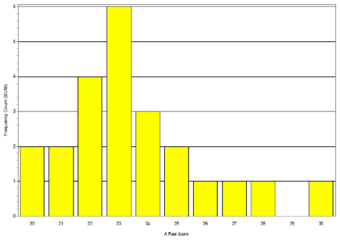
All Three Proficiency Levels Concurrently

Speaking Grades 6–8

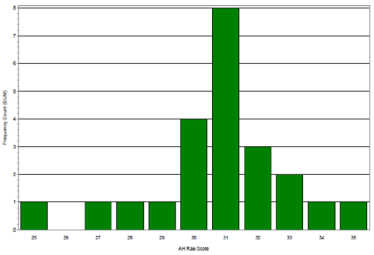
Round 1:



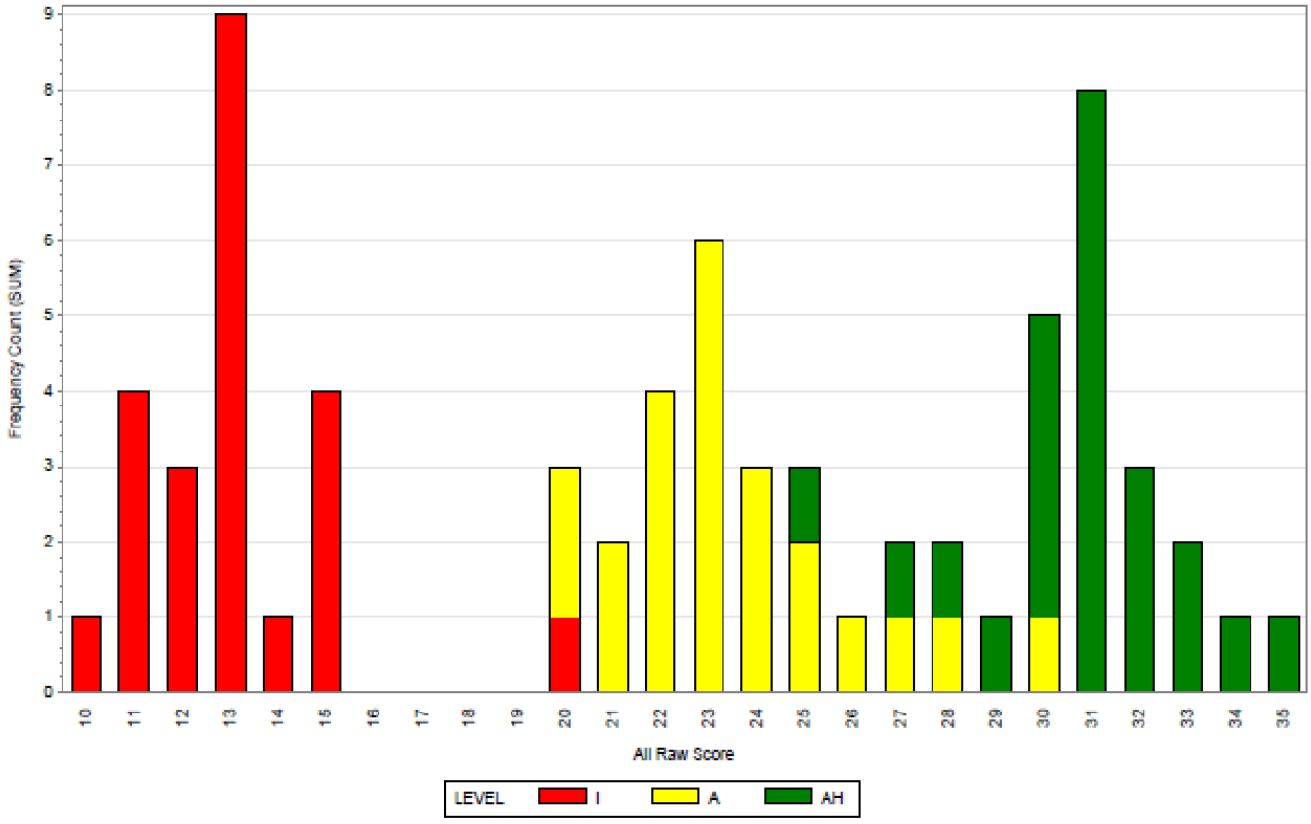
Intermediate



Advanced



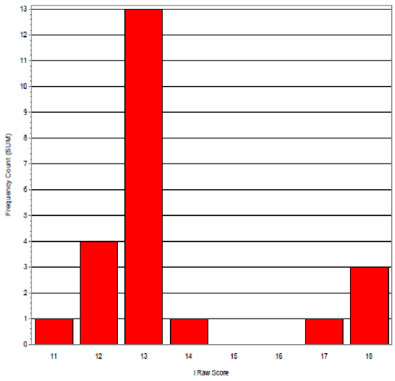
Advanced High



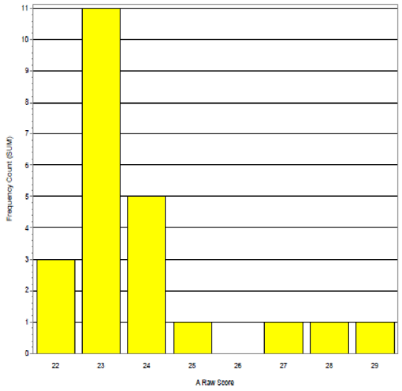
All Three Proficiency Levels Concurrently

Speaking Grades 6–8

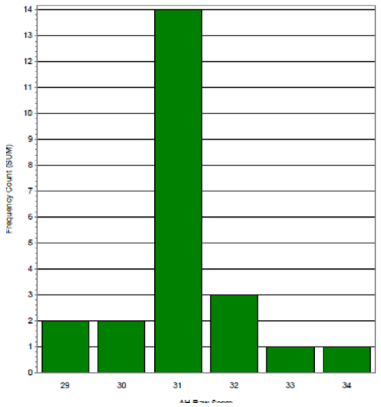
Round 2:



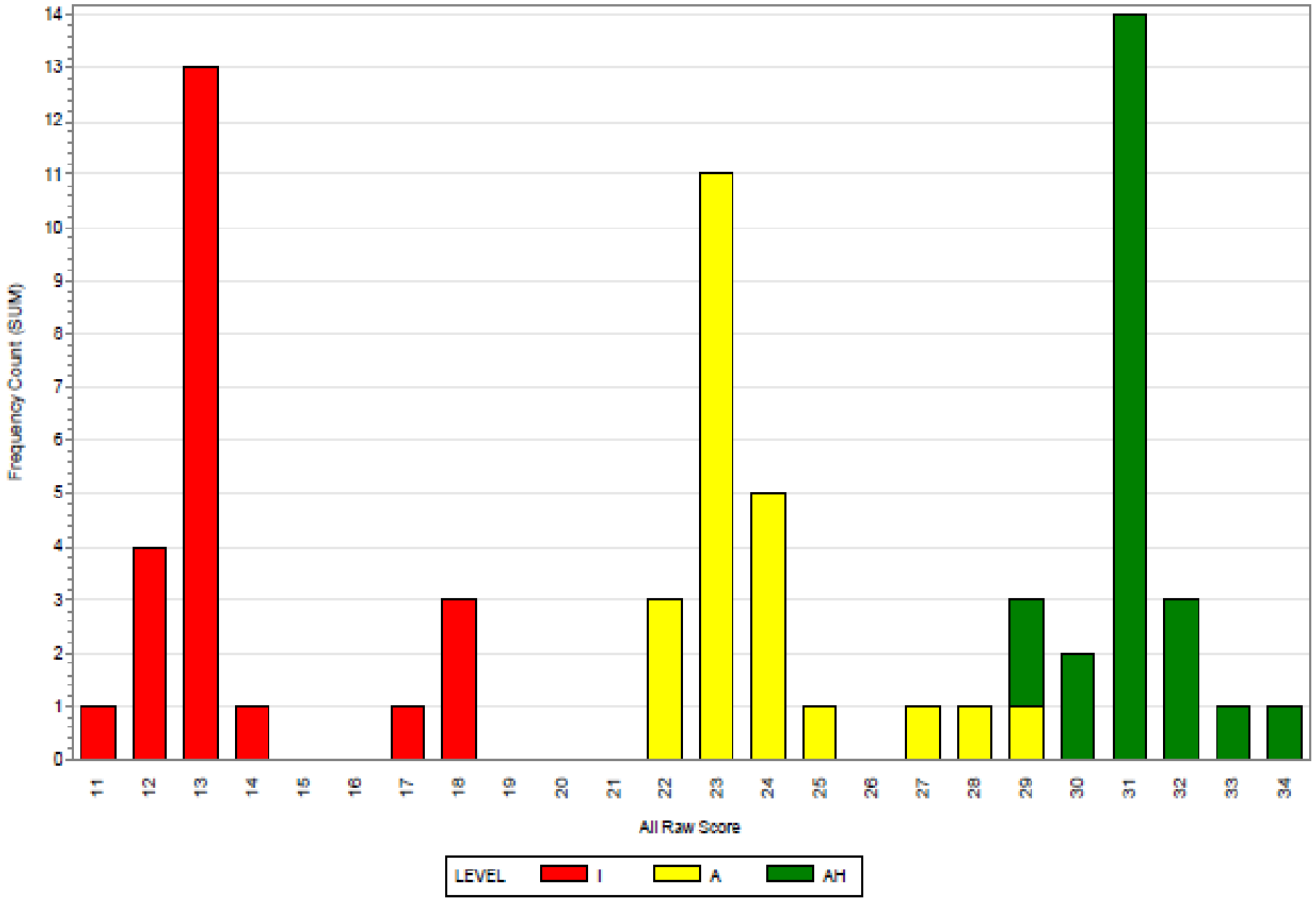
Intermediate



Advanced



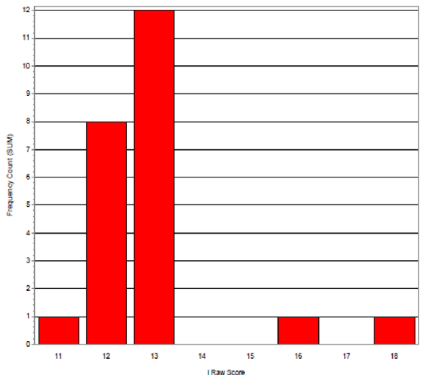
Advanced High



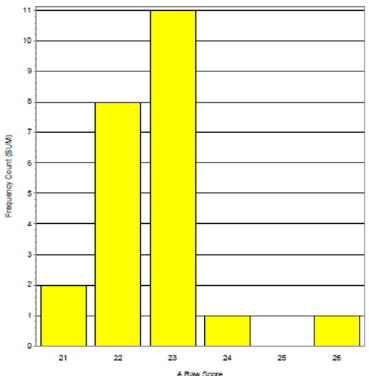
All Three Proficiency Levels Concurrently

Speaking Grades 6–8

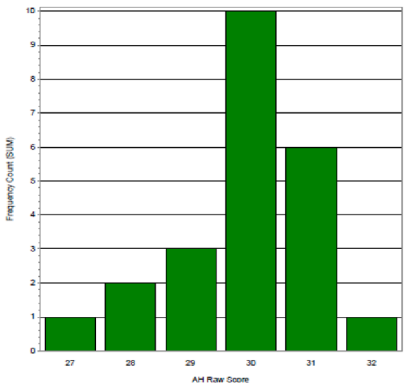
Round 3:



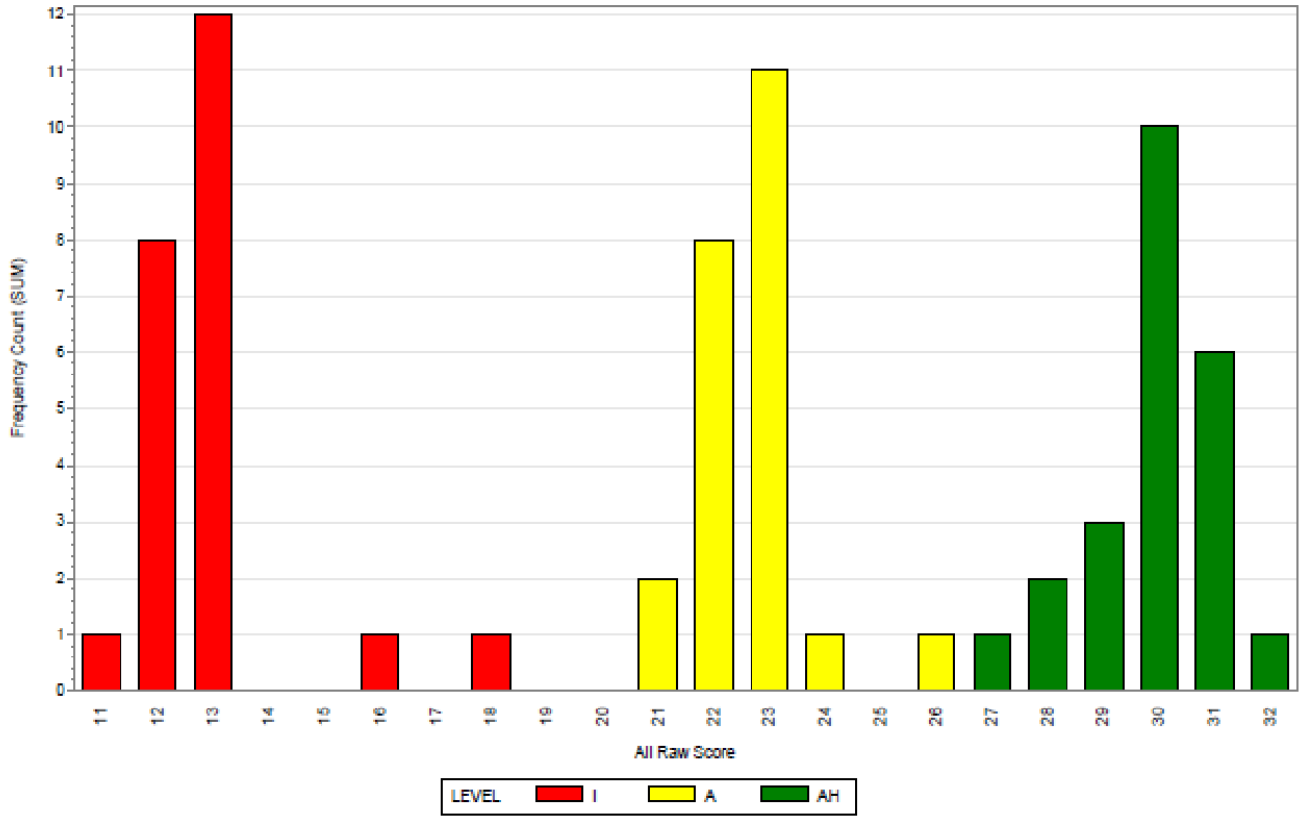
Intermediate



Advanced



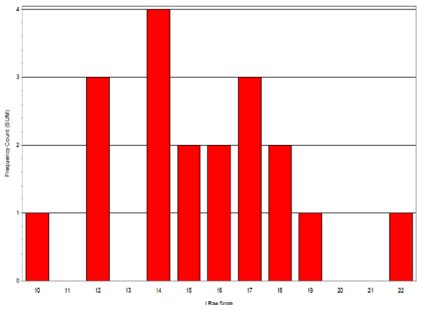
Advanced High



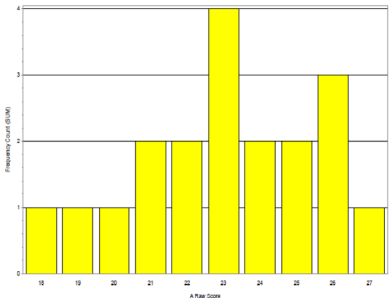
All Three Proficiency Levels Concurrently

Speaking Grades 9–12

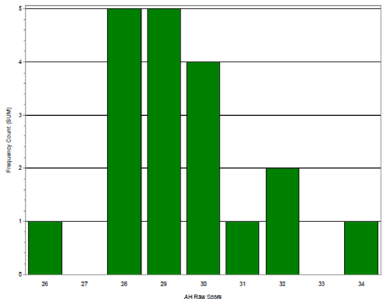
Round 1:



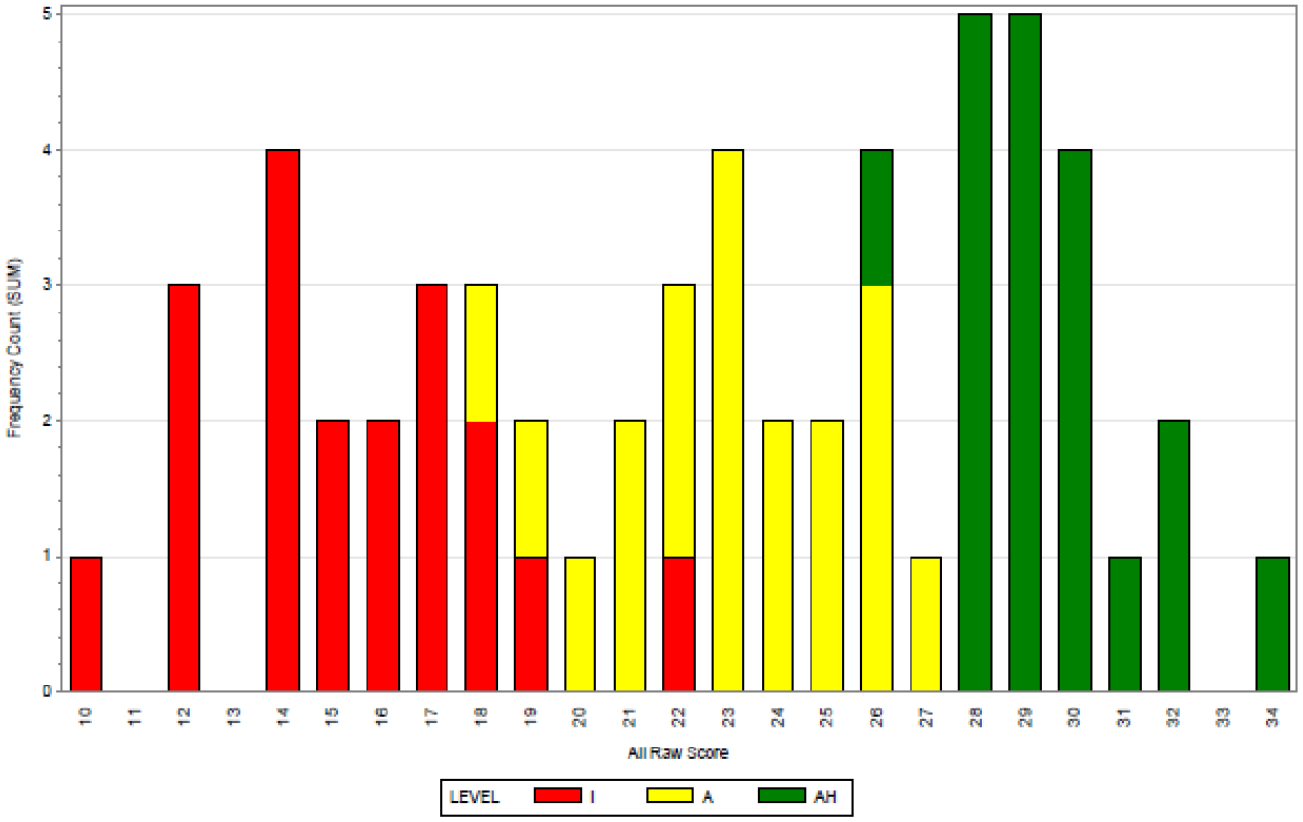
Intermediate



Advanced



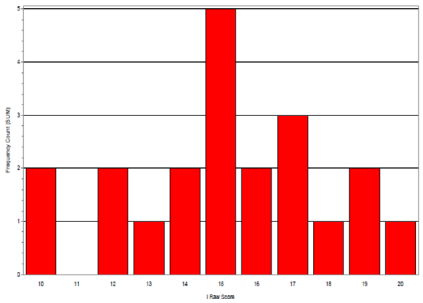
Advanced High



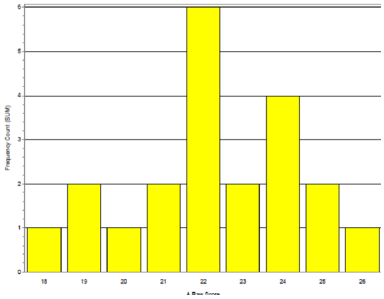
All Three Proficiency Levels Concurrently

Speaking Grades 9–12

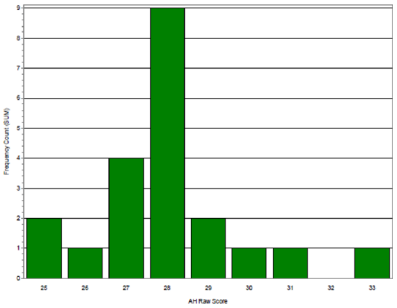
Round 2:



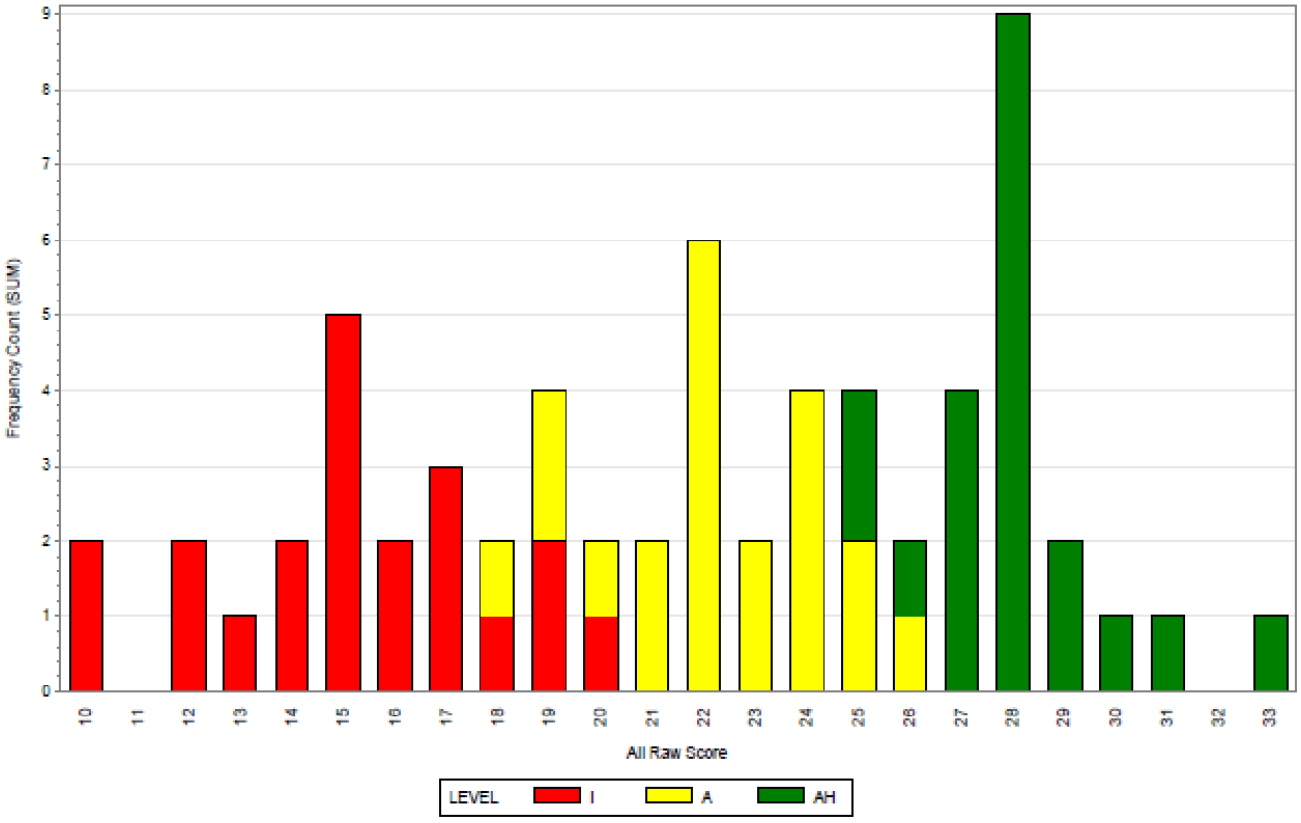
Intermediate



Advanced



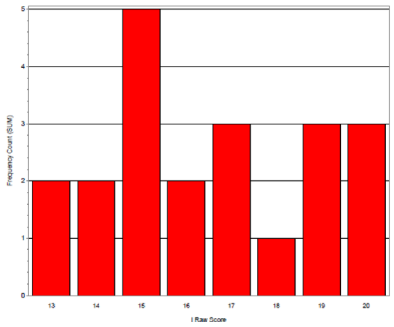
Advanced High



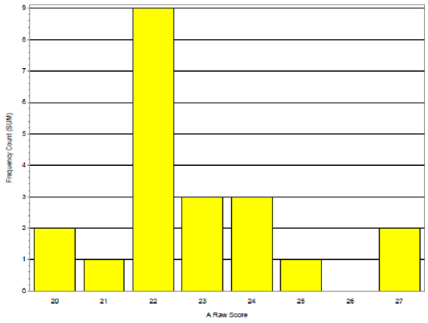
All Three Proficiency Levels Concurrently

Speaking Grades 9–12

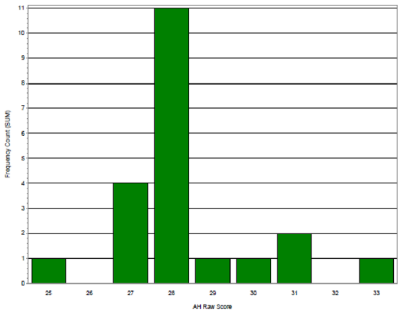
Round 3:



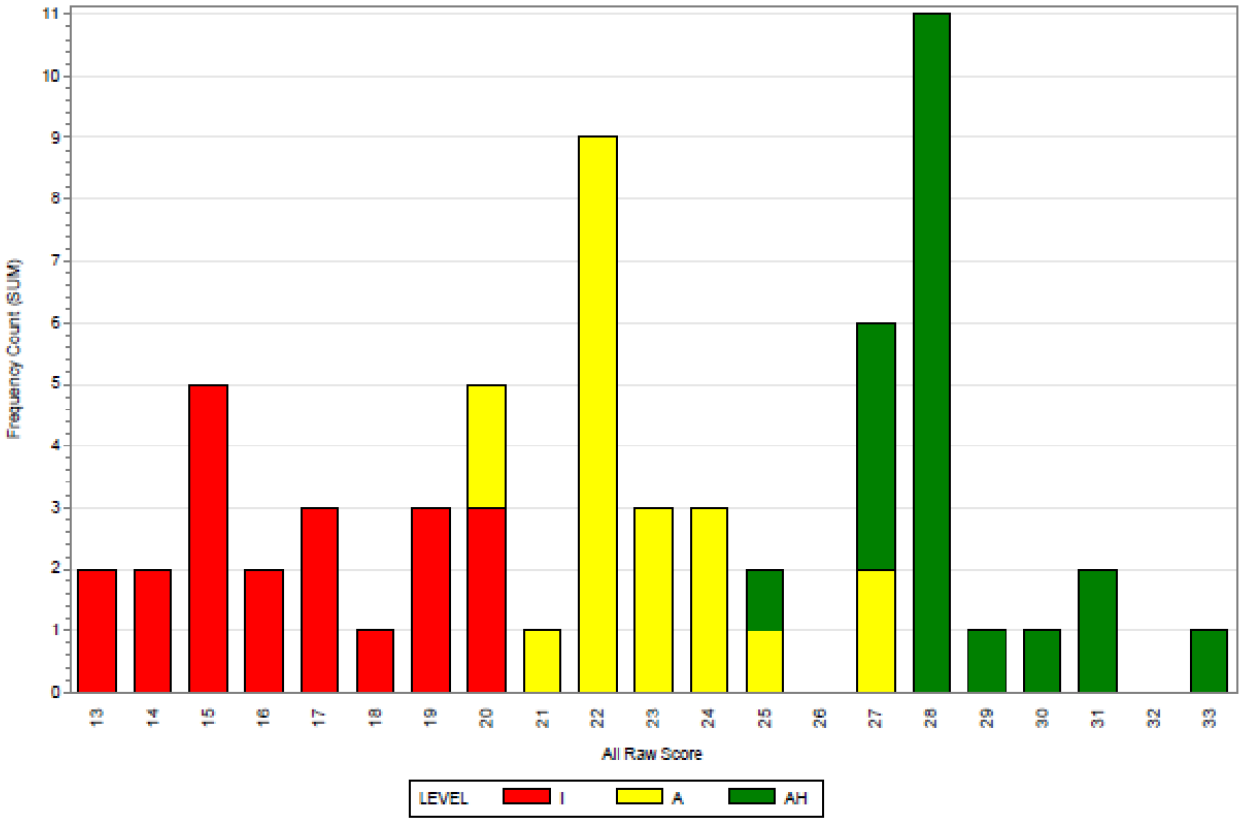
Intermediate



Advanced



Advanced High



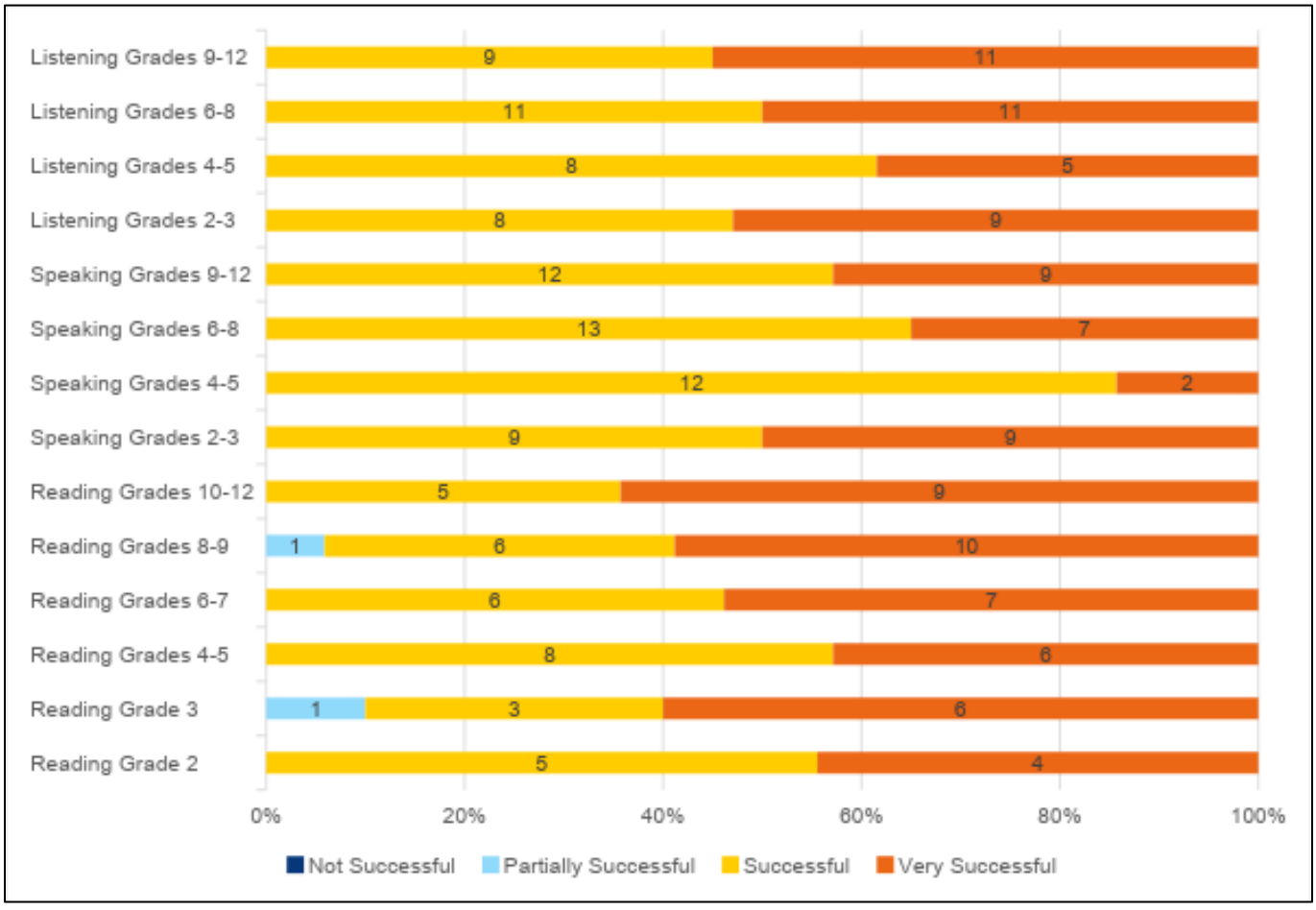
All Three Proficiency Levels Concurrently

Appendix I – Participant Evaluation Results

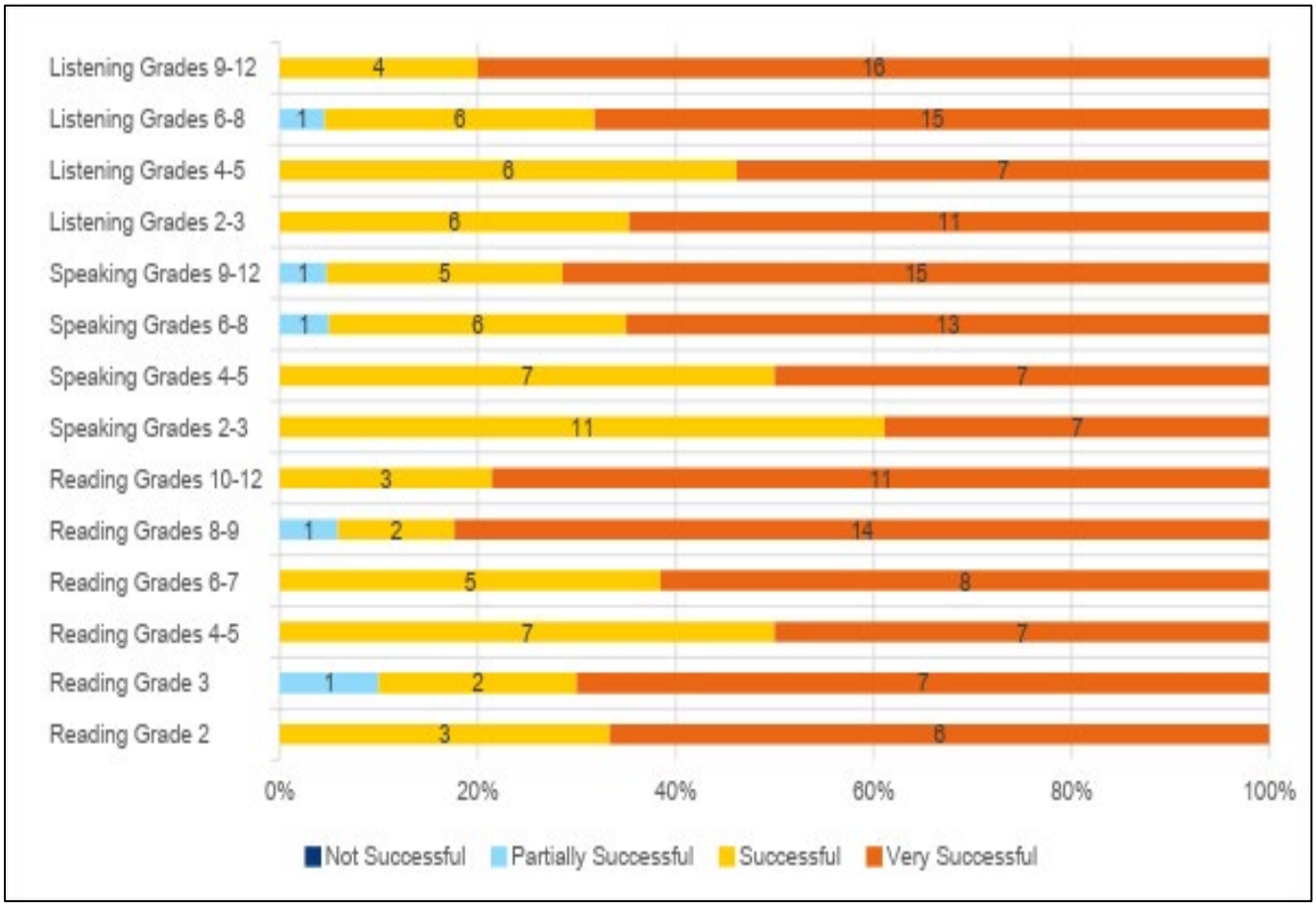
Breakout Session Process Evaluation

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

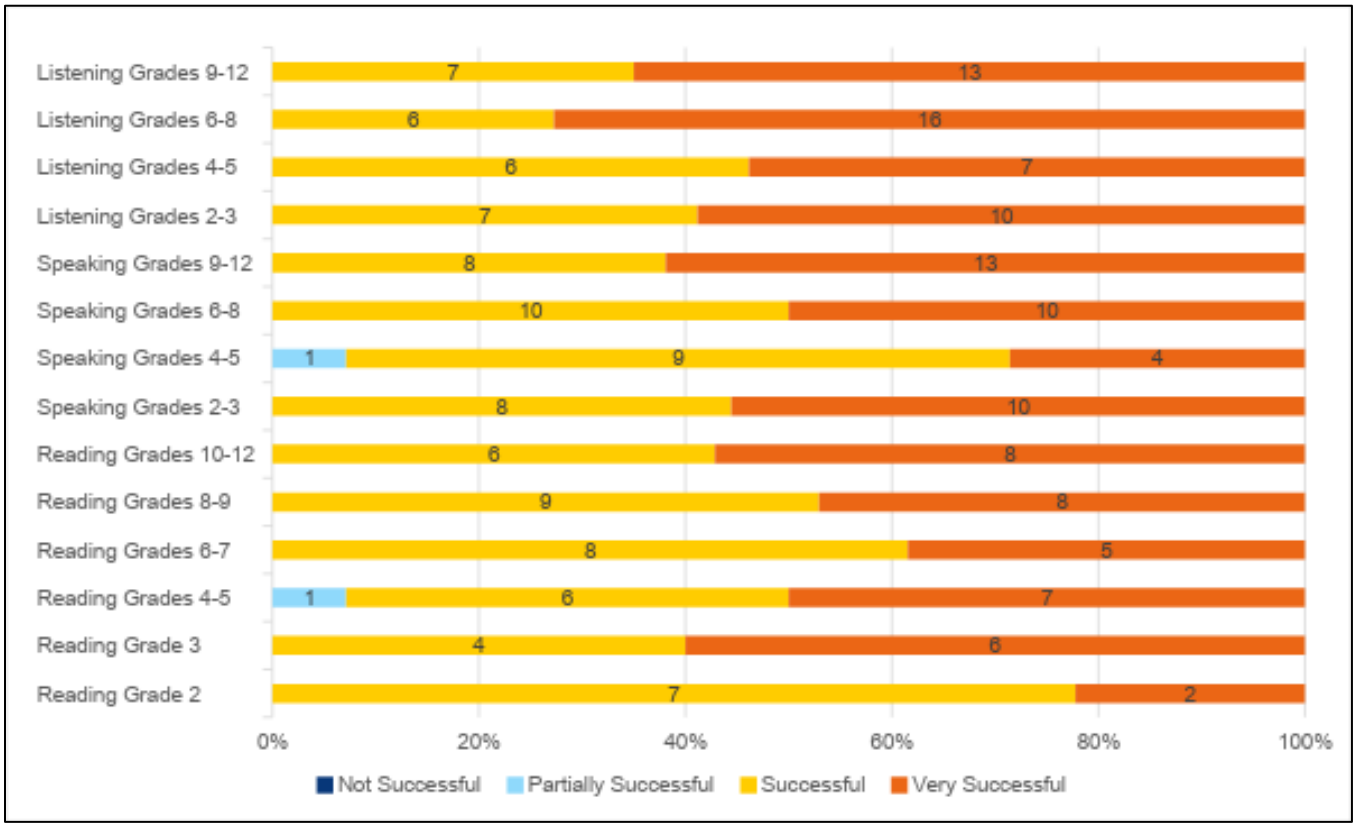
Overview of the TELPAS assessments



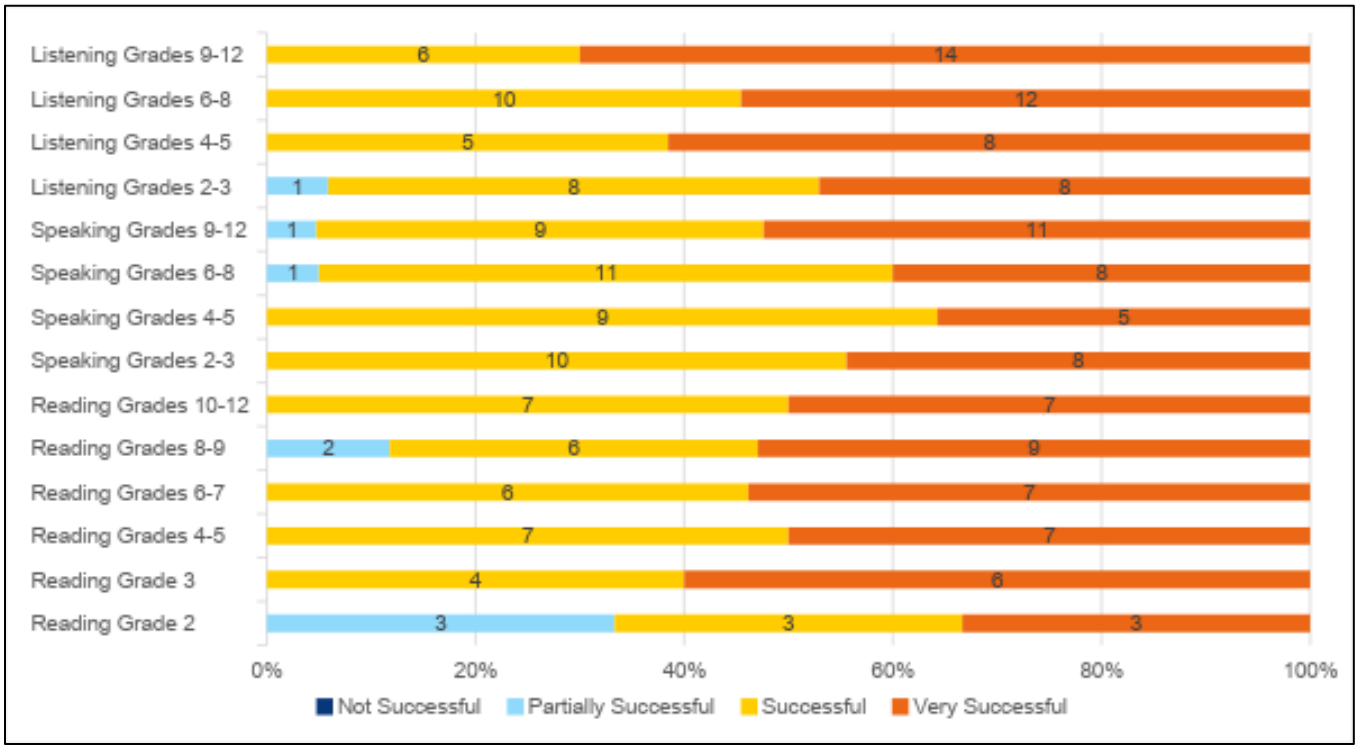
Experiencing the actual assessment



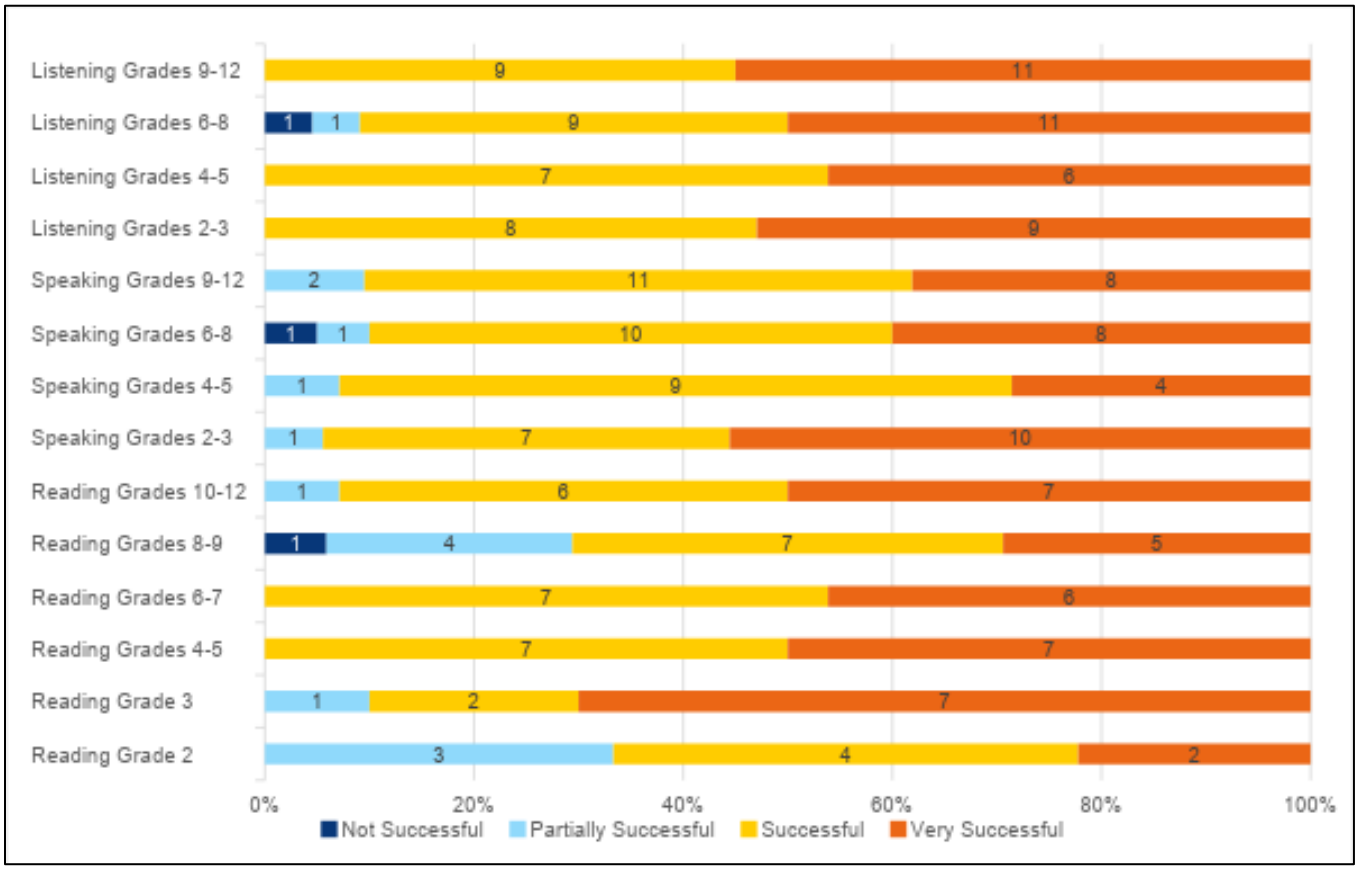
Discussion of the scoring of items on the assessment



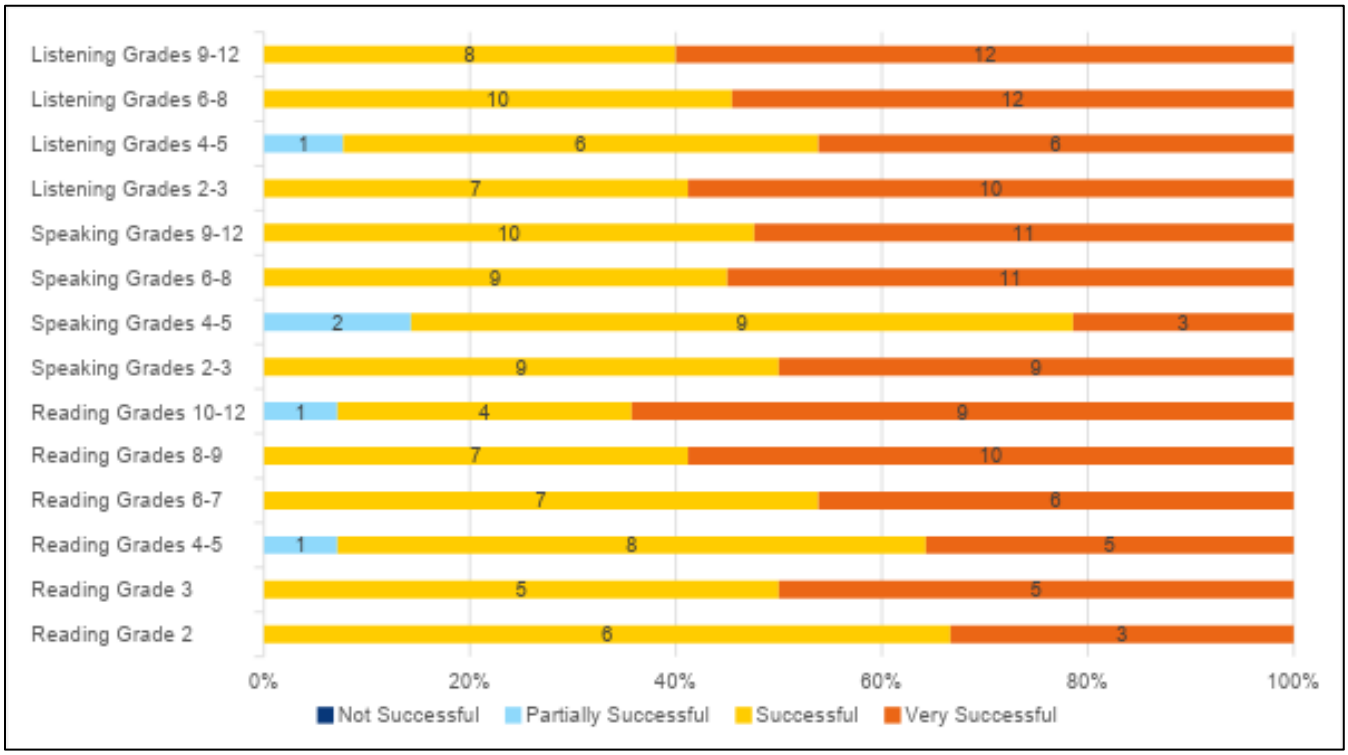
Discussion of proficiency level descriptors (PLDs)



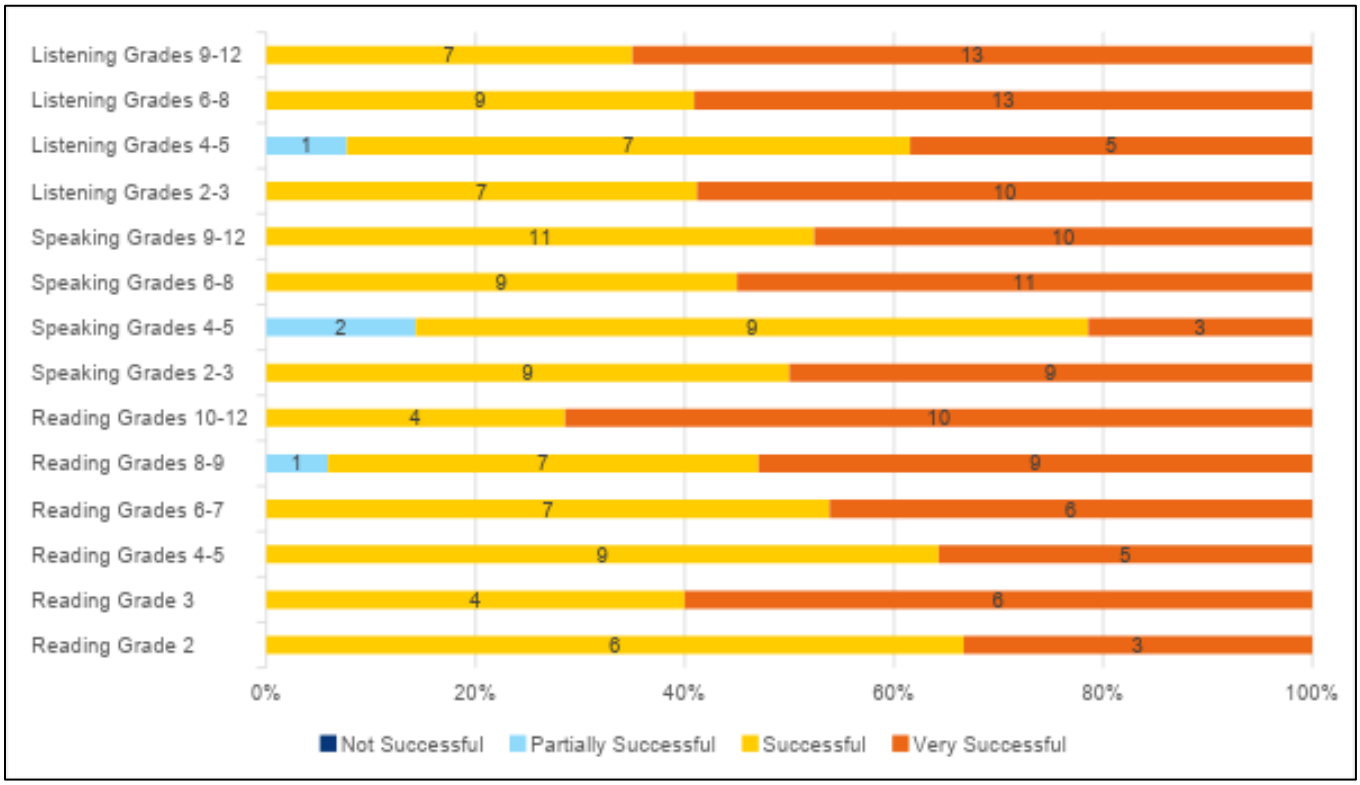
Development and discussion of the borderline proficiency level descriptors



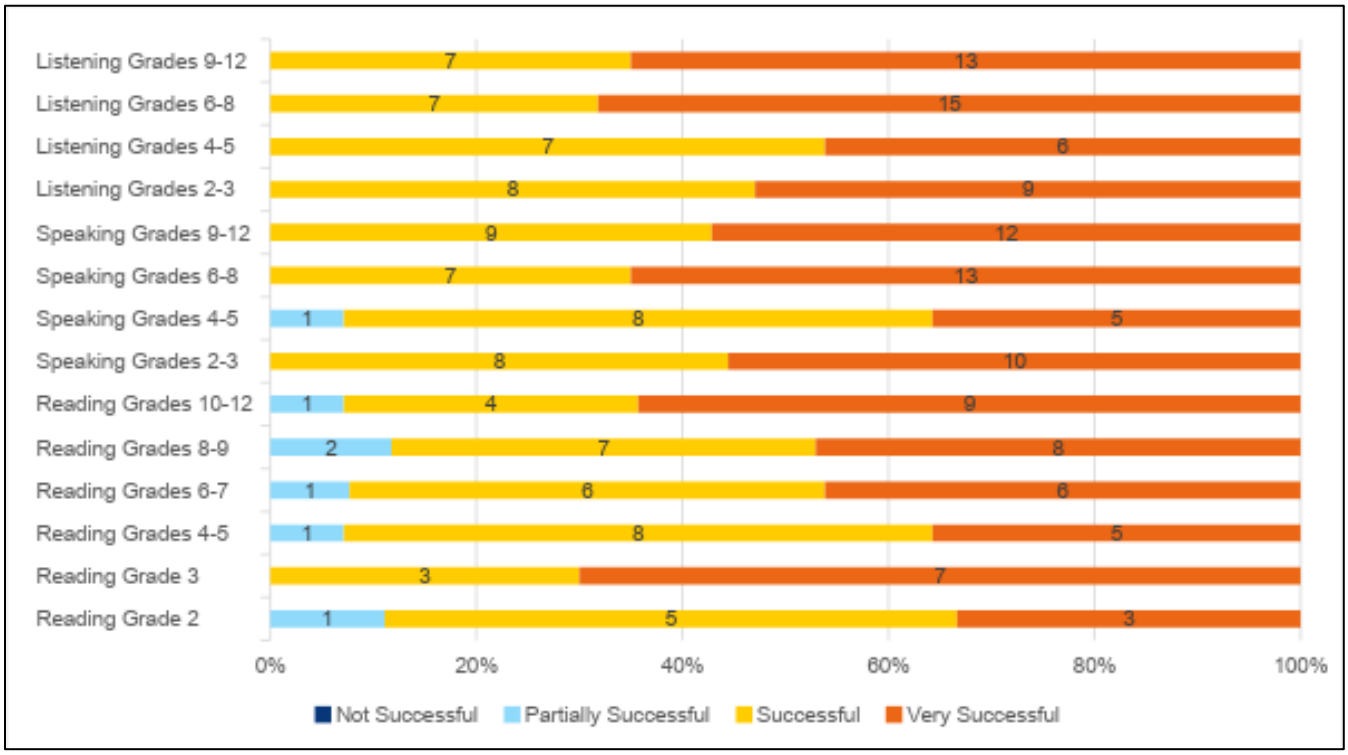
Overview of the standard-setting procedure



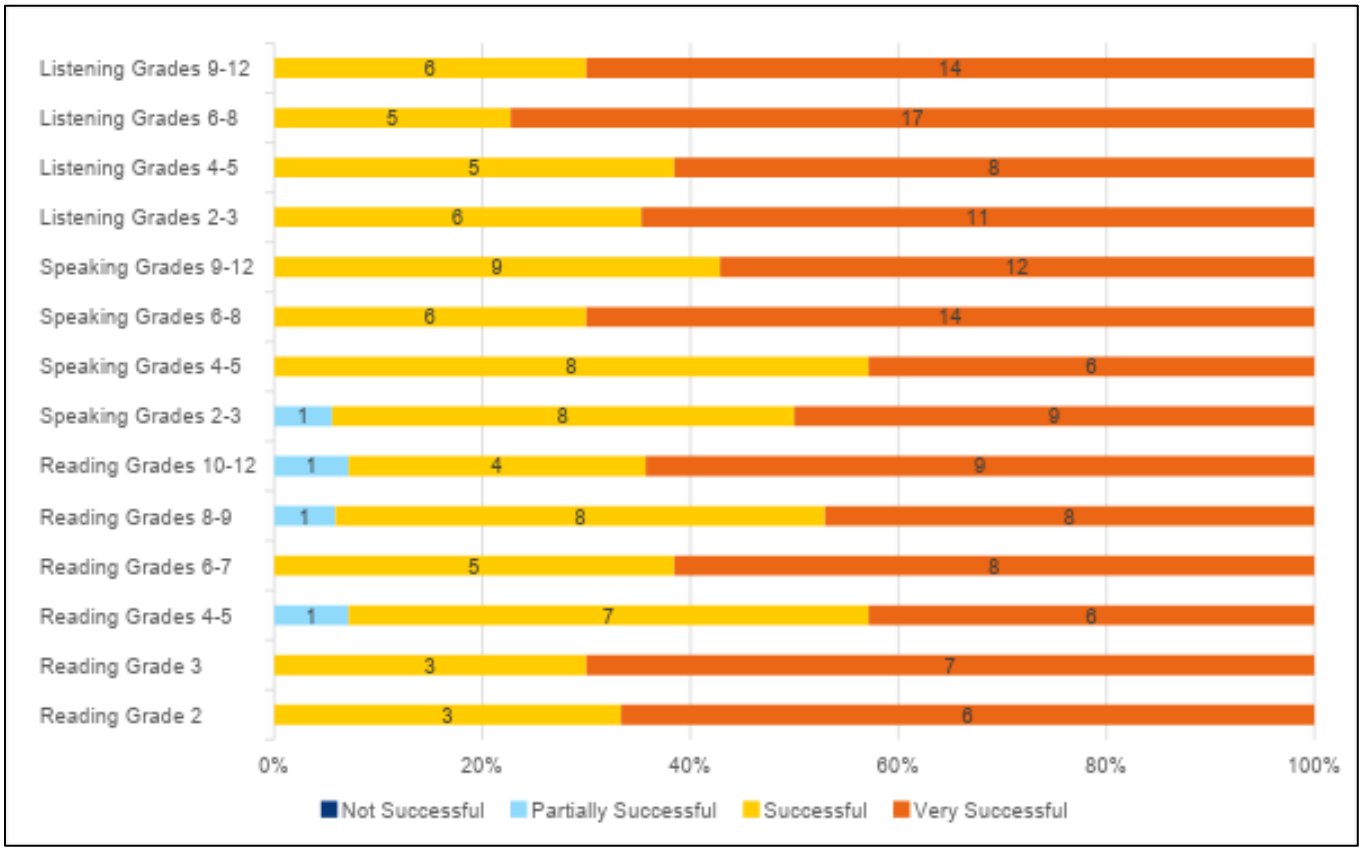
Practice exercise for the standard-setting procedure



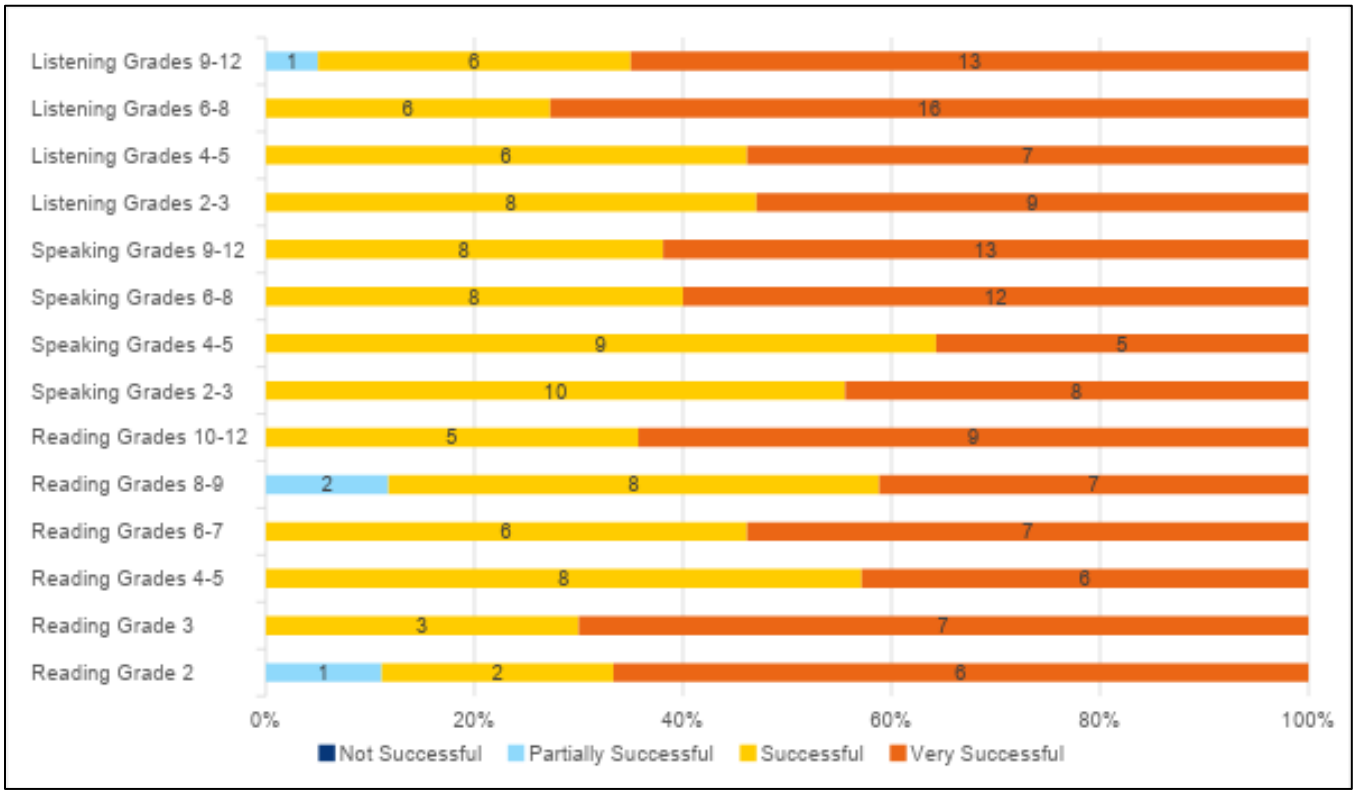
Judgment rounds



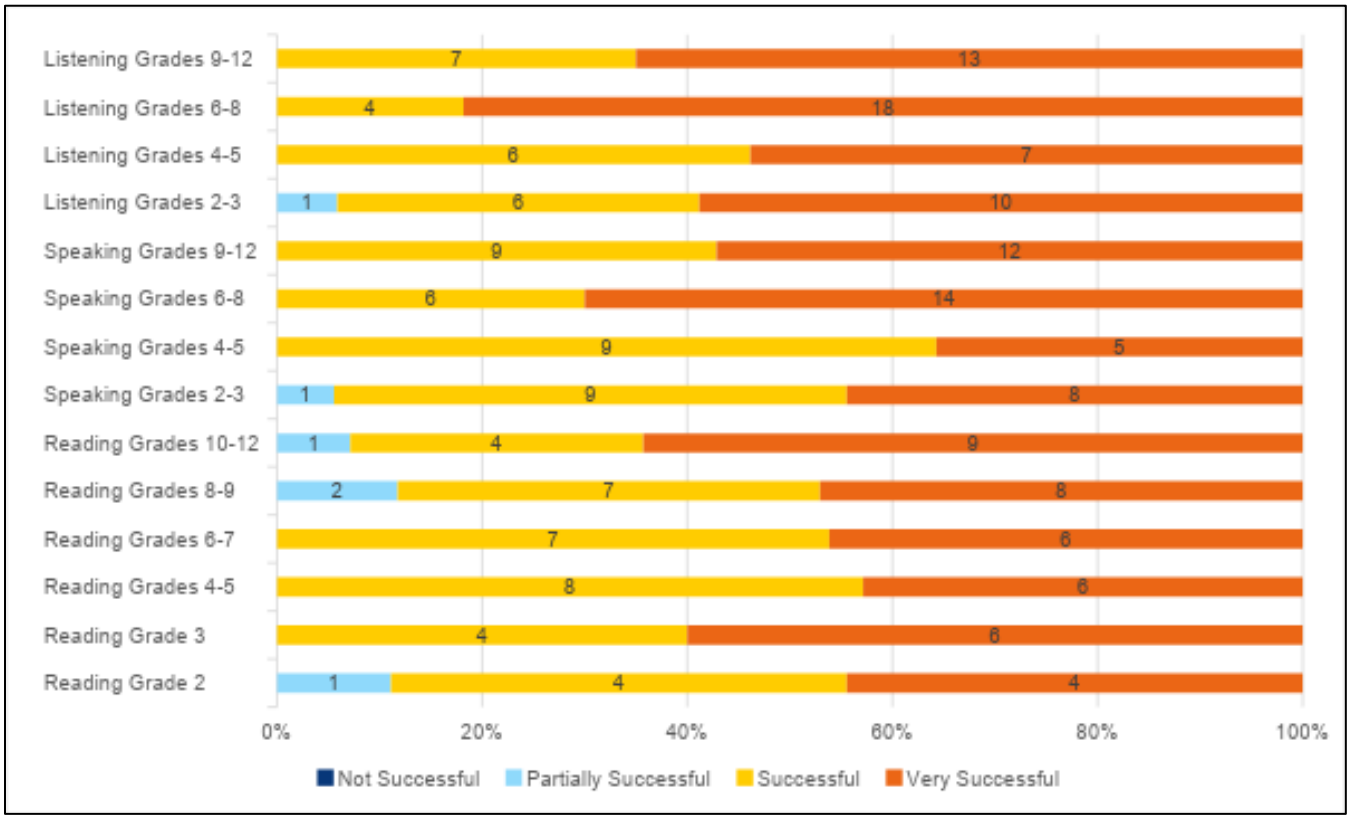
Judgment round feedback — table-level statistics



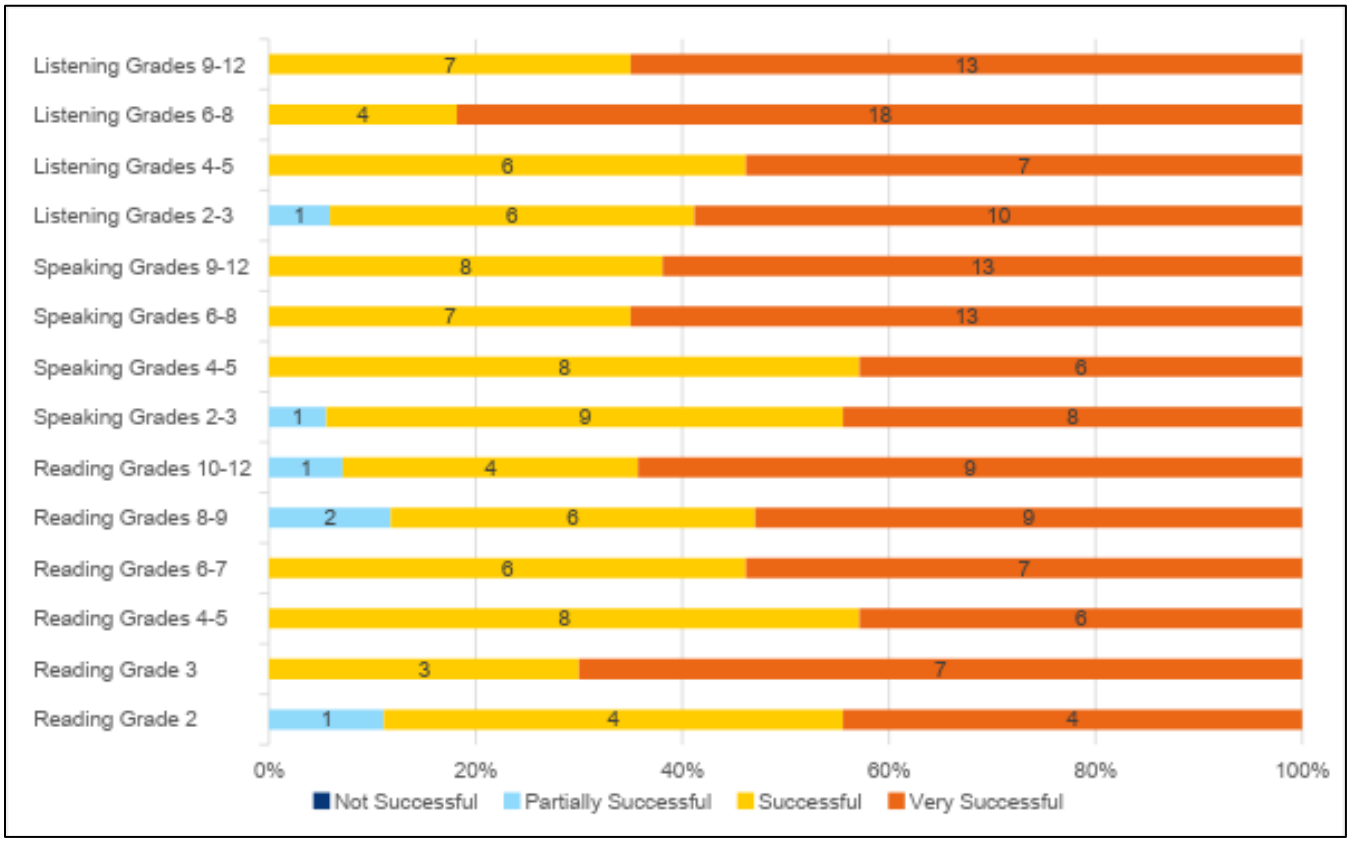
Judgment round feedback — committee-level statistics



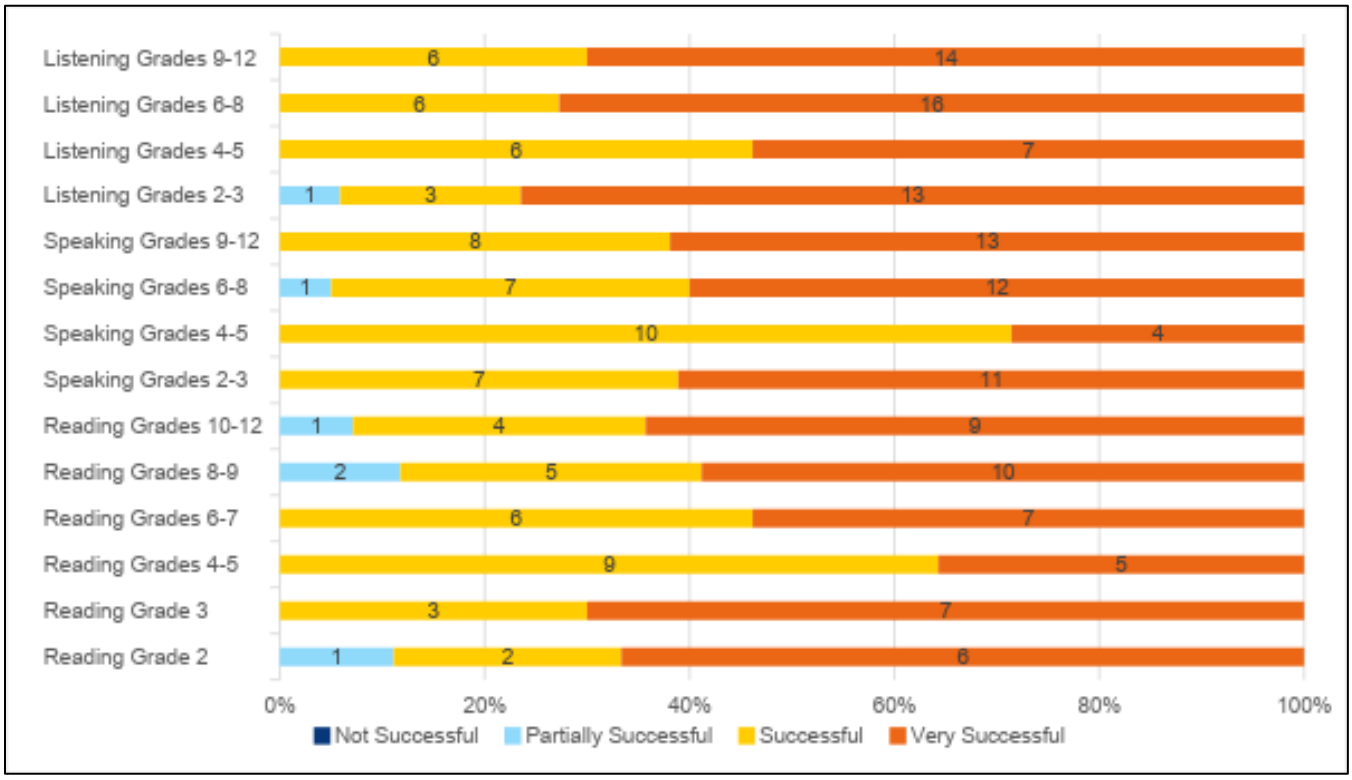
Judgment round feedback — panelist agreement data



Judgment round feedback — impact data

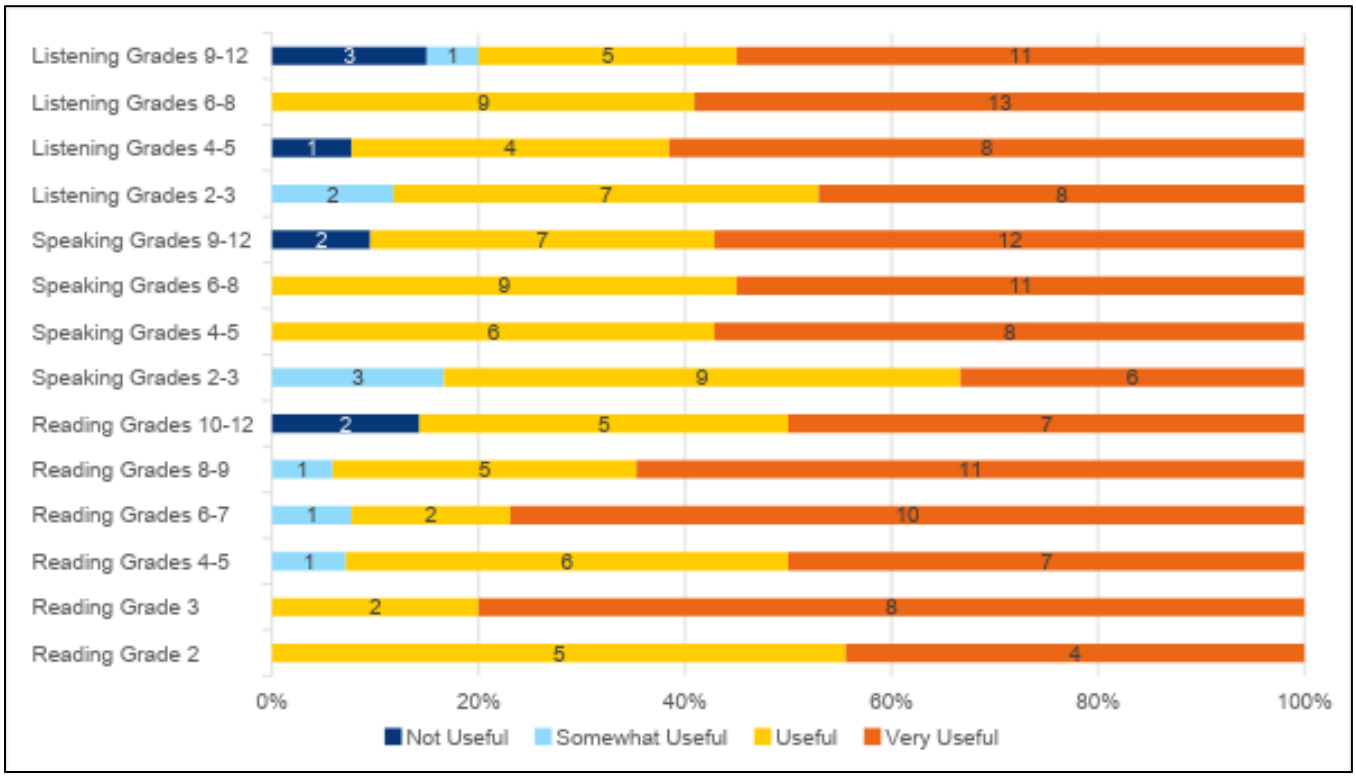


Discussions after each round



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

Proficiency level descriptors (PLDs)



Borderline proficiency level descriptors

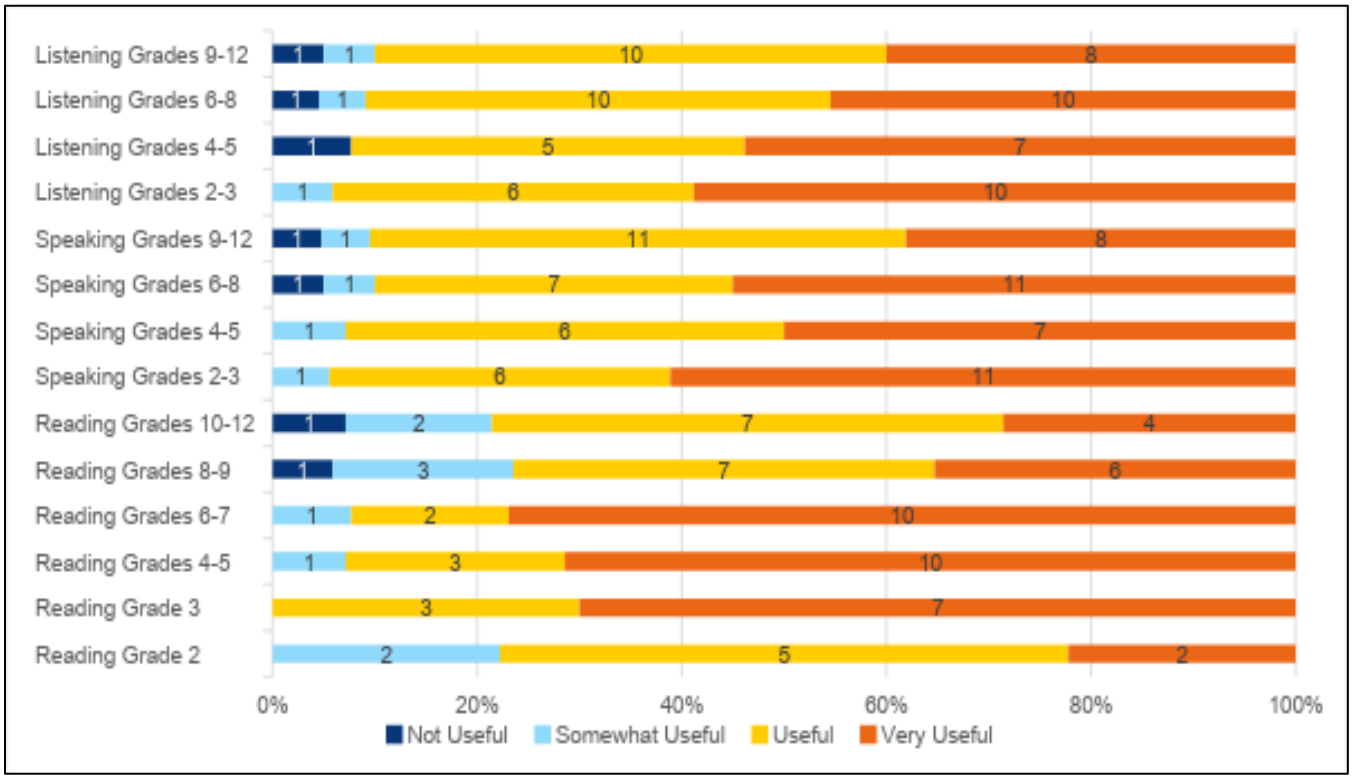
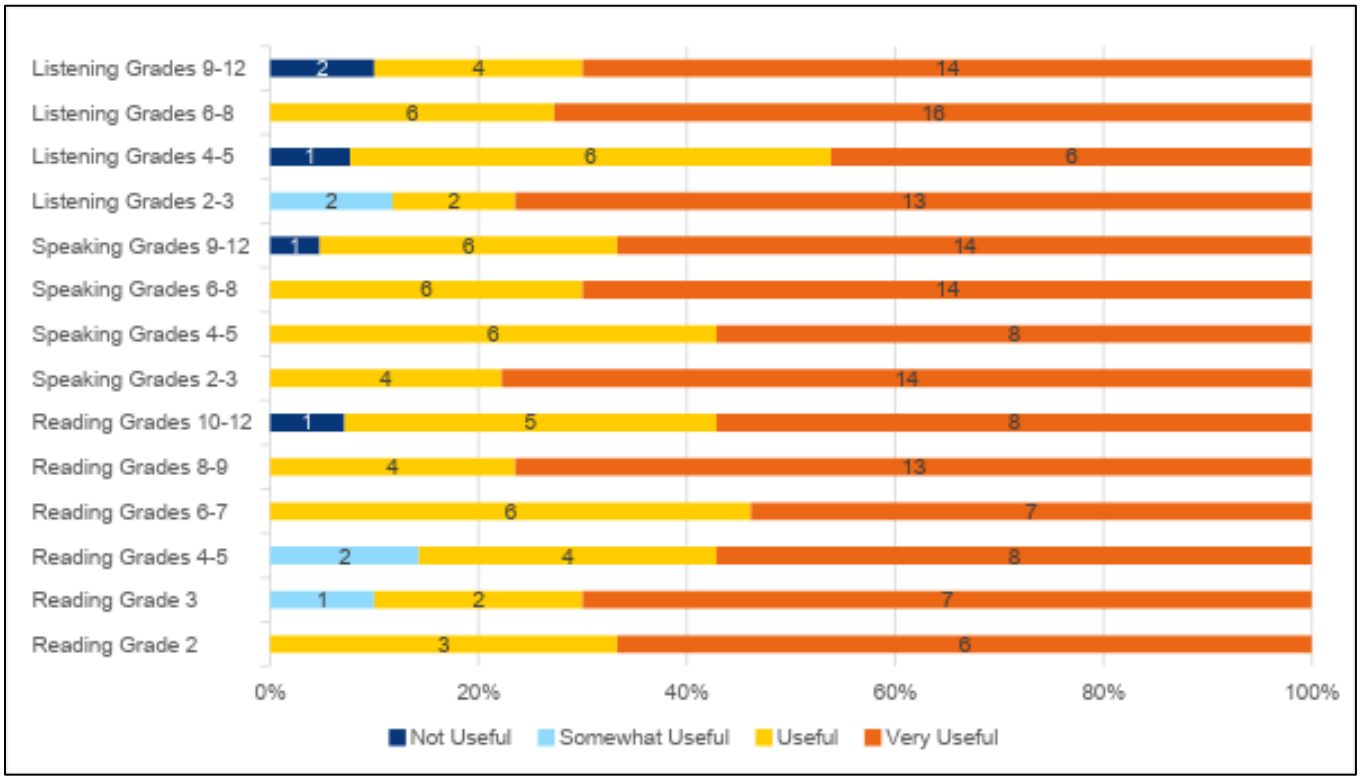
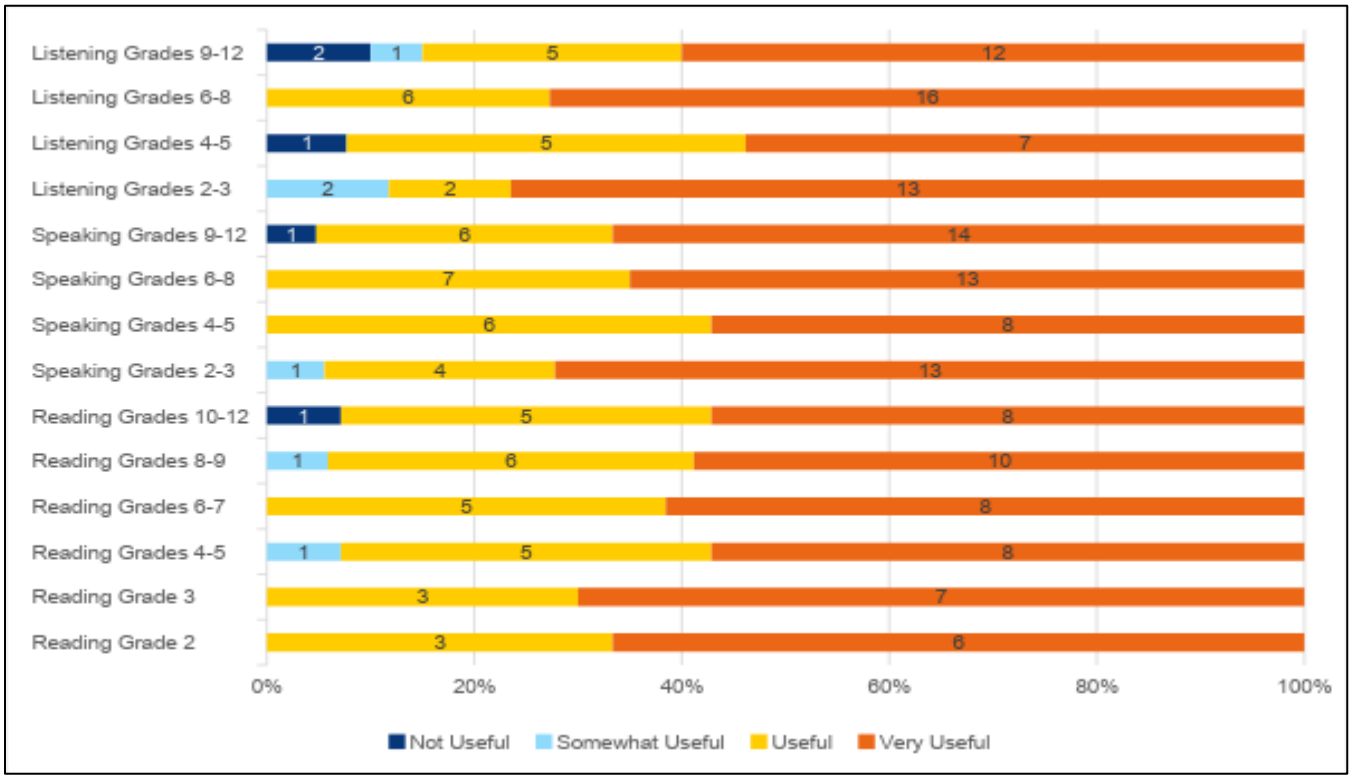


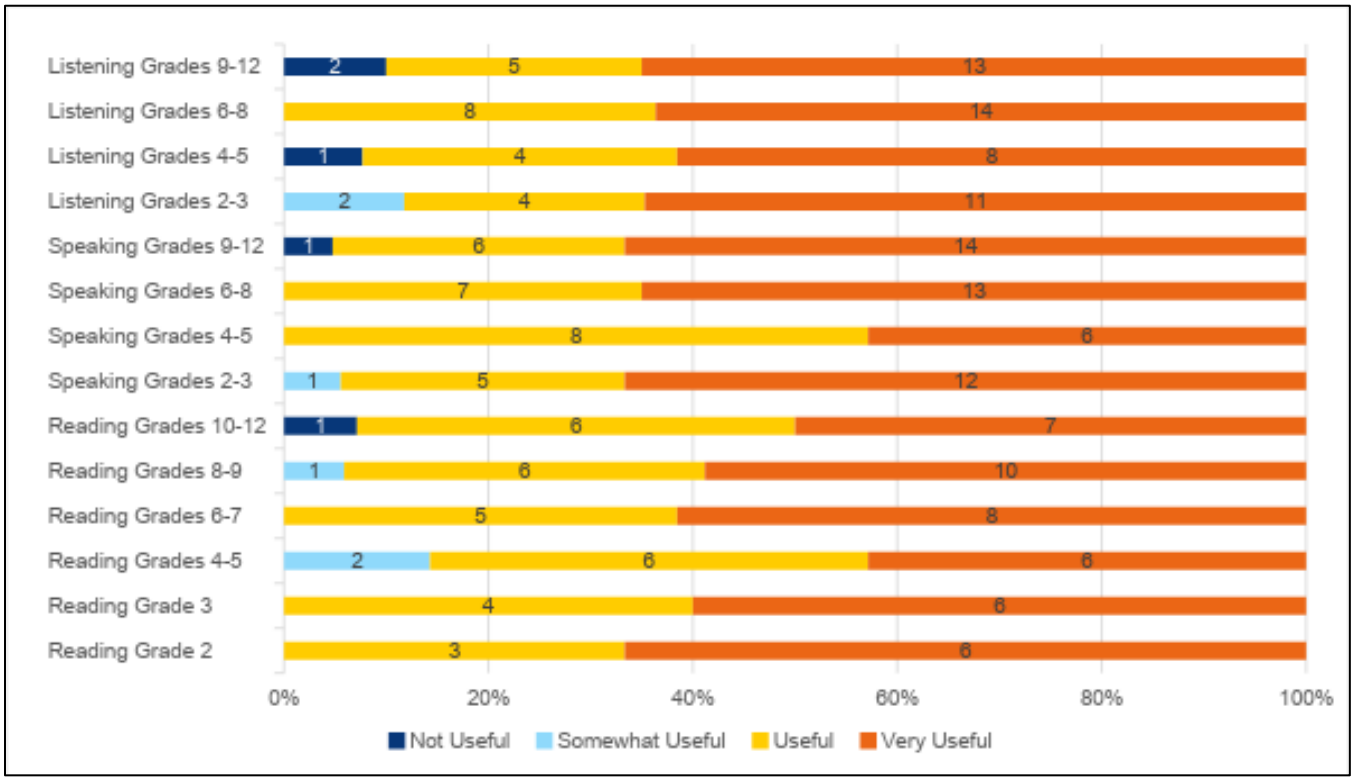
Table-level statistics after Rounds 1 and 2



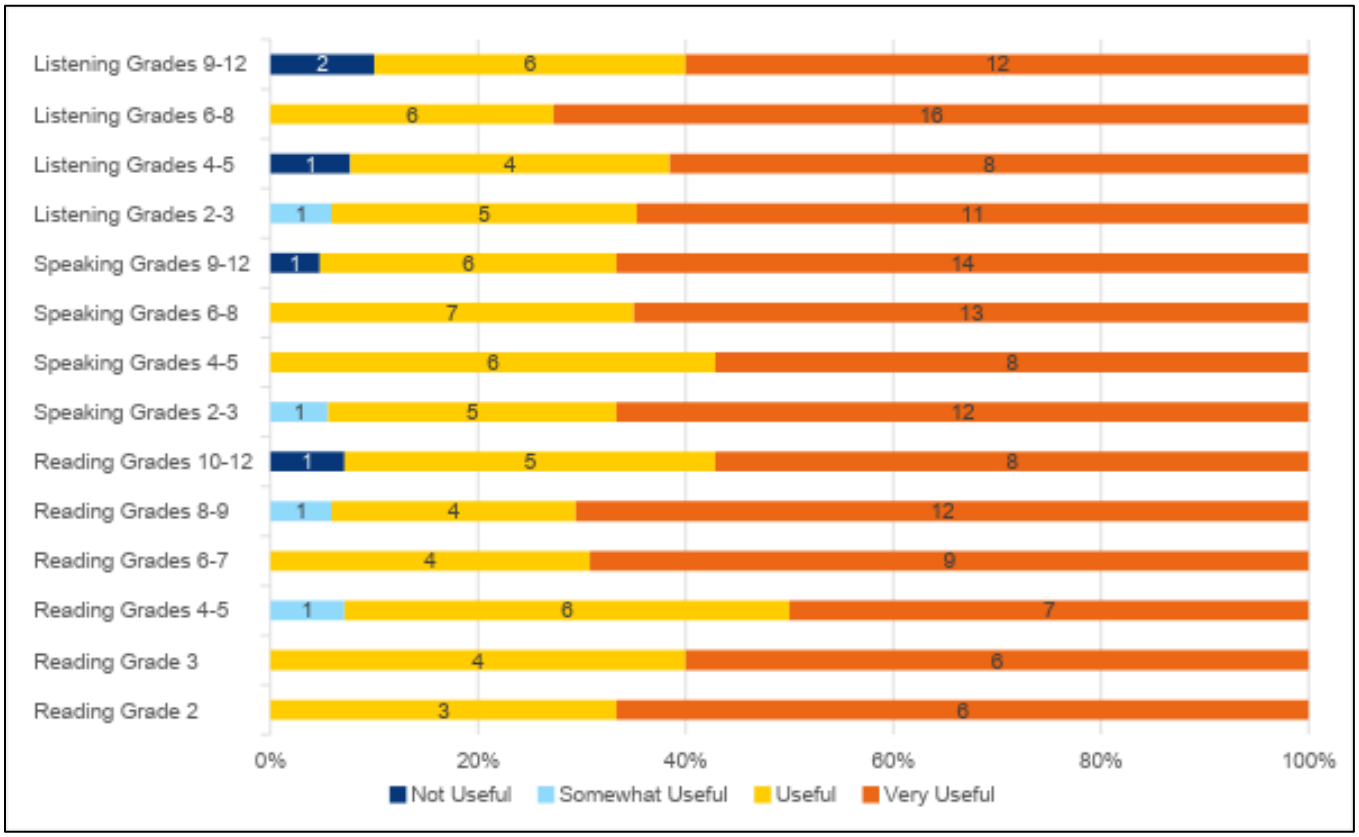
Committee-level statistics after Round 2



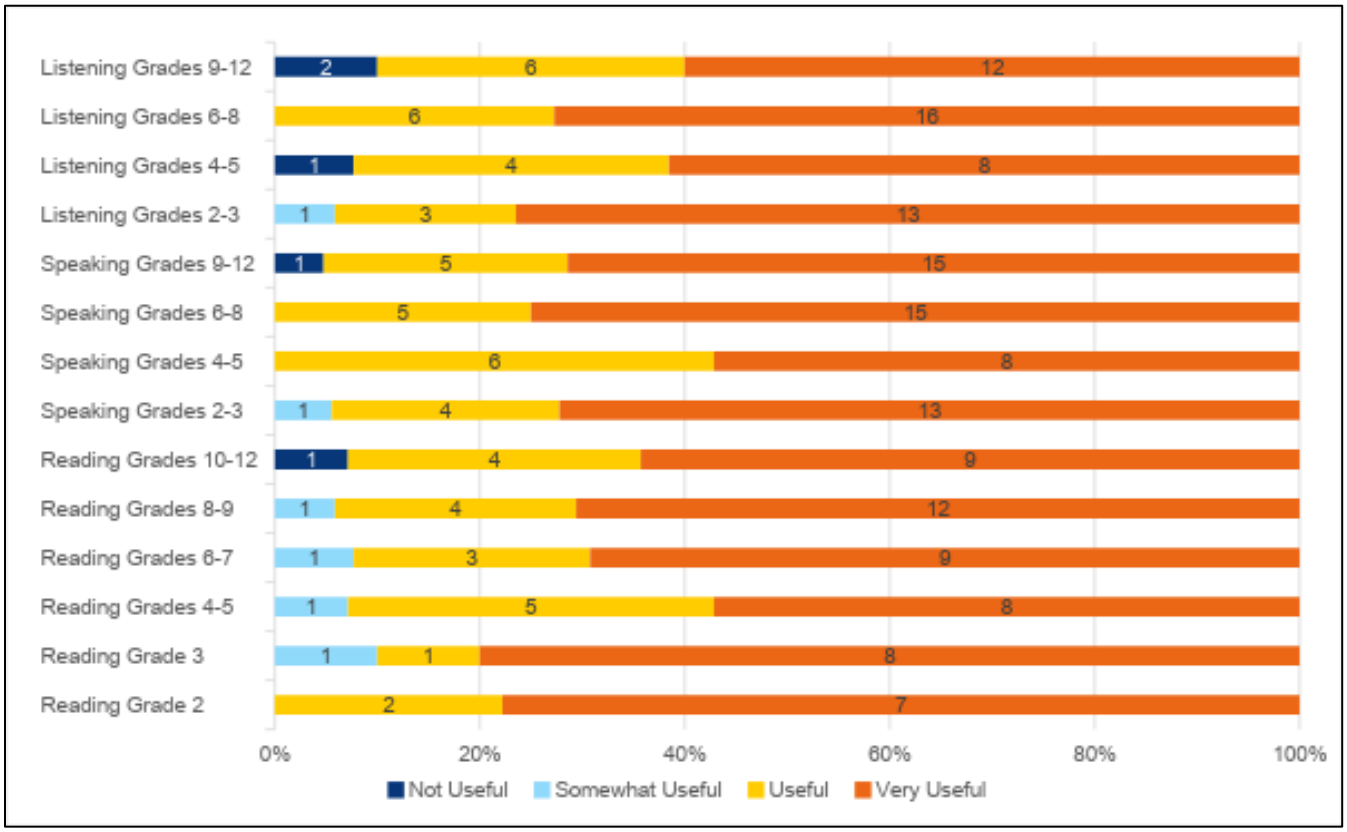
Panelist agreement data provided after Round 1



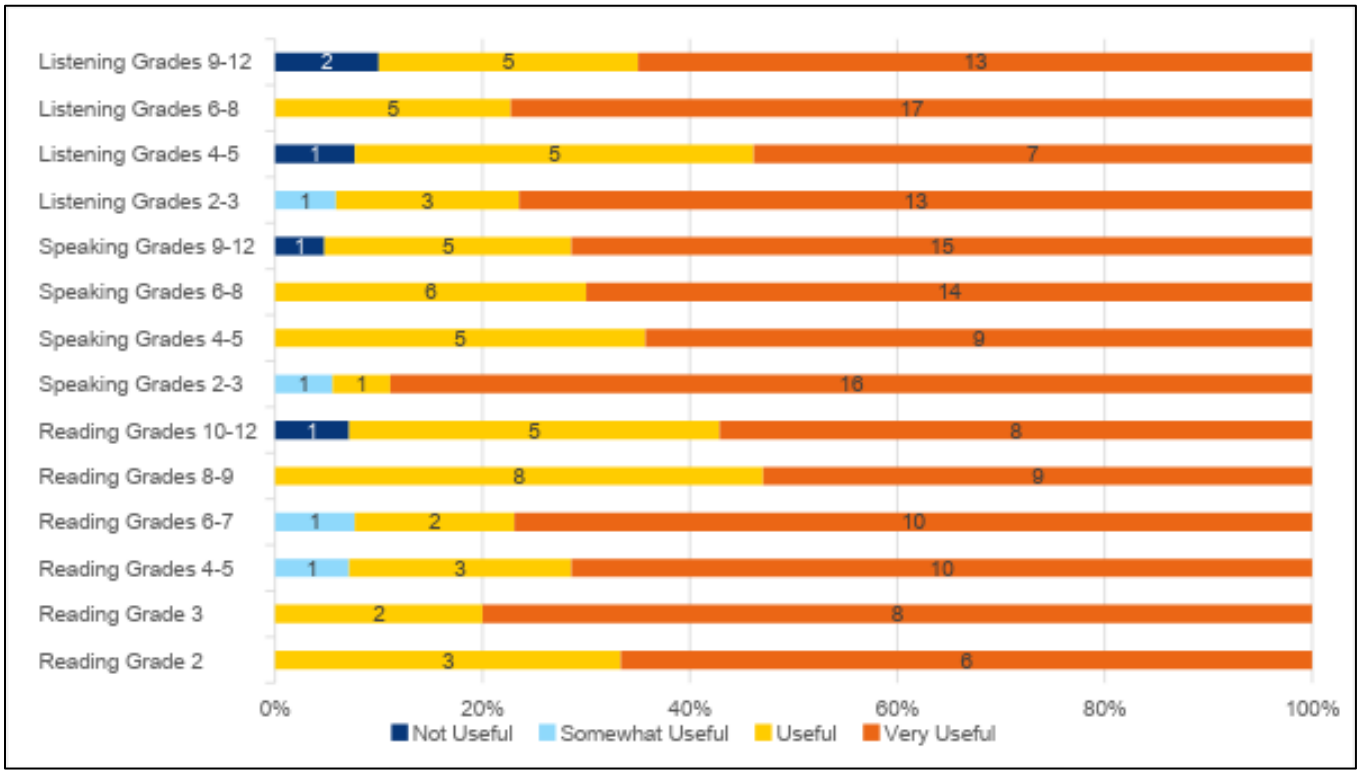
Panelist agreement data provided after Round 2



Impact data after Round 2

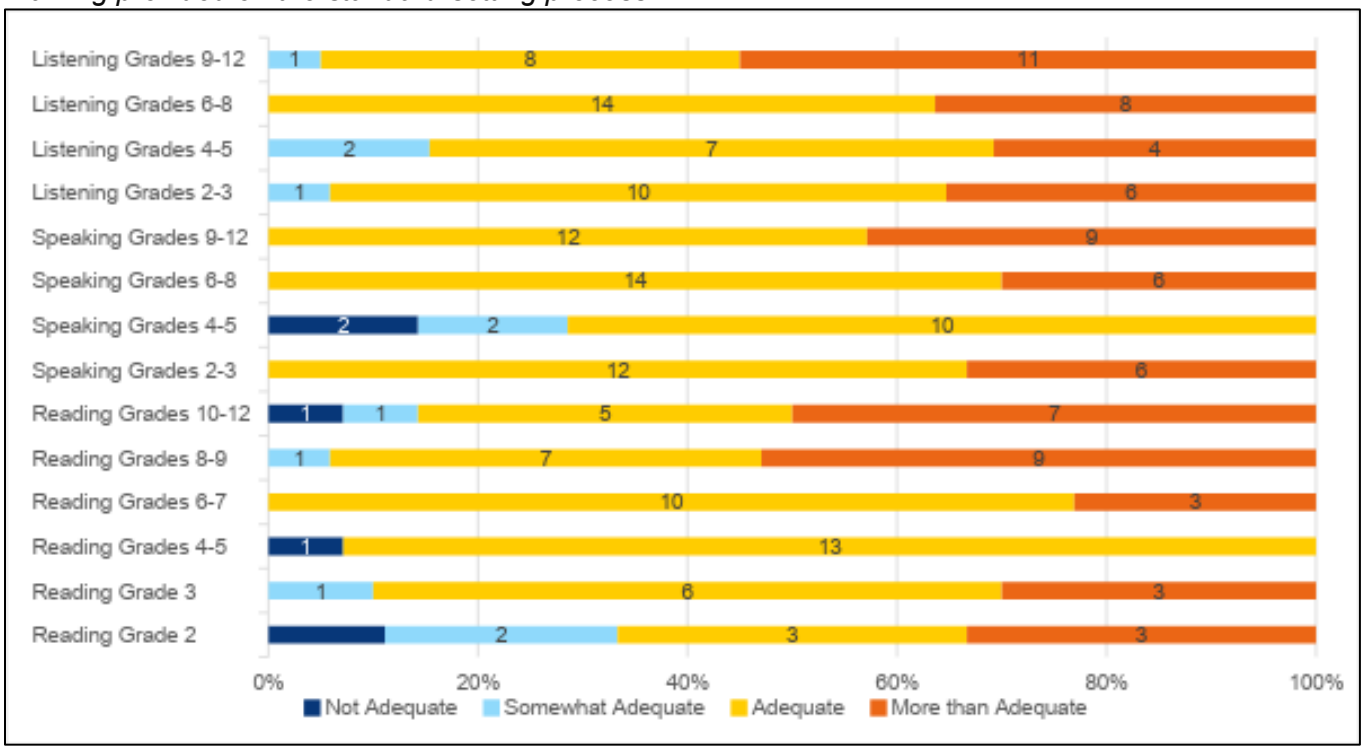


Discussion after each judgment round

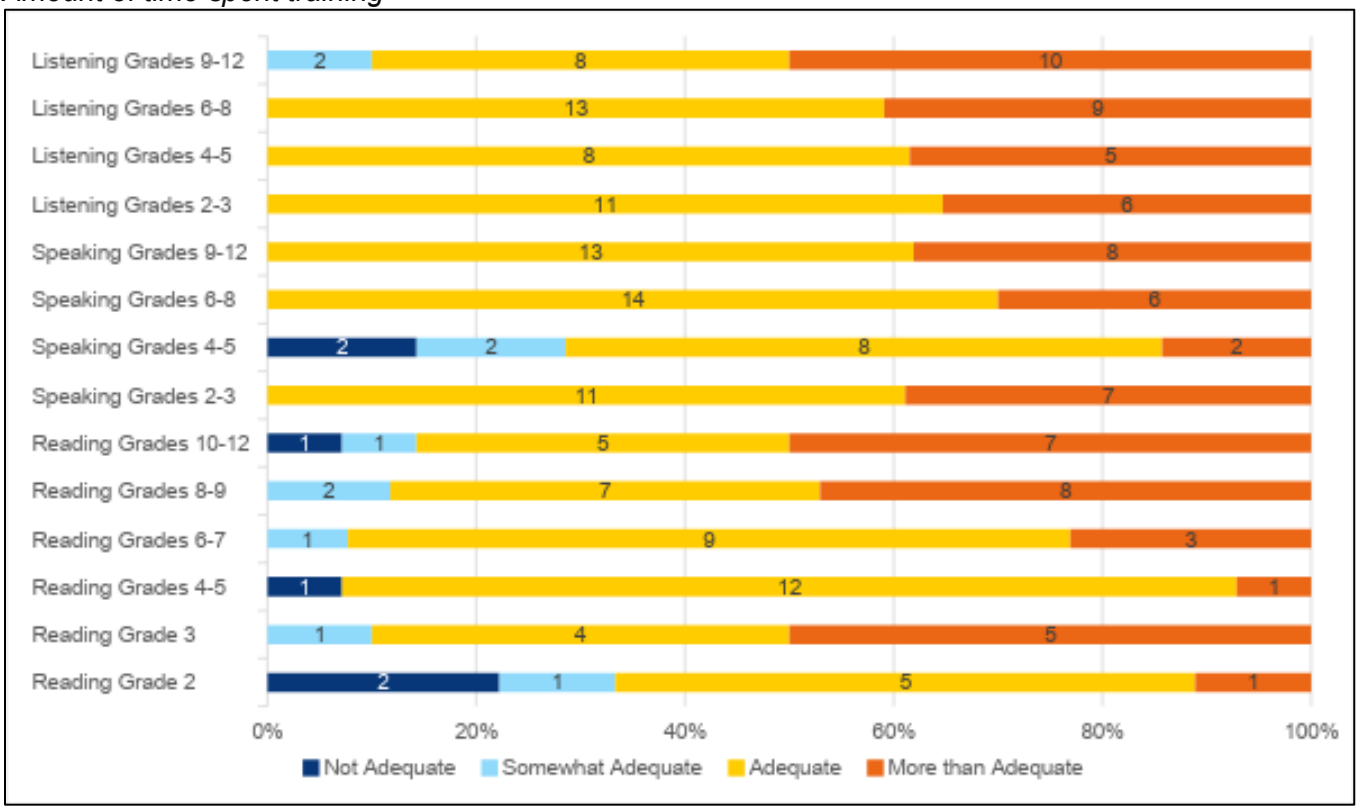


Question 3: How adequate were the following elements of the session?

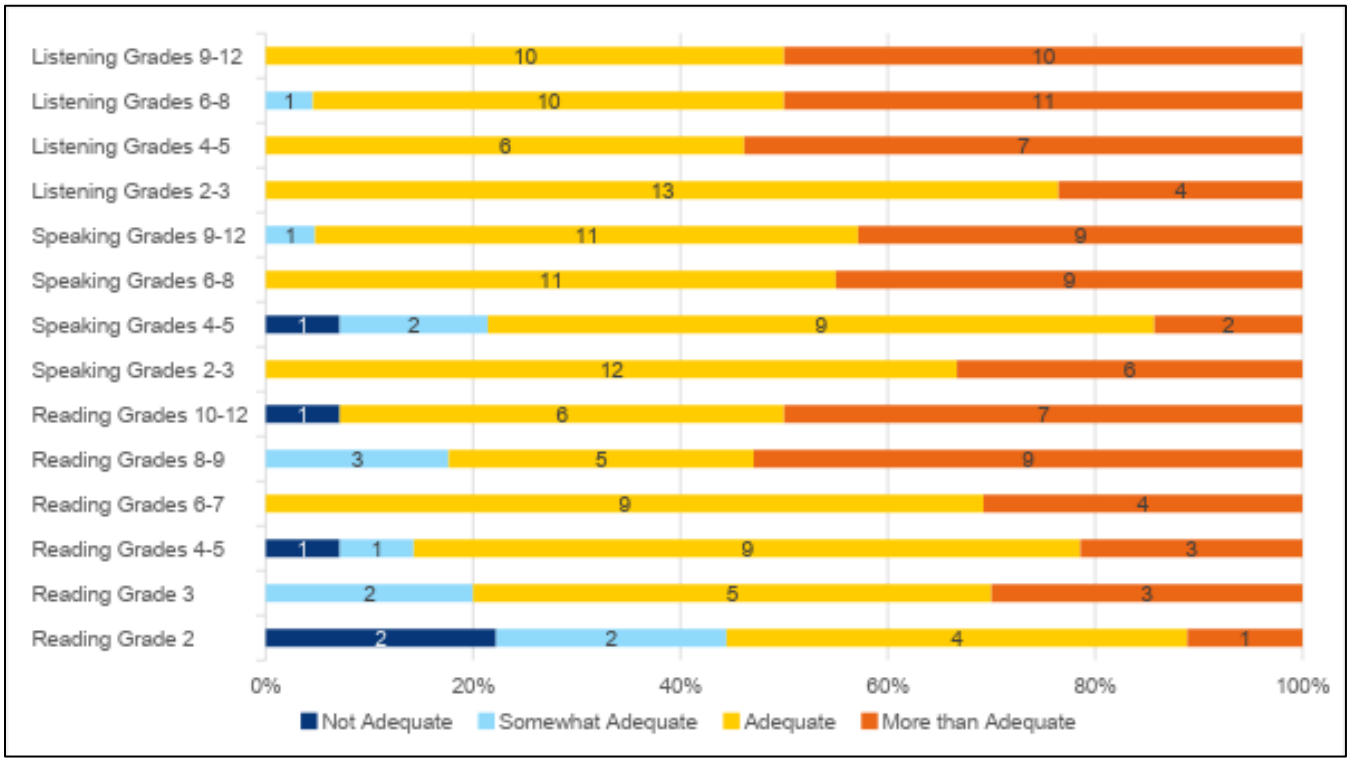
Training provided on the standard-setting process



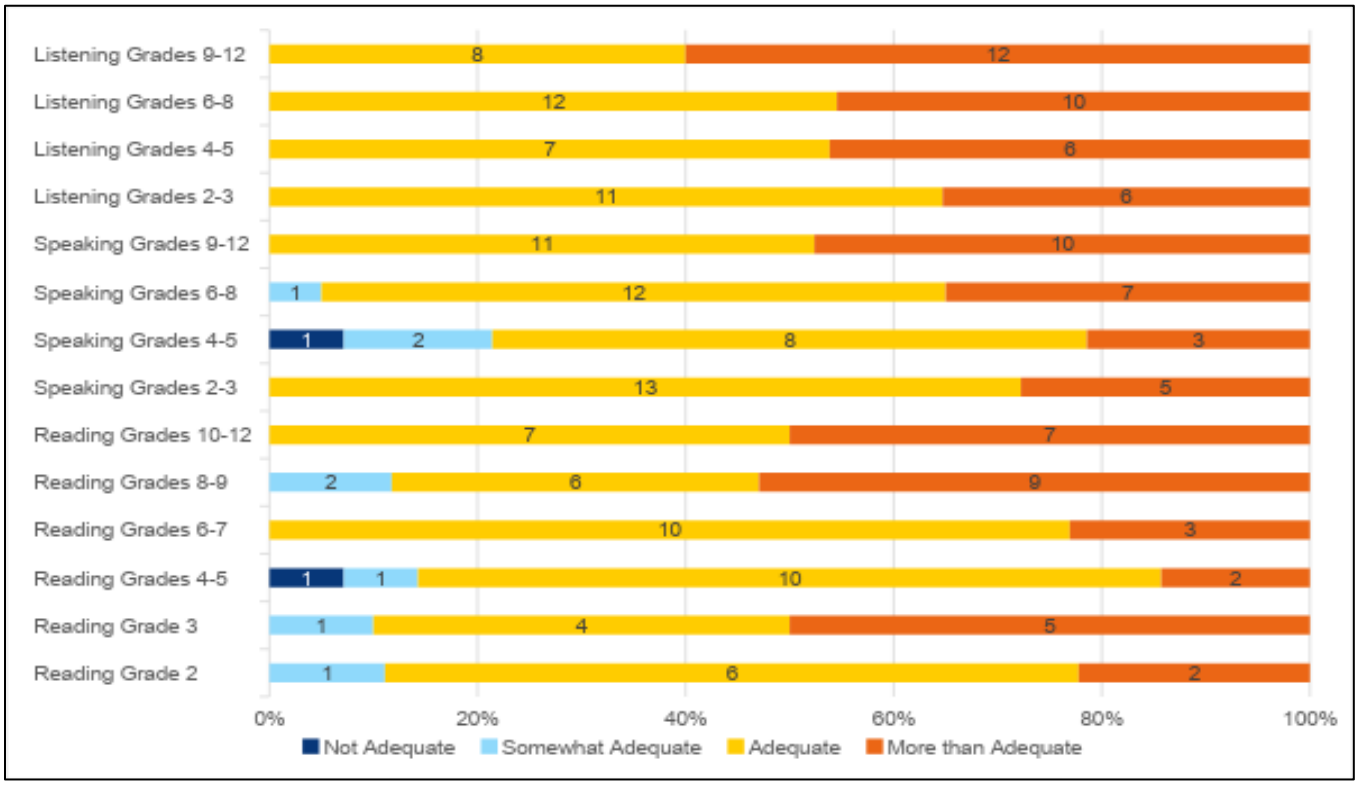
Amount of time spent training



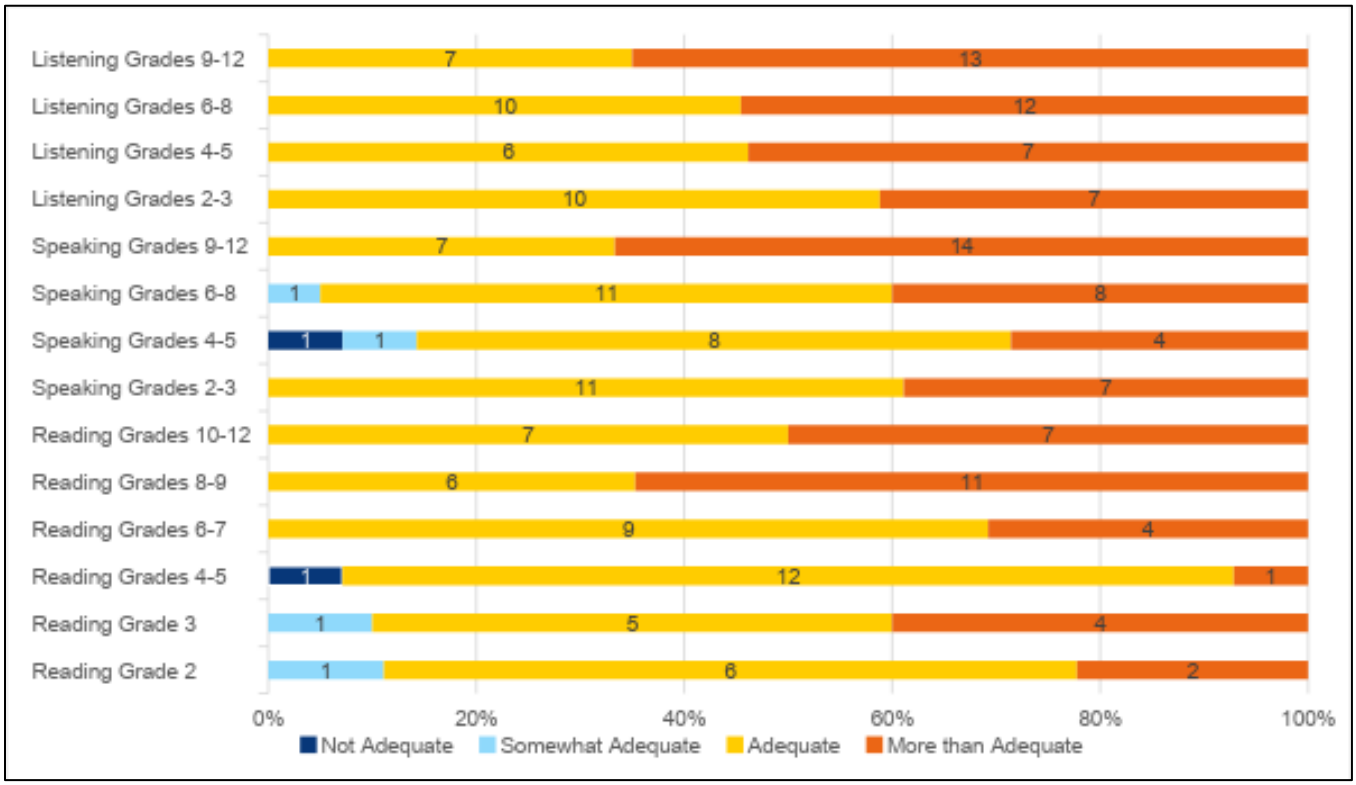
Total amount of time to review and discuss borderline proficiency level descriptors



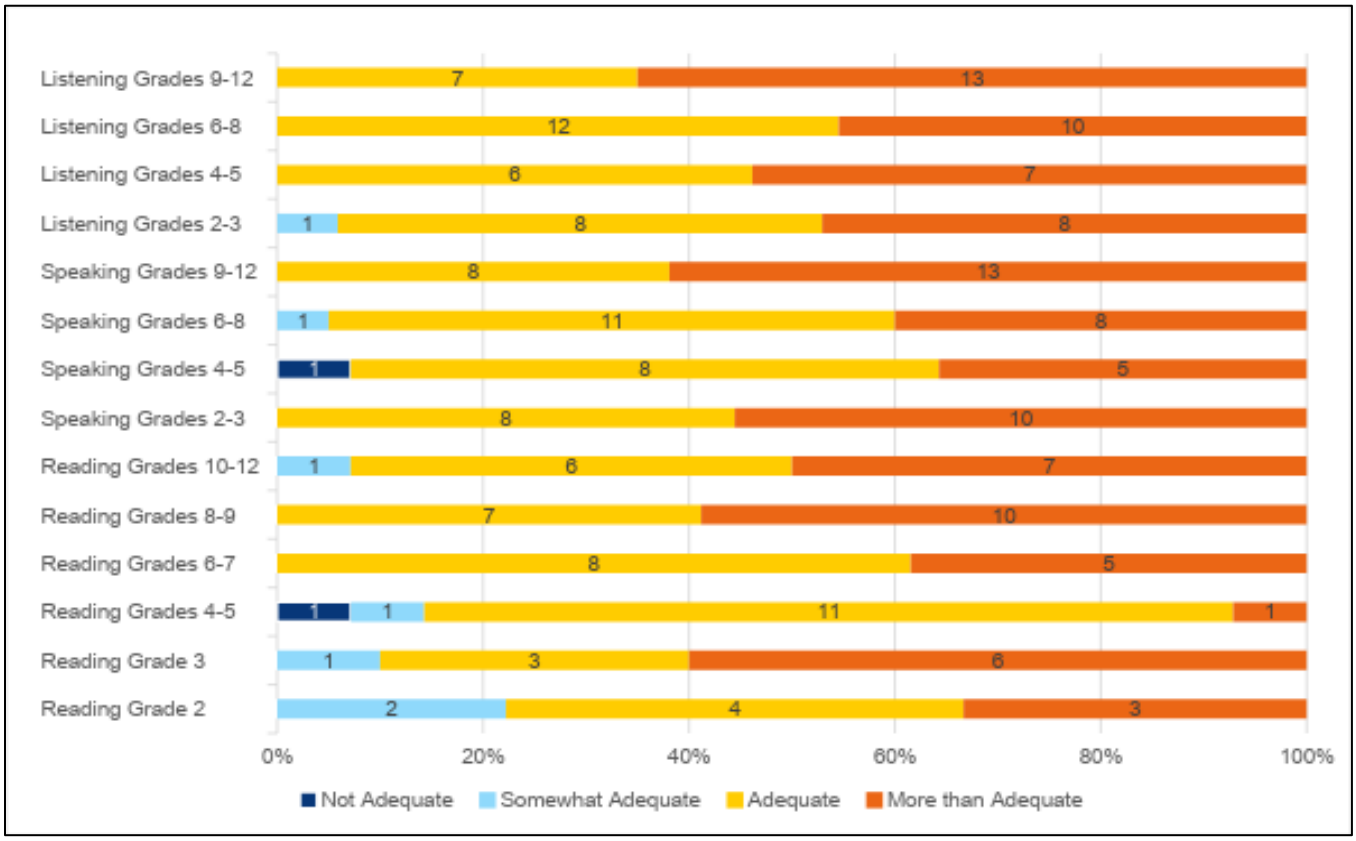
Total amount of time to discuss the practice judgments



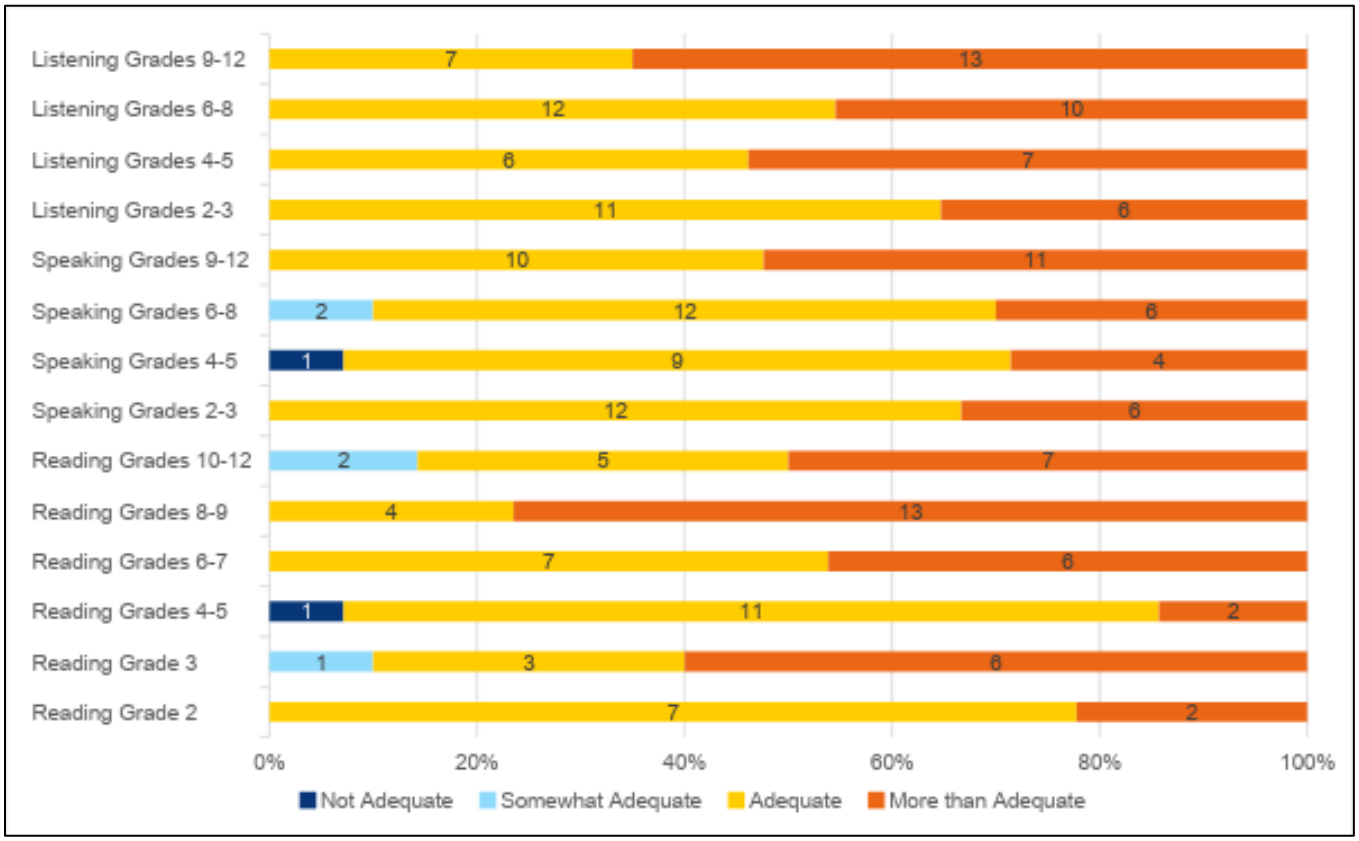
Amount of time to make judgments



Visual presentation of the feedback provided

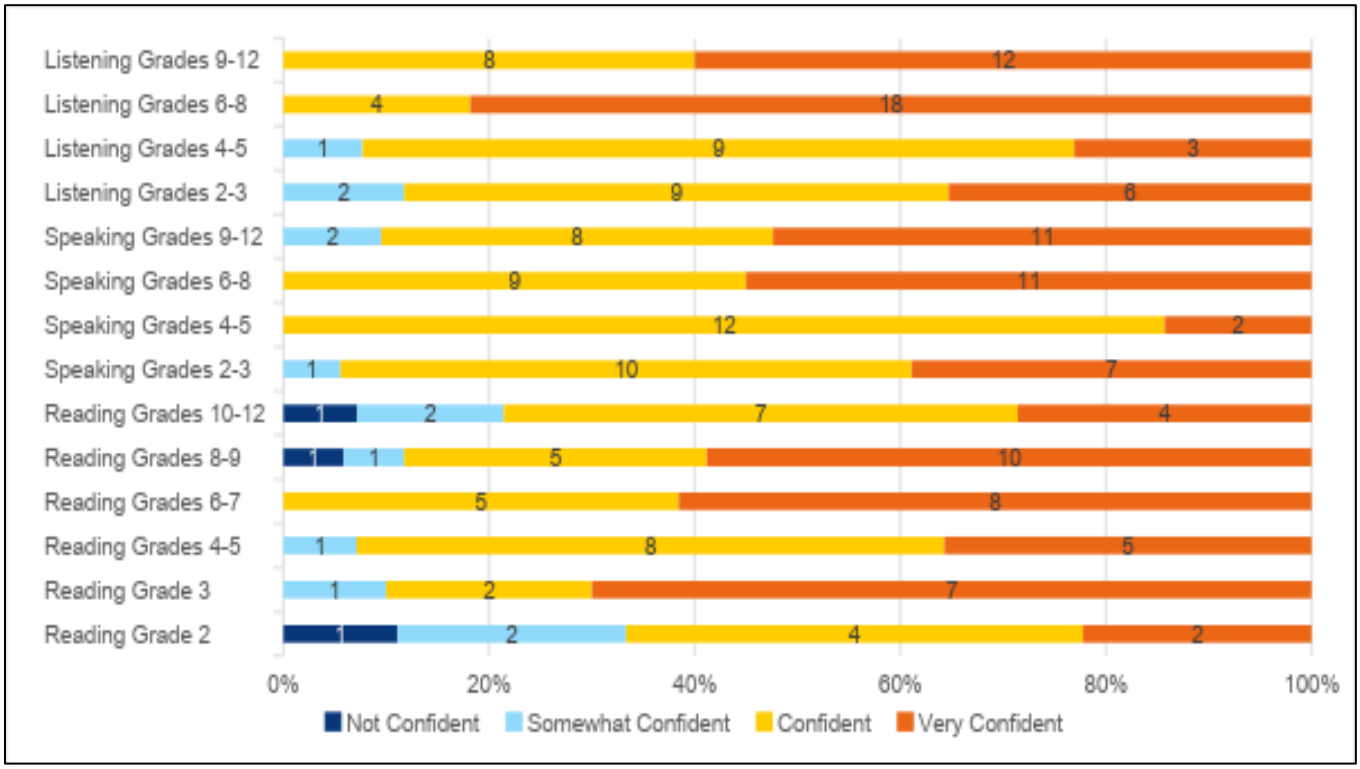


Number of judgment rounds

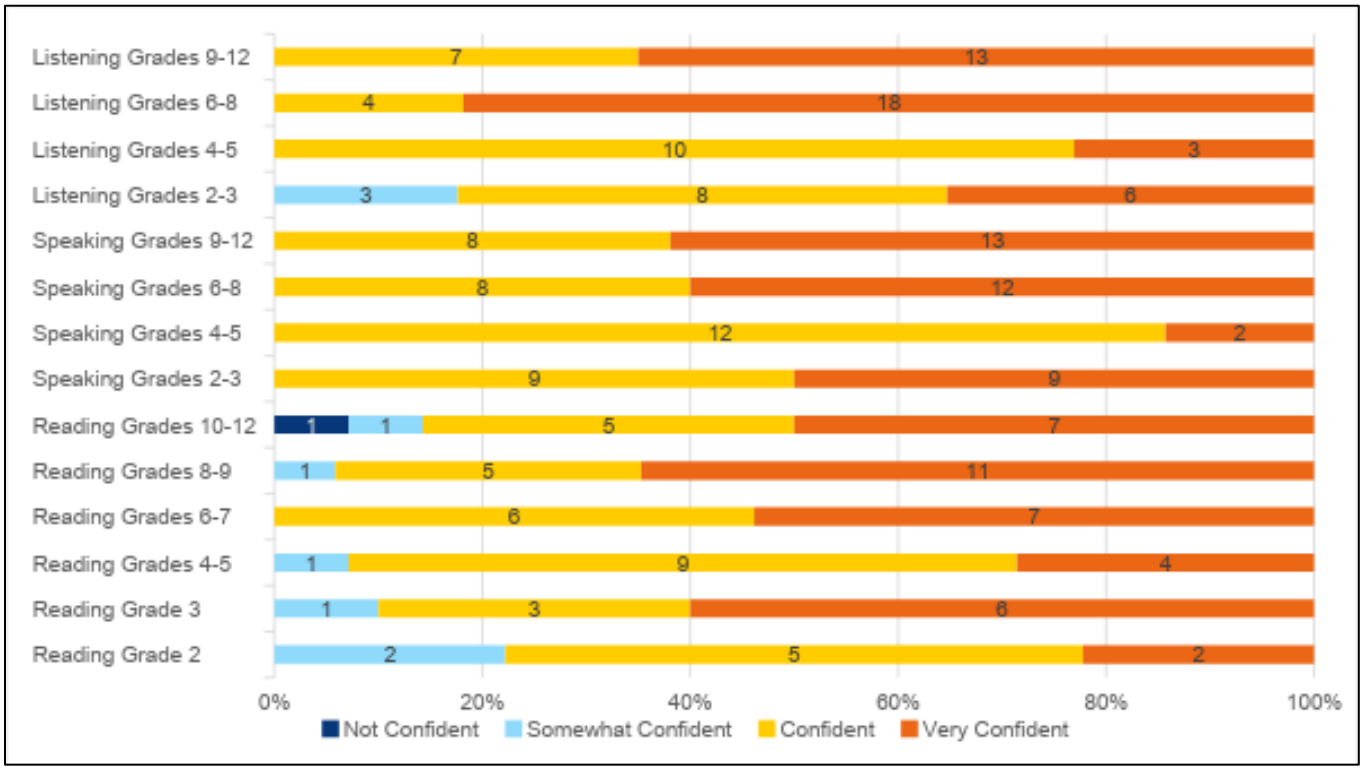


Question 4: How confident do you feel that the proficiency level descriptors (PLDs) for reading, speaking, and listening are reasonable for each student proficiency level?

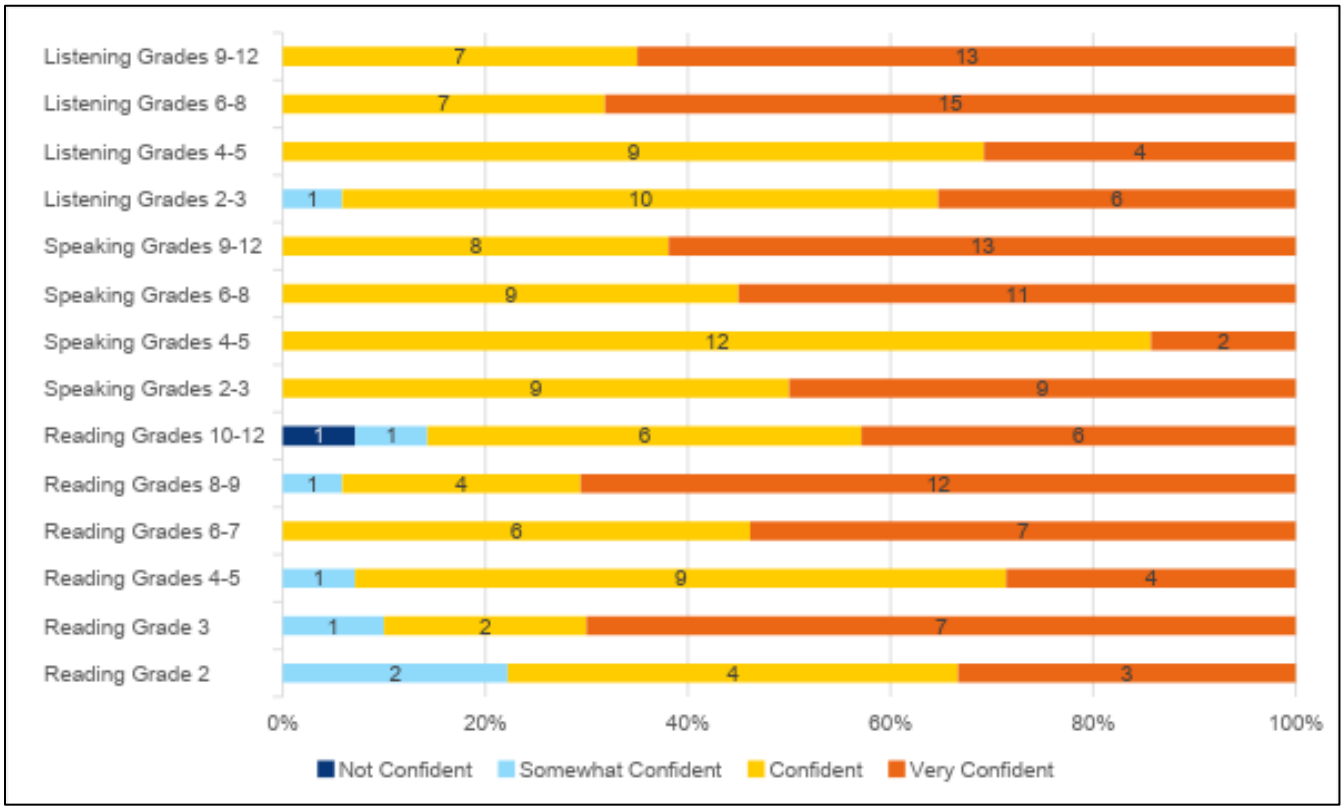
Intermediate



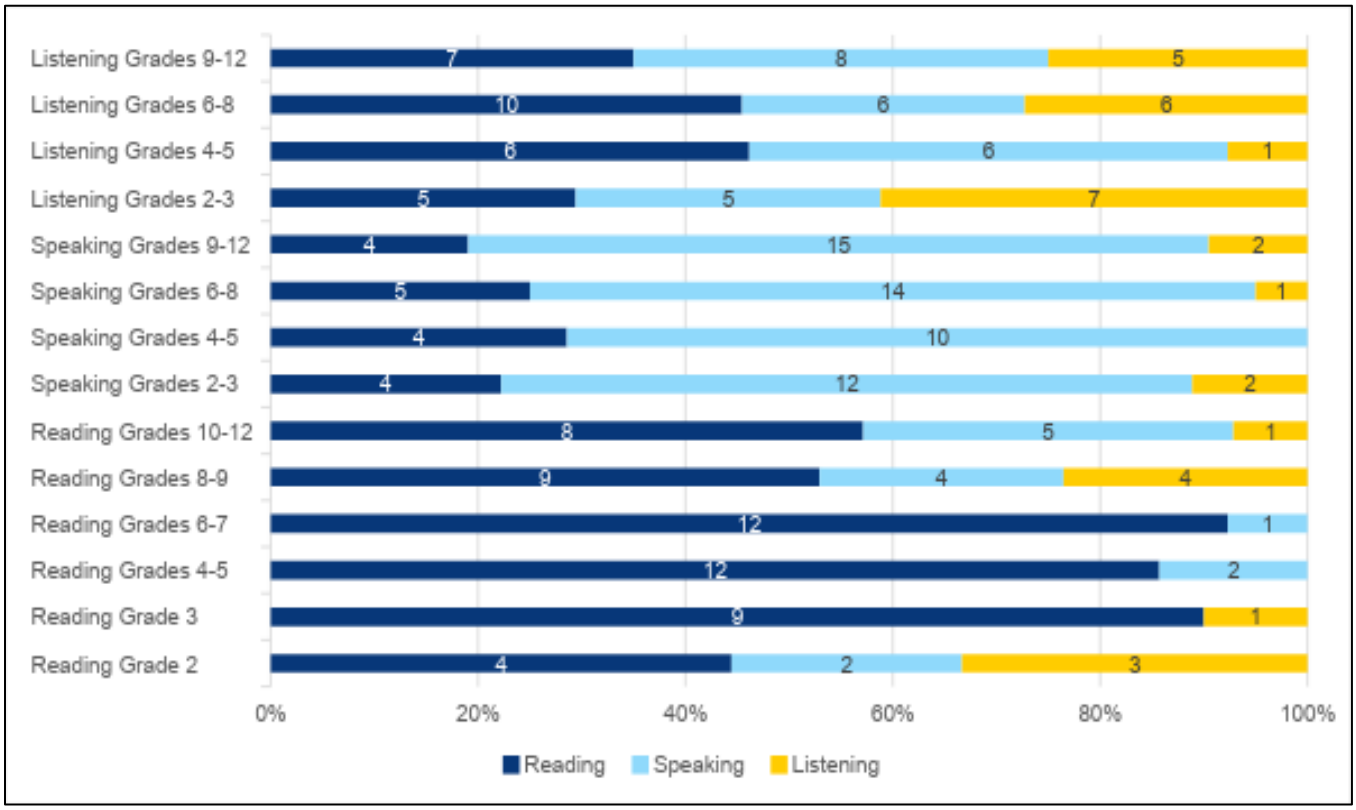
Advanced


Listening Grades 9-12
7
13
Listening Grades 6-8
4
18
Listening Grades 4-5
10
3
Listening Grades 2-3
3
8
6
Speaking Grades 9-12
8
13
Speaking Grades 6-8
8
12
Speaking Grades 4-5
12
2
Speaking Grades 2-3
9
9
Reading Grades 10-12
1
1
5
7
Reading Grades 8-9
1
5
11
Reading Grades 6-7
6
7
Reading Grades 4-5
1
9
4
Reading Grade 3
1
3
6
Reading Grade 2
2
5
2
0%
20%
40%
60%
80%
100%
Not Confident
Somewhat Confident
Confident
Very Confident

Advanced High



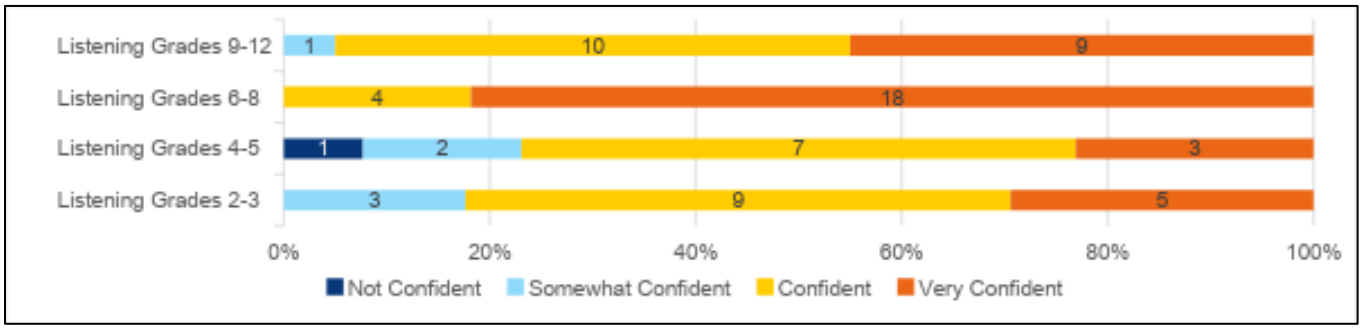
Question 5: Which borderline proficiency level group did you participate in?



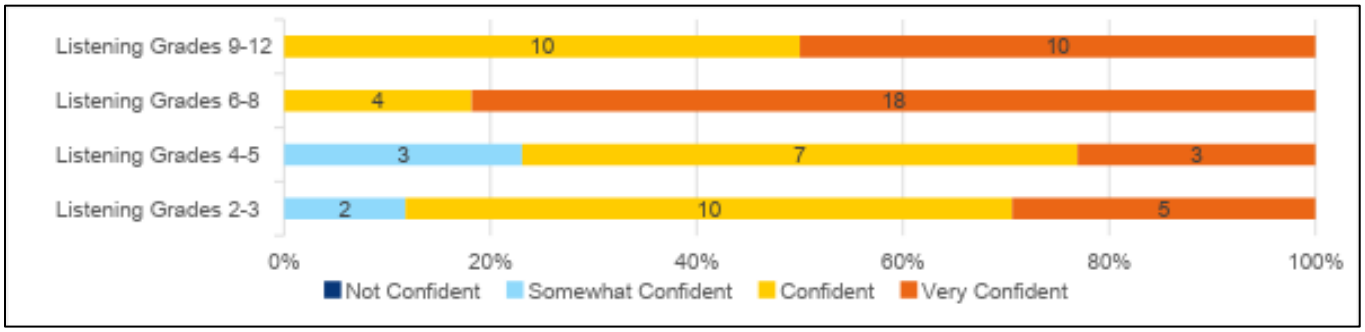
The following questions ask about overall confidence in the final cut score recommendations:

Question 6: How confident do you feel the final cut score recommendations for listening represent appropriate levels of student proficiency?

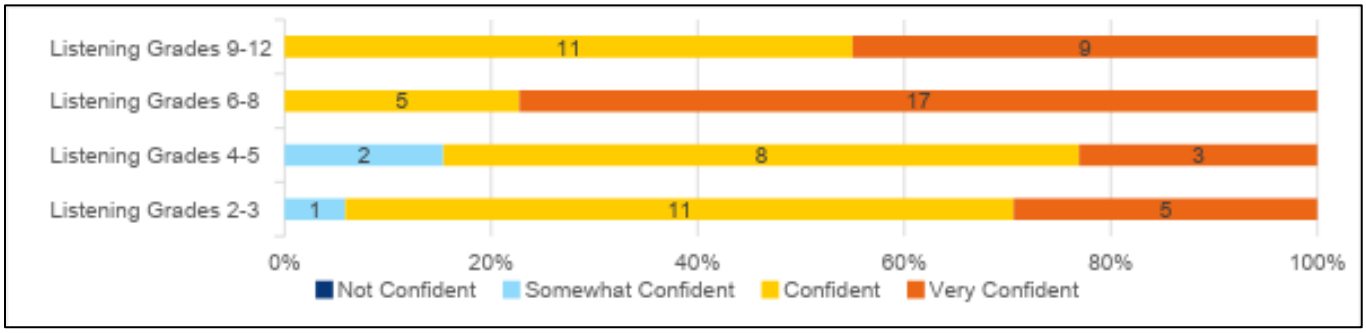
Intermediate:



Advanced:

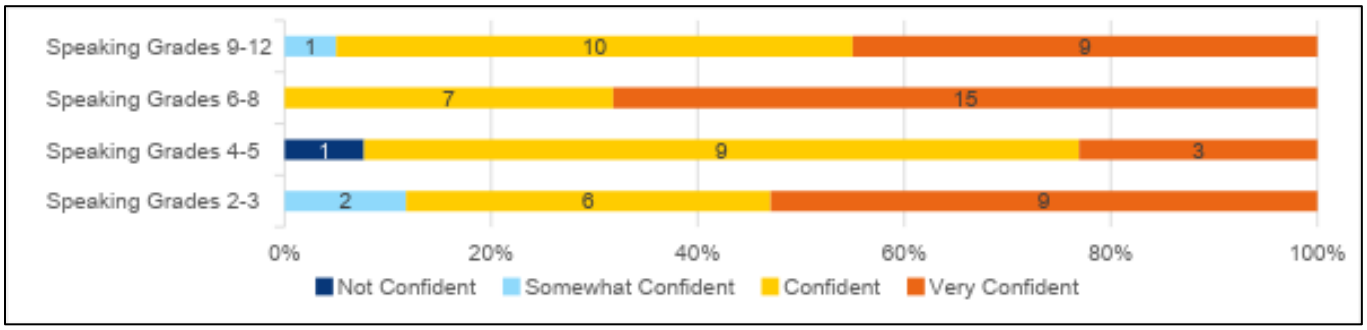


Advanced High:

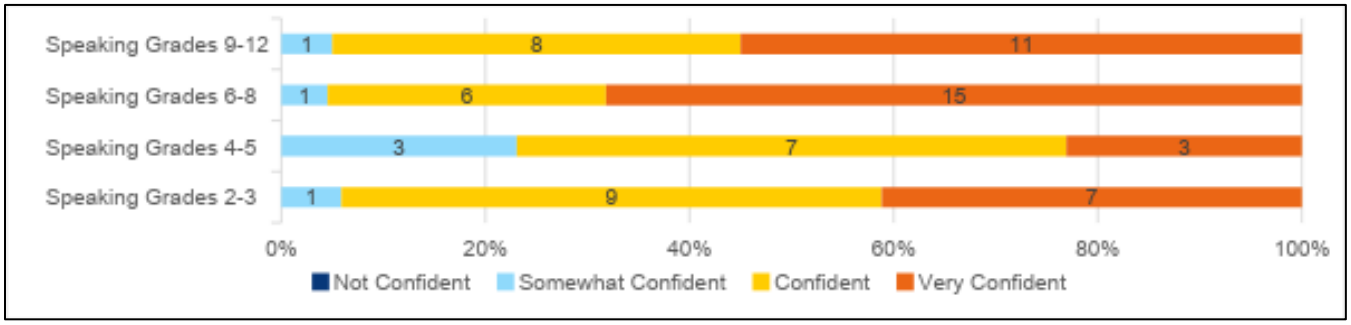


Question 7: How confident do you feel the final cut score recommendations for speaking represent appropriate levels of student proficiency?

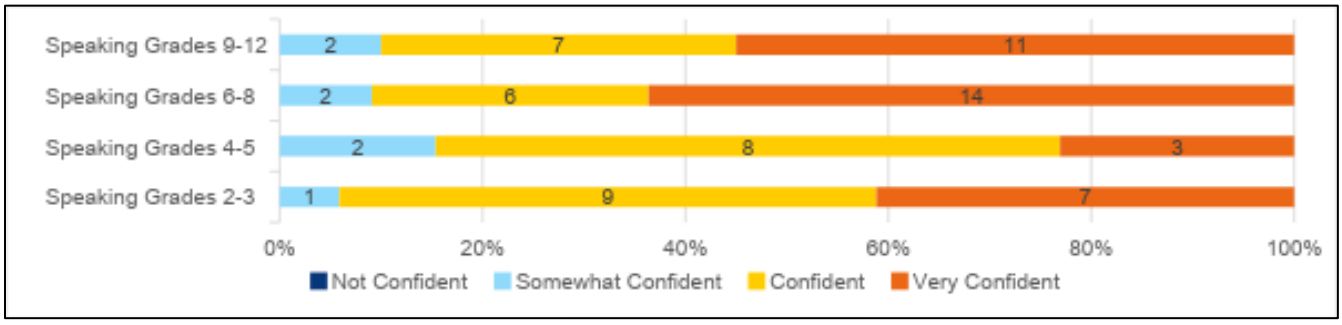
Intermediate:



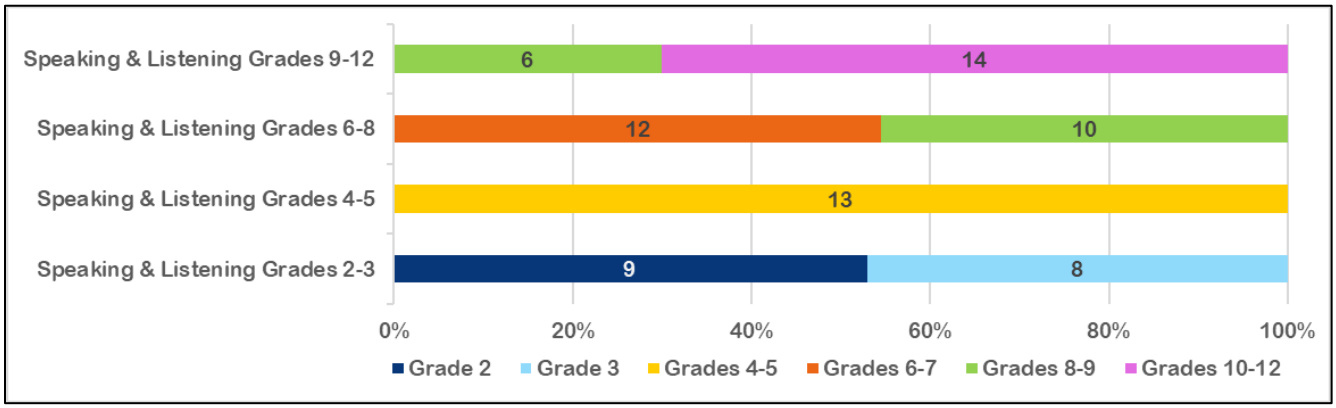
Advanced:



Advanced High:

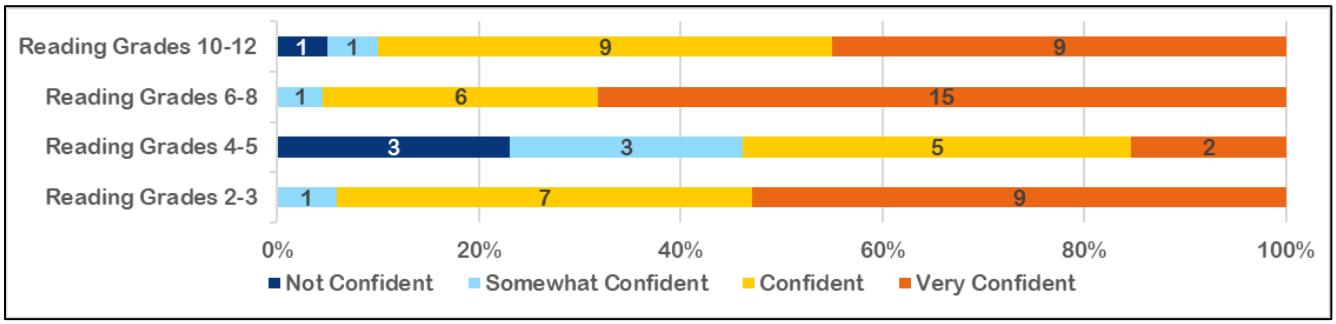


Question 8: Which speaking & listening group did you participate in?

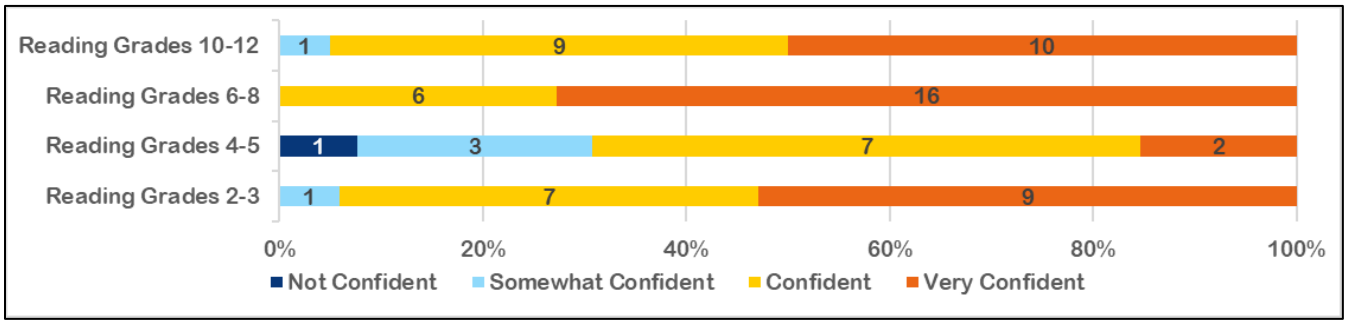


Question 9: How confident do you feel the final cut score recommendations for reading represent appropriate levels of student proficiency?

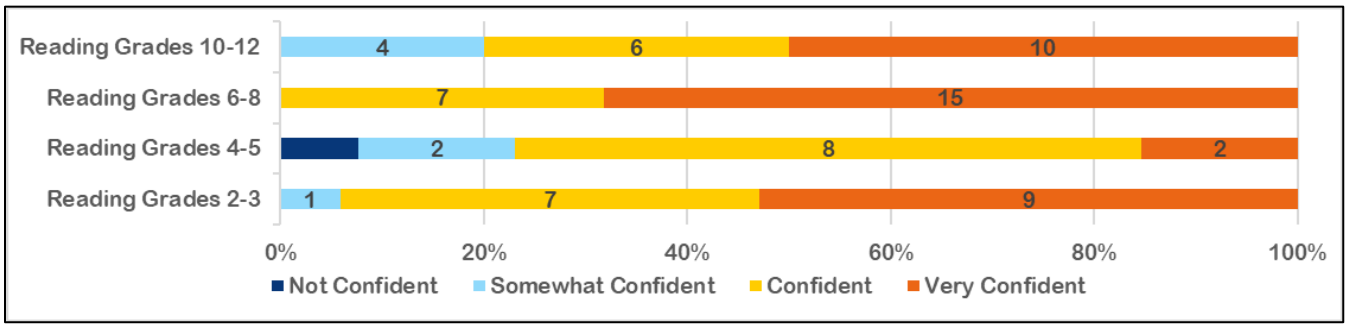
Intermediate:



Advanced:



Advanced High:

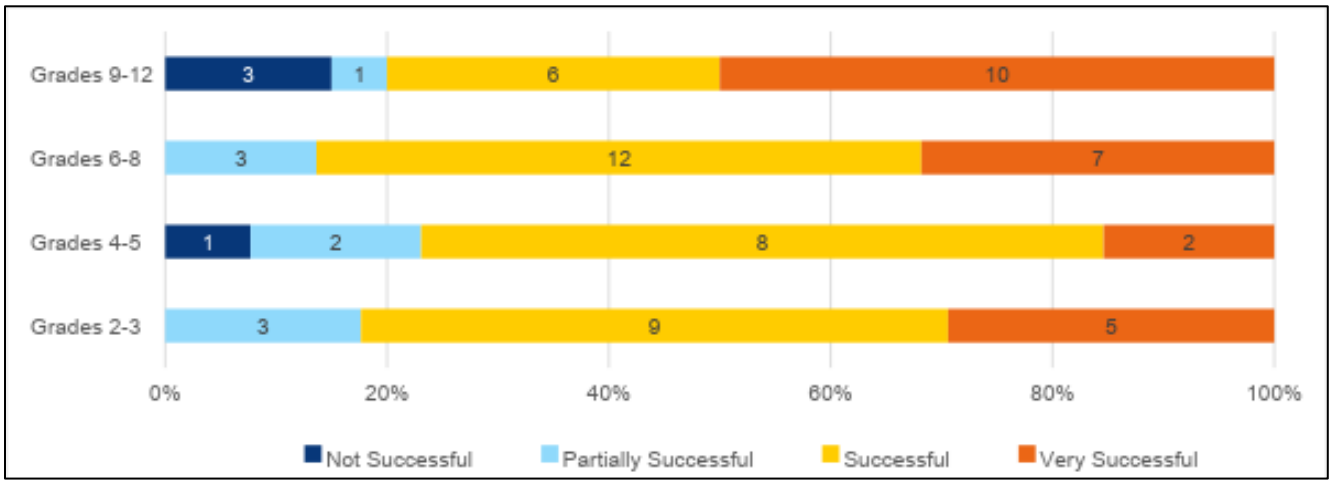


The following questions ask about overall experience at the TELPAS standard setting:

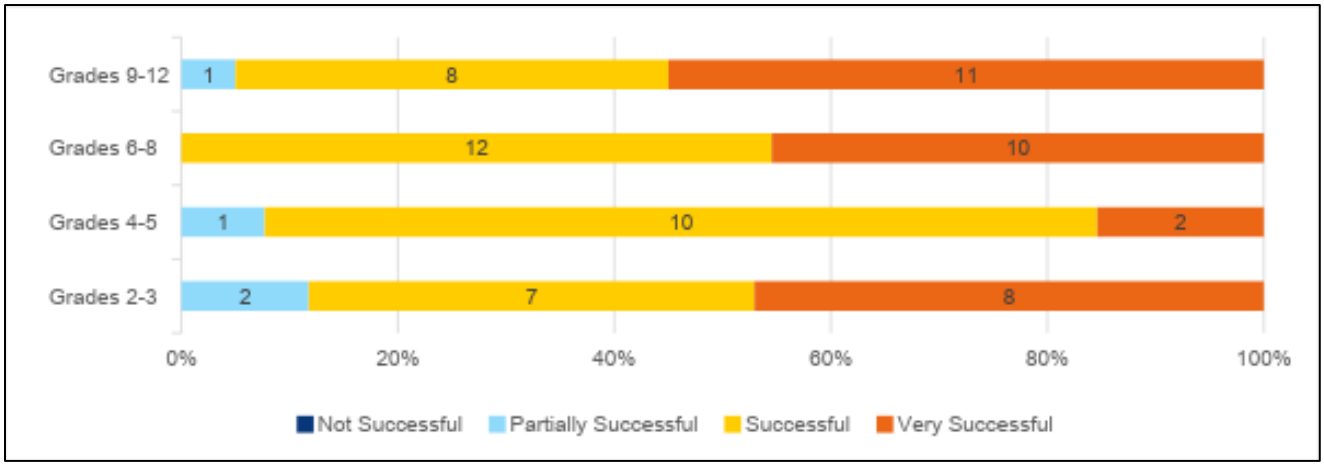
**Note*: Questions 10-13 were asked as part of the final TELPAS standard-setting evaluation, which occurred during the Listening and Speaking meeting. Therefore, the grade bands for each graph associated with Questions 10-13 are representative of the Listening and Speaking groups.

Question 10: Select the option that best reflects your opinion about the level of success of the various components of the meeting in which you participated. The activities were designed to help you both understand the process and be supportive of the recommendations made by the committee.

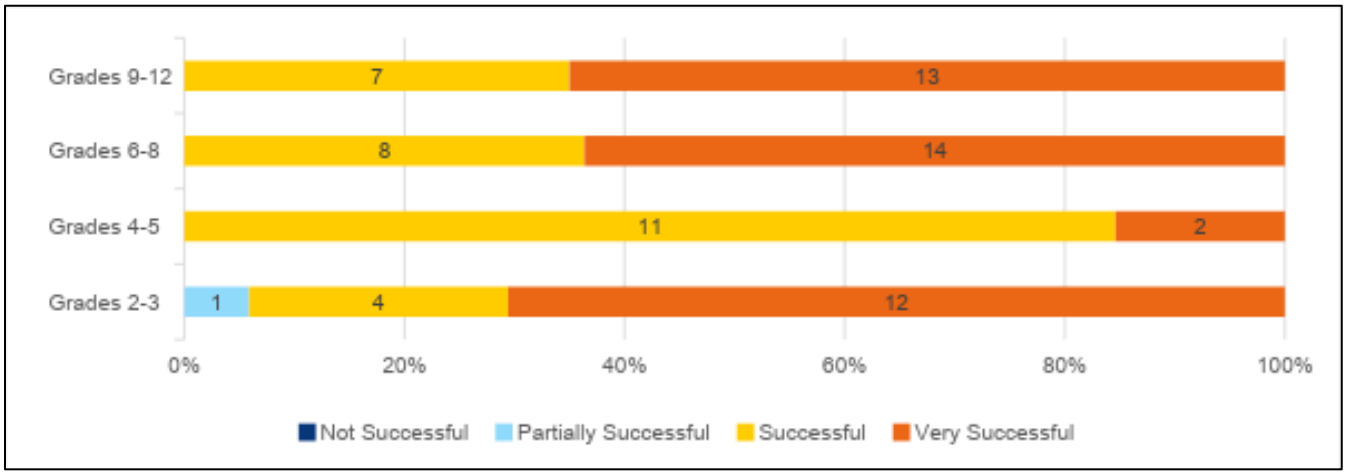
Meeting pre-work



General session training

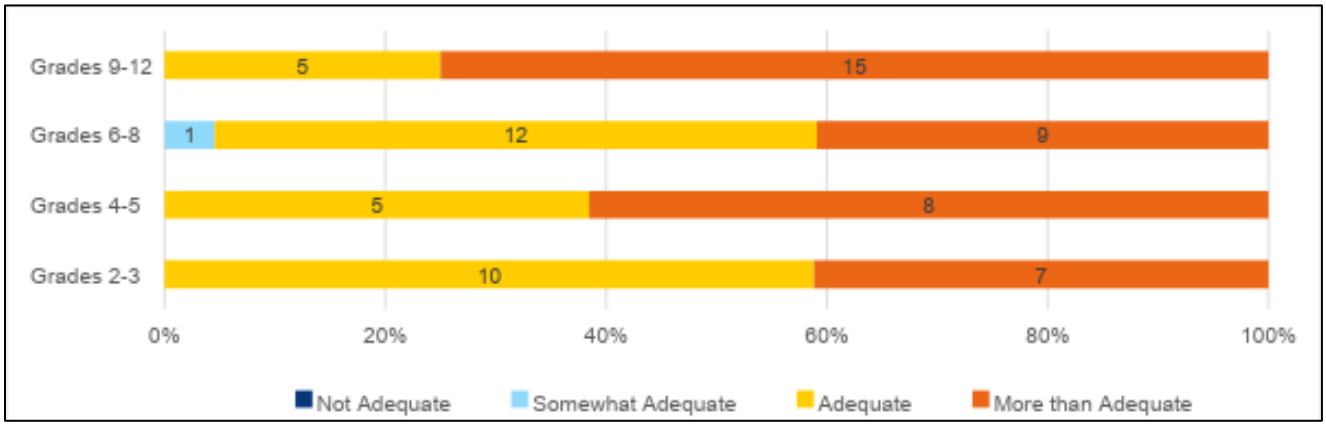


Breakout sessions

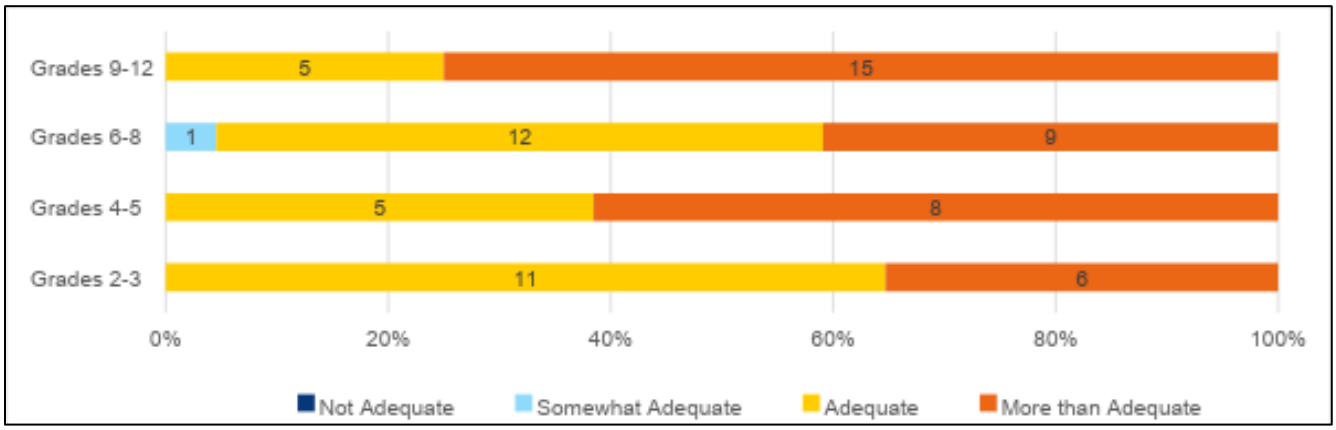


Question 11: How adequate were the following elements of the session?

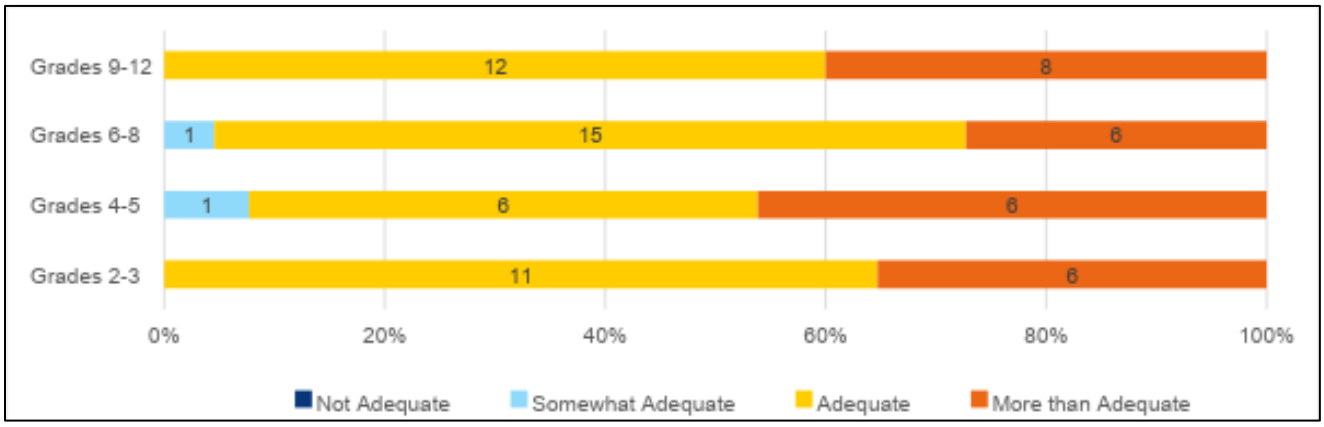
Facilities used for the general session



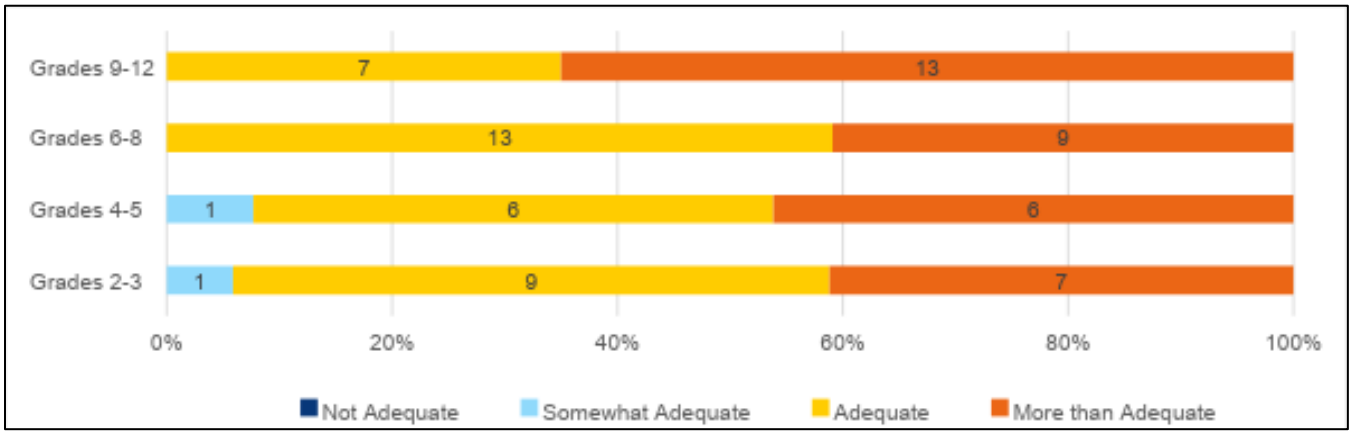
Facilities used for the breakout session



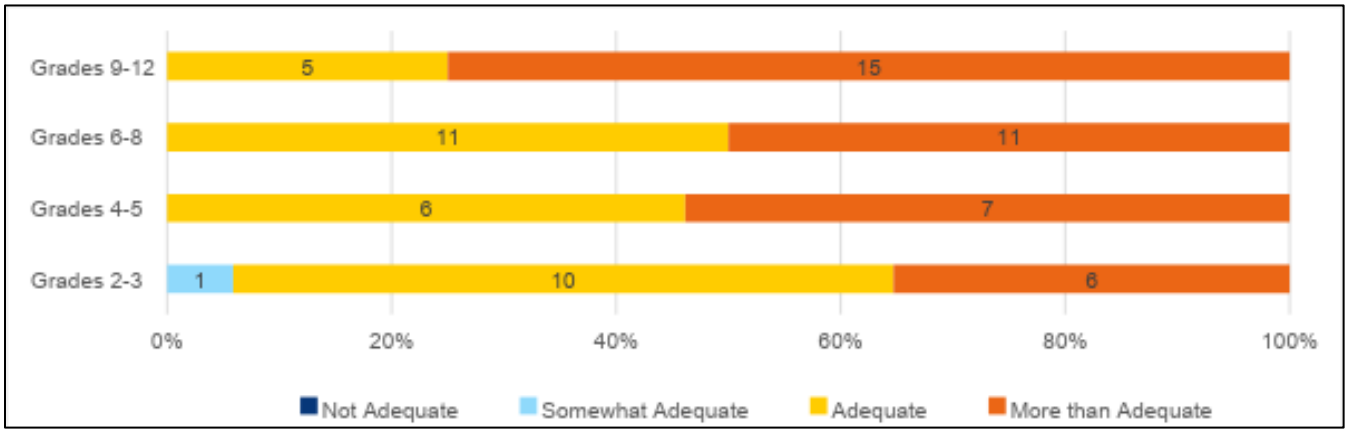
Computers used during the meetings



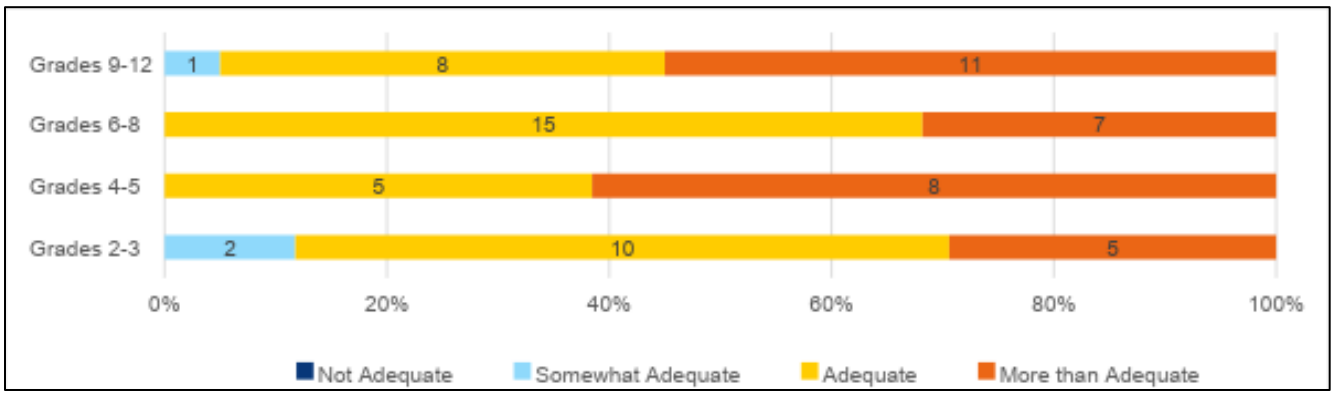
Moodle site for accessing materials and making judgments



Materials provided in the folder

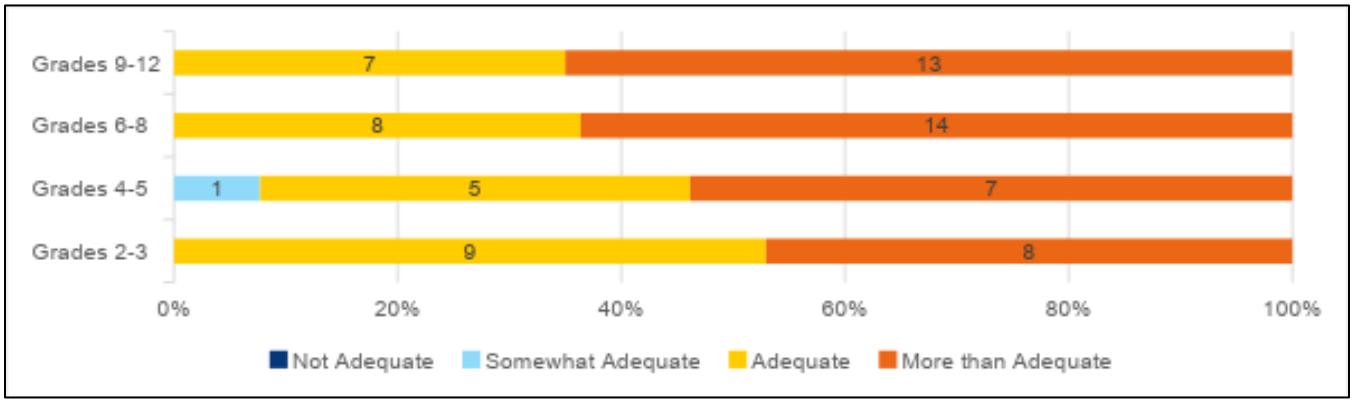


Work space in table groups during meeting

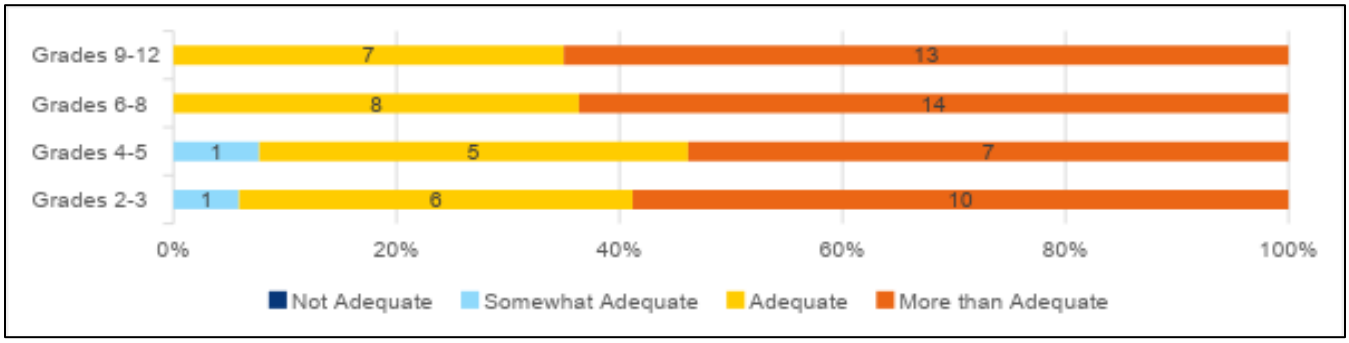


Question 12: Did you have adequate opportunities during the session to

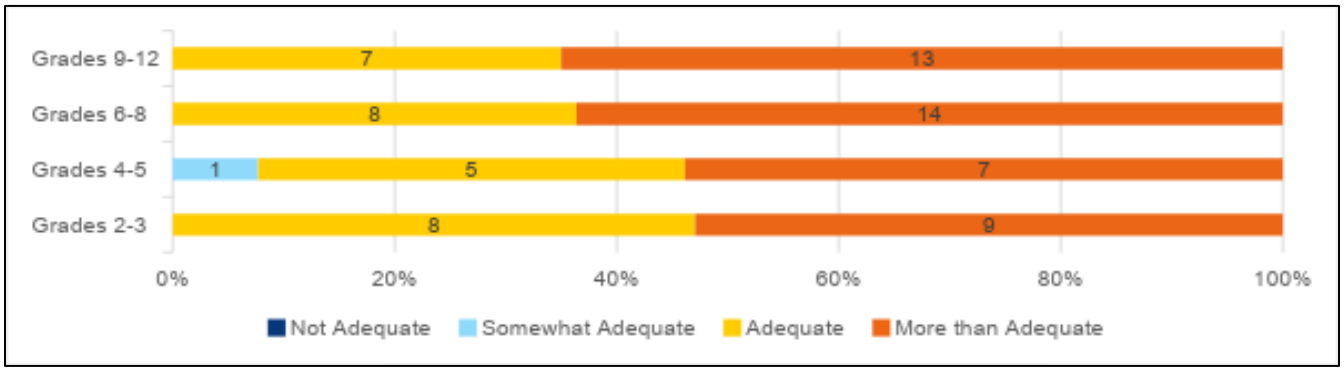
Express your opinions about student proficiency levels



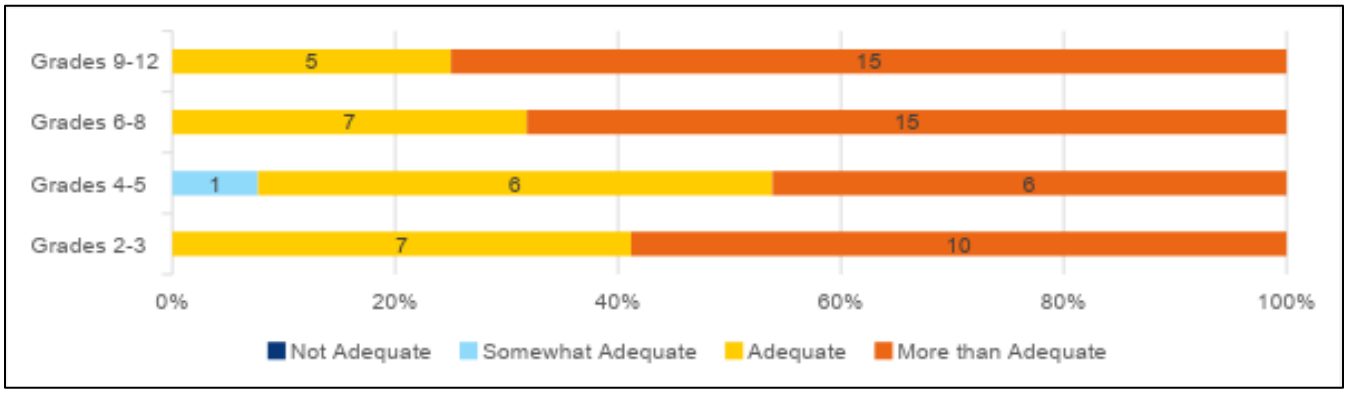
Ask about the cut scores and how they will be used



Ask about the process of making cut score recommendations

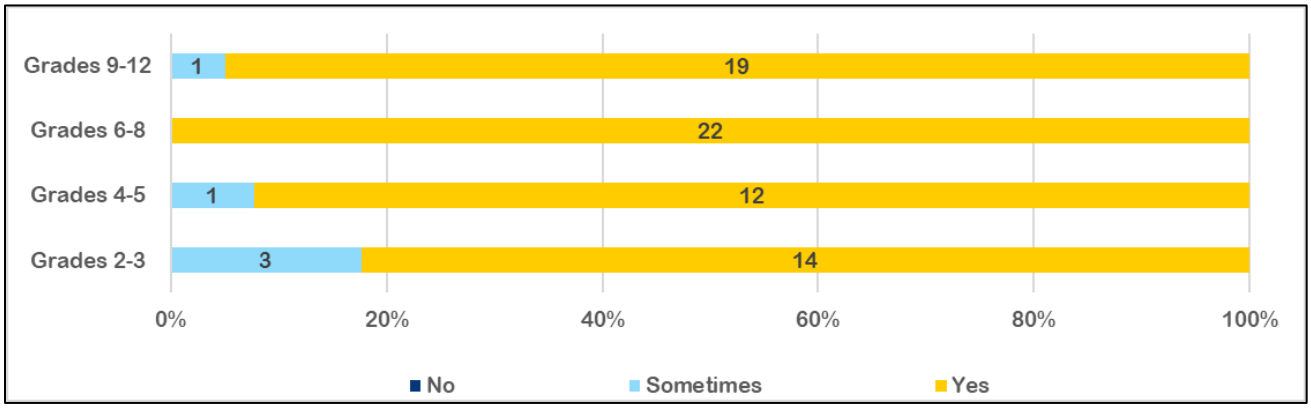


Interact with your fellow panelists

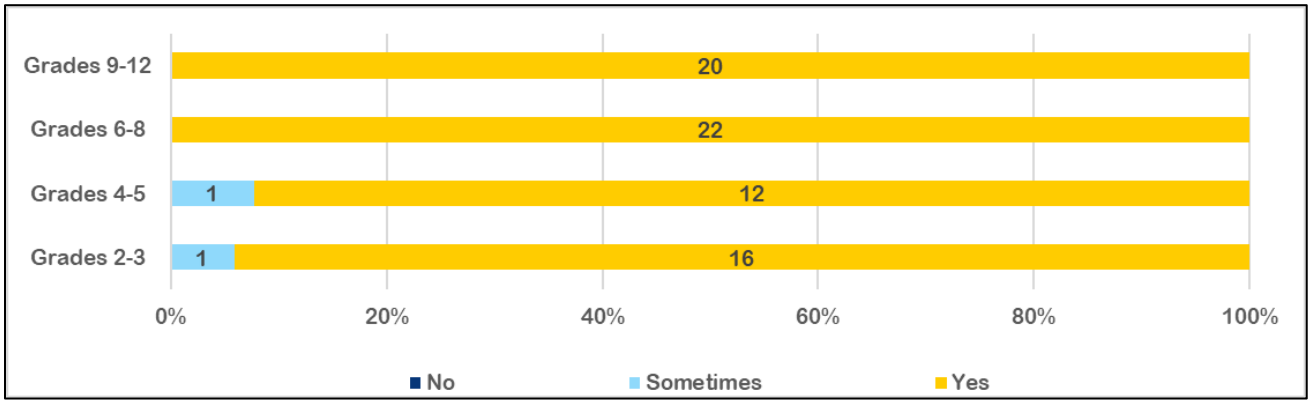


Question 13: Do you believe your opinions and judgments were treated with respect by

Fellow panelists



Facilitators



Appendix J – Composite Score Profiles

TELPAS Composite Intermediate Neighborhood (1 of 2)

Proficiency Levels				Equally Weighted Composite Score
Reading	Writing	Listening	Speaking	
1	1	1	1	1
1	1	1	2	1.25
1	1	2	1	1.25
1	2	1	1	1.25
2	1	1	1	1.25
1	1	1	3	1.5
1	1	2	2	1.5
1	1	3	1	1.5
1	2	1	2	1.5
1	2	2	1	1.5
1	3	1	1	1.5
2	1	1	2	1.5
2	1	2	1	1.5
2	2	1	1	1.5
3	1	1	1	1.5
1	1	1	4	1.75
1	1	2	3	1.75
1	1	3	2	1.75
1	1	4	1	1.75
1	2	1	3	1.75
1	2	2	2	1.75
1	2	3	1	1.75
1	3	1	2	1.75
1	3	2	1	1.75
1	4	1	1	1.75
2	1	1	3	1.75
2	1	2	2	1.75
2	1	3	1	1.75
2	2	1	2	1.75
2	2	2	1	1.75
2	3	1	1	1.75
3	1	1	2	1.75
3	1	2	1	1.75
3	2	1	1	1.75
4	1	1	1	1.75

TELPAS Composite Intermediate Neighborhood (2 of 2)

Proficiency Levels				Equally Weighted Composite Score
Reading	Writing	Listening	Speaking	
1	1	2	4	2
1	1	3	3	2
1	1	4	2	2
1	2	1	4	2
1	2	2	3	2
1	2	3	2	2
1	2	4	1	2
1	3	1	3	2
1	3	2	2	2
1	3	3	1	2
1	4	1	2	2
1	4	2	1	2
2	1	1	4	2
2	1	2	3	2
2	1	3	2	2
2	1	4	1	2
2	2	1	3	2
2	2	2	2	2
2	2	3	1	2
2	3	1	2	2
2	3	2	1	2
2	4	1	1	2
3	1	1	3	2
3	1	2	2	2
3	1	3	1	2
3	2	1	2	2
3	2	2	1	2
3	3	1	1	2
4	1	1	2	2
4	1	2	1	2
4	2	1	1	2

TELPAS Composite Advanced Neighborhood (1 of 4)

Proficiency Levels				Equally Weighted Composite Score
Reading	Writing	Listening	Speaking	
1	1	3	4	2.25
1	1	4	3	2.25
1	2	2	4	2.25
1	2	3	3	2.25
1	2	4	2	2.25
1	3	1	4	2.25
1	3	2	3	2.25
1	3	3	2	2.25
1	3	4	1	2.25
1	4	1	3	2.25
1	4	2	2	2.25
1	4	3	1	2.25
2	1	2	4	2.25
2	1	3	3	2.25
2	1	4	2	2.25
2	2	1	4	2.25
2	2	2	3	2.25
2	2	3	2	2.25
2	2	4	1	2.25
2	3	1	3	2.25
2	3	2	2	2.25
2	3	3	1	2.25
2	4	1	2	2.25
2	4	2	1	2.25
3	1	1	4	2.25
3	1	2	3	2.25
3	1	3	2	2.25
3	1	4	1	2.25
3	2	1	3	2.25
3	2	2	2	2.25
3	2	3	1	2.25
3	3	1	2	2.25
3	3	2	1	2.25
3	4	1	1	2.25
4	1	1	3	2.25
4	1	2	2	2.25
4	1	3	1	2.25
4	2	1	2	2.25
4	2	2	1	2.25
4	3	1	1	2.25

TELPAS Composite Advanced Neighborhood (2 of 4)

Proficiency Levels				Equally Weighted Composite Score
Reading	Writing	Listening	Speaking	
1	1	4	4	2.5
1	2	3	4	2.5
1	2	4	3	2.5
1	3	2	4	2.5
1	3	3	3	2.5
1	3	4	2	2.5
1	4	1	4	2.5
1	4	2	3	2.5
1	4	3	2	2.5
1	4	4	1	2.5
2	1	3	4	2.5
2	1	4	3	2.5
2	2	2	4	2.5
2	2	3	3	2.5
2	2	4	2	2.5
2	3	1	4	2.5
2	3	2	3	2.5
2	3	3	2	2.5
2	3	4	1	2.5
2	4	1	3	2.5
2	4	2	2	2.5
2	4	3	1	2.5
3	1	2	4	2.5
3	1	3	3	2.5
3	1	4	2	2.5
3	2	1	4	2.5
3	2	2	3	2.5
3	2	3	2	2.5
3	2	4	1	2.5
3	3	1	3	2.5
3	3	2	2	2.5
3	3	3	1	2.5
3	4	1	2	2.5
3	4	2	1	2.5
4	1	1	4	2.5
4	1	2	3	2.5
4	1	3	2	2.5
4	1	4	1	2.5
4	2	1	3	2.5
4	2	2	2	2.5
4	2	3	1	2.5
4	3	1	2	2.5

4	3	2	1	2.5
4	4	1	1	2.5

TELPAS Composite Advanced Neighborhood (3 of 4)

Proficiency Levels				Equally Weighted Composite Score
Reading	Writing	Listening	Speaking	
1	2	4	4	2.75
1	3	3	4	2.75
1	3	4	3	2.75
1	4	2	4	2.75
1	4	3	3	2.75
1	4	4	2	2.75
2	1	4	4	2.75
2	2	3	4	2.75
2	2	4	3	2.75
2	3	2	4	2.75
2	3	3	3	2.75
2	3	4	2	2.75
2	4	1	4	2.75
2	4	2	3	2.75
2	4	3	2	2.75
2	4	4	1	2.75
3	1	3	4	2.75
3	1	4	3	2.75
3	2	2	4	2.75
3	2	3	3	2.75
3	2	4	2	2.75
3	3	1	4	2.75
3	3	2	3	2.75
3	3	3	2	2.75
3	3	4	1	2.75
3	4	1	3	2.75
3	4	2	2	2.75
3	4	3	1	2.75
4	1	2	4	2.75
4	1	3	3	2.75
4	1	4	2	2.75
4	2	1	4	2.75
4	2	2	3	2.75
4	2	3	2	2.75
4	2	4	1	2.75
4	3	1	3	2.75
4	3	2	2	2.75
4	3	3	1	2.75
4	4	1	2	2.75
4	4	2	1	2.75

TELPAS Composite Advanced Neighborhood (4 of 4)

Proficiency Levels				Equally Weighted Composite Score
Reading	Writing	Listening	Speaking	
1	3	4	4	3
1	4	3	4	3
1	4	4	3	3
2	2	4	4	3
2	3	3	4	3
2	3	4	3	3
2	4	2	4	3
2	4	3	3	3
2	4	4	2	3
3	1	4	4	3
3	2	3	4	3
3	2	4	3	3
3	3	2	4	3
3	3	3	3	3
3	3	4	2	3
3	4	1	4	3
3	4	2	3	3
3	4	3	2	3
3	4	4	1	3
4	1	3	4	3
4	1	4	3	3
4	2	2	4	3
4	2	3	3	3
4	2	4	2	3
4	3	1	4	3
4	3	2	3	3
4	3	3	2	3
4	3	4	1	3
4	4	1	3	3
4	4	2	2	3
4	4	3	1	3

TELPAS Composite Advanced High Neighborhood

Proficiency Levels				Equally Weighted Composite Score
Reading	Writing	Listening	Speaking	
1	4	4	4	3.25
2	3	4	4	3.25
2	4	3	4	3.25
2	4	4	3	3.25
3	2	4	4	3.25
3	3	3	4	3.25
3	3	4	3	3.25
3	4	2	4	3.25
3	4	3	3	3.25
3	4	4	2	3.25
4	1	4	4	3.25
4	2	3	4	3.25
4	2	4	3	3.25
4	3	2	4	3.25
4	3	3	3	3.25
4	3	4	2	3.25
4	4	1	4	3.25
4	4	2	3	3.25
4	4	3	2	3.25
4	4	4	1	3.25
2	4	4	4	3.5
3	3	4	4	3.5
3	4	3	4	3.5
3	4	4	3	3.5
4	2	4	4	3.5
4	3	3	4	3.5
4	3	4	3	3.5
4	4	2	4	3.5
4	4	3	3	3.5
4	4	4	2	3.5
3	4	4	4	3.75
4	3	4	4	3.75
4	4	3	4	3.75
4	4	4	3	3.75
4	4	4	4	4