



# **Texas English Language Proficiency Assessment System (TELPAS) Alternate Standard-Setting Technical Report**

---

*November 2020*

# Table of Contents

<b>Chapter 1 — Overview of the Standard-Setting Process</b>	<b>3</b>
Goals of the Standard-Setting Meeting.....	3
TELPAS Alternate English Language Proficiency Levels .....	3
The TELPAS Alternate Standard-Setting Process.....	4
<b>Chapter 2 — Background Information</b>	<b>6</b>
TELPAS Alternate Curriculum Standards.....	6
TELPAS Alternate Language Domain Definitions .....	6
TELPAS Alternate Proficiency Level Descriptors.....	7
<b>Chapter 3 — Standard-Setting Meetings</b>	<b>8</b>
Purpose of the Standard-Setting Meetings.....	8
Standard-Setting Meeting Participants .....	8
Committee Panelists .....	9
Standard-Setting Meeting Staff .....	10
Facilitators .....	10
Statistical Analysts .....	10
Observers .....	10
Facilitator Training.....	10
Preparation for Data Analysis During the Meetings .....	11
Materials .....	12
Pearson Standard-Setting Website.....	12
Committee Panelist Folders .....	12
Procedure .....	13
Standard-Setting Meeting Proceedings.....	13
General Session.....	13
Breakout Sessions .....	14
<i>Introductions and Overview</i> .....	14
<i>Borderline Descriptions Development</i> .....	15
Standard-Setting Training and Judgment Rounds .....	16
Process Evaluation.....	19
Recommended TELPAS Alternate Cut Scores from Standard-Setting Committees.....	19
<b>Chapter 4 — Post-Standard Setting</b>	<b>22</b>
Composite Score Meeting.....	22
Meeting Process.....	22
Meeting Results.....	23
Reasonableness Review.....	24

<b>Chapter 5 — Evidence of Procedural Validity of the Standard-Setting Process</b>	<b>25</b>
Committee Representation .....	25
Committee Training.....	25
Perceived Validity of the Meeting .....	26
Process Standardization .....	29
<b>References</b>	<b>30</b>
<b>Appendix A — TELPAS Alternate Proficiency Level Descriptors</b>	<b>31</b>
<b>Appendix B — Participants’ Meeting Materials</b>	<b>36</b>
<b>Appendix C — Committee Participant Composition</b>	<b>44</b>
<b>Appendix D — Meeting Agenda</b>	<b>51</b>
<b>Appendix E — Examples of Feedback Data</b>	<b>54</b>
Individual Test-Level Recommendations.....	54
Overall Test-Level Recommendations.....	54
Observable Behavior-Level Judgment Agreement.....	55
Test-Level Panelist Recommendation Agreement .....	55
Observable Behavior Score Mean and Score Distribution .....	57
Impact Data .....	58
<b>Appendix F — Participant Evaluation Results</b>	<b>59</b>
<b>Appendix G — Committee Recommended Cut Scores by Round</b>	<b>84</b>
<b>Appendix H — Recommended Cut Score Summary Statistics</b>	<b>86</b>
<b>Appendix I — Test-Level Participant Judgment Agreement</b>	<b>90</b>
<b>Appendix J — Composite Score Domain Profiles</b>	<b>103</b>

# 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) Alternate and includes the following sections:

- Goals of the standard-setting meeting
- TELPAS Alternate English language proficiency levels
- TELPAS Alternate standard-setting process

## Goals of the Standard-Setting Meeting

The Texas Education Agency (TEA) developed TELPAS and TELPAS Alternate to evaluate the English language proficiency of students identified as limited English proficient (LEP/EL). English language proficiency (ELP) assessments are required under the Every Student Succeeds Act (ESSA) to evaluate the progress English learners (ELs) make in becoming proficient in the domains of listening, speaking, reading, and writing. All ELs in grades K-12 are required to participate in TELPAS or TELPAS Alternate. TELPAS Alternate was specifically developed for ELs with the most significant cognitive disabilities who cannot participate in the general ELP assessment, even with allowable accommodations.

Once students are administered an assessment, various groups—including students, parents, educators, administrators, and policy makers—want to know how 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 Alternate program, achievement on the assessment is compared to a set of predefined content standards.

The standards define a collection of knowledge, skills, and abilities students are expected to demonstrate. A cut score is the minimum score a student must achieve on an assessment to be classified in a designated proficiency level. Cut scores distinguish proficiency levels, which describe a student's competency.

## TELPAS Alternate English Language Proficiency Levels

Student proficiency for each language domain (listening, speaking, reading, and writing) on the TELPAS Alternate assessment is classified into one of five English language proficiency levels,

or stages of increasing proficiency in English. The five levels are Awareness, Imitation, Early Independence, Developing Independence, and Basic Fluency.

The global definitions for the TELPAS Alternate proficiency levels explain what it means for a student to be classified at each of the levels. These definitions are the same across language domains. The global definitions for the TELPAS Alternate proficiency levels are shown in Table 1.

**Table 1. TELPAS Alternate Proficiency Level Global Definitions**

Proficiency Level	Global Definition
<b>Basic Fluency</b>	Students who receive this rating understand and produce more detailed, complex, and elaborate messages with multiple sentences in English. These students participate independently in communication activities in English in familiar environments.
<b>Developing Independence</b>	Students who receive this rating understand longer messages of multiple sentences in English and produce simple, descriptive, original messages by combining two or more words (e.g., new red bike, big fast truck). They participate meaningfully in linguistically accommodated communication activities in English in familiar environments.
<b>Early Independence</b>	Students who receive this rating understand short, simple messages and produce messages of one or two high-need, high-frequency words (e.g., book, cafeteria, teacher). They are starting to participate in linguistically accommodated communication activities in English in familiar environments.
<b>Imitation</b>	Students who receive this rating match, imitate, or approximate some English in their environment; however, they are not able to independently understand or produce English. They participate in routine communication activities in a familiar environment when the activities are significantly linguistically accommodated.
<b>Awareness</b>	Students who receive this rating may be aware of English sounds or print; however, they have little or no functional ability to participate in communication activities in English.

## The TELPAS Alternate Standard-Setting Process

The cut score recommendations by the standard-setting committees represent the proficiency students are expected to demonstrate to be classified into each proficiency level. To establish the cut scores for each domain, a test-centered, criterion-referenced method was used to guide panelists. The procedure implemented was a hybrid of the Angoff method (Angoff, 1971) and Extended Modified (Yes/No) Angoff method (Davis & Moyer, 2015; Plake, Ferdous, Impara, &

Buckendahl, 2005). The hybrid standard-setting procedure is a systematic method that combines various considerations into the process of recommending cut scores for the different proficiency levels.

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

- Pre-meeting development — In anticipation of the standard setting meetings, various tasks were completed, including:
  - Development of TELPAS Alternate proficiency level descriptors (PLDs)
  - Sample borderline description for each domain assessed
  - 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 domain-specific borderline descriptions to make recommendations for cut scores that define the different proficiency levels for each assessment.
- Composite score and rating — The rules to determine the TELPAS Alternate composite score were established using student profiles of domain scores by proficiency level.
- Reasonableness review — TEA conducted a reasonableness review of the TELPAS Alternate 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 recommended proficiency standards.

The remaining chapters will describe background information regarding the TELPAS Alternate assessment program as well as the specific procedures and activities that occurred during each step of the standard-setting process.

# Chapter 2 — Background Information

This chapter provides an overview of TELPAS Alternate and includes the following sections:

- TELPAS Alternate curriculum standards
- TELPAS Alternate language domain definitions
- TELPAS Alternate PLDs

## TELPAS Alternate Curriculum Standards

The English Language Proficiency Standards (ELPS) are second-language-acquisition curriculum standards that support the ability of ELs to acquire academic English, while at the same time allowing them to meaningfully engage in regular, all-English academic instruction. Adopted by the State Board of Education in 2007, 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 contained in the state-mandated curriculum standards. The ELPS outline the instruction ELs must receive to support their ability to develop academic English language proficiency and acquire challenging academic knowledge and skills. To identify the ELPS that would be assessed by the TELPAS Alternate, a committee of educators determined which ELPS were accessible for this student population and capable of being assessed on this type of test. For TELPAS Alternate, PLDs that align with this set of ELPS were created to address the specific access needs of students with significant cognitive disabilities.

## TELPAS Alternate Language Domain Definitions

TELPAS Alternate assesses student proficiency for the language domains of listening, speaking, reading, and writing. The definitions of the language domains are intentionally broad for TELPAS Alternate to allow for alternate forms of expressive and receptive language, as defined and compared to TELPAS in Table 2.

**Table 2. TELPAS and TELPAS Alternate Language Domain Definitions**

Domain	TELPAS Definition	TELPAS Alternate Definition
<b>Listening</b>	The ability to understand spoken language, comprehend and extract information, and follow social and instructional discourse through which information is provided	The ability to understand spoken or signed language, comprehend and extract information, and follow social and instructional discourse through which information is provided
<b>Speaking</b>	The ability to use spoken language appropriately and effectively in learning activities and social interactions	The ability to use spoken language or alternative communication appropriately and effectively in learning activities and social interactions
<b>Reading</b>	The ability to comprehend and interpret written text at the grade-appropriate level	The ability to comprehend and interpret written text, including braille, at a modified level
<b>Writing</b>	The ability to produce written text with content and format to fulfill grade-appropriate classroom assignments	The ability to produce written text or alternative communication with content and format to fulfill classroom and community-based assignments

For TELPAS Alternate, “English” is not limited to the typical spoken or written English of other state assessments because some ELs use another method of communication, such as sign language or braille, as a substitute for traditional English in one or more domain(s).

## TELPAS Alternate Proficiency Level Descriptors

The TELPAS Alternate PLDs are domain-specific and define how well ELs in the five proficiency levels understand and use English in academic settings. The PLDs are also a critical part of the process used to set the TELPAS Alternate proficiency standards. They show the progression of second language acquisition from one proficiency level to the next and provide a common framework for understanding the language acquisition skills needed to be classified within each proficiency level. The TELPAS Alternate PLDs for grades 2–12 listening, speaking, reading, and writing can be found in [Appendix A](#).

# 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 educators' judgments and to translate them into cut scores for the proficiency levels. The purpose of the TELPAS Alternate 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 Alternate listening, speaking, reading, and writing assessments.

Student performance on each domain of the TELPAS Alternate assessment was classified at one of five proficiency levels. Each committee was asked to recommend four cut scores that defined the boundaries between the different proficiency levels. These recommended cut scores represent the performance a student would need to meet or exceed in order to be classified into the specific proficiency level.

## Standard-Setting Meeting Participants

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

- Committee panelists
- Meeting facilitators
- Observers and staff

## Committee Panelists

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 Alternate assessments. Panelists were separated into four groups (A, B, C, and D) to construct the committees for each domain, based on demographic analysis, as shown in Figure 1. Two groups at a time worked on a single domain, and the panelists were regrouped into two new groups for the other two domains.

	<b>Reading &amp; Writing</b>	<b>Listening &amp; Speaking</b>
Days 1 & 2	Reading (15 panelists) Groups A and C	Listening (13 panelists) Groups B and D
Days 2 & 3	Writing (15 panelists) Groups B and C	Speaking (13 panelists) Groups A and D

**Figure 1. Committee Construction from Panelist Groups**

The selection process of committee panelists involved considerations intended to create a sample as representative of the state as possible, including demographic variables (gender, race, etc.), geographic representation, and job title and experience (educational experience, education, etc.). When selecting panelists, TEA placed an emphasis on educators who had relevant content knowledge as well as experience with a variety of student groups.

There was a total of 28 panelists at the standard-setting meetings. The tables in [Appendix C](#) summarize the characteristics and experience of the panelists in each committee. These tables provide demographics about the panelists as well as information regarding the panelists' current positions in education, their experience working with various student populations, and the types of districts they represent. Panelists' responses to the gender and ethnicity questions were voluntary.

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 facilitator during the meeting by helping guide 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 Staff

### Facilitators

The facilitator was a member of the Pearson psychometric staff with experience overseeing standard-setting meetings and was responsible for leading panelists through the standard-setting process. The facilitator ensured 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.

Though the facilitators had prior experience leading standard-setting meetings, they 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 3 lists the facilitators for each standard-setting committee.

**Table 3. Facilitators for TELPAS Alternate Standard-Setting Committees**

Committee	Facilitator
Reading/Writing	Bob Schwartz, Ph.D.
Listening/Speaking	Malena McBride, Ph.D.

### Statistical Analysts

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

### Observers

The individuals who attended as observers consisted of TEA 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 and, in some cases, provide feedback on the process used. In addition to observing the meeting, TEA staff members addressed content, assessment, and policy questions. The number of observers in a committee meeting were kept to a maximum of ten individuals, so the committee panelists did not feel overwhelmed. Whenever possible, observers were assigned to a single committee meeting for the duration of the standard-setting meeting.

## Facilitator Training

The meetings were facilitated by a psychometrician from Pearson with knowledge and experience leading standard-setting meetings. The facilitator was responsible for ensuring that 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 Alternate — The facilitators were provided an overview of the TELPAS Alternate assessment 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.

A facilitator training meeting was held for 60 minutes on May 31, 2019. Additionally, a final training and discussion was held on-site June 9, the day before the standard-setting meetings began, to address any final topics. At the end of each day during the standard-setting meetings, a debriefing was held to discuss any potential concerns, positive outcomes from the day, and to review material planned for the next day.

Content experts from Pearson and TEA were also available as observers to answer 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

Impact data refers to the percentage of students that fall within a proficiency level based on the recommended cut scores at the given judgment round for a particular domain. Impact data was provided to panelists during the standard-setting meeting to present the expected results, based on each committee's cut score recommendations, regarding student proficiency level classifications. The analysis programs used lookup tables to create the impact data output during the meetings.

Impact data lookup tables were created using the responses from students who took the TELPAS Alternate assessment during the spring 2019 administration. A frequency distribution of student results on test was created based on the spring 2019 administration. The frequency distribution was used to determine the expected percentage of students classified into each proficiency level, based on panelists' judgments.

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

- Individual panelist feedback — the judgments for each panelist to ensure they were recorded and analyzed accurately (given to all panelists)
- 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 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 for 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 feedback report. Staff from TEA were provided the opportunity to review and suggest final revisions to the reports.

## Materials

The following section describes the materials used by committee members during the standard-setting breakout sessions. Separate materials were developed for each standard-setting meeting. A sample set of materials is provided in [Appendix B](#).

### 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.

The website was built using 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 content area. Electronic copies of the Observable Behaviors for the TELPAS Alternate listening, speaking, reading, and writing were posted to the website for panelists to review. The website also provided panelists access to online documents that included background information about the TELPAS Alternate assessments. During the meetings, each panelist was provided a laptop computer in his or her meeting room to access the online resources.

### 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 throughout the meeting. These materials included the following:

- Meeting agenda
- Non-disclosure agreement
- TELPAS Alternate reporting categories
- TELPAS Alternate PLDs
- TELPAS Alternative Observable Behaviors and classroom examples
- Judgment round record sheet

## Procedure

The Extended Modified Yes/No Angoff method (Davis & Moyer, 2015; Plake et al., 2005) was used during the TELPAS Alternate standard-setting meetings to guide committee panelists as they made their cut score recommendations. The procedure was both a content- and item-based method that leads panelists through a standardized process in which they considered student expectations, as defined by the TELPAS Alternate PLDs, and the Observable Behaviors to recommend cut scores for each proficiency level. The standardized process was used by the committees for each domain and resulted in cut score recommendations.

## Standard-Setting Meeting Proceedings

The standard-setting meetings were conducted across three days, June 10–12, 2019, in Austin, Texas. [Appendix D](#) includes the complete agenda for the standard-setting meetings. Table 4 presents a high-level agenda for the standard-setting committee meetings.

**Table 4. Standard-Setting Meeting Agenda Topics**

Meeting Dates					
June 10		June 11		June 12	
General Session	Reading & Listening Standard Setting	Writing & Speaking Standard Setting	Composite Score Review		

### General Session

The standard-setting meeting began with an orientation, during which panelists were presented an overview of the TELPAS Alternate program and the standard-setting process. The purpose of the general session was to welcome the panelists, provide background information about TELPAS Alternate, and introduce the standard-setting process. A single general session including all standard-setting panelists was conducted on June 10 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 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 Extended Modified Yes/No Angoff procedure, such as the TELPAS Alternate PLDs and standard-setting website. A clear description of the review process for the cut score recommendations was included as part

of the overview process to emphasize that the committees would be making cut score recommendations for other groups to review.

### **Breakout Sessions**

After the general session, panelists moved into 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 assigned domain. Table 5 provides an overview of the activities conducted during the breakout session for each standard-setting meeting.

**Table 5. Overview of Activities During Breakout Sessions**

Activity
Introductions and process overview*
Review of TELPAS Alternate assessment and scoring training
Review of domain-specific TELPAS Alternate PLDs
Borderline descriptions development
Standard-setting training*
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 not repeated during each breakout session.

### **Introductions and Overview**

To begin the breakout sessions, 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 Confidentiality Agreement on the standard-setting website and completed and signed the agreement for the TELPAS Alternate standard-setting meeting. Their signature acknowledged that they understood the security expectations for the meeting and agreed to follow them as described.

Next, the facilitator distributed 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. Panelists had the opportunity to ask questions before proceeding.

Finally, panelists were introduced to the TELPAS Alternate assessment system and the test blueprint for their domain. Panelists reviewed the TELPAS Alternate PLDs as part of the standard-setting meeting pre-work but spent additional time during the breakout session understanding the PLDs. Time was also devoted to the respective TELPAS Alternate domain assessment for each committee, which included the Observable Behaviors and classroom examples.

### ***Borderline Descriptions Development***

Panelists were led through a collaborative process to develop borderline descriptions, which began with a review of the TELPAS Alternate PLDs. Training on the PLDs included the facilitator reviewing the development process that was used to create the TELPAS Alternate PLDs for each domain. The panelists were informed that the TELPAS Alternate 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 student performing at the borderline of a proficiency level (i.e., a student that is just barely past the point-of-entry for that level) in order to complete the activity. For example, the following questions were shared as useful considerations for table-group discussion:

- What would a student with proficiency just barely at the proficiency level be able to do with respect to the key characteristics?
- What differentiates a student at the borderline of the proficiency level relative to a student in the middle or upper end of this level?
- What differentiates a student at the borderline of the proficiency level relative to the upper end of a lower adjacent level (e.g., borderline Imitation relative to upper Awareness)?
- How well did the TELPAS Alternate PLDs describe the students at the borderline of the proficiency level for the domain as we envisioned them?
- How were the borderline descriptions similar or different from the TELPAS Alternate PLDs for the proficiency level?

After the discussion of the TELPAS Alternate PLDs concluded, the panelists moved on to the development of borderline descriptions for four groups of students:

- Students with proficiency at the Imitation borderline
- Students with proficiency at the Early Independence borderline
- Students with proficiency at the Developing Independence borderline
- Students with proficiency at the Basic Fluency borderline

Each table group was assigned a proficiency level. In their table groups, panelists reviewed and discussed the TELPAS Alternate PLDs for their assigned proficiency level and identified general characteristics of a student with proficiency at the borderline of that level in order to create borderline descriptions. The facilitator placed the borderline descriptions from each group into a single document. The combined borderline descriptions were reviewed with the whole group for consistency in expectations and additional edits or clarifications were made as needed.

The result of the final, whole-group discussion was a single set of borderline descriptions for each domain that was provided to panelists as a reference throughout subsequent activities.

The goal of the borderline descriptions was to help panelists maintain a common understanding of the characteristics possessed by the students at the borderline of each proficiency level.

### **Standard-Setting Training and Judgment Rounds**

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

#### ***Training***

The panelists were provided thorough training regarding how they should make their cut score recommendations as part of the standard-setting meetings. For each TELPAS Alternate domain, the panelists were instructed on how to apply the Extended Modified Yes/No process (Davis & Moyer, 2015; Plake et al., 2005). The Extended Modified Yes/No method was used to support judgments for TELPAS Alternate Observable Behaviors, all of which are polytomously scored. Panelists reviewed each Observable Behavior and answered the following question for each proficiency level:

*“How many points would a student with proficiency at the borderline of the level likely earn for the Observable Behavior?”*

Facilitators modeled the thought process panelists should go through with each part of the question.

- “How many points...” — Rather than recording “yes” or “no” judgments, panelists recorded the number of points for an Observable Behavior.
- “...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 student with proficiency at the borderline of the level...” — The panelists referenced the borderline descriptions for the proficiency level to determine how a student with proficiency at the borderline would be expected to respond.
- “...likely earn for the Observable Behavior?” — In this context, likely was defined as two out of three times, or 67%. To make this concept concrete for panelists, facilitators asked them to consider three students with proficiency at the borderline of a level. The maximum score per Observable Behavior is five points. If a panelist believed two of three students with proficiency at the borderline would receive five points, their judgment should be five points. If the panelist does not believe two of three students with proficiency at the borderline would earn five points for the Observable Behavior, then the possibility of four points should be considered. If a panelist believed two of three students with proficiency at the borderline would receive four points, their judgment should be four points. If the panelist does not believe two of three students with proficiency at the borderline would earn four points for the Observable Behavior, then the possibility of three points should be considered, and so on until a point-judgment can be made.

The standard-setting training 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 Observable Behaviors used in the judgment activity, various reference resources, and the judgment survey, where panelists recorded their individual recommendations for each proficiency level.
- Online operational test items — A set of items that represented an online operational test was presented in the order items were administered. 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
  - Maximum number of possible points
- Judgment record form — The panelists recorded their judgments for each round in the standard-setting website and on the paper record form.

Beginning with the Imitation proficiency level, panelists reviewed each Observable Behavior and made a judgment for each borderline proficiency level. Because a student’s response to an Observable Behavior was expected to increase or stay the same as the proficiency level increased (i.e., the selected number of points should increase or stay the same across proficiency levels), panelists were trained to check their judgments for expected patterns across proficiency levels. The training included multiple examples of different judgment patterns, which were reviewed with panelists to help them understand the judgment task. The examples included responses that follow and do not follow the expected judgment patterns as well as floor and ceiling patterns.

The panelists also kept a record of their judgments on the paper Judgment Record Form. The form included the Observable Behavior number, reporting category, and maximum possible points for the Observable Behavior. The panelists were shown how to use the Observable Behavior number to ensure that they were referencing the correct behavior on all documents within the judgment survey and in the online system.

### ***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 indicating 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. Panelists answered the following questions:

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

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

During each round, panelists individually made judgments for each Observable Behavior by starting at the lowest proficiency level (Imitation) and using the borderline descriptions and the skills and abilities required by the behavior. The panelists then made judgments for the same Observable Behavior for the rest of the proficiency levels — Early Independence, Developing Independence, and Basic Fluency — before proceeding to the next behavior. Judgments were recorded in the website using the Judgment Survey for each specific round. Once panelists recorded judgments for all Observable Behaviors, they submitted their judgments for analysis.

After all panelists completed the judgment activity for the round, data analysts from Pearson collected the judgments from the website, performed the necessary analysis of the data, and created feedback reports that were provided to panelists.

### ***Feedback and Discussion***

Panelists received the following feedback to facilitate discussion:

- Information about panelists' cut scores for each proficiency level:
  - Individual cut scores: 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 individual judgments 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, standard deviation, 25<sup>th</sup> percentile, and 75<sup>th</sup> percentile) 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.
- Judgment agreement across panelists: Distributions of individual judgments for each Observable Behavior and proficiency level were presented.
- Observable Behavior means (p-values) and score point distributions: The average score earned for each Observable Behavior and the distribution of score points 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 2019 administration.

Certain types of information were provided only after specific rounds. Table 6 shows the feedback information provided after each judgment round. Examples of feedback data shared with panelists are available in [Appendix E](#).

**Table 6. Feedback Data Provided to Panelists, by Round**

		Round 1	Round 2	Round 3
<b>Observable Behavior-Level Feedback</b>	Panelist Agreement Data	✓	✓	
	Observable Behavior Means	✓		
	Score Point Distributions	✓		
<b>Test-Level Feedback</b>	Individual Cut Score Recommendations	✓	✓	
	Committee Cut Score Recommendations	✓	✓	
	Panelist Agreement Data	✓	✓	
	Impact Data		✓	

Before 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 discussion, which was facilitated by table-group leaders, panelists shared the rationale for their cut score recommendations with other panelists at their table. After Round 2, the panelists also participated in a whole-group discussion led by the process facilitator.

### Process Evaluation

The validity of standard-setting outcomes depends on procedural validity. Evidence of procedural validity was gathered through evaluation surveys administered during and 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 F](#).

## Recommended TELPAS Alternate Cut Scores from Standard-Setting Committees

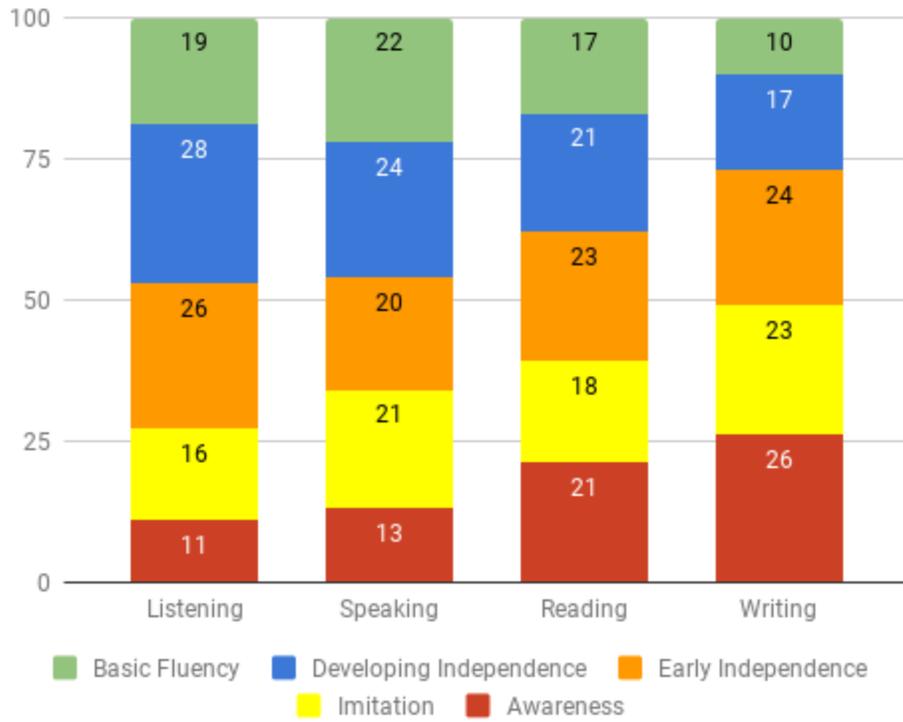
Panelists at the standard-setting meeting recommended cut scores for each proficiency level. To determine a committee's single cut score recommendation for a proficiency level, analysts used the median cut score from a set of panelists' recommendations. The Round 3 judgments

were considered the committee’s final recommendation for the standard-setting meeting. The recommended cut scores for each proficiency level are displayed by domain in Table 7.

**Table 7. Cut Score Recommendations from Standard-Setting Committees**

Domain	Maximum Possible Score	Cut Scores			
		Imitation	Early Independence	Developing Independence	Basic Fluency
Listening	50	17	26	36	45
Speaking	50	16	26	35	44
Reading	50	18	24	33	42
Writing	50	16	24	33	41

The recommended cut scores from each judgment round are presented by proficiency level as raw scores in [Appendix G](#). The summary statistics for the recommended cut scores of each judgment round are shown by proficiency level for all standard-setting committees in [Appendix H](#). The panelist agreement data for judgment Rounds 1, 2, and 3 are displayed for all standard-setting committees in [Appendix I](#). The estimated impact data after judgment Round 2 is shown by proficiency level for each standard-setting committee in Figure 2 below.



**Figure 2. Impact data from Round 2 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

## Composite Score Meeting

In addition to the individual proficiency level rating of students (Awareness, Imitation, Early Independence, Developing Independence, and Basic Fluency) for each of the four language domains assessed by TELPAS Alternate (listening, speaking, reading, and writing), student reports provide composite scores and composite proficiency ratings. The composite score and composite proficiency rating for students taking the TELPAS Alternate 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 Alternate assessments are weighted and averaged to calculate a composite TELPAS Alternate score. The domain score for each domain proficiency level is shown in Table 8. If a student does not take the assessment or does not receive a score for a domain, the domain score received is a 0.

**Table 8. Domain Scores for Domain Proficiency Levels**

	Domain Proficiency Level				
	Awareness	Imitation	Early Independence	Developing Independence	Basic Fluency
Domain Scores	1	2	3	4	5

As part of the standard-setting process, a set of participants were invited to discuss rules used to assign composite proficiency ratings for TELPAS Alternate.

### Meeting Process

The composite score process involved the following three steps:

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

Panelists examined the scoring profiles associated with each possible composite score for each composite proficiency level. The composite score profiles can be found in [Appendix J](#). Panelists then discussed 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.

Next, panelists considered the scoring profiles associated with the minimum composite score chosen for each composite proficiency level and determined if there were any profiles panelists would not consider to be classified at the same composite proficiency level.

The final step was to use the composite score profiles associated with the recommended minimum scores for each composite proficiency level to create a general 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 9.

**Table 9. Composite Proficiency Rating Profile Descriptors**

Composite Proficiency Rating	Composite Rating Minimum Requirements
<b>Basic Fluency</b>	<ul style="list-style-type: none"> <li>• A TELPAS Alternate composite score greater than or equal to 4</li> <li>• At least two domains with a proficiency level of Basic Fluency</li> <li>• All domains with a proficiency level of Early Independence or higher</li> </ul>
<b>Developing Independence</b>	<ul style="list-style-type: none"> <li>• A TELPAS Alternate composite score greater than or equal to 3.25</li> <li>• At least two domains with a proficiency level of Developing Independence or higher</li> <li>• All domains with a proficiency level of Imitation or higher</li> </ul>
<b>Early Independence</b>	<ul style="list-style-type: none"> <li>• A TELPAS Alternate composite score greater than or equal to 2.25</li> <li>• At least two domains with a proficiency level of Early Independence or higher</li> <li>• At least three domains with a proficiency level of Imitation or higher</li> </ul>
<b>Imitation</b>	<ul style="list-style-type: none"> <li>• A TELPAS Alternate composite score greater than or equal to 1.5</li> <li>• At least two domains with a proficiency level of Imitation or higher</li> </ul>

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 TEA on June 12, 2019. The presentation included a brief overview of the methodology used to obtain the cut score recommendations, the panelists' recommended cut scores for each domain 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 were provided.

Additionally, the cut score recommendations from the standard setting for the TELPAS Alternate assessment were compared to student performance on the STAAR Alternate 2 reading, writing, and English end-of-course assessments.

The goal of this process was to evaluate the reasonableness of the results of the standard-setting meeting, expected alignment across domains, and usefulness in communicating results. TEA reviewed the alignment of the cut score ranges for each of the proficiency levels across the domains. 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. Based on the Reasonableness Review, TEA did not adjust any of the cut scores recommended by the standard-setting committees.

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.

# 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 C](#).

As part of the survey, panelists were asked to report their current position (Table C.1), years of teaching experience (Table C.2), and highest level of education (Table C.7). About 67 percent of the panelists were K–12 teachers. The panelists had an extensive range of teaching experience, with greater than 75 percent of teachers possessing more than 11 years of experience in education. Additionally, all of the panelists had at least a Bachelor's or Master's degree.

A large majority of panelists indicated they had experience teaching student populations relevant to their committee (presented in Tables C.3 and C.4), which is a pertinent factor in relation to the cut score recommendations. A large majority of panelists had experience teaching general education, mainstream special education, and English learners. Nearly 80 percent of panelists had experience teaching in bilingual or ESL classrooms and 75 percent were experienced teaching in special education classrooms. A wide range of teaching experience is an important consideration that ideally increases the cohesiveness of cut score recommendations across committees.

Most panelists were currently working in school districts, as shown in Table C.11. 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.

## Committee Training

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

(Yes/No) Angoff standard-setting methodology. The panelists were trained in the standard-setting methodology during the general session and received 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.

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 are presented in [Appendix F](#).

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 the success of the initial introduction to the standard-setting process during the general session. All panelists responded that it was either *Successful* or *Very Successful*. Another question asked about the overview of the TELPAS Alternate assessments in the general session. All panelists responded that the overview was either *Successful* or *Very Successful*. Ninety-six percent of panelists reported that the overview they were provided on their specific TELPAS Alternate assessment and scoring rules was either *Successful* or *Very Successful*. Taken as a whole, the evaluation survey results indicate 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.

## Perceived Validity of the Meeting

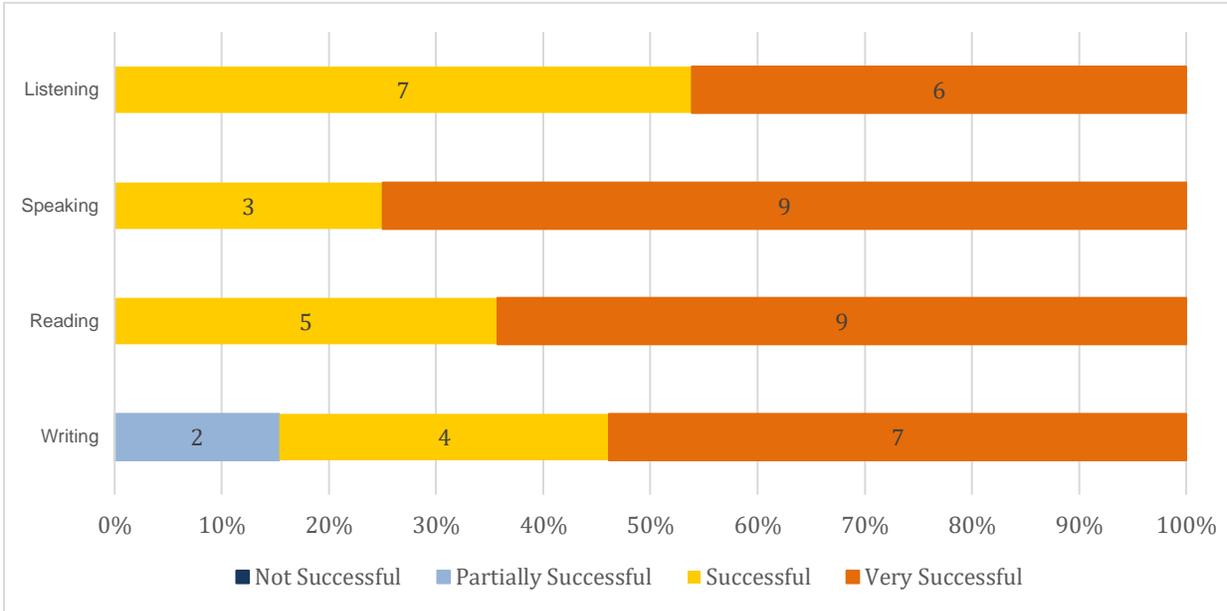
Panelists communicated their perceived validity of the meeting and the recommended cut scores as part of the standard-setting process evaluation. Evaluations are important evidence for establishing the validity of recommended cut scores for the proficiency levels. Generally, the panelists were satisfied 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 TELPAS Alternate PLDs were reasonable for each of the proficiency levels. A majority of panelists were *Confident* or *Very Confident* the PLDs were reasonable for all proficiency levels and domains. These responses provide evidence that the TELPAS Alternate PLDs, a foundation for the standard-setting process, were perceived by the panelists as reasonable expectations for each proficiency level.

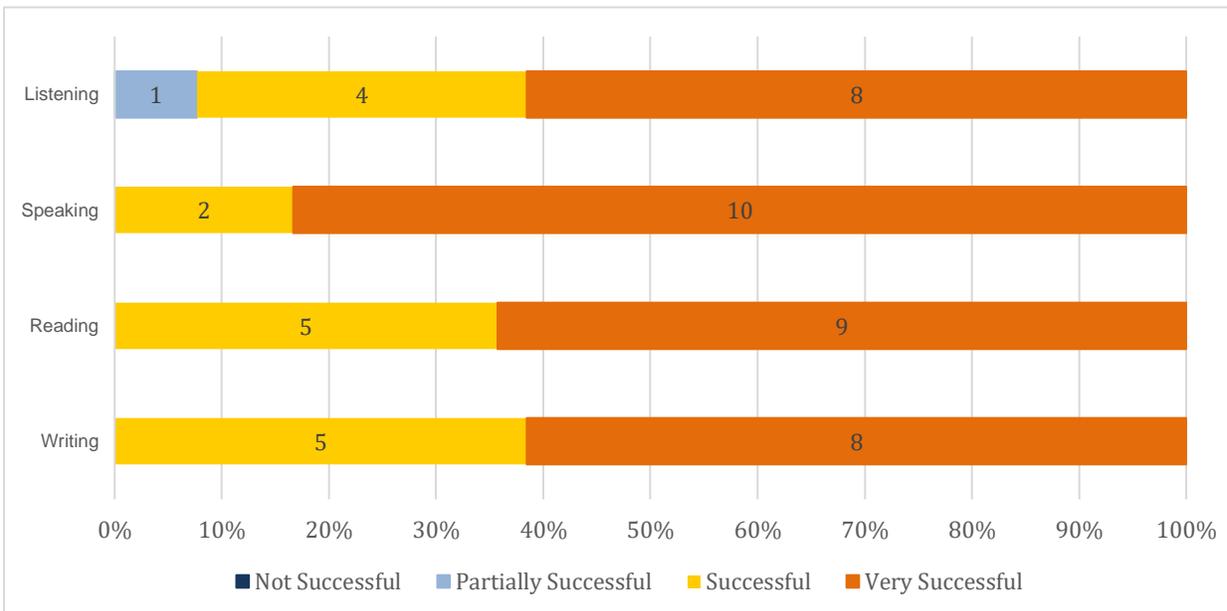
The panelists were also given an opportunity to indicate their confidence in the recommended cut scores for each domain, the results of which are shown in Figures 3 and 4.

**How confident do you feel that the final cut score recommendations for the respective domain represent appropriate levels of student performance?**

**Basic Fluency**



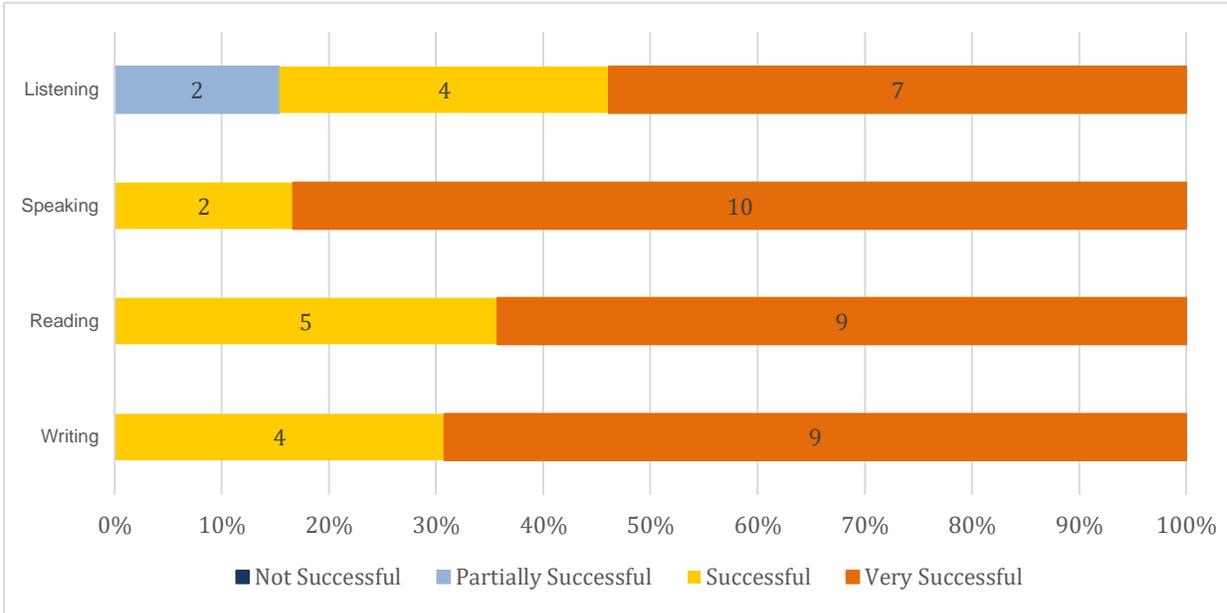
**Developing Independence**



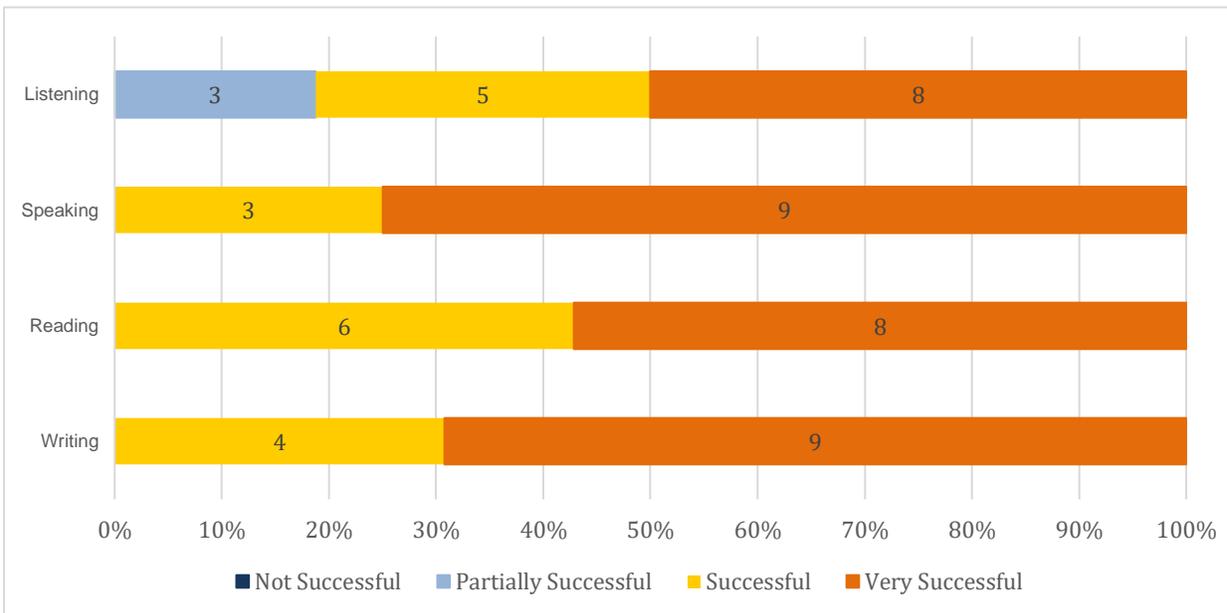
**Figure 3. Evaluation results on reasonableness of cut scores for Basic Fluency and Developing Independence proficiency levels**

**How confident do you feel that the final cut score recommendations for the respective domain represent appropriate levels of student performance?**

**Early Independence**



**Imitation**



**Figure 4. Evaluation results on reasonableness of cut scores for each Early Independence and Imitation proficiency levels**

For the Basic Fluency proficiency level, more than 95 percent of panelists believed the final cut score recommendations were *Successful* or *Very Successful* representations of appropriate student performance; two panelists in the writing committee did not believe the final cut score recommendations were successful. All but one panelist responded that the final cut score recommendations were *Successful* or *Very Successful* for the Developing Independence proficiency level. Nearly all panelists across the committees reported the final cut score recommendations were *Successful* or *Very Successful* for the Early Independence proficiency level; two panelists in the listening committee did not believe the final cut score recommendations were successful representations of student performance. Lastly, over 90 percent of panelists indicated the final cut score recommendations for the Imitation proficiency level were *Successful* or *Very Successful*.

Overall, the results from the feedback given by standard-setting panelists provide supportive validity evidence for the cut score recommendations for all proficiency levels. Full results of the evaluation survey are presented in [Appendix F](#).

## 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 Alternate standard-setting meetings, two facilitators worked with four panels (a separate panel for each of listening, speaking, reading, and writing) to determine cut scores for five proficiency levels (Awareness, Imitation, Early Independence, Developing Independence, and Basic Fluency) 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 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.

Materials were used to facilitate each of the meetings and were prepared in advance to ensure consistency of the presentation and recording of 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 use 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.

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

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 — TELPAS Alternate Proficiency Level Descriptors

## TELPAS Alternate Proficiency Level Descriptors

### Grades 2–12 Listening

Awareness	Imitation	Early Independence	Developing Independence	Basic Fluency
<p>ELs at this level have little or no functional ability to understand spoken English even when interacting with concrete symbols that are combined with highly familiar spoken words.</p>	<p>ELs at this level participate in routine listening activities in English when interacting with concrete symbols that are combined with highly familiar spoken words.</p>	<p>ELs at this level participate in listening activities in English when working with spoken short, simple messages.</p>	<p>ELs at this level demonstrate understanding of longer spoken messages of multiple sentences during listening activities in English.</p>	<p>ELs at this level are able to understand detailed and complex spoken messages in English.</p>
<p>These students may:</p> <ul style="list-style-type: none"> <li>• require full second language acquisition support (visuals, gestures, verbal cues) for social and academic listening activities in English in their daily routine</li> <li>• know too little English to understand or react to routine directions and simple conversations even when the topics are familiar and the speaker uses linguistic supports</li> <li>• alert to or show reaction to spoken English, but do not demonstrate understanding</li> </ul>	<p>These students may:</p> <ul style="list-style-type: none"> <li>• require significant second language acquisition support (visuals, gestures, verbal cues) for social and academic listening activities in English in their daily routine</li> <li>• understand simple familiar spoken words when paired with concrete symbols</li> <li>• follow single-word directions when paired with concrete symbols</li> </ul>	<p>These students may:</p> <ul style="list-style-type: none"> <li>• require moderate second language acquisition (visuals, gestures, verbal cues) for social and academic listening activities in English in their daily routine</li> <li>• understand simple familiar spoken words, but not how they fit into a larger language context</li> <li>• follow single-word directions</li> </ul>	<p>These students may:</p> <ul style="list-style-type: none"> <li>• require occasional second language acquisition support (visuals, gestures, verbal cues) for social and academic listening activities in English in their daily routine</li> <li>• understand short simple conversations as well as short simple discussions on familiar topics</li> <li>• follow single-step directions consisting of a few words</li> <li>• often identify and distinguish key words and phrases necessary to understand the general meaning during social and academic interactions</li> </ul>	<p>These students may:</p> <ul style="list-style-type: none"> <li>• require minimal second language acquisition support (visuals, gestures, verbal cues) for social and academic listening activities in English in their daily routine across settings</li> <li>• understand longer and more elaborated directions, conversations, and discussions on familiar and unfamiliar topics, but sometimes need processing time</li> <li>• follow multi-step directions</li> <li>• understand most main points, most important details, and some implicit information during social and academic interactions</li> </ul>

## TELPAS Alternate Proficiency Level Descriptors

### Grades 2–12 Speaking

Awareness	Imitation	Early Independence	Developing Independence	Basic Fluency
<p><b>ELs at this level have little or no functional use of spoken English or augmentative and alternative communication even when interacting with highly familiar words.</b></p>	<p><b>ELs at this level approximate speaking or augmentative and alternative communication activities in English when interacting with highly familiar words.</b></p>	<p><b>ELs at this level participate in speaking or augmentative and alternative communication activities when working with familiar words.</b></p>	<p><b>ELs at this level produce short, simple messages in English or with augmentative and alternative communication.</b></p>	<p><b>ELs at this level produce detailed spoken messages in English or with augmentative and alternative communication.</b></p>
<p><b>These students may:</b></p> <ul style="list-style-type: none"> <li>• require full second language acquisition support (visuals, gestures, verbal cues) for social and academic communicative activities in their daily routine</li> <li>• know too little English to communicate simple ideas even when topics are highly familiar</li> <li>• alert to or show reaction to stimuli, but do not demonstrate ability to clearly communicate thoughts</li> <li>• not initiate spoken communication in English</li> </ul>	<p><b>These students may:</b></p> <ul style="list-style-type: none"> <li>• require significant second language acquisition support (visuals, gestures, verbal cues) for social and academic communicative activities in their daily routine</li> <li>• imitate or attempt to imitate use of spoken English words after modeling</li> <li>• rarely initiate spoken communication in English independently</li> </ul>	<p><b>These students may:</b></p> <ul style="list-style-type: none"> <li>• require moderate second language acquisition support (visuals, gestures, verbal cues) for social and academic communicative activities in their daily routine</li> <li>• communicate with a very limited vocabulary of high-frequency, high-need, concrete one- or two-word responses</li> <li>• not understand how words fit into a larger language context</li> <li>• hesitate to speak in English and often give up in their attempts to communicate</li> </ul>	<p><b>These students may:</b></p> <ul style="list-style-type: none"> <li>• require occasional second language acquisition support (visuals, gestures, verbal cues) for social and academic communicative activities in their daily routine</li> <li>• combine spoken words to create simple original messages</li> <li>• pause to find words to restate or clarify meaning</li> </ul>	<p><b>These students may:</b></p> <ul style="list-style-type: none"> <li>• require minimal second language acquisition support (visuals, gestures, verbal cues) for social and academic communicative activities in their daily routine across settings</li> <li>• be able to express detailed ideas through spoken words in social and academic English</li> <li>• occasionally pause to search for words and phrases to clarify meaning</li> </ul>

## TELPAS Alternate Proficiency Level Descriptors

### Grades 2–12 Reading

Awareness	Imitation	Early Independence	Developing Independence	Basic Fluency
<p>ELs at this level have little or no functional use of the English language even when interacting with concrete symbols that are combined with highly familiar letters and words.</p>	<p>ELs at this level participate in routine reading activities that include matching when interacting with concrete symbols that are combined with highly familiar letters and words.</p>	<p>ELs at this level participate in reading activities in English when working with highly familiar words.</p>	<p>ELs at this level participate meaningfully in reading activities in English when working with words and simple texts.</p>	<p>ELs at this level are able to read and understand more detailed texts in English.</p>
<p>These students may:</p> <ul style="list-style-type: none"> <li>• require full second language acquisition support (visuals, gestures, verbal cues) for social and academic reading activities in English in their daily routine</li> <li>• alert to or show reaction to concrete symbols, letters, and/or words in English, but do not demonstrate understanding</li> <li>• alert to or show reaction to letter sounds in English, but do not demonstrate understanding</li> <li>• not react or respond to letters, words, texts, or concrete symbols</li> </ul>	<p>These students may:</p> <ul style="list-style-type: none"> <li>• require significant second language acquisition support (visuals, gestures, verbal cues) for social and academic reading activities in English in their daily routine</li> <li>• understand letters or words when paired with concrete symbols</li> <li>• recognize a limited number of concrete symbols combined with letters or high frequency/high need sight words</li> </ul>	<p>These students may:</p> <ul style="list-style-type: none"> <li>• require moderate second language acquisition support (visuals, gestures, verbal cues) for social and academic reading activities in English in their daily routine</li> <li>• understand a few single printed words, but not how they fit into a larger language context</li> <li>• identify a limited number of high frequency/high need sight words</li> <li>• demonstrate the ability to pair letters with sounds</li> </ul>	<p>These students may:</p> <ul style="list-style-type: none"> <li>• require occasional second language acquisition support (visuals, gestures, verbal cues) for social and academic reading activities in English in their daily routine</li> <li>• understand how individual words or phrases connect to other words or phrases</li> <li>• decode simple words or phrases</li> </ul>	<p>These students may:</p> <ul style="list-style-type: none"> <li>• require minimal second language acquisition support (visuals, gestures, verbal cues) for social and academic reading activities in English in their daily routine across settings</li> <li>• understand longer and more detailed social and academic texts in English</li> <li>• decode longer and less familiar words, phrases, or sentences</li> </ul>

## TELPAS Alternate Proficiency Level Descriptors

### Grades 2–12 Writing

Awareness	Imitation	Early Independence	Developing Independence	Basic Fluency
<p>ELs at this level have little or no functional use of written English or augmentative and alternative communication even when interacting with highly familiar letters or words.</p>	<p>ELs at this level approximate writing or augmentative and alternative communication activities in English when interacting with highly familiar letters and words.</p>	<p>ELs at this level participate in writing or augmentative and alternative communication activities in English when working with letters and highly familiar words.</p>	<p>ELs at this level participate meaningfully in writing or using augmentative and alternative communication activities to write simple, short original messages in English.</p>	<p>ELs at this level are able to produce detailed and complex written messages in English or with augmentative and alternative communication.</p>
<p>These students may:</p> <ul style="list-style-type: none"> <li>• require full second language acquisition support (visuals, gestures, verbal cues) for social and academic writing activities in English in their daily routine</li> <li>• know too little English to participate in shared writing activities</li> <li>• alert to or show reaction to stimuli, but do not demonstrate the ability to communicate their thoughts in writing</li> </ul>	<p>These students may:</p> <ul style="list-style-type: none"> <li>• require significant second language acquisition support (visuals, gestures, verbal cues) for social and academic writing activities in English in their daily routine</li> <li>• begin to participate in routine shared writing activities with teacher modeling</li> <li>• attempt to write symbols or letters</li> </ul>	<p>These students may:</p> <ul style="list-style-type: none"> <li>• require moderate second language acquisition support (visuals, gestures, verbal cues) for social and academic writing activities in English in their daily routine</li> <li>• actively participate in shared writing activities when the writing topic is routine and concrete</li> <li>• write a few single letters or words, but do not understand how they fit into a larger language context*</li> </ul>	<p>These students may:</p> <ul style="list-style-type: none"> <li>• require occasional second language acquisition support (visuals, gestures, verbal cues) for social and academic writing activities in English in their daily routine</li> <li>• actively participate in shared writing activities when the writing topic is familiar</li> <li>• write simple, original messages consisting of a few words or memorized phrases*</li> <li>• label, list, and copy high frequency words and phrases*</li> </ul>	<p>These students may:</p> <ul style="list-style-type: none"> <li>• require minimal second language acquisition support (visuals, gestures, verbal cues) for social and academic writing activities in English in their daily routine across settings</li> <li>• contribute written text to shared writing activities when the writing topic has been recently experienced*</li> <li>• write detailed texts in social and academic English*</li> </ul>

\*Students using augmentative and alternate communication (AAC) as expressive communication must use symbols that are combined with printed letters or words beginning with the "Early Independence" level.

# Appendix B — Participants' Meeting Materials

The materials developed for the Reading standard-setting committee are provided as an example of what was developed and provided to the participants.

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) Alternate Standard Setting Meeting

## Participant Information Survey 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 in a bilingual or ESL classroom?

- 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 special education in a functional academic classroom?

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

Which of the following certifications do you currently hold?

(Check all that apply.)

- Bilingual education certification
- ESL education certification
- Special education

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

TELPAS Alternate  
 Standard Setting Meeting  
 June 2019



**Judgment Round Record Sheet**  
**READING**

***"How many points would a student with proficiency at the borderline of the level likely earn for the observable behavior?"***

Observable Behavior	Reporting Category	Maximum Points	Judgment Round											
			1				2				3			
			I	EI	DI	BF	I	EI	DI	BF	I	EI	DI	BF
R1	Reading_01	5												
R2	Reading_01	5												
R3	Reading_01	5												
R4	Reading_01	5												
R5	Reading_01	5												
R6	Reading_02	5												
R7	Reading_02	5												
R8	Reading_02	5												
R9	Reading_03	5												
R10	Reading_03	5												

*Note: I=Imitation; EI=Early Independence; DI=Developing Independence; BF=Basic Fluency*

# TELPAS Alternate Standard Setting Meeting Reading

## Round 1 Judgment Survey

**You are now ready to begin!**

For each observable behavior in the Round 1 set, do the following for each proficiency level:

- Review the observable behaviors, behaviors and classroom examples.
- Review the borderline descriptions for the proficiency level.
- Answer the following questions:

*"How many points would a student with proficiency at the borderline of the level likely earn for the observable behavior?"*

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

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

You will now start the judgment process for the observed behaviors in the Round 1 set.

For each of the items, answer the following question:

*"How many points would a student with proficiency at the borderline of the level likely earn for the observable behavior?"*

**Item: R1**

Reporting Category: 1

R1.	The student:				
Understanding Letter-Sound Relationships	may or may not attend to a teacher modeling matching letter/sound combinations to the same letter/sound combinations <b>(A)</b>	matches familiar letter/sound combinations to the same letter/sound combinations <b>(B)</b>	pairs the initial sound of words with pictures of objects that have the same initial sound <b>(C)</b>	pairs the sound made by consonant blends (e.g., <i>bl, br, sm</i> ) with pictures of words containing that blend <b>(D)</b>	pairs the sound made by digraphs (e.g., <i>wh, sh, ch</i> ) with pictures of words containing that digraph <b>(E)</b>

- Imitation
- Early Independence
- Developing Independence
- Basic Fluency

	1 Point (A)	2 Points (B)	3 Points (C)	4 Points (D)	5 Points (E)
●	<input type="radio"/>				
●	<input type="radio"/>				
●	<input type="radio"/>				
●	<input type="radio"/>				

# TELPAS Alternate Standard Setting Meeting

## Reading

### Process Evaluation Survey #1

The purpose of this evaluation is to collect information about your experience in recommending cut scores associated with the proficiency levels for the TELPAS Alternate assessments. Your opinions provide 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 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
General session overview of the TELPAS Alternate assessments	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
General session introduction to the standard setting process	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Overview of the specific TELPAS Alternate assessment and scoring	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Discussion of the Alternate Proficiency Level Descriptors (PLDs)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Development and discussion of the borderline descriptions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Training on the standard-setting procedure	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

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
Review of the Alternate Proficiency Level Descriptors (PLDs)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Borderline descriptions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How adequate were the following elements of the session?

	Not Adequate	Somewhat Adequate	Adequate	More Than Adequate
Total amount of time to create and discuss borderline descriptions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Training provided on the standard-setting process	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Amount of time spent training	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

# Appendix C — Committee Participant Composition

**Table C.1. Participant Position**

	All Participants
Teacher (K–12)	19
Teacher (Higher Ed.)	1
Administrator (School)	0
Administrator (District)	3
Other	5
<b>Total</b>	<b>28</b>

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

	All Participants
None	0
1 to 5 years	3
6 to 10 years	3
11 to 15 years	6
16 to 20 years	3
More than 20 years	13
<b>Total</b>	<b>28</b>

**Table C.3. Years of Teaching Experience in Bilingual or ESL Classroom**

	All Participants
None	6
1 to 5 years	6
6 to 10 years	4
11 to 15 years	4
16 to 20 years	3
More than 20 years	5
<b>Total</b>	<b>28</b>

**Table C.4. Years of Teaching Experience in Special Education Classroom**

	All Participants
None	7
1 to 5 years	5
6 to 10 years	8
11 to 15 years	3
16 to 20 years	5
More than 20 years	0
<b>Total</b>	<b>28</b>

**Table C.5. Bilingual, ESL, or Special Education Teaching Certifications**

	All Participants
Bilingual Education	7
ESL Education	17
Special Education	18
<b>Total</b>	<b>35</b>

**Table C.6. Experience Teaching Languages**

	All Participants
Spanish	25
Vietnamese	0
Chinese	0
Tagalog	1
German	0
French	0
Hindi	0
Urdu	0
Korean	0
Arabic	0
Other	4
<b>Total</b>	<b>30</b>

**Table C.7. Highest Education Degree**

	All Participants
High School Diploma	0
Associate's degree	0
Bachelor's degree	9
Master's degree	19
Doctoral degree	0
<b>Total</b>	<b>28</b>

**Table C.8. Demographic: Gender**

	All Participants
Male	2
Female	26
No answer	0
<b>Total</b>	<b>28</b>

**Table C.9. Demographic: Ethnicity**

	All Participants
Hispanic or Latino	13
Not Hispanic or Latino	13
No answer	2
<b>Total</b>	<b>28</b>

**Table C.10. Demographic: Race**

	All Participants
American Indian or Alaskan Native	0
Asian	1
Black or African American	1
Native Hawaiian or Pacific Islander	0
White	23
No answer	3
<b>Total</b>	<b>28</b>

**Table C.11. *Currently Work in a School District***

	All Participants
Yes	24
No	4
<b>Total</b>	<b>28</b>

**Table C.12. *Size of School District***

	All Participants
Small	4
Medium	8
Large	12
No answer	4
<b>Total</b>	<b>28</b>

**Table C.13. Type of School District**

	All Participants
Rural	6
Metropolitan/Urban	7
Suburban	11
No answer	4
<b>Total</b>	<b>28</b>

**Table C.14. Socioeconomic Status of School District**

	All Participants
Low	10
Moderate	13
High	1
No answer	4
<b>Total</b>	<b>28</b>

# Appendix D — Meeting Agenda

## TELPAS Alternate Standard Setting Meeting June 2019 Agenda



### Day 1

8:00 – 8:30 a.m. *Breakfast*

#### **General Session**

8:30 – 9:00 a.m. Welcome and orientation  
Welcome and introduction  
Meeting agenda and security policies  
TELPAS Alternate orientation

9:00 – 10:00 a.m. Standard setting overview

#### **Breakout Session (Reading or Listening)**

10:00 – 10:30 a.m. Introductions and orientation  
Participant introductions  
Meeting security policies and Non-disclosure agreement  
Computer and materials orientation

10:30 – 11:15 a.m. Review and discuss the assessment (Reading or Listening)

11:15 – 11:45 a.m. TELPAS Alternate PLDs (Reading or Listening)  
Introduction to Alternate PLDs  
Table-group review of Alternate PLDs  
Whole-group discussion of Alternate PLDs

11:45 – 12:30 p.m. *Lunch*

12:30 – 2:15 p.m. Borderline descriptions (Reading or Listening)  
Introduction to borderline descriptions (10 min)  
Modeling the creation of borderline descriptions (20 min)  
Table-group discussion (30 min)  
Whole-group development of borderline descriptions (45 min)

2:15 – 3:00 p.m. Standard setting training

3:00 – 4:30 p.m. Round 1 judgments (Reading or Listening)

## **Day 2**

8:00 – 8:30 a.m.	<i>Breakfast</i>
	<b>Breakout Session (Reading or Listening)</b>
8:30 – 9:30 a.m.	Round 1 judgment feedback and discussion (Reading or Listening) Introduction to data (15 minutes) Table group discussion (45 minutes)
9:30 – 10:15 a.m.	Round 2 judgments (Reading or Listening)
10:15 – 10:45 a.m.	<b>Break (Data analysis)</b>
10:45 – 11:30 a.m.	Round 2 judgment feedback and discussion (Reading or Listening) Table group discussion (20 minutes) Whole group discussion (25 minutes)
11:30 – 12:15 p.m.	<i>Lunch</i>
12:15 – 1:00 p.m.	Round 3 judgments (Reading or Listening)
	<b>Breakout Session (Writing or Speaking)</b>
1:00 – 1:15 p.m.	<b>Break/Transition</b>
1:15 – 1:30 p.m.	Introductions
1:30 – 2:00 p.m.	Review and discuss assessment (Writing or Speaking)
2:00 – 2:30 p.m.	TELPAS Alternate PLDs (Writing or Speaking) Table group review of the Alternate PLDs (10 minutes) Whole group discussion of Alternate PLDs (20 minutes)
2:30 – 4:00 p.m.	Borderline descriptions (Writing or Speaking) Table-group development of borderline descriptions (30 min) Whole-group development of borderline descriptions (45 min)
4:00 – 5:00 p.m.	Round 1 judgments (Writing or Speaking)

### **Day 3**

8:00 – 8:30 am	<i>Breakfast</i>
	<b>Breakout Session (Writing or Speaking)</b>
8:30 – 9:15 a.m.	Round 1 judgment feedback and discussion (Writing or Speaking) Table-group discussion (45 minutes)
9:15 – 10:00 a.m.	Round 2 judgments (Reading or Listening)
10:00 – 10:30 a.m.	<b>Break (Data analysis)</b>
10:30 – 11:15 a.m.	Round 2 judgment feedback and discussion (Writing or Speaking) Table-group discussion Whole-group discussion
11:15 – 11:45 a.m.	Round 3 judgments (Writing or Speaking)
11:45 – 12:00 p.m.	Close-out and evaluations
12:00 – 1:00 p.m.	<i>Lunch</i>
	<b>TELPAS Alternate Articulation and Composite Score (Table Leaders)</b>
1:00 – 1:15 p.m.	Introductions and orientation
1:15 – 1:45 p.m.	Review of TELPAS Alternate domain PLDs
1:45 – 2:30 p.m.	Review of domain impact data and articulation discussion
2:30 – 2:45 p.m.	<b>Break</b>
2:45 – 3:15 p.m.	Introduction to TELPAS Alternate domain composite score
3:15 – 4:00 p.m.	Proficiency level cut score discussion and judgment
4:00 – 4:30 p.m.	Close-out and evaluation

# Appendix E — Examples of Feedback Data

Feedback data were provided to participants after each judgment round. The following are examples of feedback data provided to participants in the TELPAS Alternate Reading committee. All domain committees received the same types of feedback data.

## Individual Test-Level Recommendations

Participants were provided their individual cut score recommendations, which were based on their judgments for the Imitation, Early Independence, Developing Independence, and Basic Fluency proficiency levels.

### Reading - Individual Cut Scores - Round 1

Table=1 Name=

I Raw Score	EI Raw Score	DI Raw Score	BF Raw Score
15	22	34	46

## Overall Test-Level Recommendations

Participants received the aggregate test-level cut score recommendations for the entire committee, which included the number of participants, mean recommendation, median recommendation, rounded median recommendation, minimum and maximum recommendation, and the first and third quartiles for each proficiency level.

### Reading Round 1 Summary Statistics - Overall

	N	Mean	Median	Roundup Median	Min	Max	Q1	Q3
I Raw Score	15	18.67	19.00	19	15.00	23.00	16.00	20.00
EI Raw Score	15	25.07	25.00	25	21.00	30.00	23.00	27.00
DI Raw Score	15	33.07	33.00	33	27.00	40.00	30.00	35.00
BF Raw Score	15	41.40	41.00	41	34.00	46.00	39.00	44.00

### Observable Behavior-Level Judgment Agreement

Observable Behavior-level judgment distributions were provided to the committee members for each behavior. Additionally, for each proficiency level, the Observable Behaviors with the greatest level of judgment disagreement were identified.

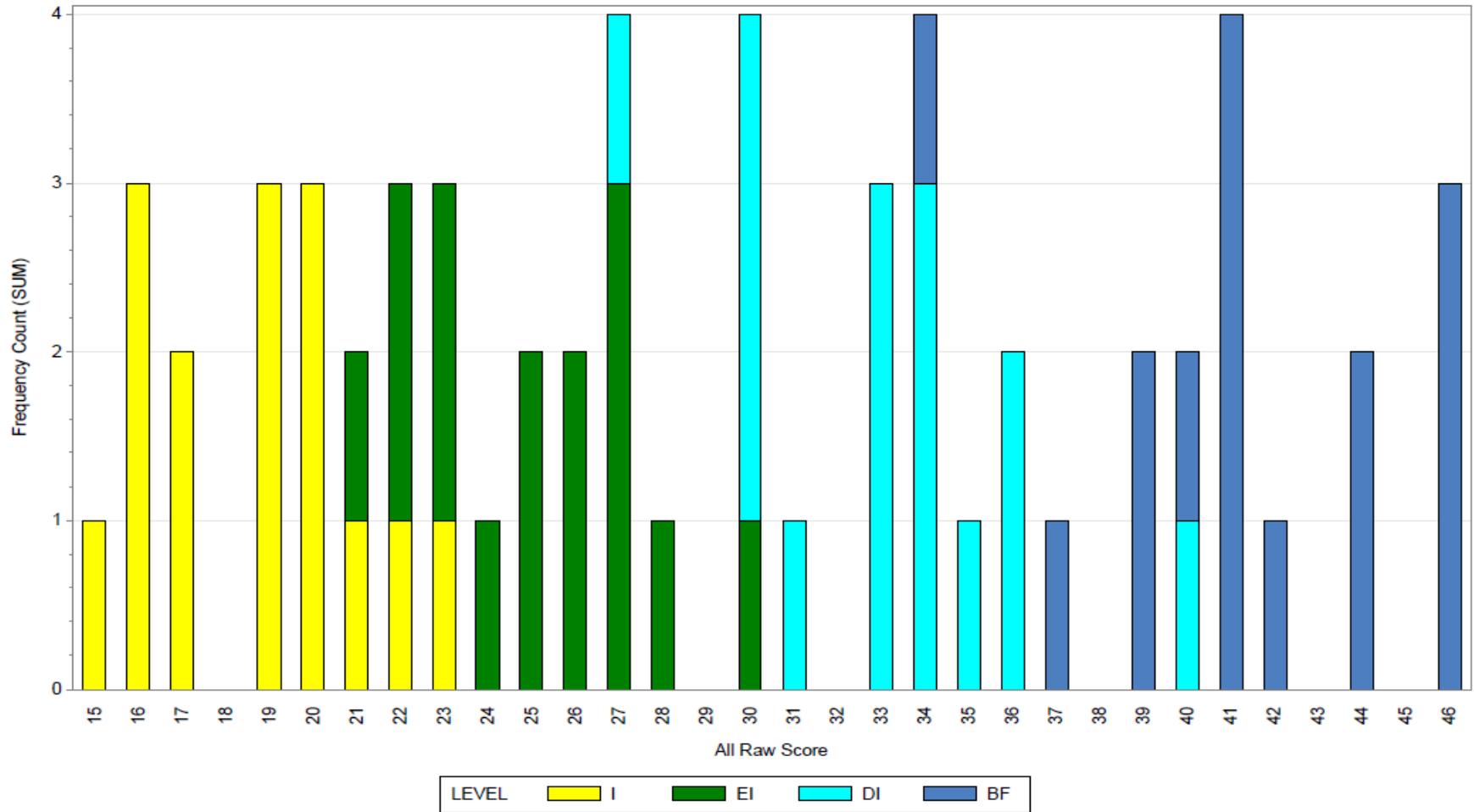
#### Reading Round 1 Level I

UIN	Max Points	1	2	3	4	5
R01	5	13%	87%	.	.	.
R02	5	13%	73%	13%	.	.
R03	5	7%	93%	.	.	.
R04	5	.	87%	13%	.	.
R05	5	33%	67%	.	.	.
R06	5	27%	73%	.	.	.
R07	5	13%	80%	7%	.	.
R08	5	27%	53%	20%	.	.
R09	5	33%	60%	7%	.	.
R10	5	33%	60%	7%	.	.

### Test-Level Panelist Recommendation Agreement

The facilitator presented feedback to participants that displayed bar graph distributions of panelist recommendations for the cut score, by raw score, for each proficiency level: Imitation, Early Independence, Developing Independence, and Basic Fluency. Graphs that showed panelist agreement across all proficiency levels were also presented.

### Reading Panelist Agreement at Level I, EI, DI and BF - Round 1



### Observable Behavior Score Mean and Score Distribution

The mean and distribution of scores was provided to participants for each item, using student data from the spring 2019 TELPAS Alternate administration. The results presented were based on the sample of data used to create the impact data.

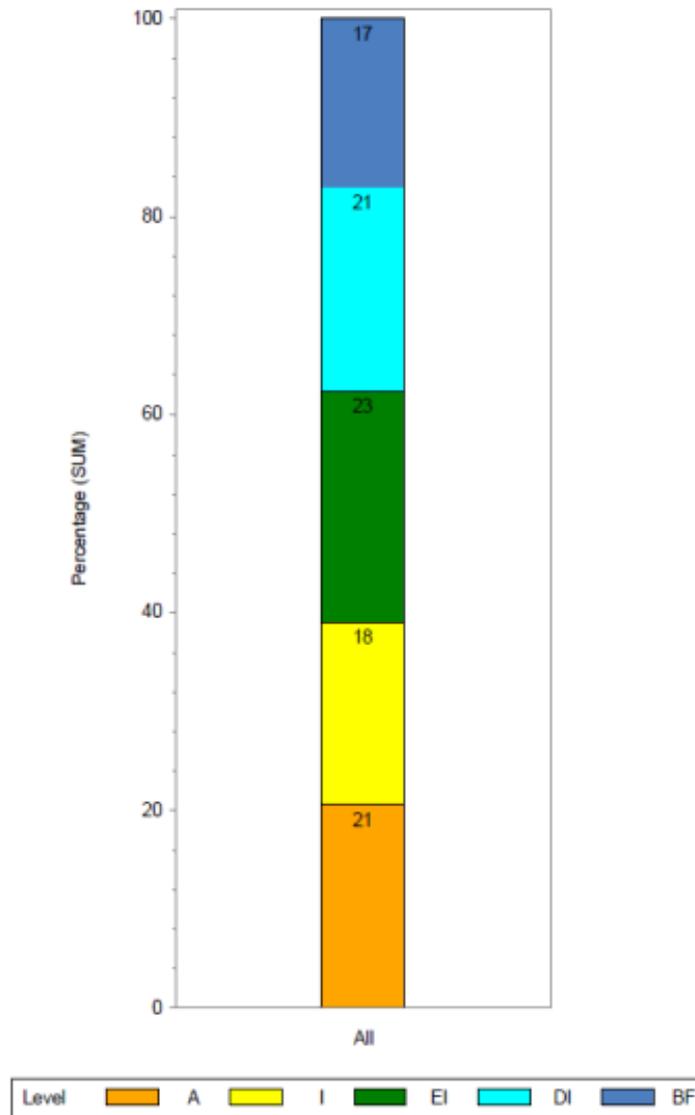
### Item Score Distribution Reading

Observable Behavior	Reporting Category	Maximum Points	Average Score	Score				
				1 (A)	2 (B)	3 (C)	4 (D)	5 (E)
R1	Reading_01	5	3.0	18.2%	19.0%	27.7%	14.4%	20.6%
R2	Reading_01	5	3.0	13.3%	29.0%	20.7%	21.2%	15.8%
R3	Reading_01	5	3.1	14.4%	25.0%	17.3%	24.7%	18.6%
R4	Reading_01	5	3.1	14.9%	23.9%	18.9%	25.5%	16.8%
R5	Reading_01	5	2.7	18.5%	28.2%	25.7%	16.8%	10.8%
R6	Reading_02	5	2.7	19.9%	29.6%	22.7%	12.1%	15.8%
R7	Reading_02	5	2.7	17.3%	34.4%	22.1%	16.5%	9.7%
R8	Reading_02	5	2.8	18.9%	23.1%	26.6%	23.4%	8.1%
R9	Reading_03	5	2.7	22.1%	27.9%	21.0%	18.4%	10.6%
R10	Reading_03	5	2.6	20.3%	30.4%	24.6%	17.3%	7.4%

## Impact Data

Impact data were provided to participants that illustrated the percentage of students expected to be classified into each proficiency level, based on the committee's test-level cut score recommendations for that round. The results were based on the sample of student data from the spring 2019 administration and were shared with participants after Round 2 judgments.

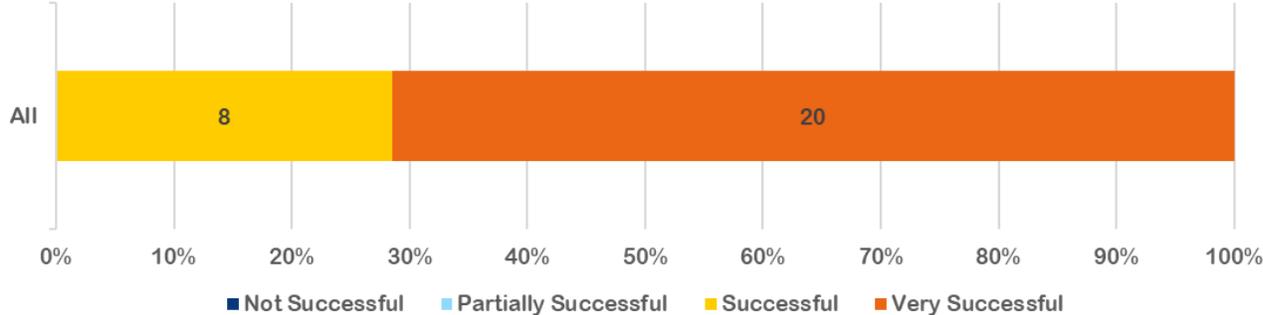
**Reading Round 2 Impact Data - Whole Group**



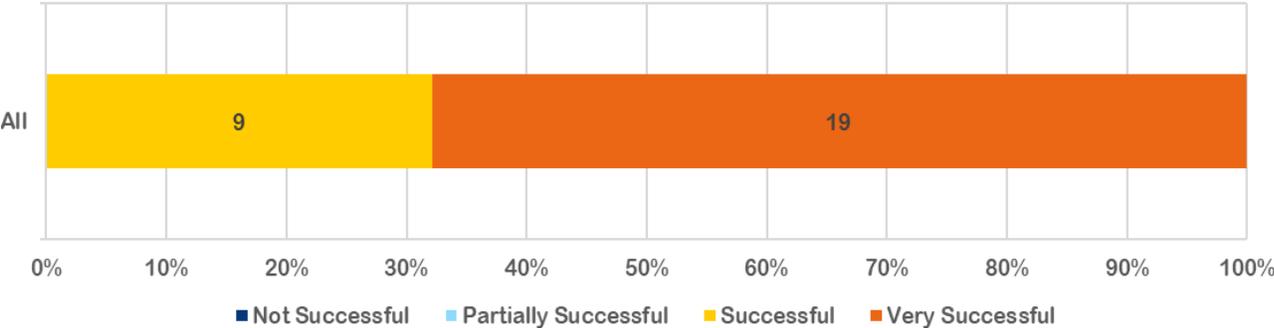
# Appendix F — Participant Evaluation Results

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.

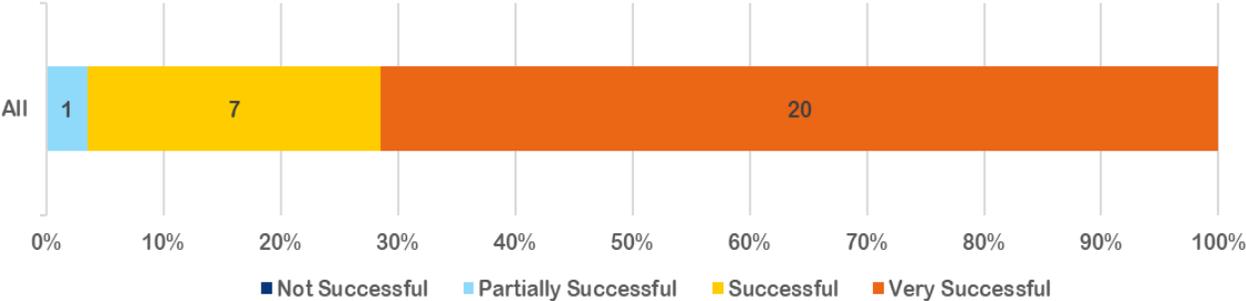
*General session overview of the TELPAS Alternate assessments*



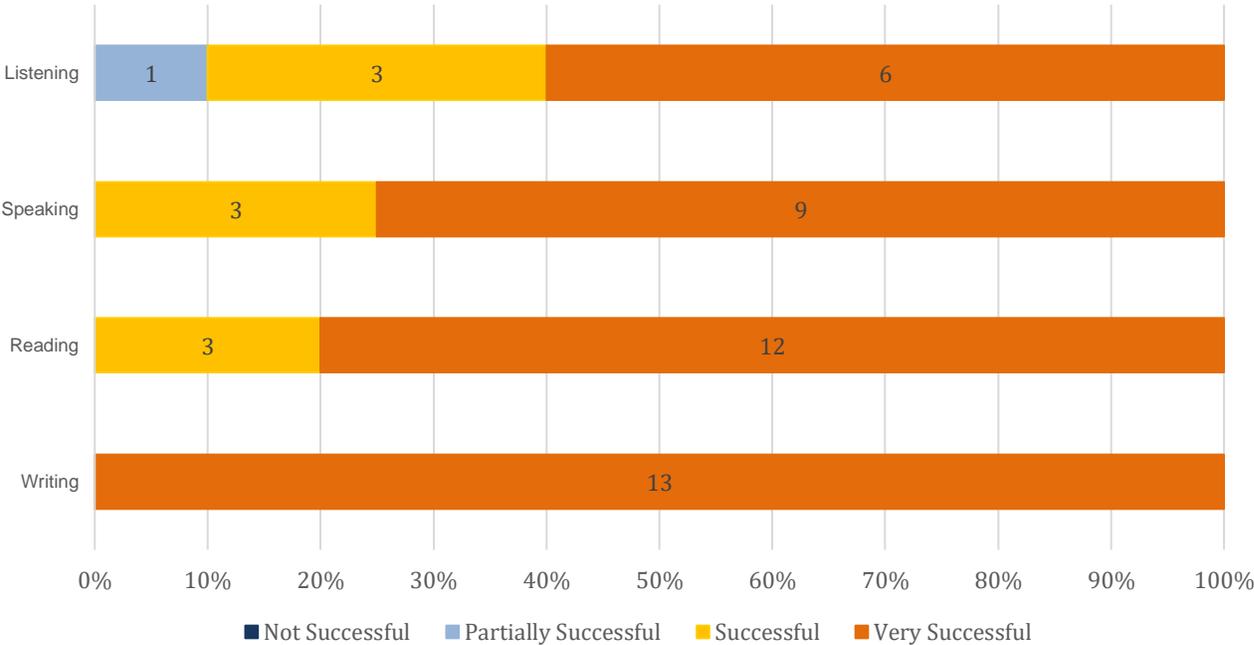
*General introduction to the standard-setting process*



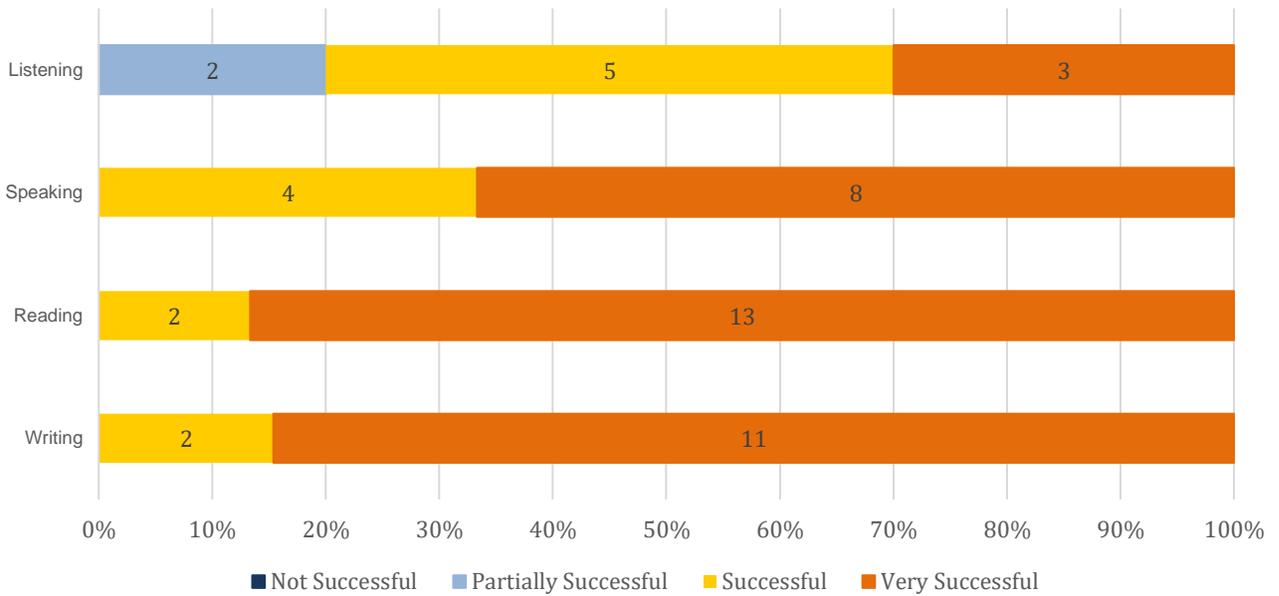
Overview of the specific TELPAS Alternate assessment and scoring



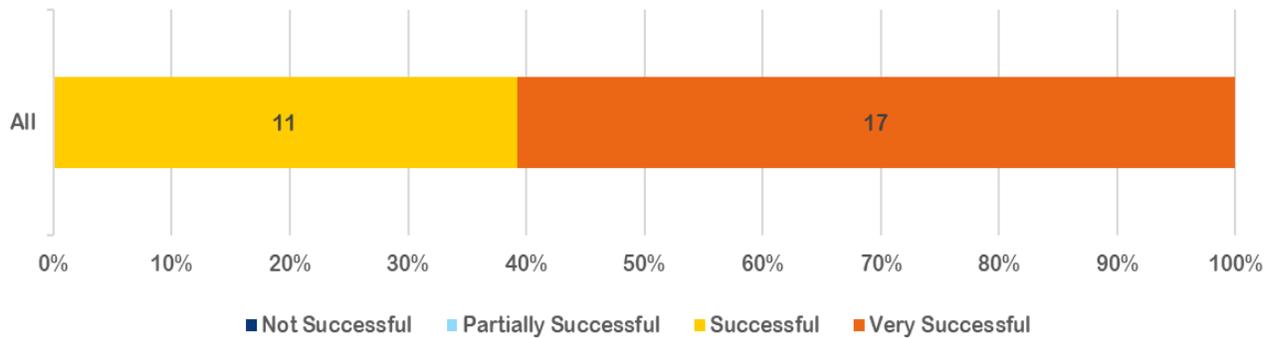
Discussion of the Alternate Proficiency Level Descriptors (PLDs)



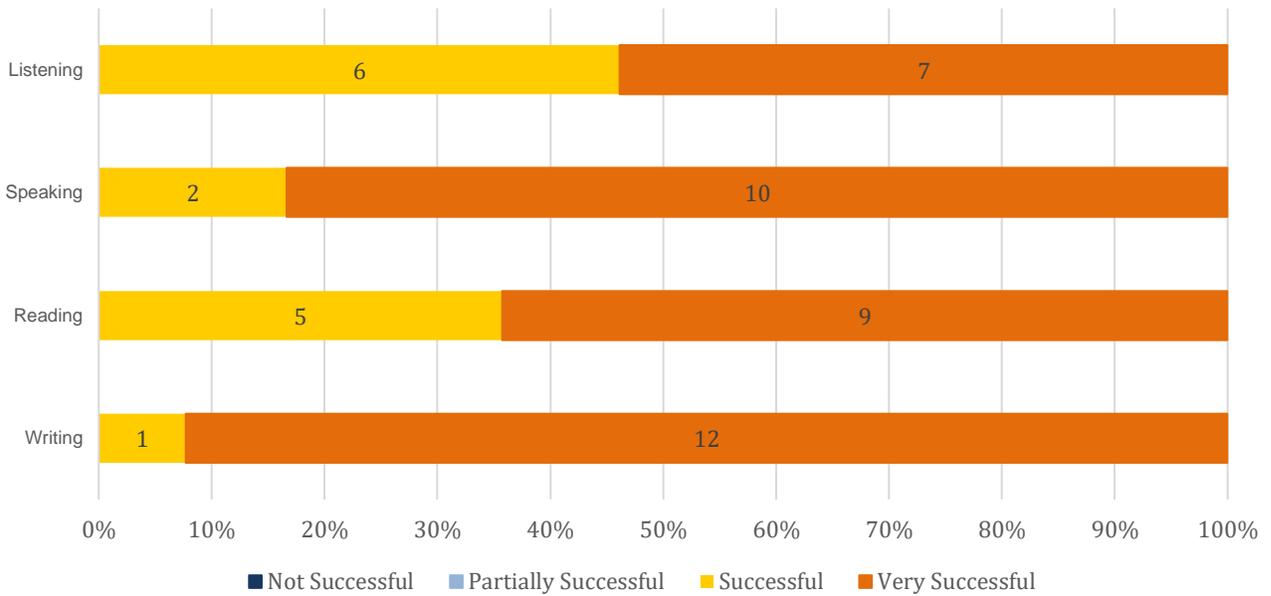
*Development and discussion of the borderline descriptions*



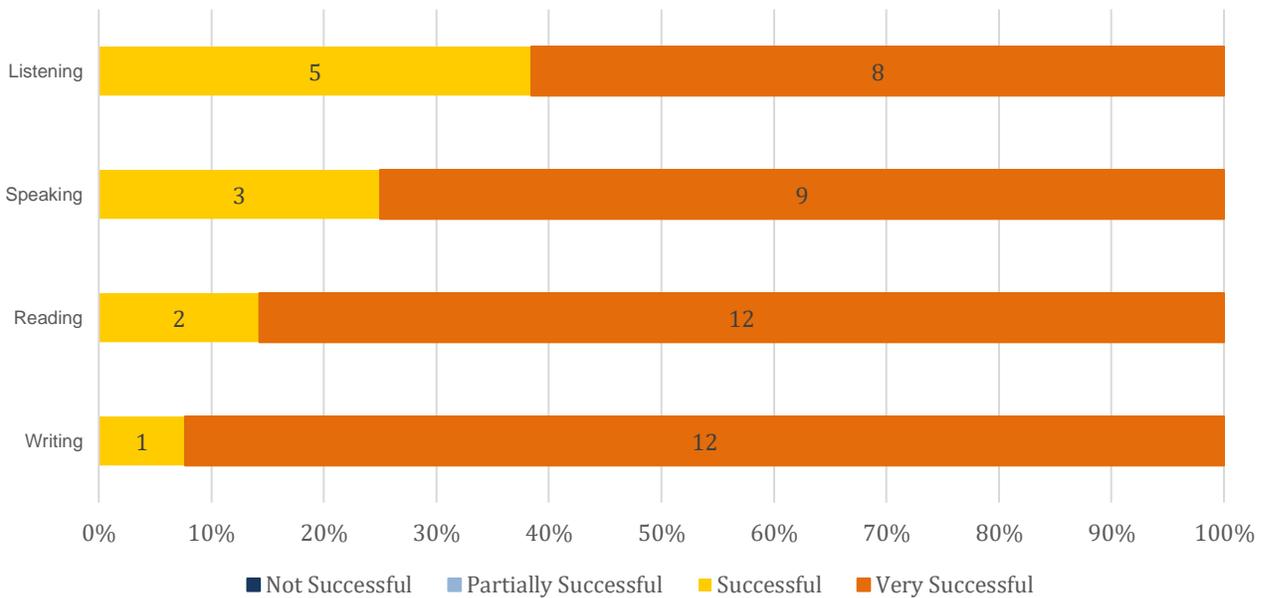
*Training on the standard-setting procedure*



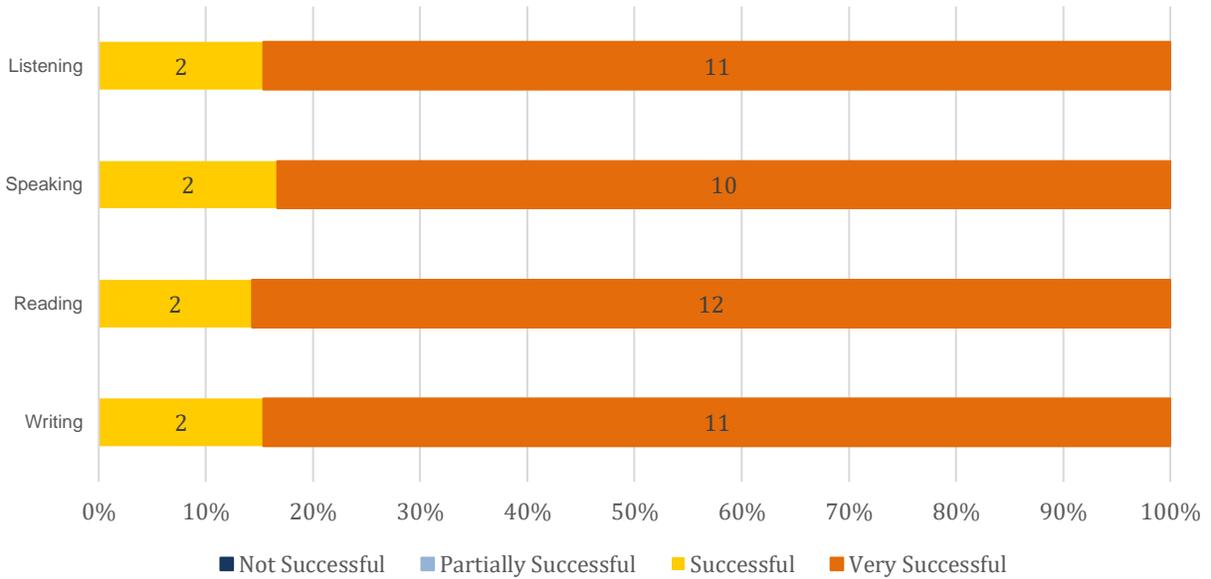
*Individual judgment round activity*



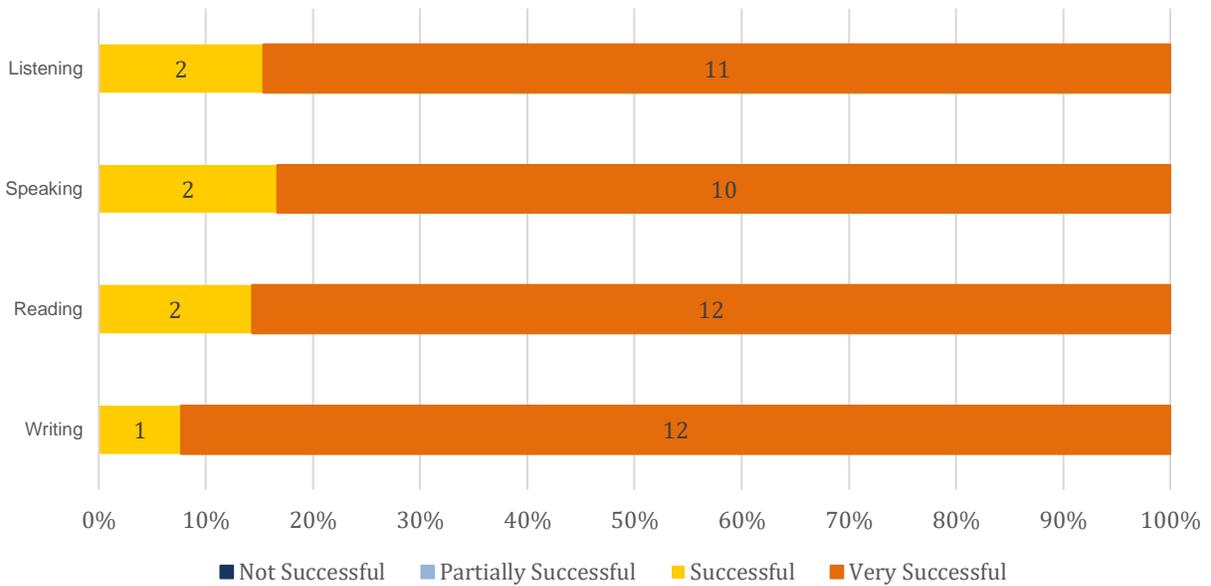
*Judgment round feedback – committee-level statistics*



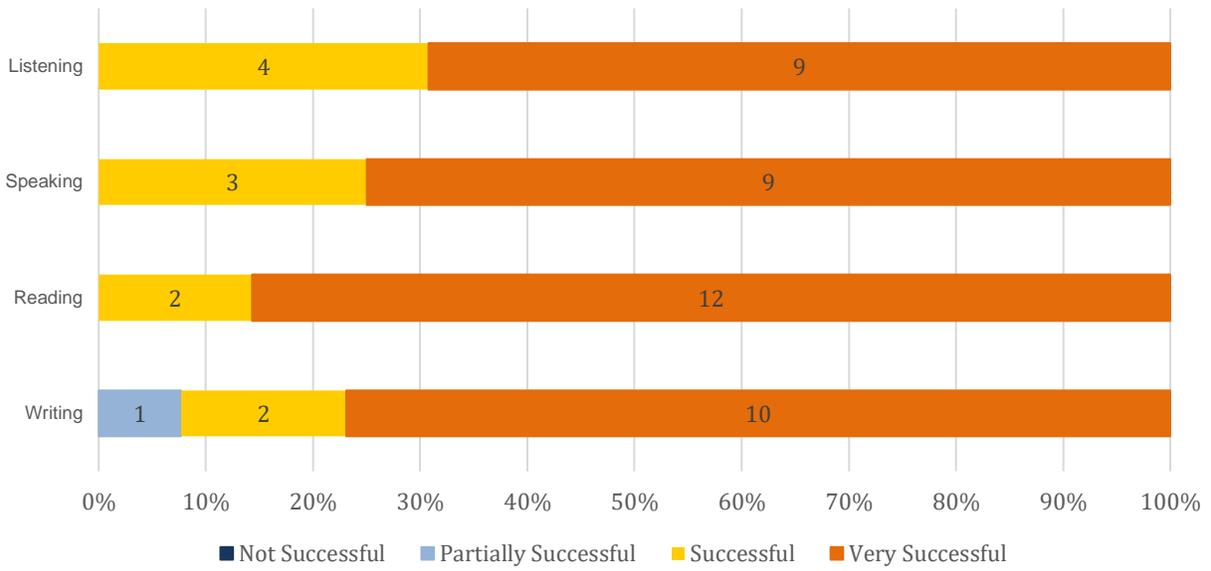
*Judgment round feedback – panelist agreement data*



*Judgment round feedback – impact data*

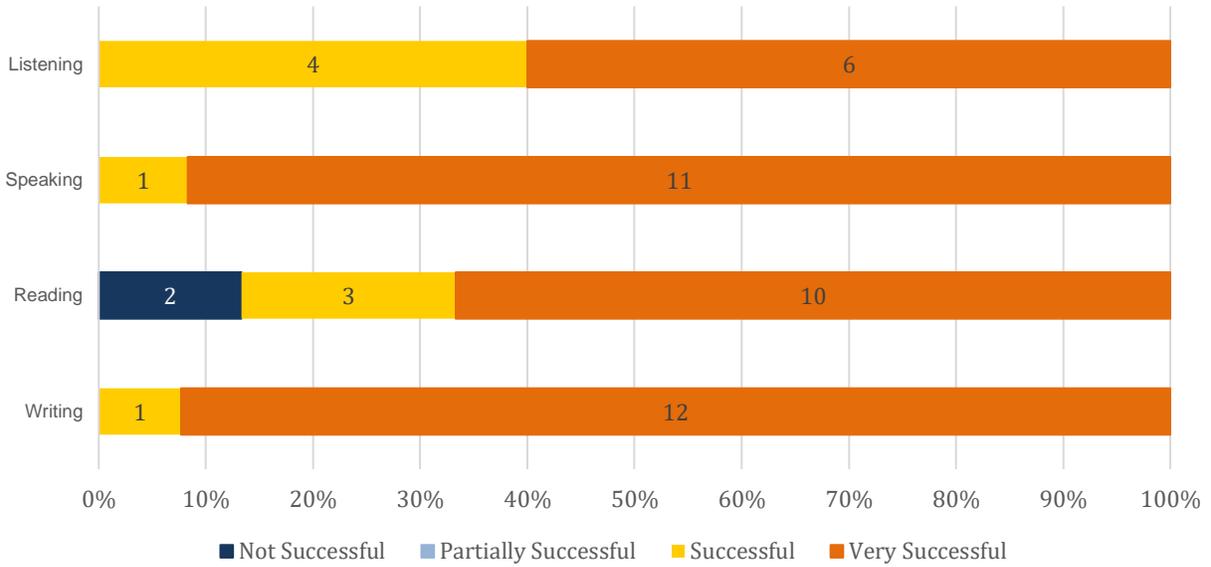


*Discussions after each round*

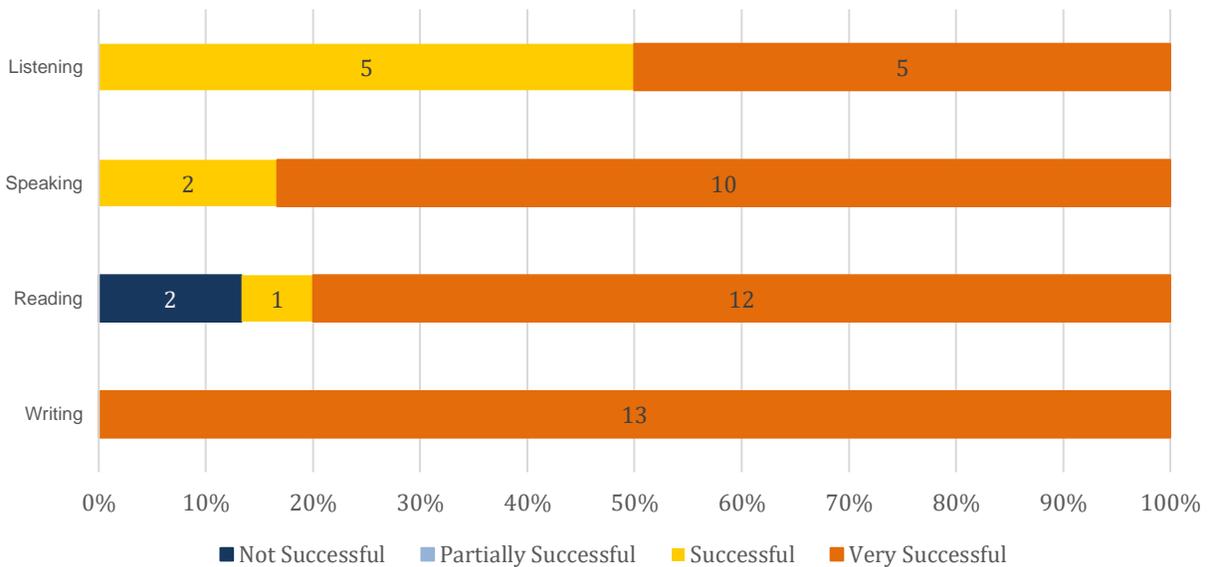


Question 2: How useful do you feel the following activities or information were in assisting you to make your recommendations?

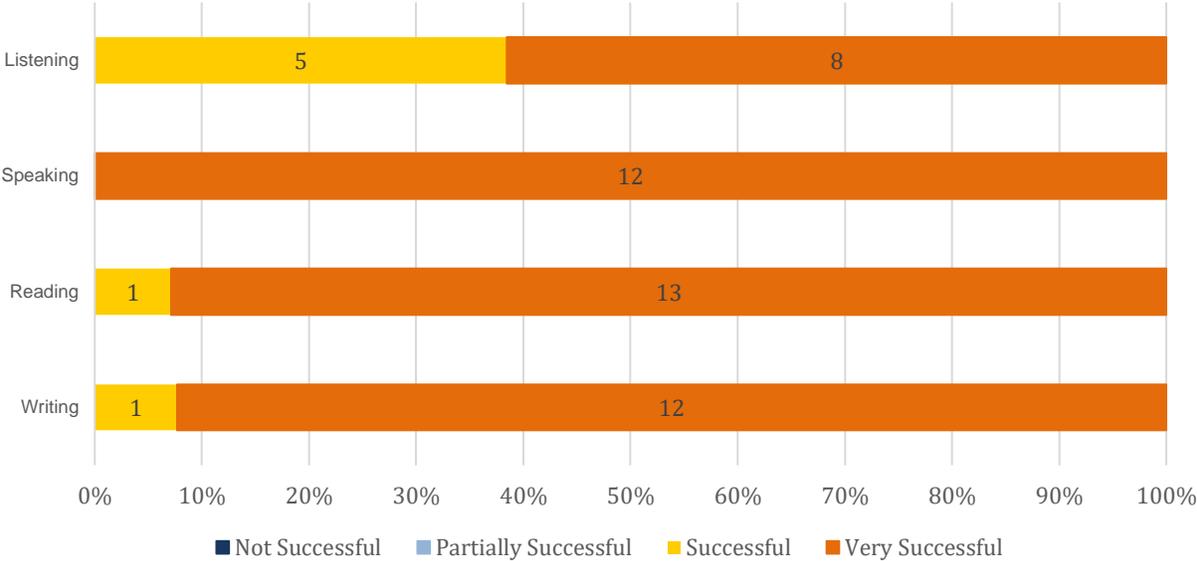
*Review of the Alternate Proficiency level descriptors (PLDs)*



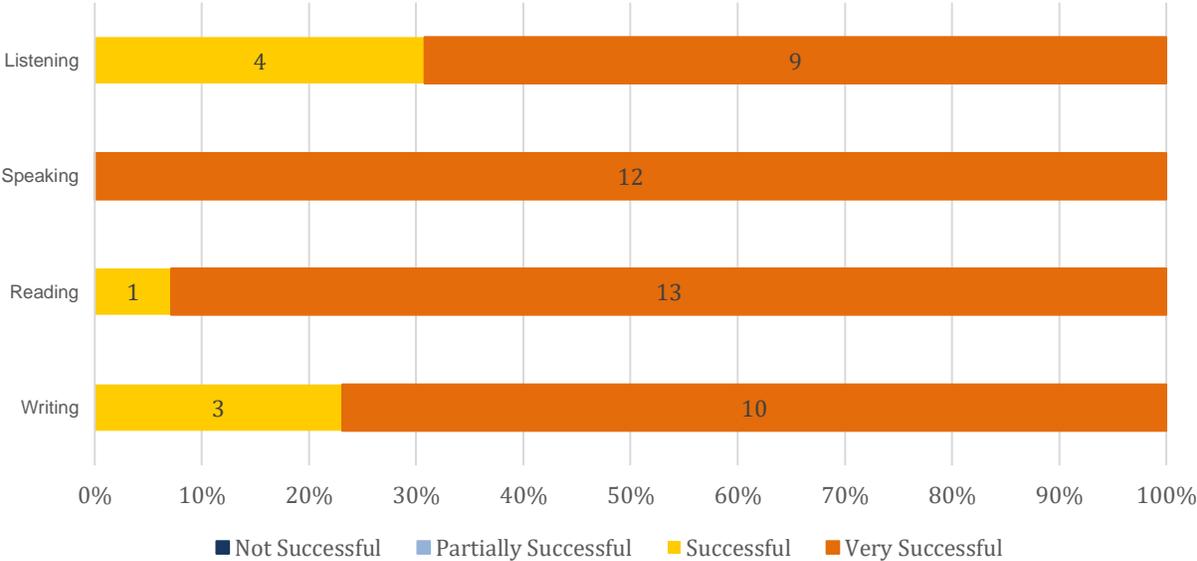
*Borderline descriptions*



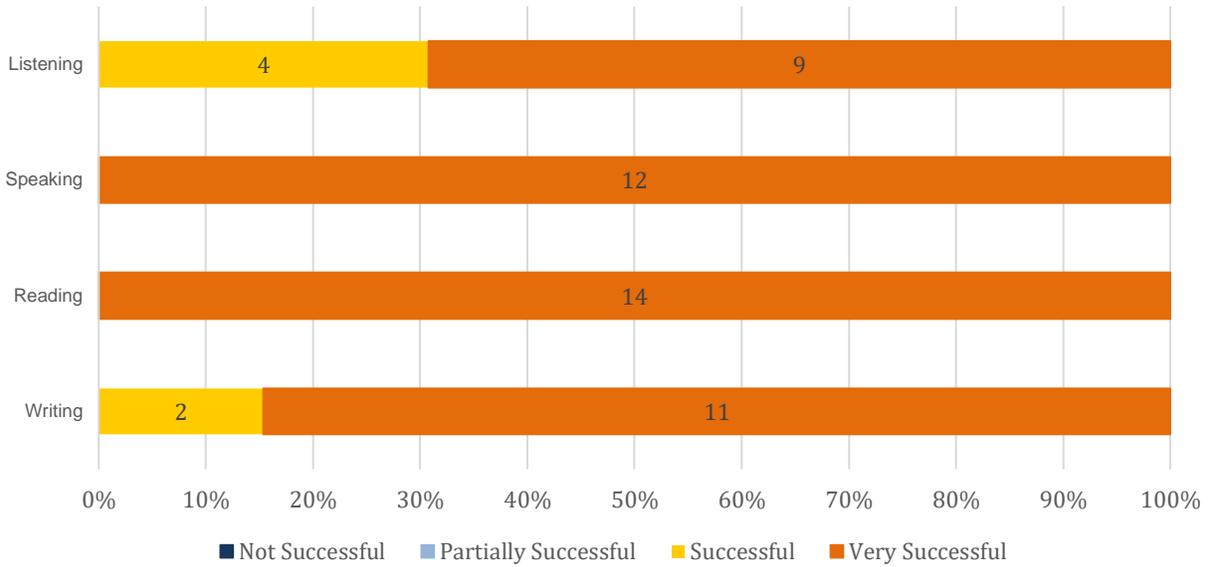
*Committee-level statistics after Rounds 1 and 2*



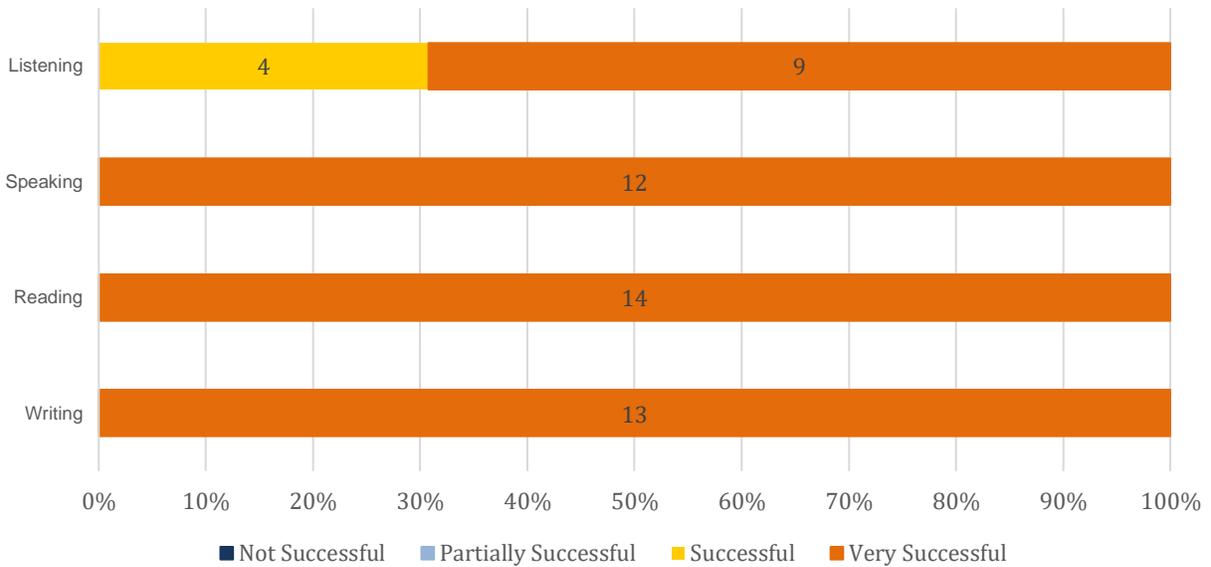
*Panelist agreement data provided after Round 1*



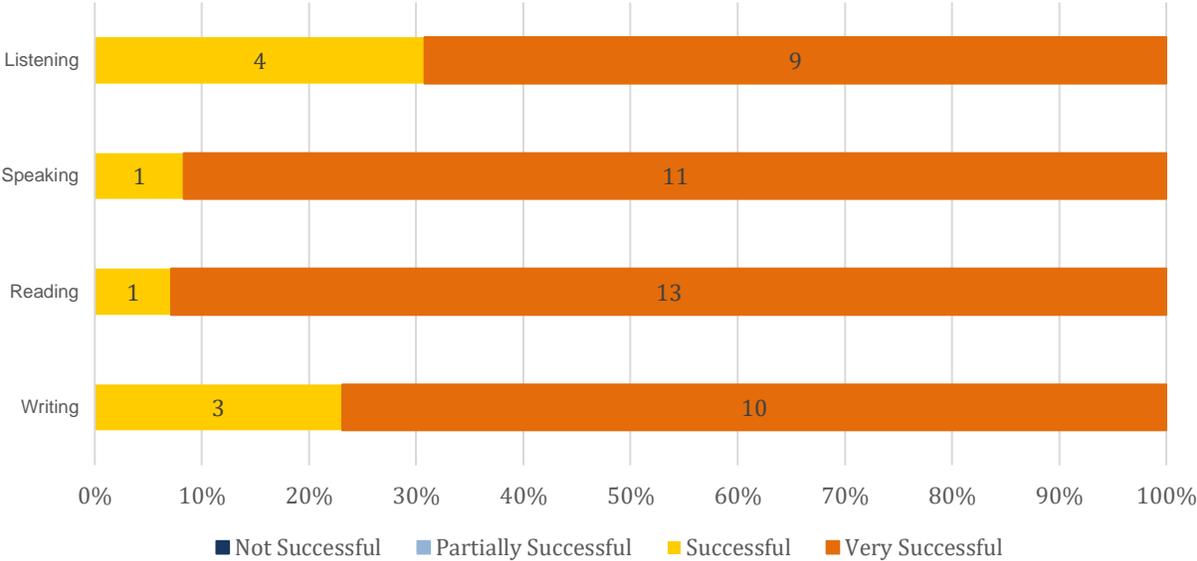
*Panelist agreement data provided after Round 2*



*Impact data after Round 2*

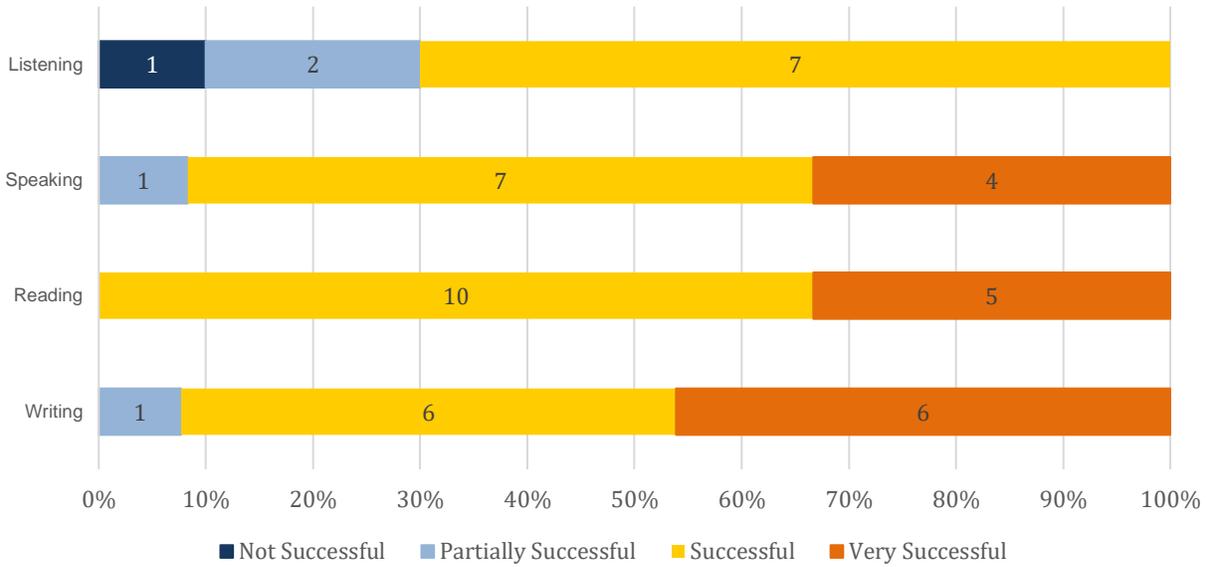


*Discussion after each judgment round*

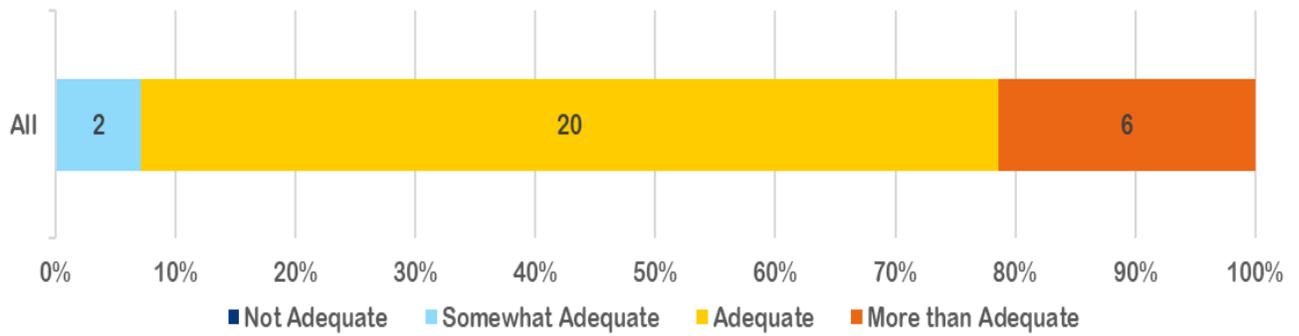


Question 3: How adequate were the following elements of the session?

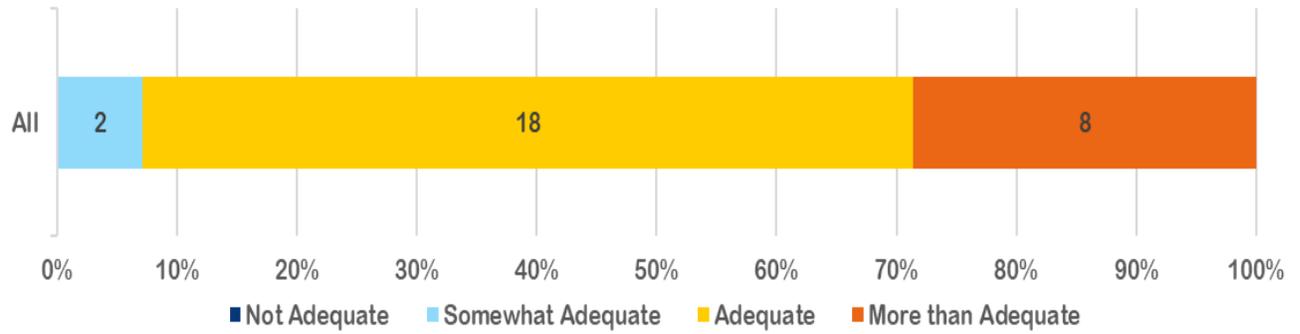
*Total amount of time to create and discuss borderline descriptions*



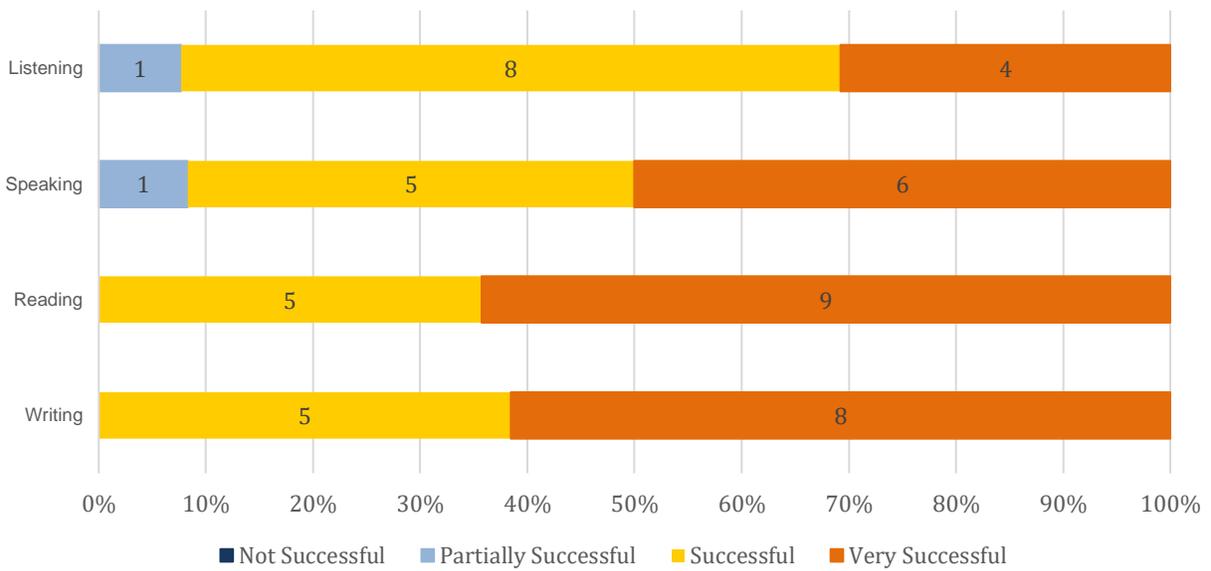
*Training provided on the standard-setting process*



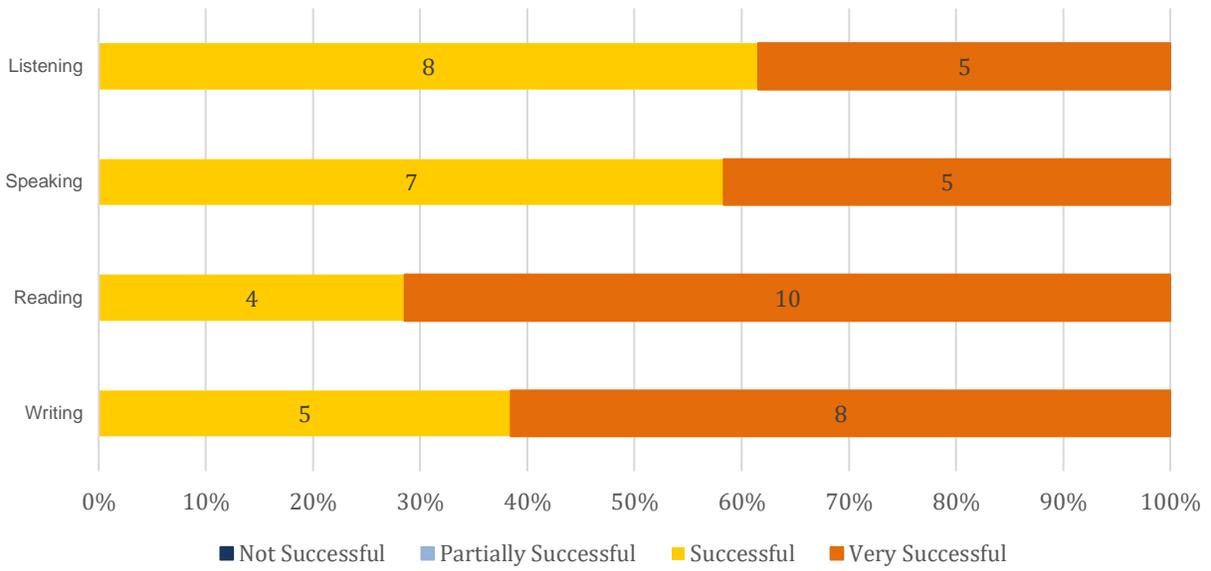
*Amount of time spent training*



*Amount of time to make judgments*

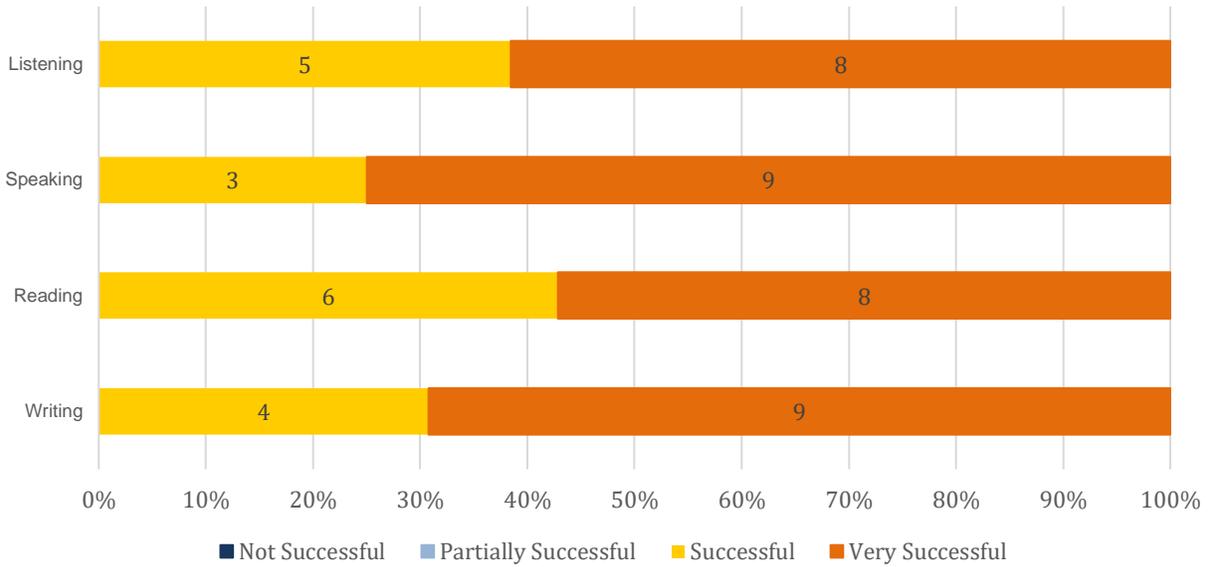


*Number of judgment rounds*

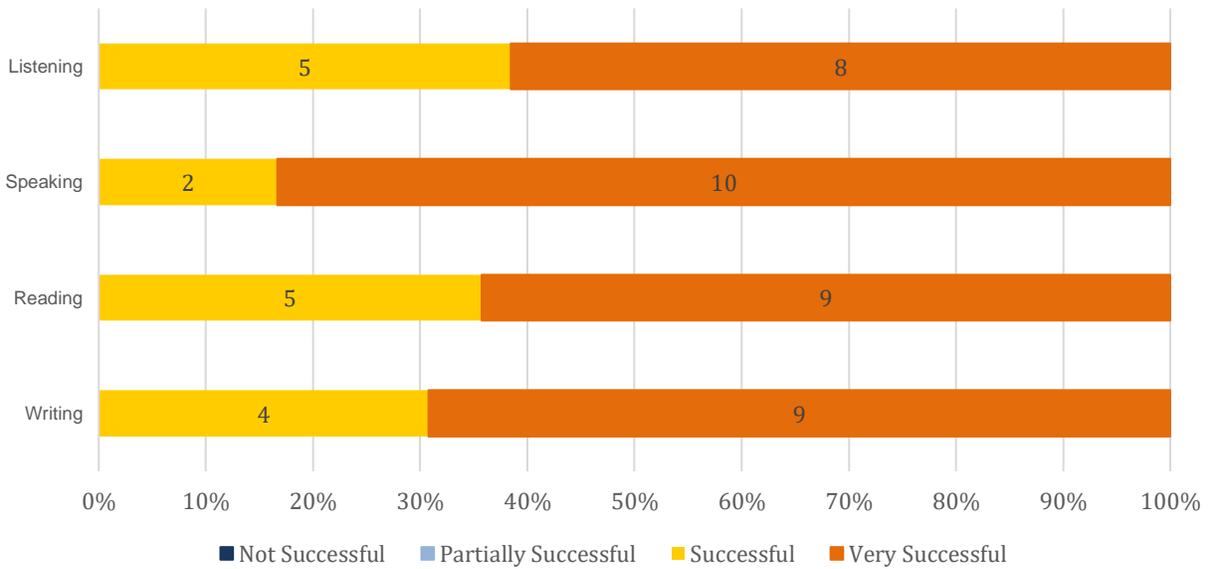


Question 4: In applying the standard-setting method, you were asked to recommend cut scores (separating five proficiency levels) for student performance on the TELPAS Alternate assessments. How confident do you feel that the Alternate Proficiency Level Descriptors (PLDs) are reasonable for each proficiency level?

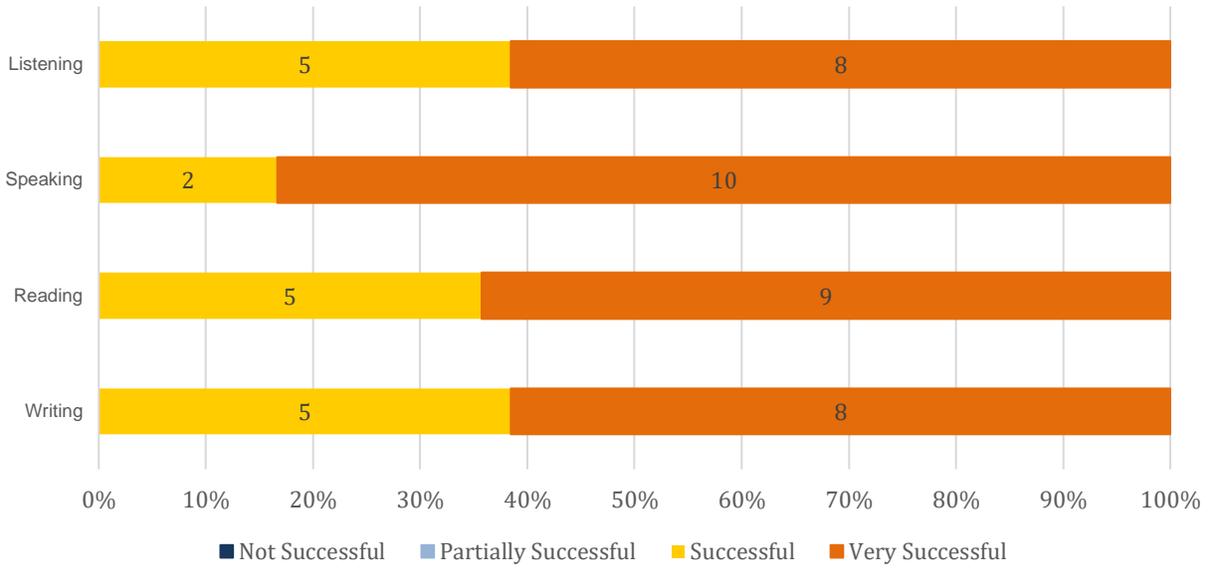
*Imitation*



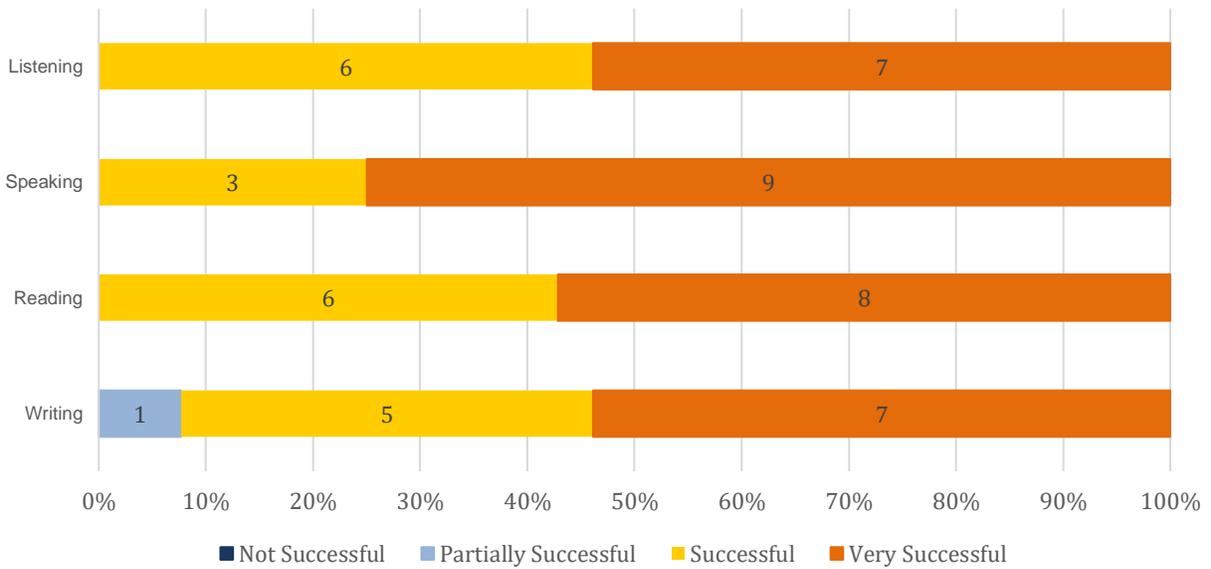
*Early Independence*



### Developing Independence

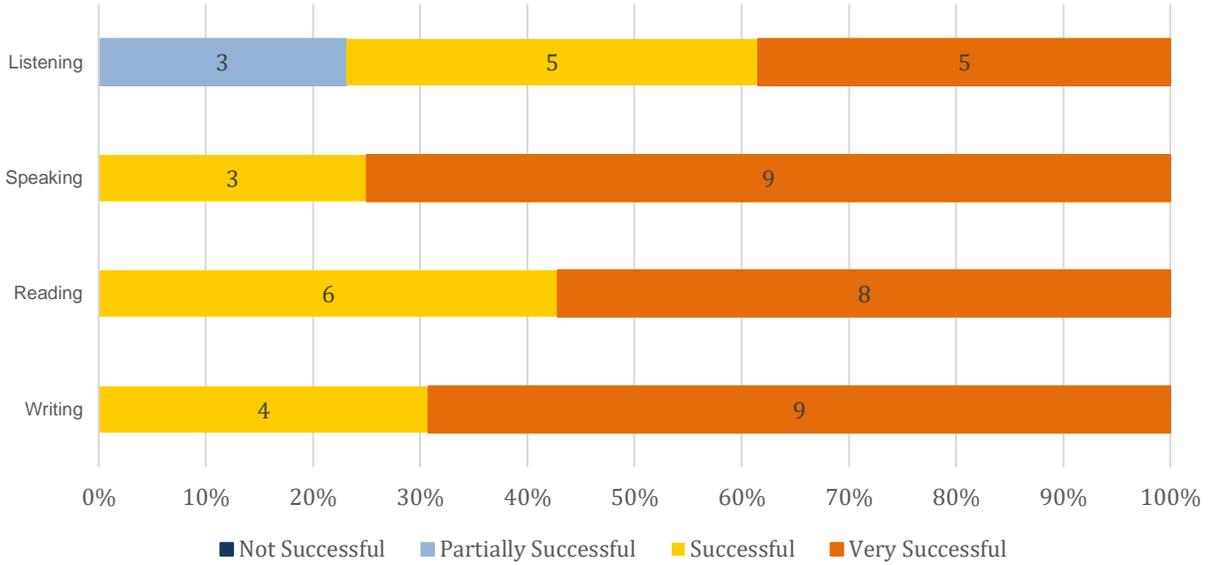


### Basic Fluency

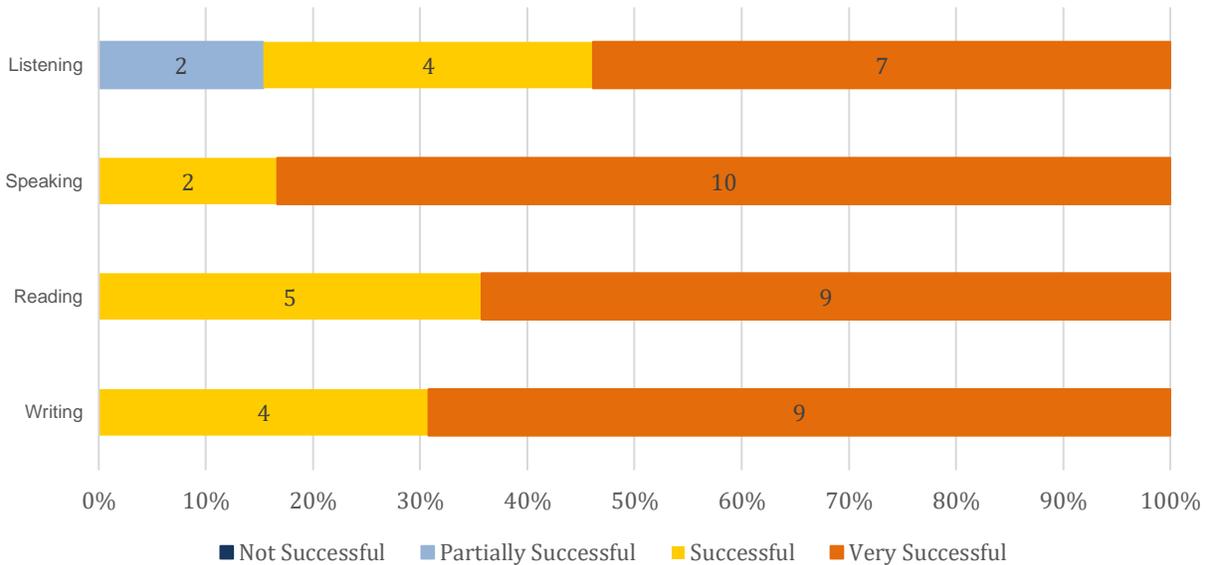


Question 5: How confident do you feel that the final cut score recommendations represent appropriate levels of student proficiency?

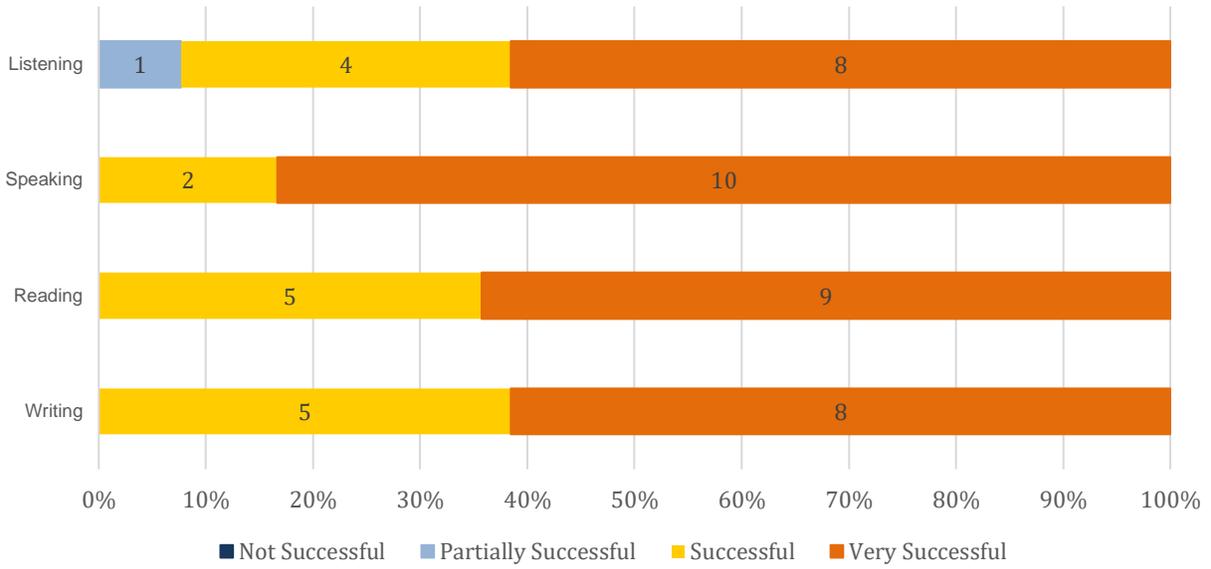
*Imitation*



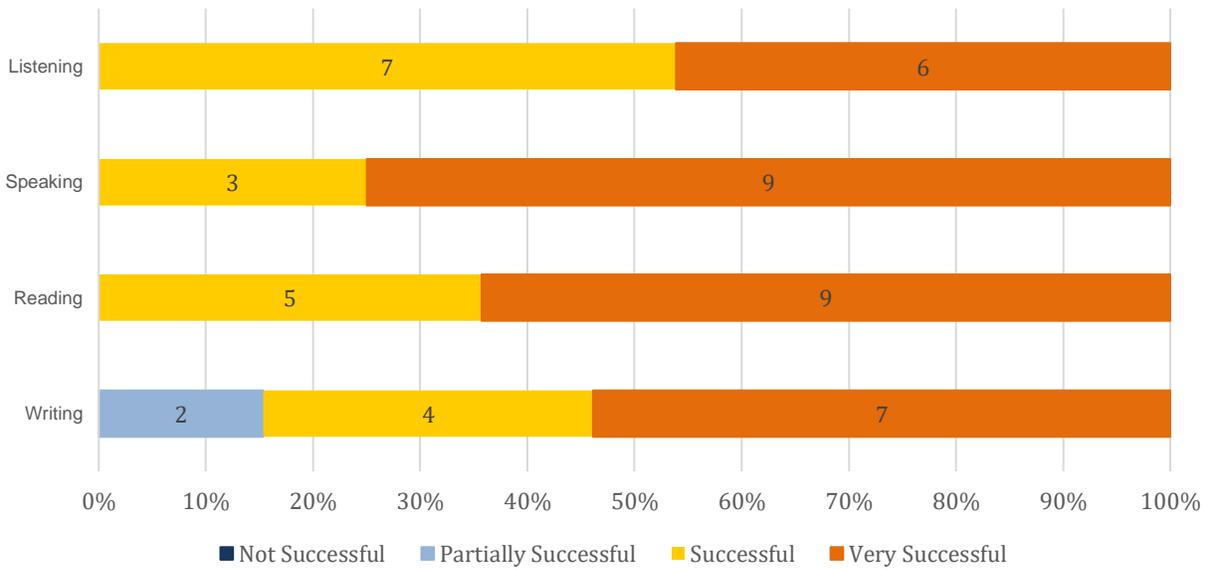
*Early Independence*



### Developing Independence



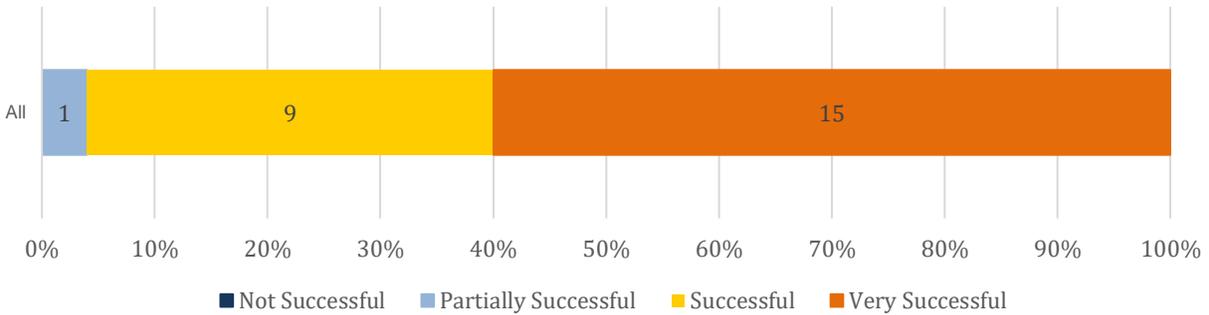
### Basic Fluency



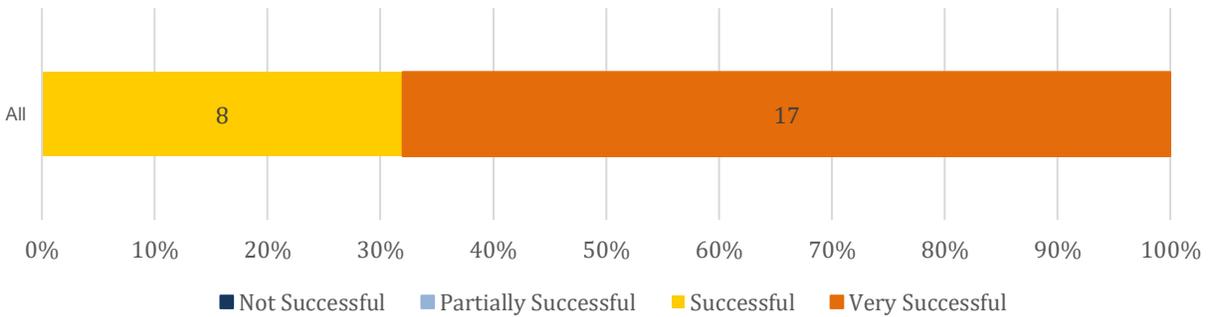
The following questions ask about overall experience at the TELPAS standard setting:

Question 6: 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.

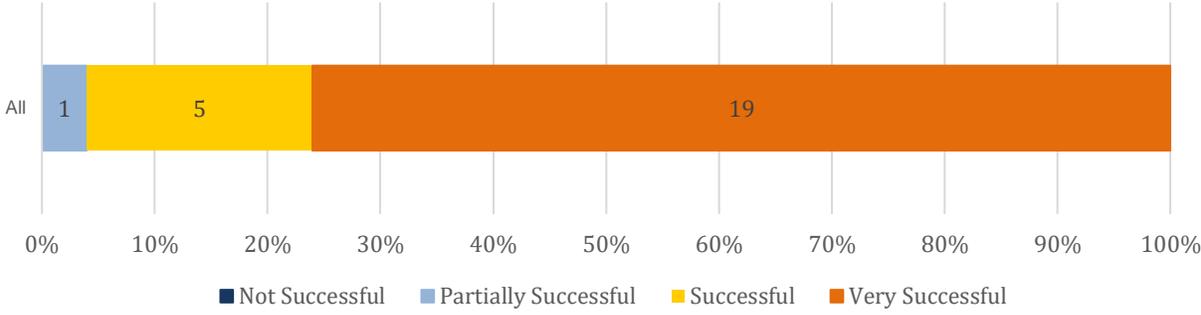
*Meeting pre-work*



*General session*

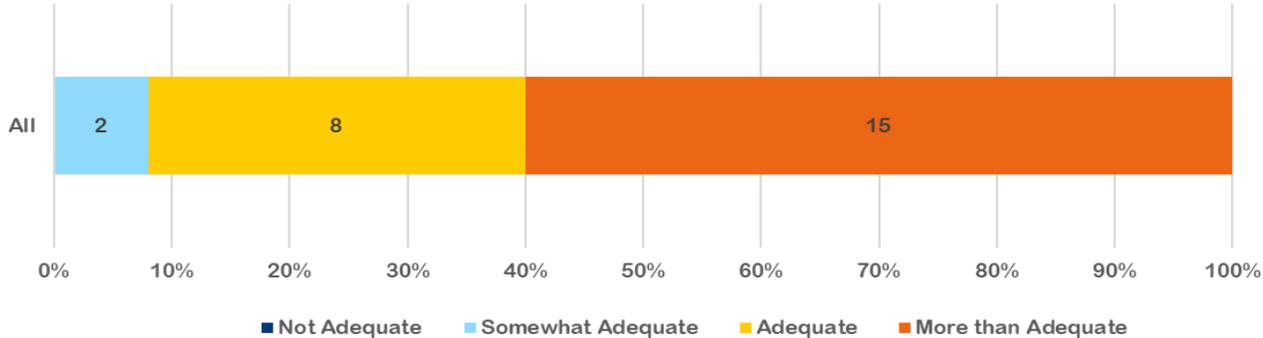


*Breakout sessions*

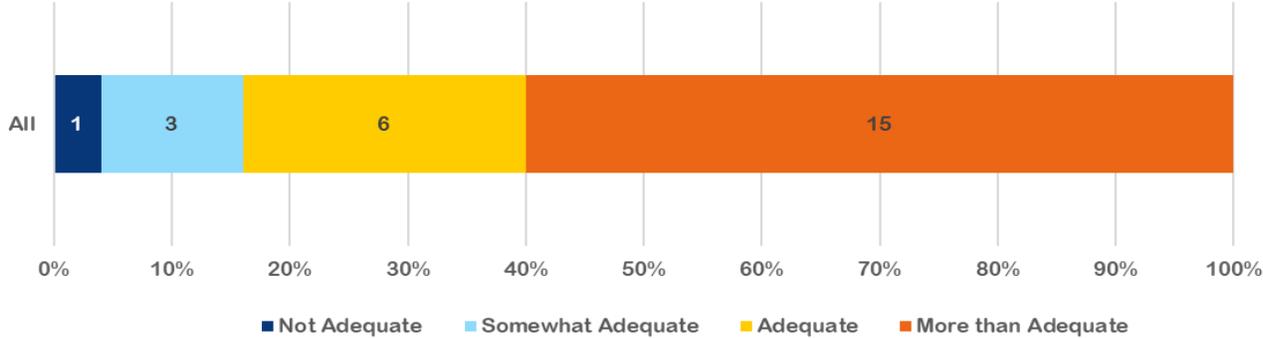


Question 7: How adequate were the following elements of the session?

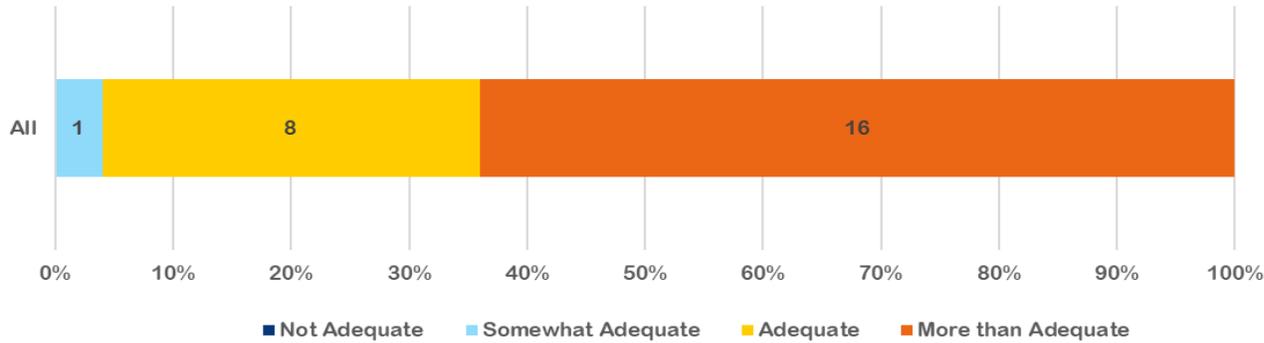
*Facilities used for the general session*



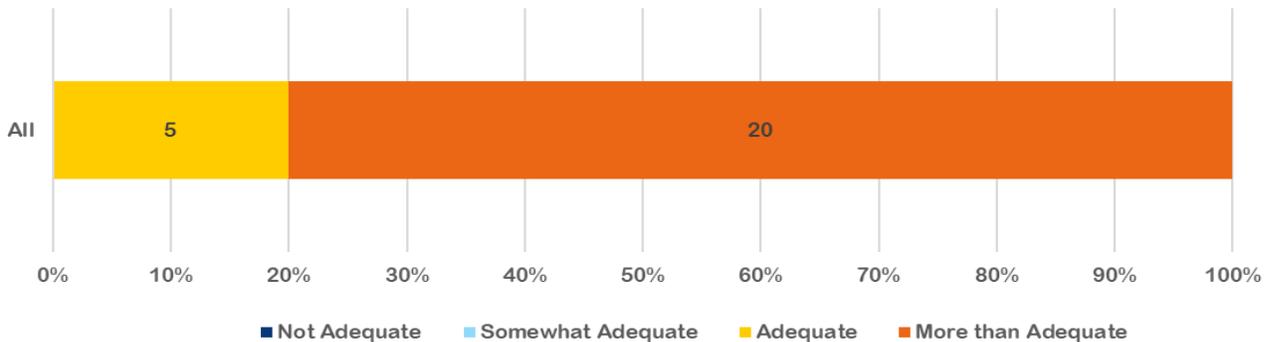
*Facilities used for the breakout session*



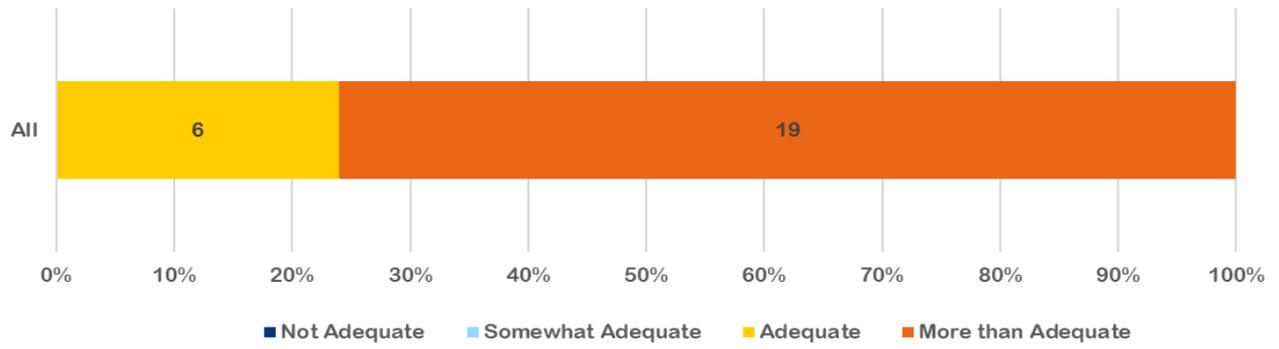
*Computers used during the meetings*



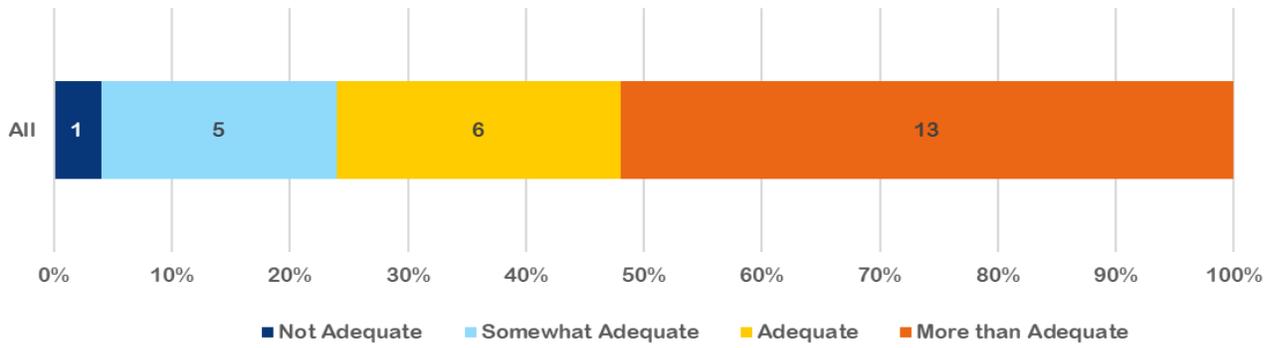
*Standard setting website for accessing materials and making judgments*



*Materials provided in the folders*

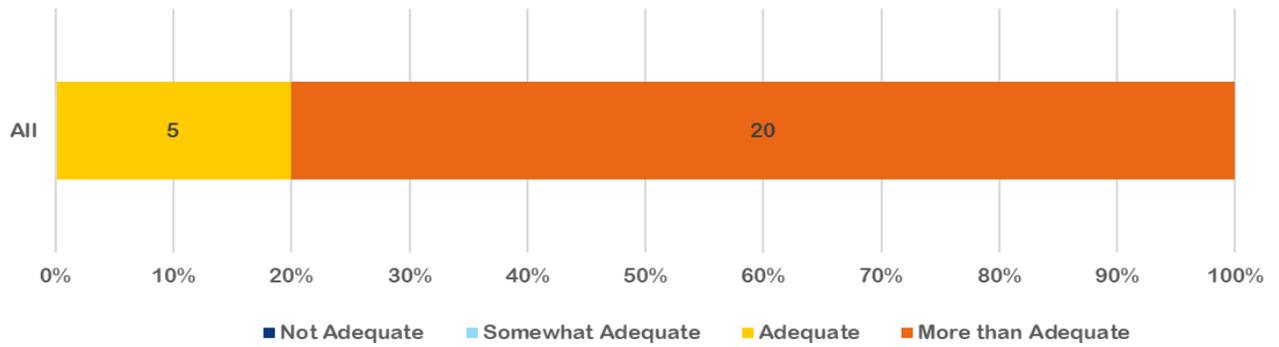


*Work space in table groups during meeting*

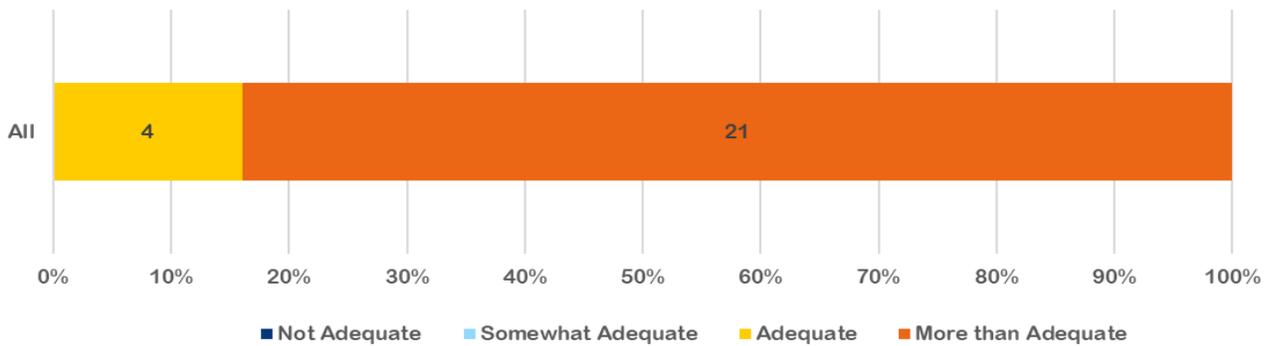


Question 8: Did you have adequate opportunities during the session to:

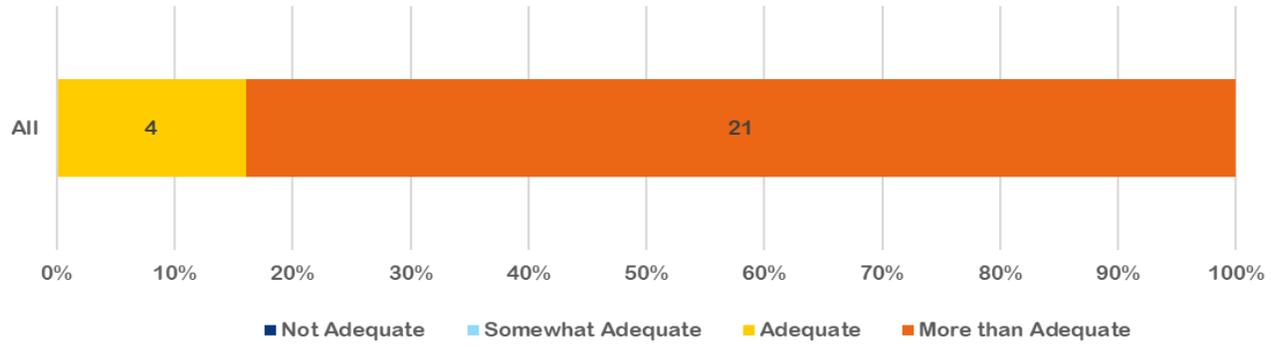
*Express your opinions about student proficiency levels*



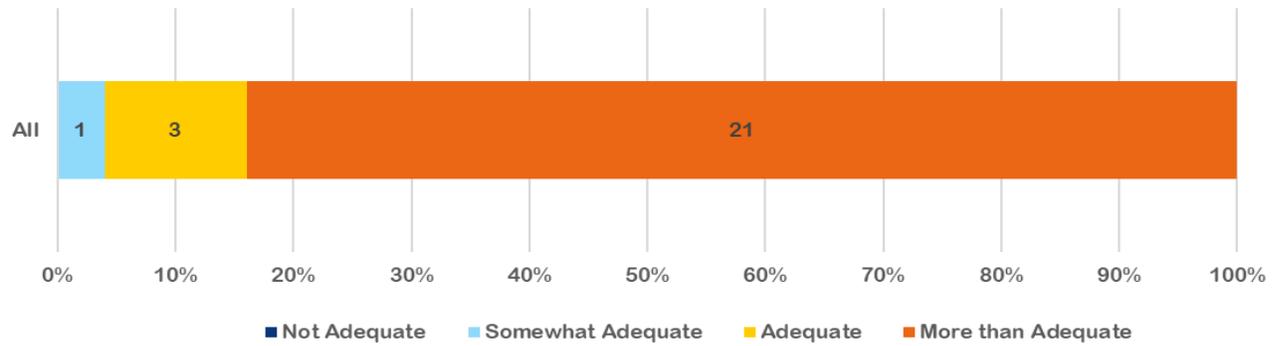
*Ask questions about the cut scores and how they will be used*



*Ask questions about the process of making cut score recommendations*

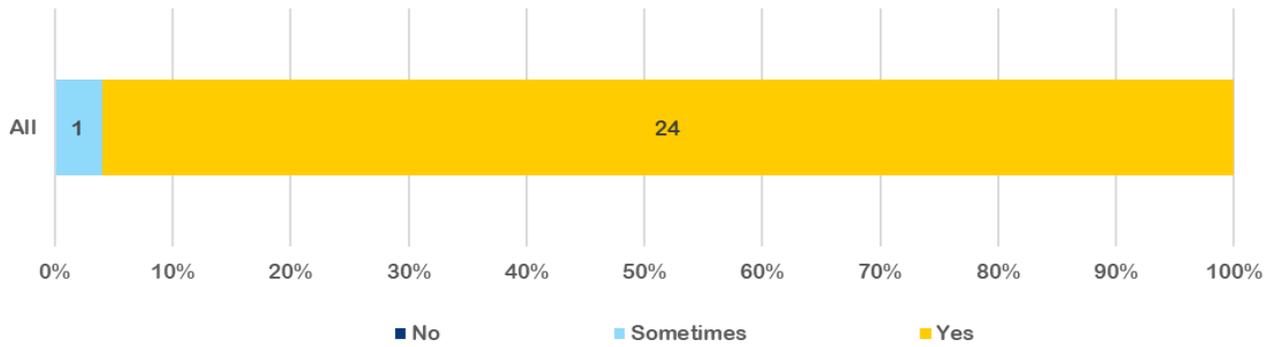


*Interact with your fellow panelists*

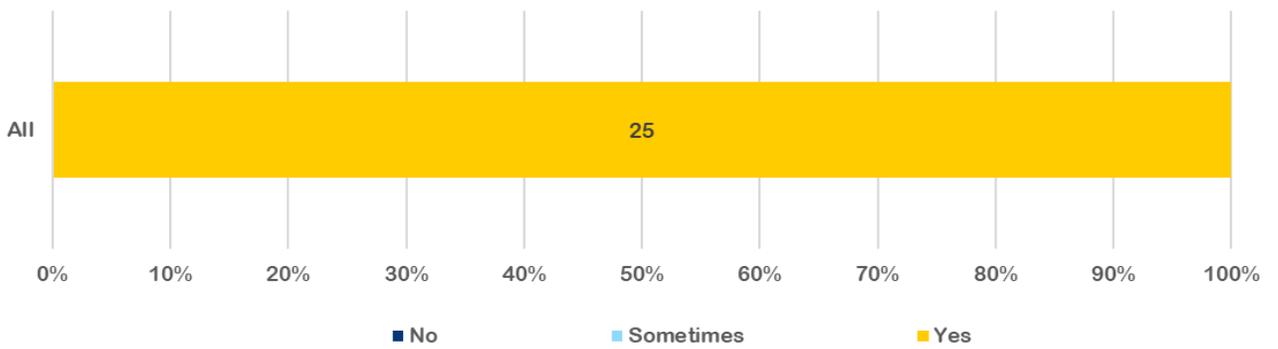


Question 9: Do you believe your opinions and judgments were treated with respect by:

*Fellow panelists*



*Facilitators*



# Appendix G — Committee Recommended Cut Scores by Round

**Table G.1. Listening**

Proficiency Level	Maximum Score	Rounds			Final
		1	2	3	
Imitation	50	17	19	17	17
Early Independence		25	26	26	26
Developing Independence		36	35	36	36
Basic Fluency		45	45	45	45

**Table G.2. Speaking**

Proficiency Level	Maximum Score	Rounds			Final
		1	2	3	
Imitation	50	17	16	16	16
Early Independence		25	26	26	26
Developing Independence		36	36	35	35
Basic Fluency		45	44	44	44

**Table G.3. Reading**

Proficiency Level	Maximum Score	Rounds			Final
		1	2	3	
Imitation	50	19	18	18	18
Early Independence		25	24	24	24
Developing Independence		33	33	33	33
Basic Fluency		41	42	42	42

**Table G.4. Writing**

Proficiency Level	Maximum Score	Rounds			Final
		1	2	3	
Imitation	50	16	16	16	16
Early Independence		24	24	24	24
Developing Independence		34	34	33	33
Basic Fluency		41	41	41	41

# Appendix H — Recommended Cut Score Summary Statistics

**Table H.1. Listening**

Round	Statistic	Imitation	Early Independence	Developing Independence	Basic Fluency
1	Mean	17.15	24.85	35.00	44.23
	Minimum	13	19	26	30
	Q1	15	23	35	44
	<b>Median</b>	<b>17</b>	<b>25</b>	<b>36</b>	<b>45</b>
	Q3	19	27	37	46
	Maximum	24	32	39	48
2	Mean	18.62	25.92	35.69	44.54
	Minimum	17	22	34	42
	Q1	17	25	35	44
	<b>Median</b>	<b>19</b>	<b>26</b>	<b>35</b>	<b>45</b>
	Q3	20	27	36	46
	Maximum	21	28	38	46
3	Mean	17.46	26.08	35.69	45.15
	Minimum	16	25	34	42
	Q1	17	25	35	45
	<b>Median</b>	<b>17</b>	<b>26</b>	<b>36</b>	<b>45</b>
	Q3	18	27	36	46
	Maximum	19	28	37	47

**Table H.2. Speaking**

Round	Statistic	Imitation	Early Independence	Developing Independence	Basic Fluency
1	Mean	15.92	24.58	35.08	43.92
	Minimum	11	20	31	38
	Q1	14.50	22	33	42
	<b>Median</b>	<b>17</b>	<b>25</b>	<b>36</b>	<b>45</b>
	Q3	17.50	27.50	36	45
	Maximum	20	28	39	48
2	Mean	16.38	25.85	35.69	44.38
	Minimum	15	24	34	42
	Q1	16	25	35	44
	<b>Median</b>	<b>16</b>	<b>26</b>	<b>36</b>	<b>44</b>
	Q3	17	26	36	45
	Maximum	18	28	37	47
3	Mean	16.23	26.15	35.08	43.77
	Minimum	15	25	34	42
	Q1	16	25	34	43
	<b>Median</b>	<b>16</b>	<b>26</b>	<b>35</b>	<b>44</b>
	Q3	17	27	36	45
	Maximum	17	28	37	45

**Table H.3. Reading**

Round	Statistic	Imitation	Early Independence	Developing Independence	Basic Fluency
1	Mean	18.67	25.07	33.07	41.40
	Minimum	15	21	27	34
	Q1	16	23	30	39
	<b>Median</b>	<b>19</b>	<b>25</b>	<b>33</b>	<b>41</b>
	Q3	20	27	35	44
	Maximum	23	30	40	46
2	Mean	17.93	24.40	33.00	41.87
	Minimum	15	21	30	39
	Q1	17	23	31	40
	<b>Median</b>	<b>18</b>	<b>24</b>	<b>33</b>	<b>42</b>
	Q3	20	26	34	43
	Maximum	20	28	36	45
3	Mean	17.07	24.47	33.13	42.07
	Minimum	14	21	30	40
	Q1	15	24	33	41
	<b>Median</b>	<b>18</b>	<b>24</b>	<b>33</b>	<b>42</b>
	Q3	18	26	34	43
	Maximum	20	27	35	45

**Table H.4. Writing**

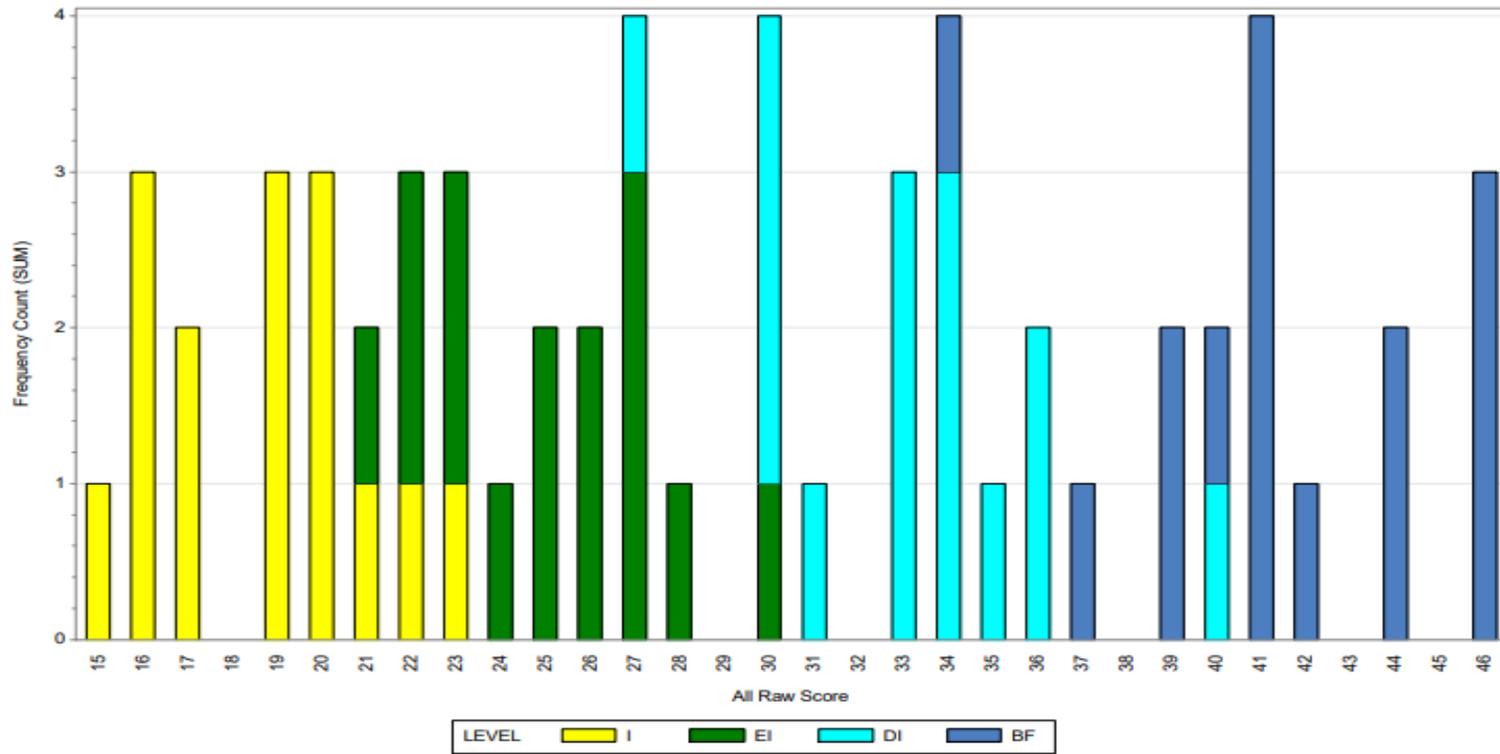
Round	Statistic	Imitation	Early Independence	Developing Independence	Basic Fluency
1	Mean	15.80	23.40	32.93	41.87
	Minimum	13	18	25	36
	Q1	15	22	30	41
	<b>Median</b>	<b>16</b>	<b>24</b>	<b>34</b>	<b>41</b>
	Q3	17	25	36	44
	Maximum	20	26	38	46
2	Mean	15.67	23.40	33.47	41.60
	Minimum	14	20	30	40
	Q1	15	23	33	41
	<b>Median</b>	<b>16</b>	<b>24</b>	<b>34</b>	<b>41</b>
	Q3	16	24	34	42
	Maximum	18	25	36	44
3	Mean	15.73	23.73	33.27	41.33
	Minimum	15	20	30	40
	Q1	15	23	33	41
	<b>Median</b>	<b>16</b>	<b>24</b>	<b>33</b>	<b>41</b>
	Q3	16	24	34	42
	Maximum	17	27	37	42

# Appendix I — Test-Level Participant Judgment Agreement

All graphs shown in Appendix I include the TELPAS Alternate cut score recommendation by panelists on the x-axis and the frequency of those recommendations on the y-axis. Abbreviations in the legend are: I = Imitation; EI = Early Independence; DI = Developing Independence; BF = Basic Fluency.

Reading

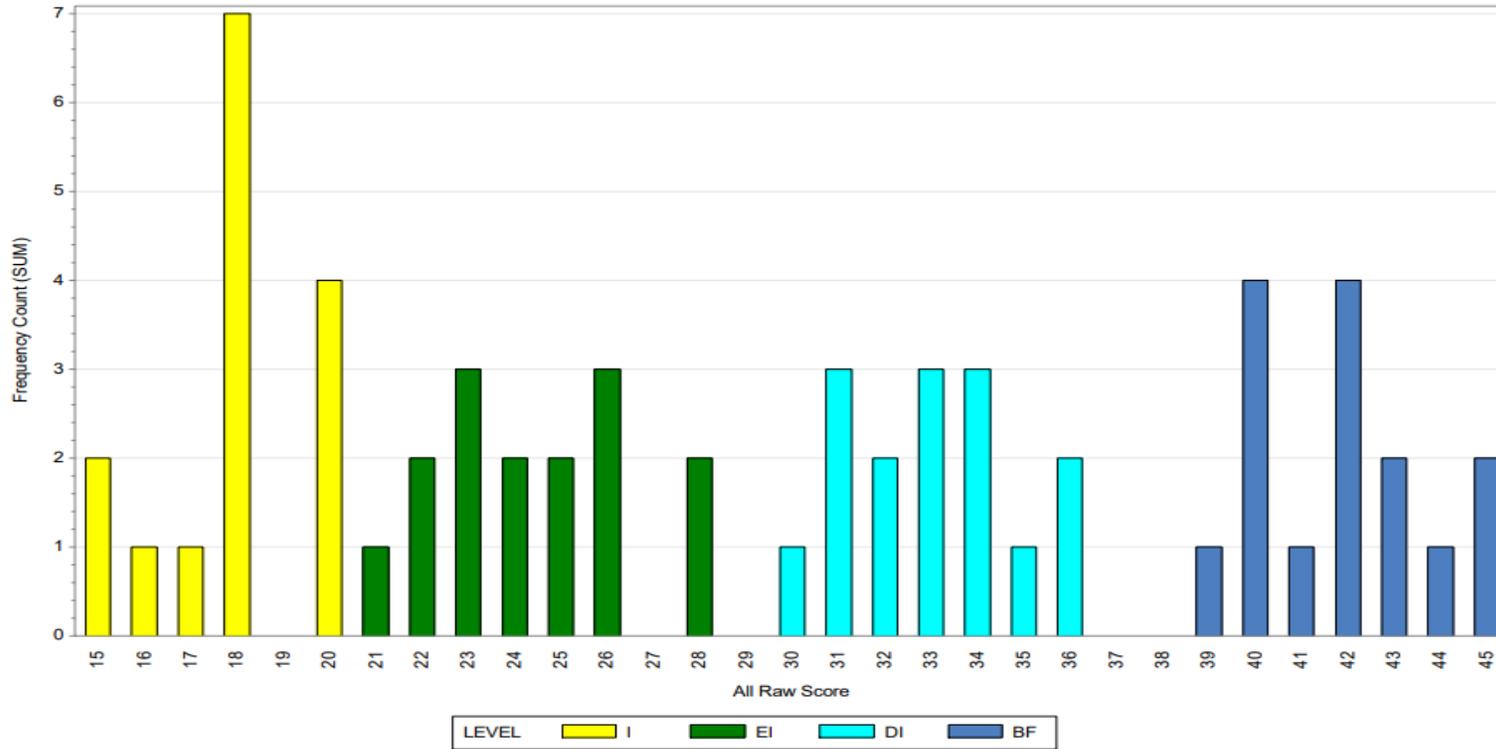
Round 1:



All Four Proficiency Levels Concurrently

Reading

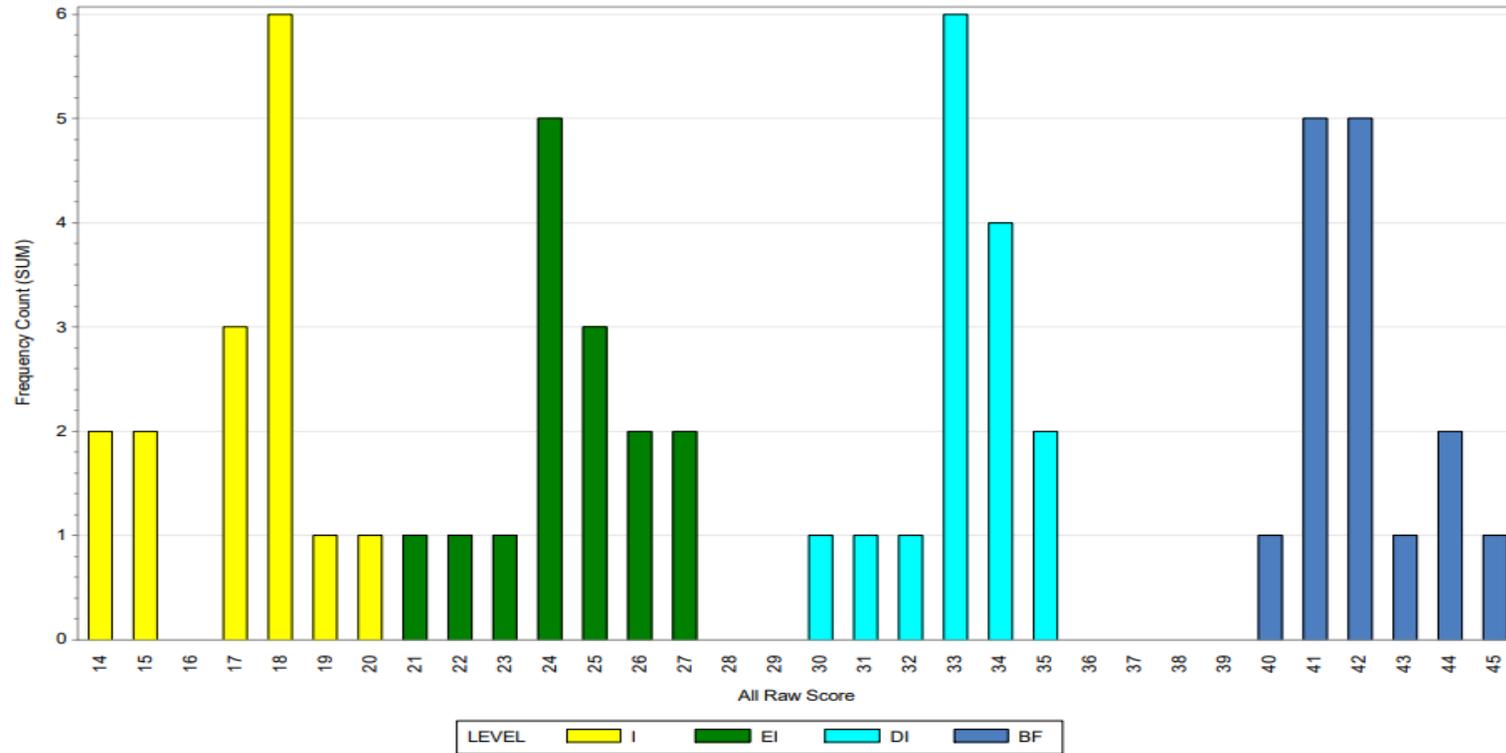
Round 2:



*All Four Proficiency Levels Concurrently*

Reading

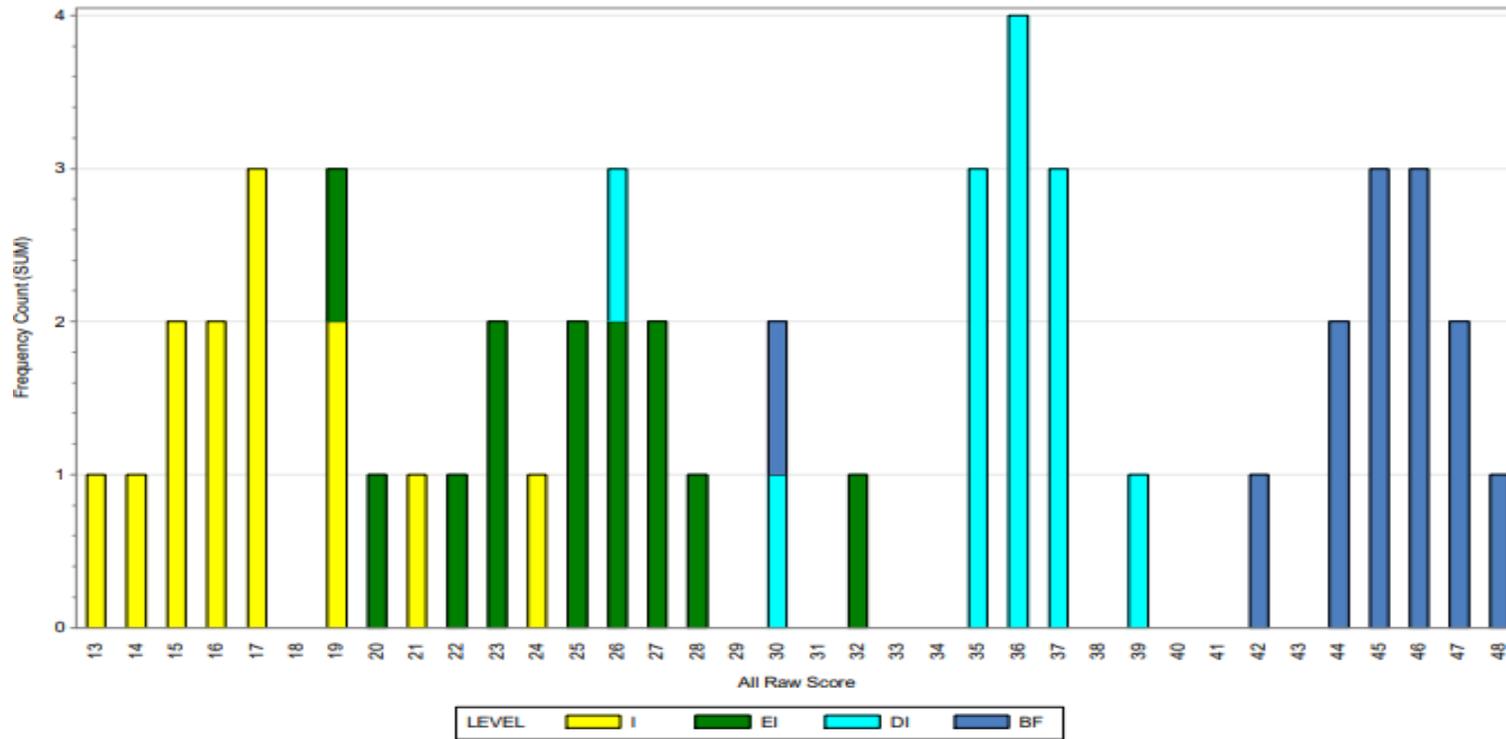
Round 3:



*All Four Proficiency Levels Concurrently*

Listening

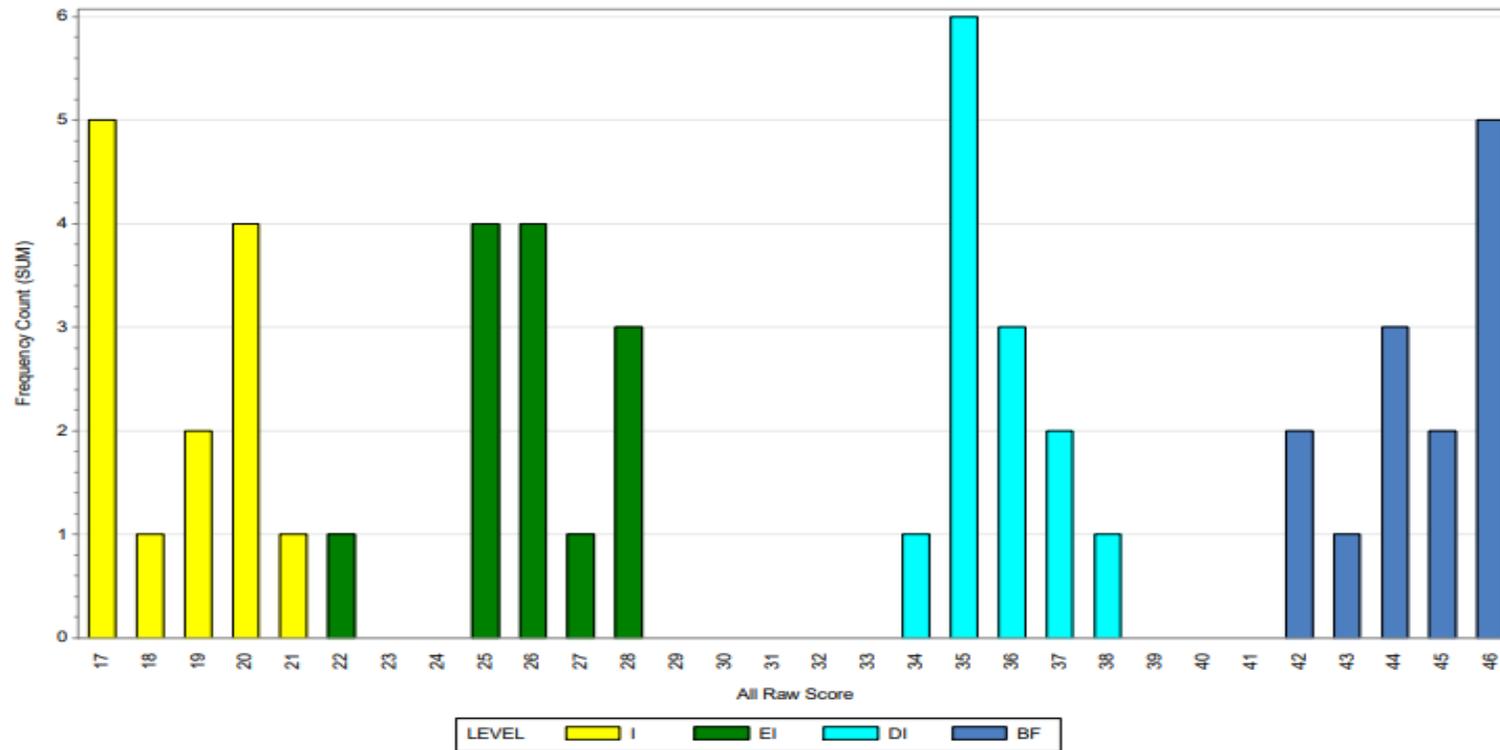
Round 1:



*All Four Proficiency Levels Concurrently*

Listening

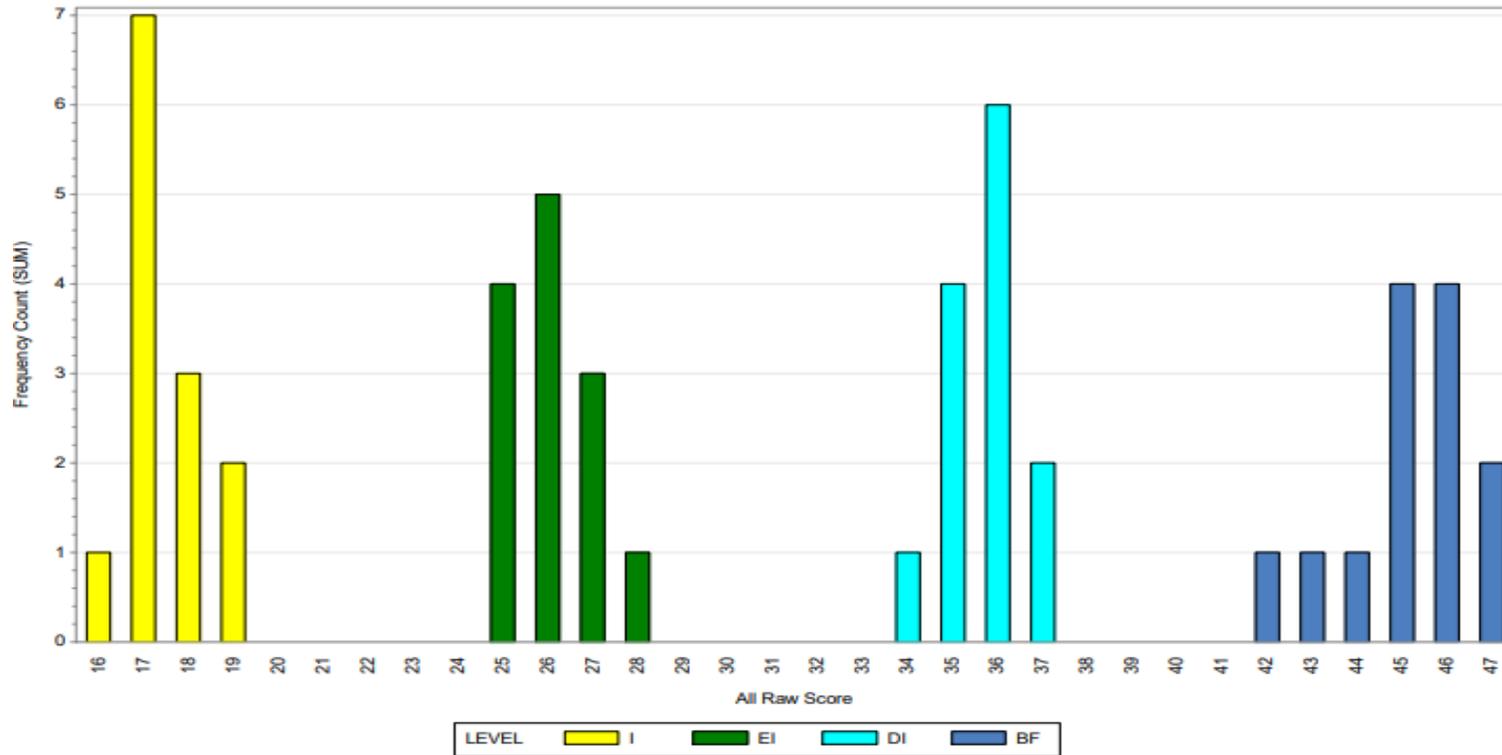
Round 2:



*All Four Proficiency Levels Concurrently*

Listening

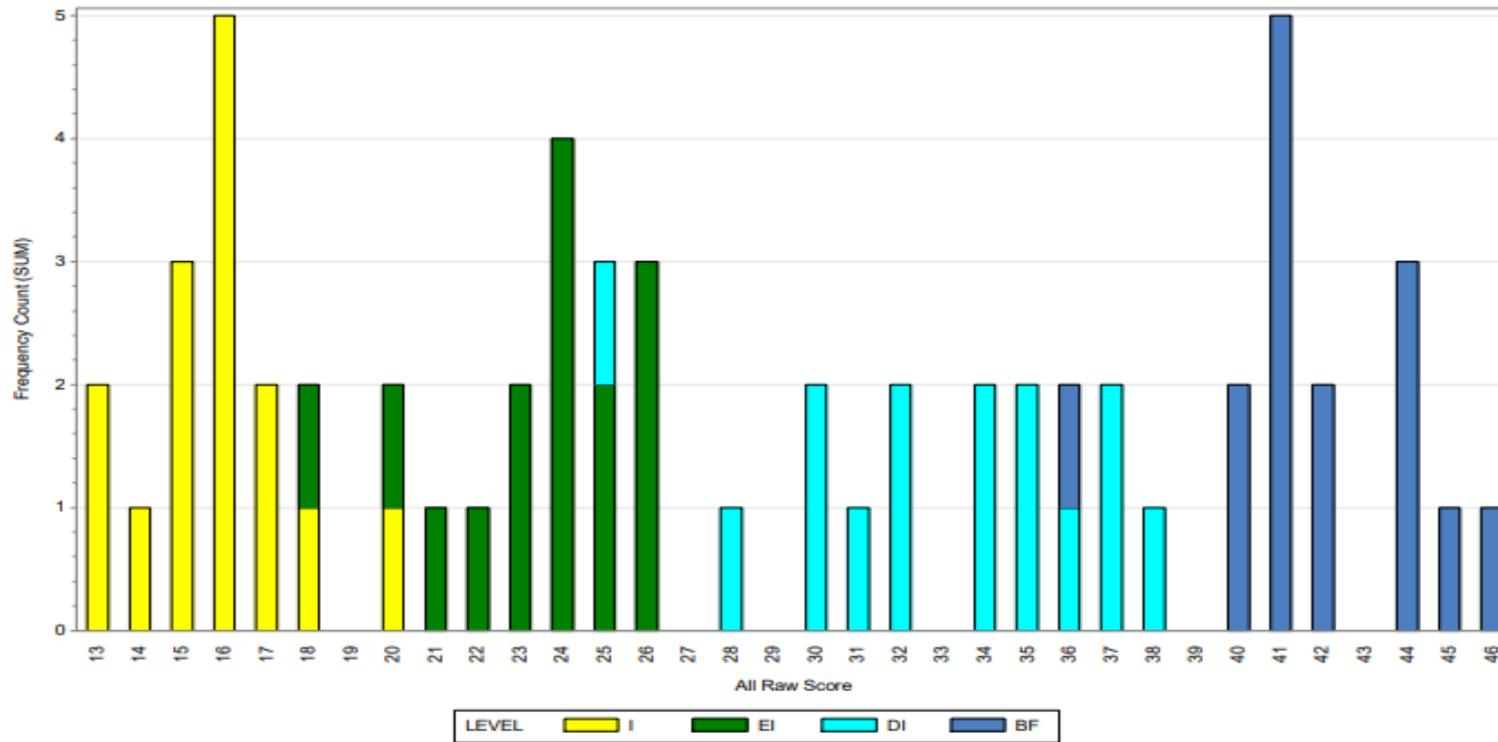
Round 3:



*All Four Proficiency Levels Concurrently*

Writing

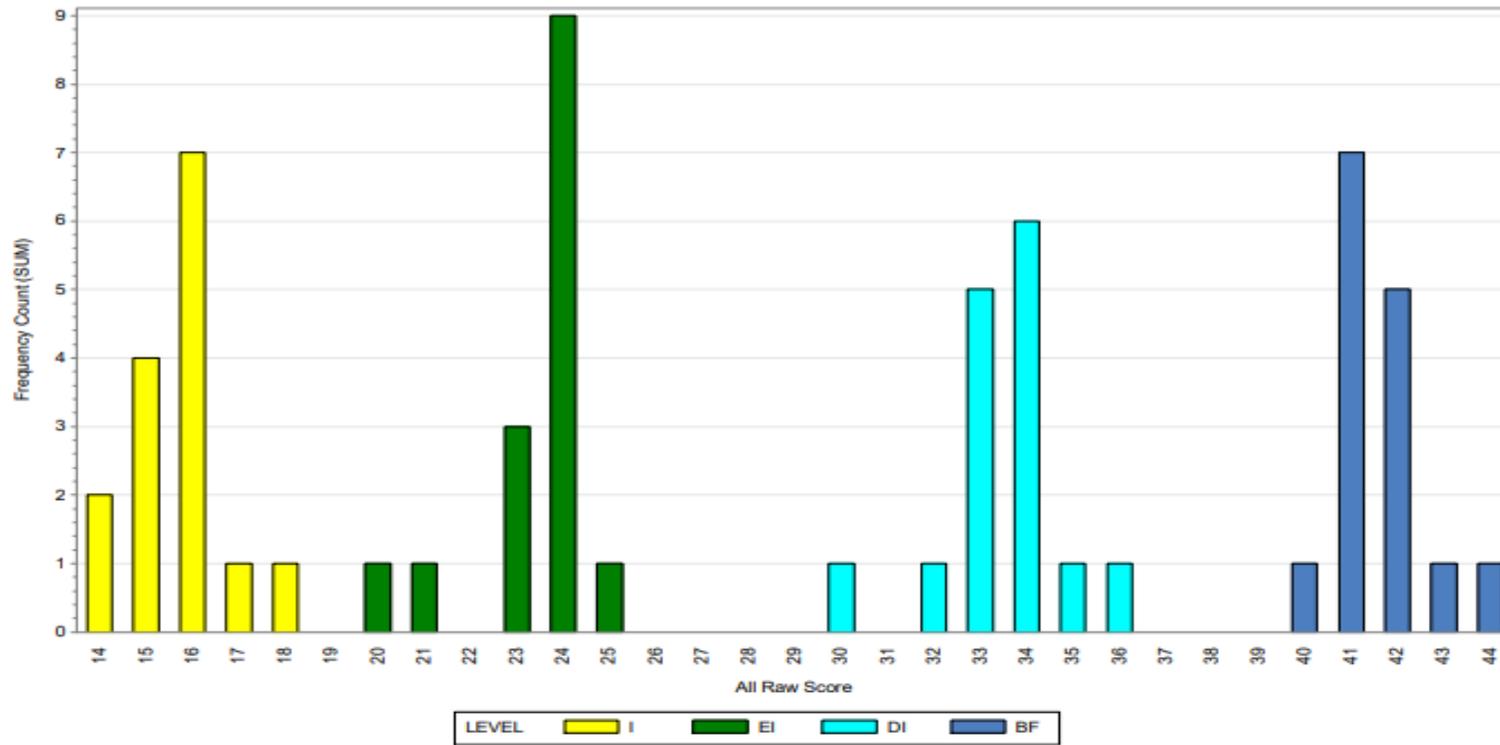
Round 1:



All Four Proficiency Levels Concurrently

Writing

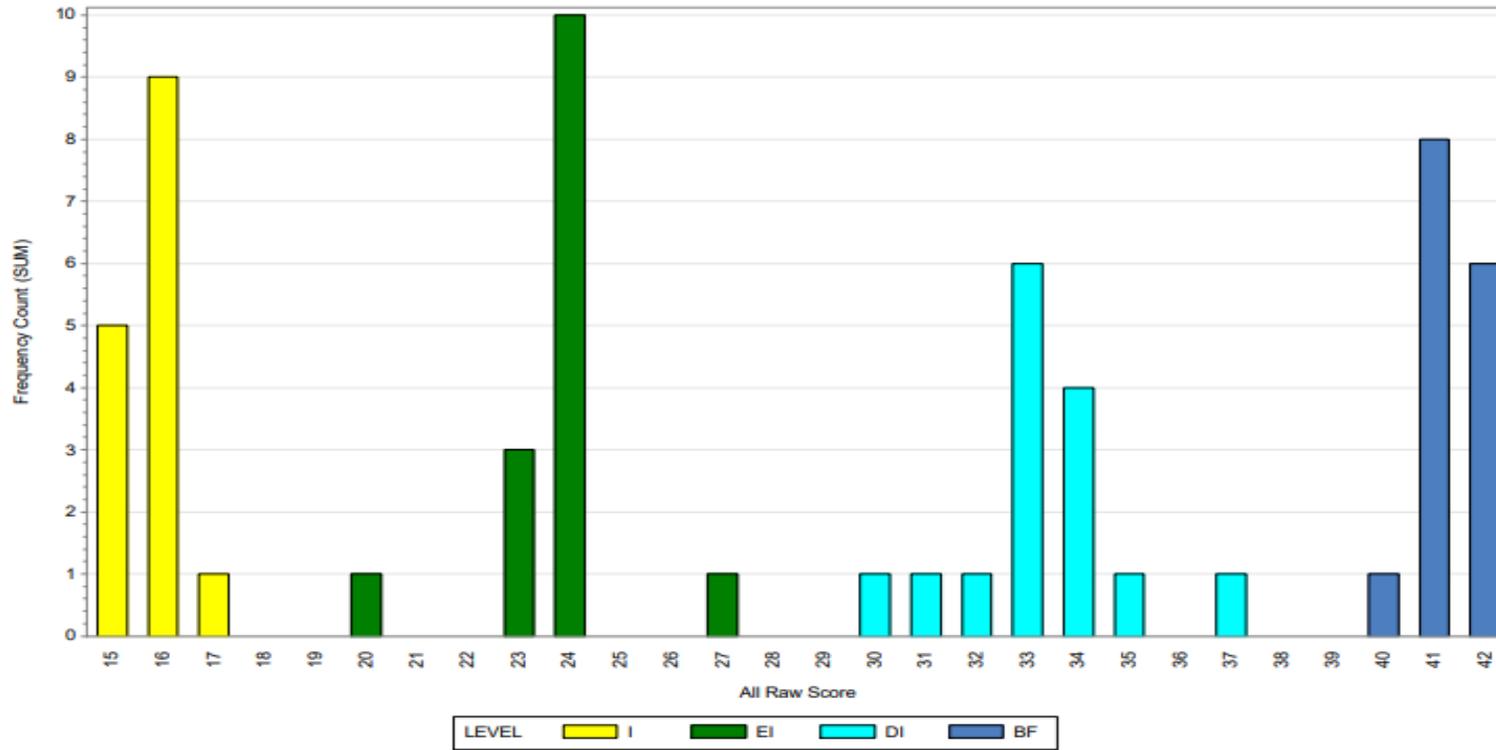
Round 2:



All Four Proficiency Levels Concurrently

Writing

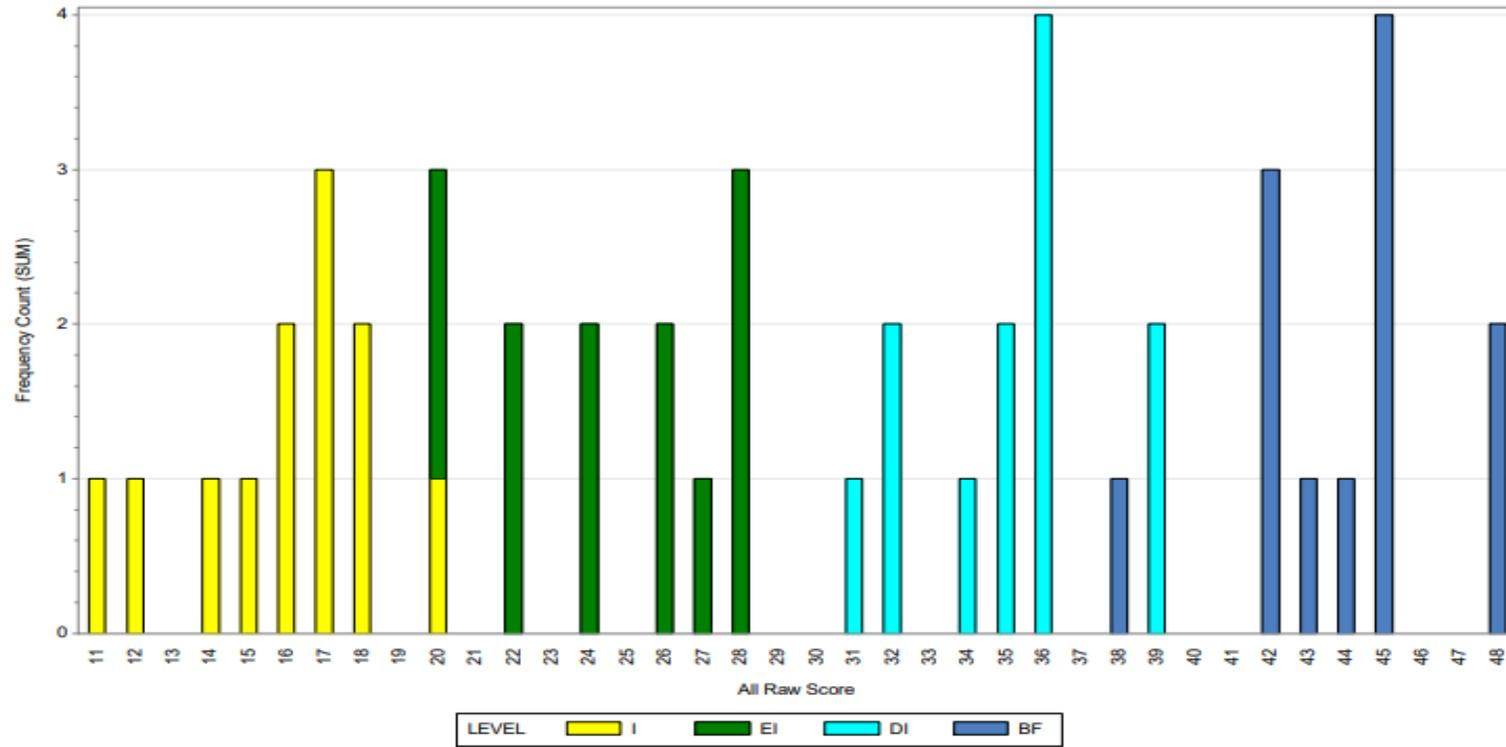
Round 3:



*All Four Proficiency Levels Concurrently*

Speaking

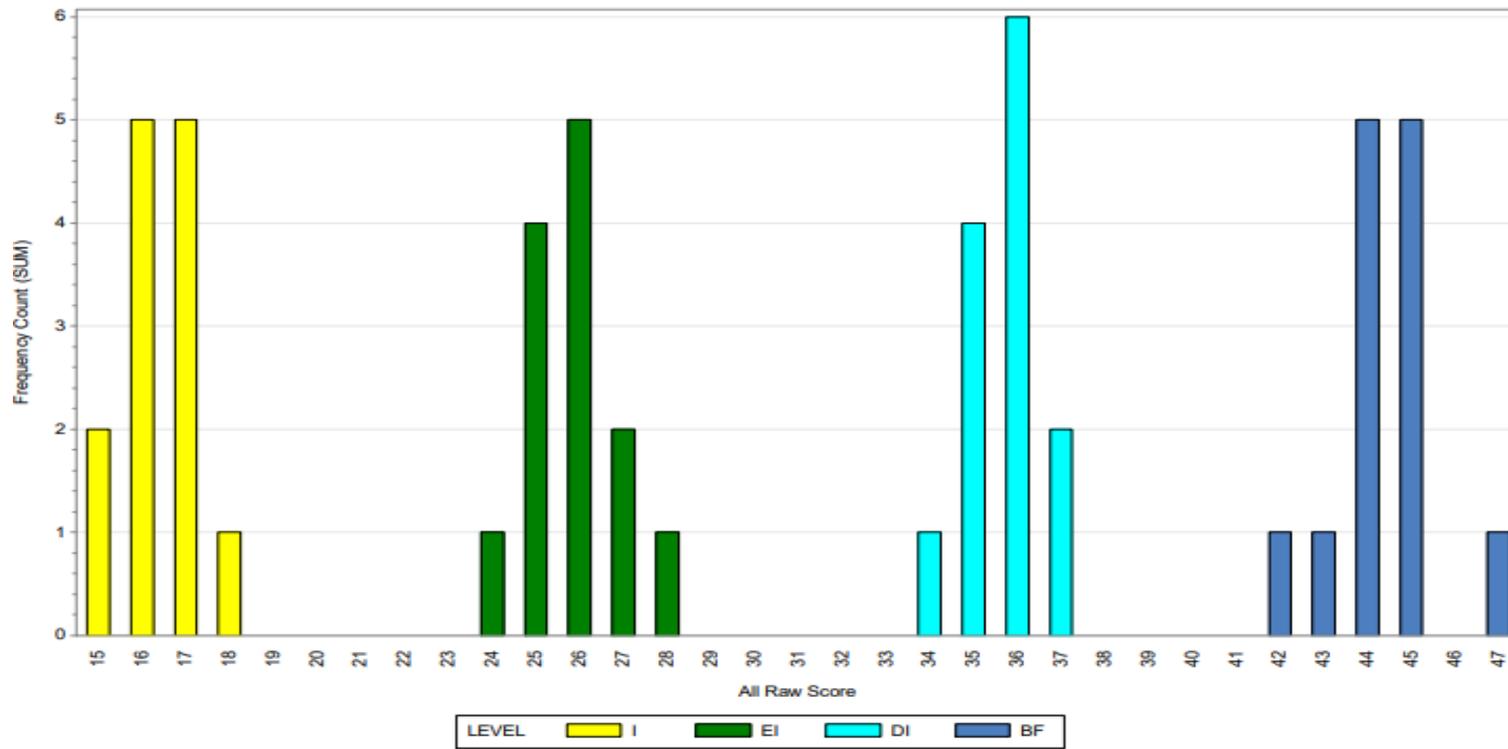
Round 1:



All Four Proficiency Levels Concurrently

Speaking

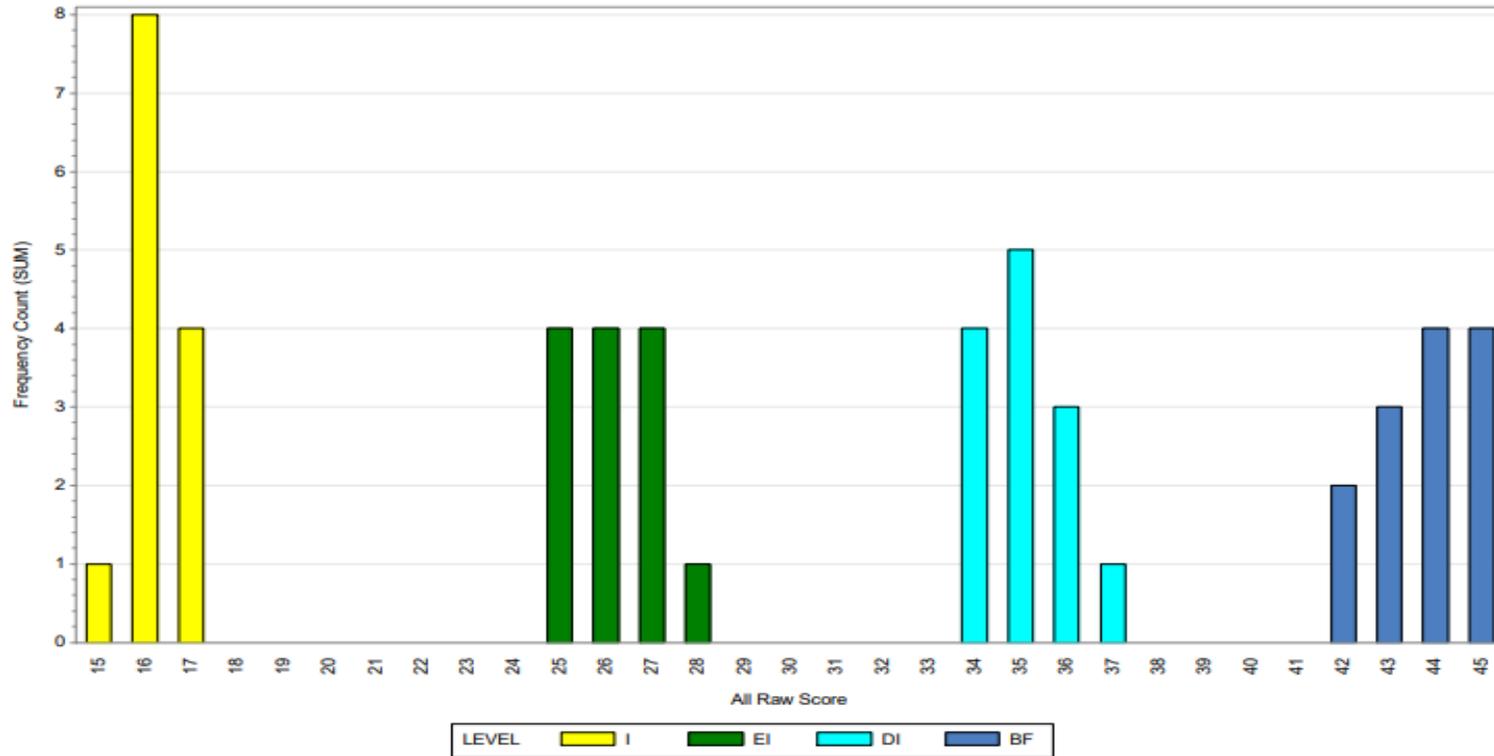
Round 2:



*All Four Proficiency Levels Concurrently*

Speaking

Round 3:



*All Four Proficiency Levels Concurrently*

# Appendix J — Composite Score Domain Profiles

TELPAS Alternate  
Standard Setting Meeting  
June 2019



## Composite Score Domain Profiles

Frequency of Domain Scores					Composite Score
1	2	3	4	5	
0	0	0	0	4	5
0	0	0	1	3	4.75
0	0	1	0	3	4.5
0	0	0	2	2	
0	1	0	0	3	4.25
0	0	1	1	2	
0	0	0	3	1	4
1	0	0	0	3	
0	1	0	1	2	
0	0	2	0	2	
0	0	1	2	1	3.75
0	0	0	4	0	
1	0	0	1	2	
0	1	1	0	2	
0	1	0	2	1	3.5
0	0	2	1	1	
0	0	1	3	0	
1	0	1	0	2	
0	2	0	0	2	3.25
1	0	0	2	1	
0	1	1	1	1	
0	1	0	3	0	
0	0	2	2	0	3
1	0	3	1	0	
1	1	0	0	2	
1	0	1	1	1	
0	2	0	1	1	2.75
0	1	2	0	1	
0	1	0	3	0	
0	0	2	2	0	
2	1	0	0	1	2.5
2	0	1	1	0	
1	2	0	1	0	
1	1	2	0	0	
0	3	1	0	0	2.25
0	2	2	0	0	
2	1	0	0	1	
2	0	1	1	0	
1	2	0	1	0	2
1	1	2	0	0	
0	3	1	0	0	
3	0	0	0	1	
2	1	0	1	0	1.75
2	0	2	0	0	
1	2	1	0	0	
0	4	0	0	0	
3	0	0	1	0	1.5
2	1	1	0	0	
1	3	0	0	0	1.25
3	0	1	0	0	
2	2	0	0	0	1
3	1	0	0	0	
4	0	0	0	0	

Frequency of Domain Scores					Composite Score
1	2	3	4	5	
2	0	0	0	2	3
1	1	0	1	1	
1	0	2	0	1	
0	2	1	0	1	
1	0	1	2	0	2.75
0	2	0	2	0	
0	1	2	1	0	
0	0	4	0	0	
2	0	0	1	1	2.5
1	1	1	0	1	
0	3	0	0	1	
1	0	1	2	0	
1	0	2	1	0	2.25
0	2	1	1	0	
0	1	3	0	0	
2	0	1	0	1	
1	2	0	0	1	2
2	0	0	2	0	
1	1	1	1	0	
0	3	0	1	0	
1	0	3	0	0	1.75
0	2	2	0	0	
2	1	0	0	1	
2	0	1	1	0	
1	2	0	1	0	1.5
1	1	2	0	0	
0	3	1	0	0	1.25
3	0	0	0	1	
2	1	0	1	0	1
2	0	2	0	0	
1	2	1	0	0	
0	4	0	0	0	
3	0	0	1	0	
2	1	1	0	0	
1	3	0	0	0	
3	0	1	0	0	
2	2	0	0	0	
3	1	0	0	0	
4	0	0	0	0	