



Arizona
Arizona English Language Learner Assessment
(AZELLA)

Evidence-Based Standard Setting Review Report:
Kindergarten Placement Test

Arizona Department of Education
August 2014

Table of Contents

Executive Summary	1
General Evidence-Based Standard Setting Review Procedures	3
Panel	3
Performance Level Descriptors	4
Methodology Overview.....	4
Data	5
Security.....	6
Staff	6
Detailed Evidence-Based Standard Setting Review Procedures	7
Opening Session	7
Introduction	7
Experience the Test	7
Original Standard Setting	7
Review of the Ordered Item Booklet	8
Test Validity	8
Presentation of Evidential Data.....	8
Standard Setting Methodology.....	10
Round 1 - Proficient	10
Round 2 - Proficient	11
Round 1 and Round 2 – Basic/Intermediate.....	11
Standard Setting Evaluation Form	12
Standard Setting Results	13
References.....	16
Appendix A Evidence-Based Standard Setting Review Participants	18
Appendix B Presentation Slides.....	19
Appendix C Performance Level Descriptors	45
Appendix D Evidential Data Graphs	48
Appendix E Page Number Recording Sheet.....	52
Appendix F Standard Setting Evaluation Forms with Responses	55

List of Tables

Table 1. Final Recommended Raw Score Cuts	2
Table 2. Final Scale Score Ranges	2
Table 3. Proficiency Level Distribution for All PHLOTE Students	2
Table 4. Median Raw Score Cuts by Round	13
Table 5. Minimum, Median, and Maximum Raw Score Cuts by Round	15
Table 6. Scale Score Cuts by Round	15
Table 7. Impact Data for PHLOTE Students Based on Median Cuts by Round.....	15

List of Figures

Figure 1. Raw score cut distributions of cuts at Round 1	13
Figure 2. Raw score cut distributions of cuts at Round 2	14
Figure D1. Proficient cut 2013 scatterplot.....	49
Figure D2. Proficient cut 2013 bar graph	49
Figure D3. Proficient cut 2014 scatterplot.....	50
Figure D4. Proficient cut 2014 bar graph	50
Figure D5. Basic cut 2013 scatterplot.....	51
Figure D6. Basic cut 2013 bar graph	51
Figure D7. Basic cut 2014 scatterplot.....	52
Figure D8. Basic cut 2014 bar graph	52

Executive Summary

This document provides information about the procedures that were implemented to review and revise the established performance standards for the Arizona English Language Learner Assessment (AZELLA) Kindergarten Placement Test (Placement Test), developed by the Arizona Department of Education (ADE). The evidence-based standard setting review took place on May 30, 2014 at the Arizona Department of Education offices in Phoenix, Arizona. Standard setting participants were provided with data that compared students' AZELLA Kindergarten Placement Test results to their spring AZELLA Reassessment results.

The Placement Test was developed to fairly assess incoming kindergarten students who reported a primary or home language other than English upon registration in their school. The test was designed to take approximately 20 minutes and focuses primarily on the Listening and Speaking sections of the English Language Proficiency Standards (ELPS). The Placement Test classifies the students into the following English proficiency levels:

- 1) Pre-emergent/Emergent
- 2) Basic/Intermediate
- 3) Proficient

Students who score *Proficient* on the test are determined to have sufficient English proficiency to do regular classwork in English. These students will be placed in mainstream classrooms. Students who score less than *Proficient* on the test are determined to not have sufficient English proficiency to do regular classwork in English. These students will be provided English language acquisition services.

The following activities were used with the panelists for the evidence-based standard setting review meeting. These activities were designed to address Standards 5.21 and 5.22 as put forth in *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014). Please refer to the *Detailed Evidence-Based Standard Setting Review Procedures* section for the details of each activity.

1. Opening session
2. Experience the test
3. Review Performance Level Descriptors
4. Review of the content of the test items
5. Review of Fiscal 2013 and Fiscal 2014 assessment data
6. Evidence-based standard setting review methodology training
7. Round 1 ratings
8. Round 1 feedback and discussion
9. Round 2 ratings
10. Round 2 feedback and discussion
11. Complete evidence-based standard setting review evaluation

To align with the standard setting procedure used for the rest of AZELLA assessment system in spring of 2013, the evidence-based standard setting review for the Placement Test utilized the Bookmark method (Mitzel et al., 2001; Cizek & Bunch, 2007). The Bookmark method was augmented in the review by the use of both past and current data. The final results after two rounds of ratings are presented below. The final recommended raw score cuts and percentage of points required for the *Basic/Intermediate* and *Proficient* levels are summarized in Table 1. To achieve the *Basic/Intermediate cut*, students must score at least 21 points out of 42 maximum possible points, which is 50% of the points. For the *Proficient* cut, students must score 35 points or more (approximately 83% of the points).

Table 1.

Final Recommended Raw Score Cuts

Basic/Intermediate		Proficient	
Raw Score	% of Points	Raw Score	% of Points
21	50%	35	83%

The final scale score ranges for each proficiency level are presented in Table 2. The scale score for the Placement Test ranges from 100 to 300.

Table 2.

Final Scale Score Ranges

Pre-emergent/Emergent	Basic/Intermediate	Proficient
100-205	206-244	245-300

The expected percentage of students at each proficiency level using the recommended cut scores is summarized in Table 3. The percentages in the table are based on the scores of students with a Primary Home Language Other Than English (PHLOTE) who took their first Placement Test between July 15 and December 21, 2013. Based on the changed cut score ranges, approximately 42% of PHLOTE students would be *Proficient*. The percentage of PHLOTE students at *Pre-emergent/Emergent* and *Basic/Intermediate* would be 24% and 34%, respectively.

Table 3.

Proficiency Level Distribution for All PHLOTE Students

Pre-emergent/Emergent	Basic/Intermediate	Proficient
24%	34 %	42%

General Evidence-Based Standard Setting Review Procedures

Panel

The ADE sought out Arizona educators who had never worked with the ADE on the English Language Proficiency Standards or on the AZELLA so that they would not have a previously established bias. The ADE invited panelists with credentials that demonstrated their subject matter expertise, skills as an educator, and awareness of the consequences of the panel's task. Since the goal of standard setting is to "enable participants to bring to bear their judgments in such a way as to *translate* policy decisions ... into locations on a score scale" (Cizek, Bunch, Koons, 2004, p. 32), the empanelling of a set of highly qualified participants with diverse policy perspectives was critical.

The panel for the meeting consisted of 8 participants. The panelists were chosen not only based on the level of awards and certifications they had earned (e.g., National Board Certified, Rodel Exemplary Teacher, Early Childhood endorsement, Spanish Bilingual endorsement, Reading endorsement), but also on their personal experience teaching in both English language learners (ELL) classrooms and mainstream classrooms. ADE expressly empanelled kindergarten and Grade 1 teachers as well as a school and a district administrator. All panelists were currently working in a school/district with a high percentage of ELL students. Additionally, in an attempt to allow the panel to bring a fresh perspective to the review of the standards, none of the panelists had served on an ADE assessment committee previously.

A summary of the panelists' backgrounds are provided below with a more extensive description included in Appendix A:

- An Assistant Superintendent from Arizona's largest school district. She had both Elementary and Administrative Certificates with a Bilingual Education Endorsement, held various positions in multiple inner-city schools in multiple districts, and had taught Grade 1 students.
- A Principal from a large unified school district. She was a Rodel Exemplary Principal in 2014, had an Administrative certificate and Elementary certificate with endorsements in both Spanish Bilingual and Structured English Immersion instruction, had taught bilingual education in a K-2 classroom, and worked as an instructional specialist prior to moving to the role of principal.
- Two Grade 1 teachers:
 - From the Arizona district with the largest percentage of English language learners, the first Grade 1 teacher is National Board Certified with 16 years of experience and has certificates in English as a New Language and Early and Middle Childhood as well as Elementary with Early Childhood and Structured English Immersion Endorsements.

- From a school district with approximately 9% ELL and 75% low SES students, the second Grade 1 teacher is a Rodel Exemplary Teacher from 2007. She has taught for 18 years and has an Elementary Certificate with Endorsements in English as a Second Language, Structured English Immersion, and Early Childhood.
- Four Kindergarten teachers:
 - From a school district with approximately 17% ELL and 84% low SES students, the first Kindergarten teacher has an Elementary Certificate with Early Childhood, Structured English Immersion, and Reading Endorsements. She also completed ADE's Kindergarten Academy.
 - The second Kindergarten teacher was a Rodel Exemplary Teacher in 2013. She has an Elementary certificate with both Early Childhood and Structured English Immersion Endorsements.
 - The third Kindergarten teacher is from Arizona's second largest school district. She also was a Rodel Exemplary Teacher in 2013. In addition to her certificate in Elementary Education, she has Endorsements in Early Childhood, Structured English Immersion, and Reading.
 - The fourth Kindergarten teacher is from a school district with approximately 11% ELL and 73% low SES students. She is currently serving as a Reading Specialist. In addition to her certificate in Elementary Education, she has Endorsements in English as a Second Language, and Early Childhood. She also has completed ADE's Kindergarten Academy.

Performance Level Descriptors

The Performance Level Descriptors (PLDs) for the Placement Test were preliminarily developed prior to the original standard setting meeting in July 2012 and were refined during the final phase of that meeting (Appendix C). There are three English language proficiency levels for this document:

- 1) *Pre-emergent/Emergent*
- 2) *Basic/Intermediate*
- 3) *Proficient*

The PLD's top half, which is also on the student score report, provides a concise description about the skills a student at each proficiency level should have. The bottom half provides a bulleted list of Performance Indicators (PIs) by domain for each proficiency level.

Methodology Overview

To align the evidence-based standard setting review for the Placement Test procedure to that used for the rest of AZELLA assessment system, the Bookmark method (Mitzel et al., 2001; Cizek & Bunch, 2007) was utilized. The Bookmark method (Mitzel et al., 2001; Cizek & Bunch, 2007), is a widely-used and well-known standard setting method.

It was augmented with evidential data to review the performance standards established in 2012 for the Placement Test. As recommended by ADE's Technical Advisory Committee (TAC) and by Standards 5.21 and 5.22 (AERA, APA, NCME, 2014), the panelists were asked to make decisions about the appropriate placement of raw cut scores. The panelists made decisions about both the Proficient and Basic/Intermediate raw cut scores using both the content of the items on the test and the results from two years of student data. Based on their knowledge of the English skills needed by students in mainstream classrooms, the panelists were asked to consider the content of the test items presented in an ordered item booklet and the results of kindergarten students' scores on both the Placement Test and the Spring Reassessment test when making decisions about the cut scores.

The panelists engaged in two rounds of ratings for each proficiency level. The median raw score cut for the panel after the second round was used as the committee's recommended cut for the proficiency level. During both Rounds 1 and 2, the panelists selected the page number of the item corresponding to the lowest acceptable cut score for that proficiency. The recommended raw score cuts for the *Proficient* level and the *Basic/Intermediate* level were then mapped on to the reporting scale developed during the original standard setting using the raw score to theta table. This mapping was used to determine the recommended cuts on the reporting scale. Please refer to the *Detailed Evidence-Based Standard Setting Review Procedures* for a detailed description of the process used.

Data

Data collected from the operational Placement Test administration during the fall of school year 2013-2014 was used as a basis for analyses. Data was restricted to the first placement test for students given during the window from July 15, 2013 to December 21, 2013 and was submitted by April 25, 2014. It was further restricted to students who had a valid score on the Spring 2014 AZELLA Reassessment Test. The data included 22,393 first tests for unique students. The restrictions resulted in the elimination of 1,252 tests from the data used. Overall, the results of over 94.7% of the Placement Tests for kindergarten submitted during all of fiscal year 2014 were used.

The Placement Test consists of 38 items with 42 points possible. The Rasch model (Rasch, 1960) was used to scale the test for one-point items, including one-point oral response items and one-point physical response items, while the Partial Credit model (Masters, 1982) was used for multiple-point items. These included two physical response items that were worth two points and a three-point oral response item. The raw score to theta conversion table developed during the original standard setting for the test was maintained during the evidence-based standard setting review. The raw score frequency distributions for students who took their first placement test during the fall semester of school year 2013-2014, as described above, were used to determine the percentage of students expected to fall into each proficiency level of the Placement Test if the cut score were changed. The evidentiary data was presented to the panelists prior to Round 1 of the

evidence-based standard setting review meeting. Please refer to *Detailed Evidence-Based Standard Setting Review Procedures* for more details about the impact data.

Security

As the panelists arrived, ADE staff registered them and had them sign a security agreement form. Upon registration, each panelist received a unique identification number which was used on their Page Number Recording Sheet. All materials received throughout the evidence-based standard setting review meeting remained within sight of one of the two facilitators at all times during the meeting. All secure materials were collected at the end of the day from each panelist to make sure that they were returned to ADE staff prior to the dismissal of the meeting.

Staff

The following content and psychometric staff supported the Placement Test evidence-based standard setting review:

Ms. Irene Hunting received her M.Ed. in Secondary Education Curriculum and Instruction from Keene State College (Keene, New Hampshire) and her B.S. in Mathematics from the University of Arizona. Ms. Hunting has worked for ADE for 9 years. She has previously served as the Coordinator of State Testing and the Director of State Test Administration. In her current role as Deputy Associate Superintendent of Assessment, she oversees the development, administration, and reporting of statewide assessments including AZELLA. She has 17 years of experience teaching mathematics in high school and community college. Ms. Hunting developed the training for the evidence-based standard setting review meeting and co-facilitated the meeting with Ms. Scott.

Ms. Lietta Scott received her M.A. in Educational Psychology: Measurement, Statistics, and Methodological Studies with specializations in Measurement and Evaluation from Arizona State University and is completing a Ph.D. in that program. Ms. Scott has worked for ADE for 8 years. She served the first four years in the Research and Evaluation Section and the last four years in the Assessment Section. Her functions in the Assessment Section range from content specialist to psychometric analysis of assessments. Prior to moving to ADE, she taught secondary mathematics for 12 years, mainly in inner city schools with a high proportion of ELL students. Ms. Scott oversaw the data analysis in support of the evidence-based standard setting review meeting and co-facilitated the meeting with Ms. Hunting.

Detailed Evidence-Based Standard Setting Review Procedures

In this section, a more detailed description of activities that took place throughout the evidence-based standard setting review meeting is provided. The co-facilitators led all activities. Please refer to Appendix B for the presentation slides. All activities were designed to comply with Standards 5.21 and 5.22 (AERA, APA, NCME, 2014).

Opening Session

The meeting began by welcoming the panelists, which was followed by the meeting facilitator introducing her co-facilitator and explaining the purpose, logistics, and security of the meeting.

Introduction

The panelists were asked to introduce themselves by describing their educational and professional backgrounds. After the introductions, the agenda of the meeting was presented. The facilitator noted that the day might flow rather fluidly dependent on the pace of the activities and additional discussions that might be needed at certain points during the day.

Next, the panelists were provided with a brief history of ELL placement tests for kindergarten students in Arizona. Included in this history was the definition of what AZELLA is, as well as what it is intended to do. Also included was Arizona's statutory definition of what it means to be proficient in English. Additionally, the panelists were given the history surrounding the reasons for, and timeline of, the development of the Placement Test. This was followed by a description of the test's Blueprint, both by test activity grouping and content domain.

Experience the Test

After introducing the Student Response Sheet/Scoring Rubric, the panelists were supplied with a copy of the Placement Test Book to review as they watched the Demonstration Video of a teacher administering the test. This video provided an example of "real-life" responses of a kindergarten-aged actor to the questions on the test. It was intended as an efficient way to help panelists become familiar with test content and gain an appreciation for the test. A brief discussion of the panelist's impressions about the test and its administration followed the presentation.

Original Standard Setting

After a short break, the panelists were provided with information about the July 2012 AZELLA Kindergarten Placement Test Standard Setting. The information included what standard setting is as well as described the people who were involved and the processes used to arrive at the current cut-scores. It was during this section of the meeting that the

Performance Level Descriptors were reviewed. The statistics for the prior committee's cut scores for each round were provided as well as the final cut score recommendations.

Review of the Ordered Item Booklet

To better understand the standard setting process used originally and to more deeply understand the content of the test, the ordered item book was reviewed. The panelists were asked to share insights about the additional information each item or set of items provided about the students' English language skills. Specifically, the panelists were asked to think about:

- The proficiency levels as they reviewed the items.
- When the students have demonstrated enough English language proficiency to be in the mainstream.
- Developing (but not sharing) a preliminary opinion about the appropriateness of the existing cut scores.

It was during this section of the meeting that they were advised that they were free to recommend the *Basic/Intermediate* cut anywhere below that of the *Proficient* cut; however, based on the recommendation of ADE's TAC, the *Proficient* cut would need to be set within the 32 to 38 raw score point range, inclusive. Within the ordered item booklet, the current cut for *Basic/Intermediate* was printed on blue paper. The items in the 32 to 38 raw score point range were printed on pink paper. (Please note: The raw score of 32 was chosen as the minimum of the range because it was the raw score cut currently in use. The raw score of 38 was chosen as the maximum of the range because it was the maximum chosen by any member of the original 13 member standard setting panel.)

Test Validity

Following the discussion of the original standard setting, some of the test's validity evidence was presented. This was followed by a reminder that it was not the charge of the committee to discuss or determine the validity of the Placement Test. Rather, the committee was charged with using the process recommended to ADE by its TAC to review the cut scores and possibly make recommendations to the State Superintendent for changes to those cut scores.

Presentation of Evidential Data

TAC had recommended that the committee use the data from the second year administration (school year 2013-2014) for both the Placement Test and the Spring AZELLA Reassessment Test (2014 Spring Test) to inform the review. ADE chose to present the data from school year 2012-2013 to the panel as additional historical background information. Data were presented first for the *Proficient* cut relating to student results during school year 2012-2013 and then for school year 2013-2014. This

was followed by data for the *Basic* cut for each year. Annual, sequential scatterplots and bar graphs were presented.

While both the Placement Test and the Spring Test are aligned to the assessable Arizona ELPS and scaled using the Rasch model for their dichotomously scored items and the Partial Credit model for their polytomous items, their scales are completely independent. Therefore, the raw scores for each test were used for scatterplots presented to the panelists.

Another difference between the two tests concerned the criteria for an examinee scoring within the *Proficient* performance level. Since no incoming kindergarten student is expected to be able to read or write, proficiency on the Placement Test is determined solely on Total Score. Consistent with the rest of the AZELLA assessments, in order to reach a *Proficient* level on the Spring Test, students not only had to obtain a scale score above the *Proficient* cut for Total Score, but also on both the Reading and the Writing subtests. Moreover, any student who had a Total Score above the *Proficient* cut but did not have a *Proficient* Reading and/or Writing score was assigned an overall proficiency level of *Intermediate*. Percentages within the bar graphs were based on overall proficiency levels as determined by the scoring rules for each test.

The scatterplots showed how student scores on the Placement Test related to those of the Spring Test. In these graphs, the raw score cut for the Spring Test was indicated by a black line and the current raw score cut for the Placement Test was indicated by a red line. For the *Proficient* Cut graphs, an additional red line indicated the top raw cut score that the TAC could endorse. Students with scores in each quadrant were indicated by a different color, with those who scored above the cut on both tests being identified in dark green, and those who scored above the cut on the Placement Test, but not on the Spring Test, being shown in red. Two additional colors, light green and blue, were used for students who scored less than the cut on the Placement Test. The colors used for each group of students in the scatterplots were maintained in the bar graphs for ease of use.

The bar graphs provided to the panelists showed the percentage of kindergarten students at each of the Spring Test proficiency levels by their score above or below the cut score for a selected set of raw score points on the Placement Test. The raw score points selected for the *Proficient* Cut graphs were those that the TAC had indicated it could endorse. Those for the *Basic* Cut graphs were the three above and three below the current *Basic/Intermediate* raw score cut for the Placement Test.

Using the 2013 graphs, the panelists were asked to share insights regarding how they felt about the performance of students in each of the quadrants. Then using the 2014 *Proficient* graphs, they were asked to compare the graphs to those for 2013 and report any changes they saw. The complete set of graphs provided to the panelists is presented as Figure D1 through D8 in Appendix D.

Standard Setting Methodology

The facilitator briefly trained the panelists on the methodology that they would use during standard setting (See Appendix B page 35). The panelists were told that there would be two rounds for the selection of the *Proficient* raw cut score and then two rounds for the selection of the *Basic/Intermediate* raw cut score. While Round 1 would be accomplished without discussion with others, the panelists would be asked to share the reasons for their choice after the round was complete. The panelists would also have the chance to change their choice, if they chose to, during Round 2. Prior to each round, any questions the panelists had were addressed.

The panelists were told that they would use the 2014 graphs in conjunction with the content of items in the ordered item book and the Performance Level Descriptors to make their selection. They were expected to make their decision based on both data and the content of the items. Both facilitators emphasized that they should choose their cut on the page where they would expect most students above the cut to get the item correct and most students below the cut to get the item wrong. They were asked to place a sticky note on the lowest possible score for the proficiency level and to record this page number on their Page Number Recording Sheet (see Appendix E). The panelists were told that it is their recorded page number that would be used by ADE as their recommended cut score for the round.

Round 1 - Proficient

For the proficient cut, the panelists were to place their sticky note on the pink page that they would advise ADE to place the cut score. The facilitators monitored the panelists as they completed their task, and when they had all recorded their selections on their Page Number Recording Sheet, they took a break which included a group lunch.

After lunch, one of the facilitators read the page selections of each of the panelists, while the other facilitator recorded their choice in a frequency graph on a flip-chart. Then the panelists were asked to explain the factors they considered and the reasons they picked that page. The panelists were reminded that each of them was an expert and that the purpose of explaining their decision was to allow other panelists to understand their rationale. It was reiterated that consensus was not necessary and that they were not trying to persuade others.

The main reason for selecting the item on page 34 as the skill that would be representative of the barely English language proficient student was its requirement that the student hear and repeat the “-ed” sound in a word embedded in a short sentence. Several of the panelists expressed the opinion that this was a necessary skill to be successful in a mainstream classroom, and all of them chose to place their bookmark at or above this page. This item corresponds to the PLD description of skills for students scoring Proficient on the Placement Test which states that they generally demonstrate that they can use “basic verbs and adjectives” and consistently use “correct pronunciation.”

Six of the eight panelists, however, placed their bookmark on page 35, indicating that this item contained the skill required for a student to be considered barely proficient in the English language. They explained that this item required that the student listen to a brief story, then recall and retell at least two key ideas about the story to achieve two of the three points possible for the item. They expressed belief that the complexity of the task was a strong indication that the student would be able to access instruction and be successful in a mainstream classroom. This item corresponds to the PLD description of skills for students scoring Proficient on the Placement Test which states that they generally demonstrate that they can use “pictures or words to retell events from a story heard.”

After all of the panelists’ page numbers had been recorded, the minimum, maximum, and median were tabulated for the panelists to see.

Round 2 - Proficient

For Round 2, the panelists were advised that based on the information gained from the explanations of the other panelists, they now had the opportunity to change the page of their advised cut, if they so choose. Again they were asked to place a sticky note on the pink page where they would advise ADE to insert the cut score.

The facilitators monitored the panelists as they completed their task, and when they had all recorded their selections on their Page Number Recording Sheet, one of the facilitators again read the page selections of each of the panelists, while the other facilitator recorded their choice in a frequency graph on a flip-chart. After all of the panelists’ page numbers had been recorded, the minimum, maximum, and median were tabulated for the panelists to see. They were then asked how they felt about the median, which would be used to present to the Superintendent of Public Instruction as the recommended cut score. None of the panelists expressed any negative feelings about its use.

Round 1 and Round 2 – Basic/Intermediate

Following the procedure laid out under Round 1 – Proficient and Round 2 – Proficient, the panelist chose cut scores for the *Basic/Intermediate* proficiency level.

In Round 1, one-half of the panelists chose to place their bookmark on the item on page 20. The item on this page requires that the students appropriately respond to a social request for information. This item corresponds to the PLD description of skills for students scoring Basic/Intermediate on the Placement Test which states that they generally demonstrate that they can “respond to social interactions with gestures and simple words.” The other four panelists each chose an item on separate pages (21, 22, 23, and 25). Each of these panelists expressed an opinion that for a student to be barely Basic/Intermediate, he or she needs to follow a one-step direction, recall and retell at least one key idea from a story, appropriately use the plural of a noun, and correctly repeat the “-ing” sound in a word embedded in a short sentence, respectively.

After the discussion following Round 1, none of the panelist chose the item on page 25 in Round 2. Moreover, the highest frequency of panelist (3) chose the item on page 21 which requires the student to follow a one-step direction. This item corresponds to the PLD description of skills for students scoring Basic/Intermediate on the Placement Test which states that they generally demonstrate that they can follow “single-step directions.”

Standard Setting Evaluation Form

The panelists filled out the standard setting evaluation form upon the completion of the meeting. The questions and responses to the evaluation form are summarized in Appendix F.

Standard Setting Results

In this section, the results from each round are presented.

The median raw score cuts for the *Basic/Intermediate* and *Proficient* levels are summarized in Table 4 by round. The median raw score cuts for the *Basic/Intermediate* and *Proficient* levels for both rounds were 21 and 35 (out of 42 maximum possible points), respectively. Although the distributions of recommended cut raw scores were somewhat different across the three rounds (Figures 1 and 2), the recommended cuts did not change at Round 2

Table 4.

Median Raw Score Cuts by Round

Raw Score Cuts	Basic/Intermediate	Proficient
Round 1	21	35
Round 2	21	35

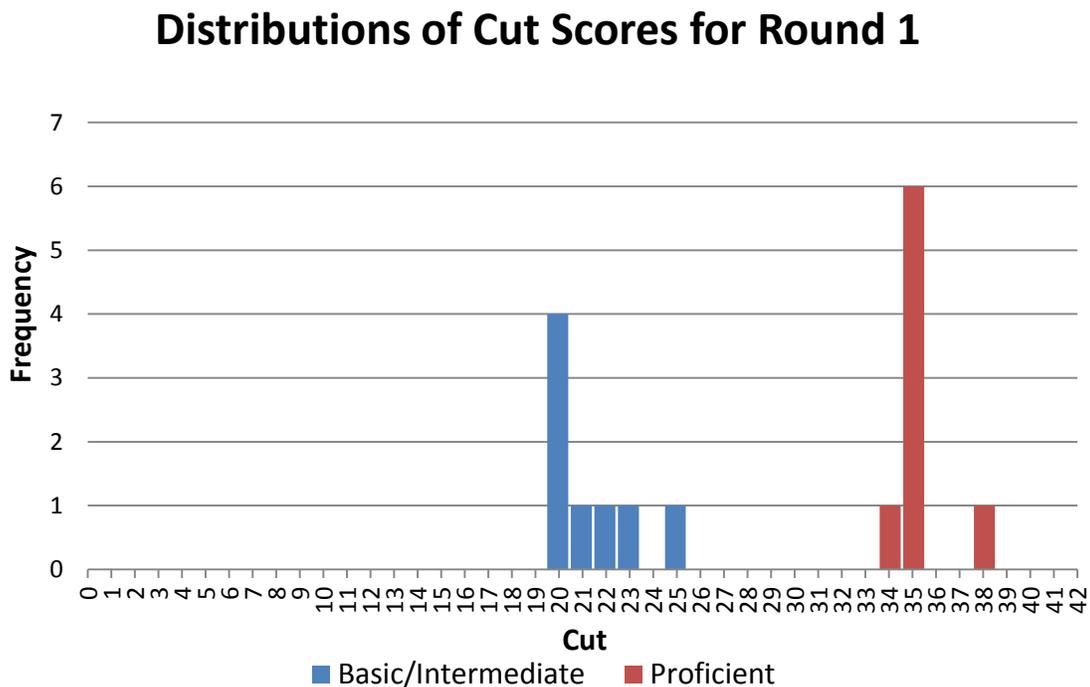


Figure 1. Raw score cut distributions of cuts at Round 1

Distributions of Cut Scores for Round 2

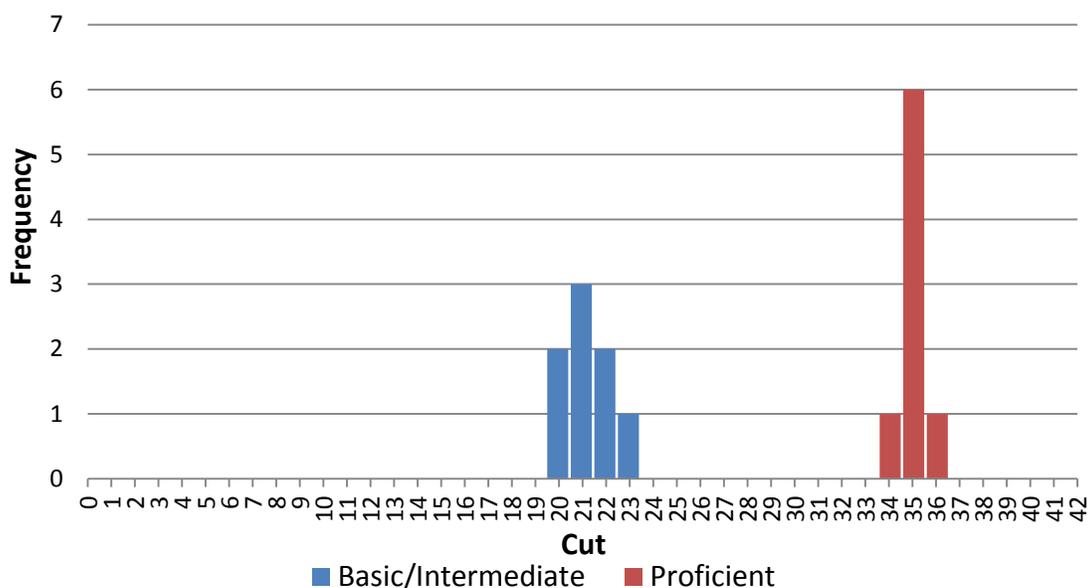


Figure 2. Raw score cut distributions of cuts at Round 2

The minimum, median, and maximum raw score cuts of the panel for each round are presented in Table 5. For each proficiency level, the range between of the raw score cuts decreased in Round 2.

Table 5.

Minimum, Median, and Maximum Raw Cut Scores by Round

	Basic/Intermediate	Proficient
	<u>Round 1</u>	
Minimum	20	34
Median	21	35
Maximum	25	38
	<u>Round 2</u>	
Minimum	20	34
Median	21	35
Maximum	23	36

The scale score cuts, based on the median raw score cuts for the *Proficient* and *Basic/Intermediate* levels by round, are presented in Table 6. Since the median raw score cuts were the same for both rounds, the scale score cuts stayed the same. Note that Placement Test scale scores range from 100 to 300.

Table 6.

Scale Score Cuts by Round

Scale Score Cuts	Basic/Intermediate	Proficient
Round 1	206	245
Round 2	206	245

The proficiency level distribution for PHLOTE students by round is presented in Table 7. Approximately, 24%, 34%, and 42% of PHLOTE students who took their first Placement Test during the fall semester of fiscal year 2014 would be classified as *Pre-emergent/Emergent*, *Basic/Intermediate*, and *Proficient*, respectively. Again, the impact data did not change from Round 1 to Round 2.

Table 7.

Impact Data for PHLOTE Students Based on Median Cuts by Round

Impact	Pre-emergent/Emergent	Basic/Intermediate	Proficient
Round 1	23.5%	34.4 %	42.1%
Round 2	23.5%	34.4 %	42.1%

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage Publications.
- Cizek, G. J., Bunch, M. B., & Koons, H. (2004). An NCME instructional module on setting performance standards: Contemporary methods. *Educational measurement: Issues and practice*, 23(4), 31-50.
- Masters, G.N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149-174.
- Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 249-281). Mahwah, NJ: Lawrence Erlbaum.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Denmark Pedagogiske Institute. (Reprinted 1980 by University of Chicago).

Appendix A

Evidence-Based Standard Setting Review Participants

Appendix A

Evidence-Based Standard Setting Review Participants

Name	District	School	Current Position	Certifications, Endorsement, and Experience	Awards
Heather Anguiano	Chandler Unified	Hartford Sylvia Encinas Elementary	Principal	Elementary with both Spanish Bilingual and Structured English Immersion Endorsement. Masters in Curriculum and Instruction. Masters in Educational Leadership. Administrative Certificate. K-2 bilingual teacher in inner-city Houston. Taught at San Marcos Elementary for 4 years and then as instructional specialist. Assistant Principal for 1 year. Principal for 7 years, 20 years in education.	Rodel Exemplary Principal 2014
Cindy Cochran	Fowler Elementary	Western Valley Elementary	Kindergarten Teacher	Elementary with Early Childhood, Structured English Immersion and Reading Endorsement. Has taught for 3 years. ADE Kindergarten Academy participant.	
Clarissa Cristofori	Flowing Wells Unified	J Robert Hendricks Elementary	Kindergarten Teacher	Elementary with Early Childhood Endorsement. Structured English Immersion Full Endorsement. 8 years in education including 5 years as a kindergarten teacher and 6 months as a second-grade teacher. Currently serving as grade-level chair, Career Ladder Instructional Coordinator, active PTO member, New Teacher Mentor and on Teacher Assistance Team.	Rodel Exemplary Teacher 2013
Nichole Handford	Tucson Unified	C. E. Rose Elementary	Kindergarten Teacher	Elementary with Early Childhood Endorsement. Structured English Immersion Endorsement. Reading Endorsement. 5 1/2 yrs as a kindergarten teacher. Currently serving as a member of the school professional development committee.	Rodel Exemplary Teacher 2013
Arlinda Mann	Mesa Unified		Assistant Superintendent	Elementary and Administrative Certificates with Bilingual Education Endorsement. Taught in multiple inner-city schools in multiple districts. Taught Grade 1, 3, 4, 5, and 6. Assistant Principal at two schools and Principal at another school with a high population of ELL students. Currently an Area Assistant Superintendent in Arizona's largest school district, in charge of the area with the highest percentage of ELL students. Has 14 years in education.	
Kelsie Pennington	Pendergast Elementary	Villa de Paz Elementary	Kindergarten Teacher	Elementary with English as a Second Language and Early Childhood Endorsements. Serves as Reading Specialist. Has taught for 13 years. ADE Kindergarten Academy participant.	
Victoria Perkins	Cartwright Elementary	G. Frank Davidson Elementary	Grade 1 Teacher	Elementary with Early Childhood and Structured English Immersion Endorsements. Certificates in English as a New Language and Early and Middle Childhood. Has taught for 16 years.	National Board Certified
Lori Ramirez-Jaquez	Laveen Elementary	Rogers Ranch Elementary	Grade 1 Teacher	Elementary with English as a Second Language, Structured English Immersion, and Early Childhood Endorsements. Has taught for 18 years.	Rodel Exemplary Teacher 2007

Appendix B
Presentation Slides

Arizona English Language Learner Assessment (AZELLA) Kindergarten Placement Test

Cut Score Review Committee

May 30, 2014



Today's Task

- Review the AZELLA Kindergarten Placement Test cut scores for “Proficient” and “Basic/Intermediate” and provide a recommendation regarding the possible revision of those cut scores.
- Your recommendations will be provided to Superintendent Huppenthal who will make the final determination for the cut scores.



Non-Disclosure

- You **may not** discuss the content of the AZELLA Kindergarten Placement Test outside of this committee meeting
- You **may** discuss this committee's task and the processes used to reach a final recommendation outside of this committee meeting
- You **may not** take any AZELLA Kindergarten Placement Test materials when you leave
- You **may** take the PowerPoint handout when you leave



Agenda

- What is the AZELLA?
- Why a Kindergarten Placement Test?
- Development of the AZELLA Kindergarten Placement Test
- Standard Setting Process
- How We Know the Test is Valid
- Data to Inform Decisions
- Decision Making Process





What is the AZELLA?

ARIZONA'S ASSESSMENTS

- Definition

The Arizona English Language Learner Assessment (AZELLA) is Arizona's test of English language proficiency that is aligned to the Arizona English Language Proficiency Standards (ELPS) and is administered to students with a primary or home language other than English (PHLOTE).



What is the AZELLA?

ARIZONA'S ASSESSMENTS

- Uses of AZELLA

- Placement Test

- Determine if a PHLOTE student is an English Language Learner (ELL)
- Determine placement for ELL students

- Annual Reassessment

- Measure progress of ELL students towards English language proficiency



What is the AZELLA?

ARIZONA'S ASSESSMENTS

- What does “Proficient” mean?
“Once English learners have acquired a good working knowledge of English and are able to do regular school work in English, they shall no longer be classified as English learners and shall be transferred to English language mainstream classrooms.” (A.R.S. 15-752)

A “good working knowledge of English” is operationalized as scoring Proficient on AZELLA.



Why a Kindergarten Placement Test?

ARIZONA'S ASSESSMENTS

- Prior to school year 2012-13, the same AZELLA test was used for both Kindergarten Placement and Reassessment
 - Test was too long, too difficult, and too academic for students just entering Kindergarten
 - More than 95% of Kindergarten PHLOTE students failed the placement test – over identifying Kinders as ELLs



Why a Kindergarten Placement Test?

ARIZONA'S ASSESSMENTS

- A test that was appropriate for students just entering Kindergarten was needed
 - Short administration time (less than 20 minutes)
 - Administered one-on-one by a trained, qualified test administrator
 - Primarily a listening and speaking test with some pre-reading and pre-writing items
 - Aligned to the Arizona ELPS



Development of the AZELLA Kindergarten Placement Test

ARIZONA'S ASSESSMENTS

- In 2011,
 - Arizona ELPS revised
 - AZELLA Advisory Committee recommended key features of the Kindergarten Placement Test
 - Initial Test Blueprint developed
 - Item writing began



Development of the AZELLA Kindergarten Placement Test

ARIZONA'S ASSESSMENTS

- In 2012,
 - Item Pilot Study in January with about 100 students
 - Test Blueprint and items revised based on pilot
 - Field Test in April/May with about 1000 students (both PHLOTE and non-PHLOTE students)
 - Final Test form developed in June/July
 - Standard Setting in July
 - Operational Test available in August



Development of the AZELLA Kindergarten Placement Test

ARIZONA'S ASSESSMENTS

- Test Blueprint by Activity

Activity	Content Descriptions
(1) Hello! Let's Play!	A Answering Social Questions B Expressing One's Thoughts C Following One-Step Directions D Following Two-Step Directions
(2) Picture Time!	A Naming Familiar Objects in a Picture B Adding Details to Drawings C Predicting a Story from a Book Cover D Sequencing Pictures About a Read Aloud
(3) Let's Talk!	A Repeating Words B Using Plural Words C Repeating Rhyming Words D Repeating Grammatical Structures in Sentences E Identifying Pictures with the Same First Sound
(4) Story Time!	A Listening to a Story and Answering Questions B Using Pictures to Retell a Story



Development of the AZELLA Kindergarten Placement Test

ARIZONA'S ASSESSMENTS

- Test Blueprint by Domain

Domain/Standard	Kindergarten Placement
Reading	17%
Print Concepts/ Phonemic Awareness/ Decoding	43%
Comprehension	57%
Writing	5%
Applications	0%
Conventions	0%
Process/Elements/Research	100%
Listening	26%
Comprehension	100%
Speaking	48%
Delivery	50%
Repeats	50%



AZELLA Kindergarten Placement Test

ARIZONA'S ASSESSMENTS

- AZELLA Kindergarten Placement Test Book
- Student Response Sheet/Scoring Rubric
- Demonstration Video

AZELLA Kindergarten Placement Test

- At this point, you should feel familiar with
 - Why there is an AZELLA Kindergarten Placement Test
 - How the AZELLA Kindergarten Placement Test was developed
 - The content of the AZELLA Kindergarten Placement Test
- Next topic – Standard Setting



Standard Setting

ARIZONA'S ASSESSMENTS

- What
 - Process used to establish cut scores for the various proficiency levels of a test
- Who
 - Committee of subject matter experts
- How
 - Develop common understanding of the proficiency levels
 - Recommend cut scores that match proficiency level descriptions



Standard Setting

ARIZONA'S ASSESSMENTS

- Kindergarten Placement Test Standard Setting Committee
 - Held July 2012
 - 13 members
 - 9 Kinder teachers, 3 ELL Coordinators/Specialists, 1 Speech language pathologist
 - 9 with a Master's degree
 - 3 – 37 years experience
 - Urban, suburban, rural schools
 - Small, medium, large districts/LEAs
 - Low to Upper Middle SES schools



Standard Setting

ARIZONA'S ASSESSMENTS

- Review and discuss the AZELLA Kindergarten Placement Test
- Develop a common understanding of the test's proficiency levels and the borderline students
 - Proficient
 - Basic/Intermediate
 - Pre-Emergent/Emergent



Standard Setting

ARIZONA'S ASSESSMENTS

Basic/Intermediate Performance Level Descriptor

Students at this level generally understand spoken English, but do not have the vocabulary to respond consistently. They orally communicate basic needs and ideas with gestures and isolated English words. They use pictures to recall objects from a story heard, repeat words that begin with the same first sound and add minimal details to drawings.

This student does not demonstrate the sufficient skills in English to access mainstream curriculum and demonstrates the need for specific support in English Language Development instruction.



Standard Setting

ARIZONA'S ASSESSMENTS

Proficient Performance Level Descriptor

Students at this level listen and respond appropriately to spoken English. They have an expanded English vocabulary to orally communicate basic needs and ideas with English words, phrases, and sentences with correct pronunciation. They use pictures or words to retell events from a story heard, identify pictures with the same first sound, and add relevant details to drawings.

This student demonstrates the skills necessary to access mainstream curriculum.



Standard Setting

ARIZONA'S ASSESSMENTS

- Recommend cut scores
 - Independent, individual recommendations
 - Not consensus
 - Three rounds of recommendations, more data provided with each round



Standard Setting

ARIZONA'S ASSESSMENTS

- Round 1
 - Cut score recommendation based solely on content of test items
- Round 2
 - Item difficulty provided
 - Distribution of Round 1 cut scores
 - Discussion
- Round 3
 - Impact data based on median Round 2 cut scores
 - Distribution of Round 2 cut scores
 - Discussion



Standard Setting

ARIZONA'S ASSESSMENTS

- Cut score recommendations

	Round 1	Round 2	Round 3
Proficient (Max)	38	38	38
Proficient (Median)	32	32	32
Proficient (Min)	22	25	30
Basic/Inter (Max)	32	30	30
Basic/Inter (Median)	22	22	22
Basic/Inter (Min)	14	15	18



Standard Setting

ARIZONA'S ASSESSMENTS

- Adopted cut scores
 - 22 points for Basic/Intermediate
 - 32 points for Proficient



Standard Setting

ARIZONA'S ASSESSMENTS

- Review of the ordered item booklet.
- What does each item tell you about the student's language skills?
- Think about the proficiency levels as we review the booklet. When have students demonstrated enough English language proficiency to be in the mainstream?
- It is ok to develop an opinion about the appropriateness of the existing cut scores while reviewing the content of the booklet but don't share that opinion now.

Standard Setting

- At this point, you should have a basic understanding of
 - AZELLA Kindergarten Standard Setting process
 - AZELLA Kindergarten Placement Test ordered item booklet
 - Relationship between the ordered item booklet and "typical" PHLOTE student performance on the Placement Test
- Next Topic – How We Know the Test is Valid





How We Know the Test is Valid

ARIZONA'S ASSESSMENTS

- Developed following the *Standards for Educational and Psychological Testing*
- Standard psychometrics used for determining which items should be on the final operational test form
- Several additional studies have been conducted
- Arizona's Technical Advisory Committee confirmed the validity of the AZELLA Kindergarten Placement Test in December 2013



How We Know the Test is Valid

ARIZONA'S ASSESSMENTS

The *preLAS* Study produced strong evidence.

- July and August of 2013, a representative sample of about 1000 incoming Kindergarten PHLOTE students took both the AZELLA Kindergarten Placement Test and the commercially available *preLAS* English language proficiency test.
- The correlation coefficient for these two tests was found to be 0.861, which is quite high.



How We Know the Test is Valid

ARIZONA'S ASSESSMENTS

preLAS Study (continued)

- The Spring 2014 Stage I Reassessment results for these same students was studied. The AZELLA Kindergarten Placement Test was a better predictor of performance on the Stage I test than was the *preLAS*.
- Both of these outcomes support the validity of the AZELLA Kindergarten Placement Test for measuring English language proficiency.



How We Know the Test is Valid

ARIZONA'S ASSESSMENTS

- We are confident of the validity of the placement test.
- This committee is not charged with discussing or determining the validity of the AZELLA Kindergarten Placement Test.
- The AZELLA Kindergarten Placement Test, taken as a whole, is a valid English language proficiency test that distributes students on a continuum from least proficient to most proficient.



Today's Task

ARIZONA'S ASSESSMENTS

- However, there has been some concern from AZ educators that this new placement test might be under identifying ELL students.
- This committee is charged with the task of reviewing the AZELLA Kindergarten Placement Test cut scores and providing recommendations
 - Proficient cut score: stay the same or increase
 - Basic/Inter cut score: stay the same or increase or decrease



Today's Task

ARIZONA'S ASSESSMENTS

- Our Technical Advisory Committee's recommended process for the review of the cut scores:
- Empanel an impartial committee of subject-matter experts
 - Use school year 2013-14 AZELLA Kindergarten Placement Test and the Stage I Reassessment data
 - review the Stage I performance of IFEP students at each Placement Test raw score
 - advise ADE on the most appropriate cut score
 - Review both the Proficient and Basic/Inter cut scores
 - A Proficient cut score between 32 and 38 is supported by the Technical Advisory Committee



Today's Task

ARIZONA'S ASSESSMENTS

- This committee's recommendations will be provided to the Superintendent who will make the final cut score decisions
- Any changes the Placement Test cut scores will be implemented for school year 2014-15
- Kinder IFEPs will not longer participate in the Stage I Reassessment beginning with school year 2014-15

Today's Task

- At this point, you should
 - Be confident that the AZELLA Kindergarten Placement Test is a valid assessment of English language proficiency
 - Know that you will review the Proficient cut score and recommend that it stay the same or be increased
 - Know that you will review the Basic/Intermediate cut score and recommend that it stay the same, be increased, or be decreased





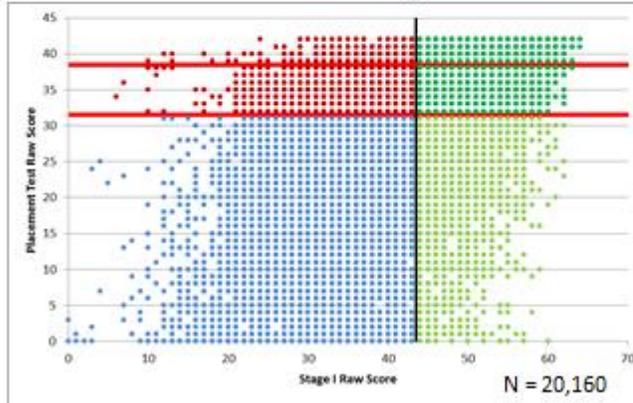
Data to Inform Decisions

ARIZONA'S ASSESSMENTS

2013 Scatterplots



Stage I Proficient cut = 44



Maximum cut = 38

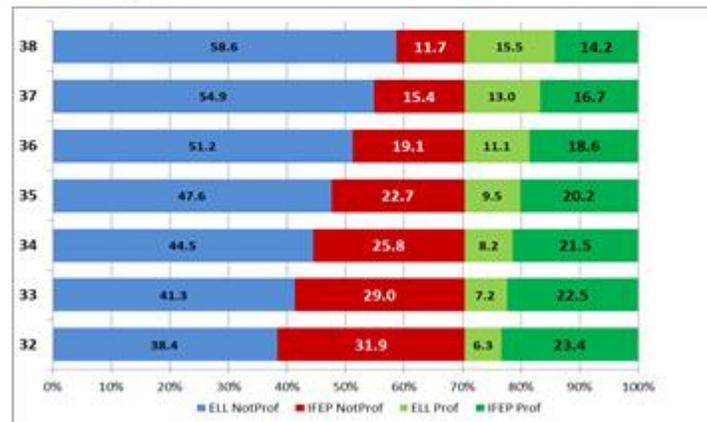
Current Placement Proficient cut = 32



Data to Inform Decisions

ARIZONA'S ASSESSMENTS

2013 Bar Graphs





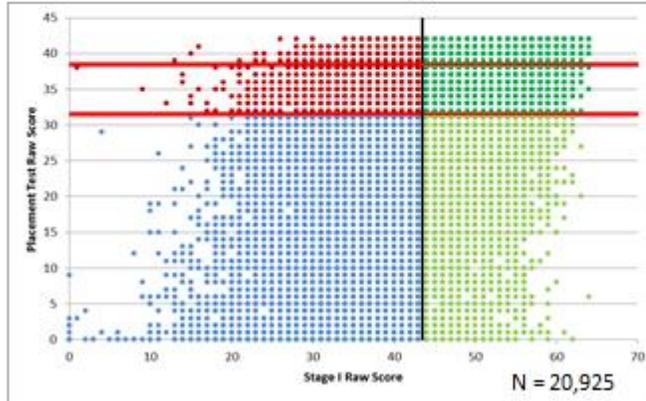
Data to Inform Decisions

ARIZONA'S ASSESSMENTS

2014 Scatterplots



Stage I Proficient cut = 44



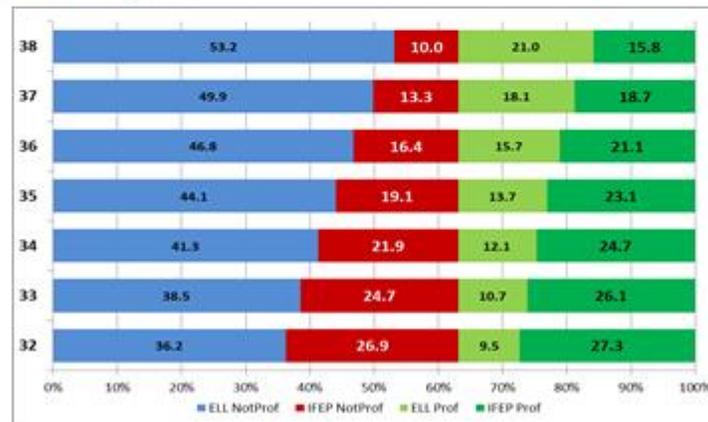
← Placement Test Proficient cut = 32



Data to Inform Decisions

ARIZONA'S ASSESSMENTS

2014 Bar Graphs





Round 1 – Proficient Cut Score

- Use the 2014 Proficient graphs, the content of the items in the ordered item book, and the Performance Level Descriptors to determine the cut score you would recommend.
- The cut should be between 32 and 38 (pink pages). Your decision should be based on both data and content.
 - Most students who score Proficient should be able to get this item correct.
 - Most students who score below Proficient should get this item wrong.
- This round will be accomplished without discussion with others.



Round 1 – Proficient Cut Score

- Place your Panelist Number on the line at the top of your Recording Sheet
 - Place a sticky note on the pink page where you would advise ADE to place the Proficient cut score (the lowest possible “Proficient” score).
 - Be prepared to share the reasons you placed it on that page. You will have a chance to change your choice during Round 2.
 - Enter the page number of the pink page with your sticky note on your Recording Sheet next to Round 1 under Proficient Cut Page Number.
-
- Do you have any questions?
 - Do you know what to do?

Round 1 Results - Proficient

- Round 1 results posted – distribution, min, median, max
- Round robin discussion – Why did you pick your cut score? What factors did you consider?
- Consensus is not necessary. You are not trying to persuade others. You are just explaining your decision.



Decision Making Process

ARIZONA'S ASSESSMENTS

Round 2 – Proficient Cut Score

- Based on the discussion of the Round 1 results you are better informed. If you want to change your recommended cut score, now is your chance.
- Place a sticky note on the pink page where you would advise ADE to place the cut score.
- Enter the page number of the pink page with your sticky note on your Recording Sheet next to Round 2 under Proficient Cut Page Number.
- Do you have any questions?
- Do you know what to do?

Round 2 Results - Proficient

- Round 2 results posted – distribution, min, median, max
- The median score is the recommendation that will be provided to the Superintendent.



Decision Making Process

ARIZONA'S ASSESSMENTS

Round 1 – Basic/Intermediate Cut Score

- Use the 2014 Basic graphs, the content of the items in the ordered item book, and the Performance Level Descriptors to determine the cut score you would recommend.
- The cut should be between 19 and 25 (the blue page is the current cut score). Your decision should be based on both data and content.
 - Most students who score at least Basic should be able to get this item correct.
 - Most students who do not score at least Basic should get this item wrong.
- This round will be accomplished without discussion with others.



Round 1 – Basic/Intermediate Cut Score

- Place a sticky note on the page where you would advise ADE to place the Basic/Intermediate cut score (lowest possible “Basic/Intermediate” score).
- Be prepared to share the reasons you placed it on that page. As with the Proficient Cut, you will have a chance to change your choice during Round 2.
- Enter the page number of the page with your sticky note on your Recording Sheet next to Round 1 under Basic Cut Page Number.
- Do you have any questions?
- Do you know what to do?

Round 1 Results – Basic/Inter

- Round 1 results posted – distribution, min, median, max
- Round robin discussion – Why did you pick your cut score? What factors did you consider?
- Consensus is not necessary. You are not trying to persuade others. You are just explaining your decision.





Round 2 – Basic/Intermediate Cut Score

- Based on the discussion of the Round 1 results you are better informed. If you want to change your recommended cut score, now is your chance.
- Place a sticky note on the page where you would advise ADE to place the cut score.
- Enter the page number of the page with your sticky note on your Recording Sheet next to Round 2 under Basic Cut Page Number.

- Do you have any questions?
- Do you know what to do?

Round 2 Results – Basic/Inter

- Round 2 results posted – distribution, min, median, max
- The median score is the recommendation that will be provided to the Superintendent.





Confidential material supplied during this committee needs to be returned. Please:

- Place all loose documents except your Recording Sheet in the front pocket of your folder.
- Hand your completed Recording Sheet to Irene. They must all be collected prior to adjournment.
- Complete exit survey.

Thank you for sharing
your expertise
and
for being
Ambassadors
of the
Assessment!



Appendix C

Performance Level Descriptors

Arizona English Language Learner Assessment Performance Level Descriptors Kindergarten Placement Test

Proficient—Students at this level listen and respond appropriately to spoken English. They have an expanded English vocabulary to orally communicate basic needs and ideas with English words, phrases, and sentences with correct pronunciation. They use pictures or words to retell events from a story heard, identify pictures with the same first sound, and add relevant details to drawings. This student demonstrates the skills necessary to access mainstream curriculum.

Basic/Intermediate—Students at this level generally understand spoken English, but do not have the vocabulary to respond consistently. They orally communicate basic needs and ideas with gestures and isolated English words. They use pictures to recall objects from a story heard, repeat words that begin with the same first sound and add minimal details to drawings. This student does not demonstrate the sufficient skills in English to access mainstream curriculum and demonstrates the need for specific support in English Language Development instruction.

Pre-Emergent/Emergent—Students at this level lack the English skills to communicate, retell stories heard, or add details to drawings. This student does not demonstrate sufficient skills in English to access mainstream curriculum and demonstrates the need for specific support in English Language Development instruction.

Pre-Kindergarten	Pre-Kindergarten
Students scoring Proficient at this grade generally demonstrate the following skills, knowledge, and abilities drawn from the Kindergarten ELP standards.	Students scoring Basic/Intermediate at this grade generally demonstrate the following skills, knowledge, and abilities drawn from the Kindergarten ELP standards.
<ul style="list-style-type: none"> • Speaking/Listening <ul style="list-style-type: none"> ○ Responds appropriately to social interactions ○ Consistently uses correct pronunciation ○ Uses correct developmentally–appropriate grammatical structures ○ Uses basic verbs and adjectives ○ Follows multiple-step directions ○ Sequences events in a story 	<ul style="list-style-type: none"> • Speaking/Listening <ul style="list-style-type: none"> ○ Responds to social interactions with gestures and simple words ○ Repeats individual words and short phrases ○ Uses basic nouns ○ Follows single-step directions ○ Identifies objects from a story

<ul style="list-style-type: none">● Pre-Reading<ul style="list-style-type: none">○ Uses pictures or words to retell events from a story heard○ Identifies pictures that begin with the same first sound○ Uses pictures to make predictions● Pre-Writing<ul style="list-style-type: none">○ Adds relevant details to drawings	<ul style="list-style-type: none">● Pre-Reading<ul style="list-style-type: none">○ Uses pictures to recall people or objects from a story heard○ Repeats words that begin with the same first sound● Pre-Writing<ul style="list-style-type: none">○ Adds minimal details to drawings
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Appendix D
Evidential Data Graphs

Proficient Cut – 2013 Data

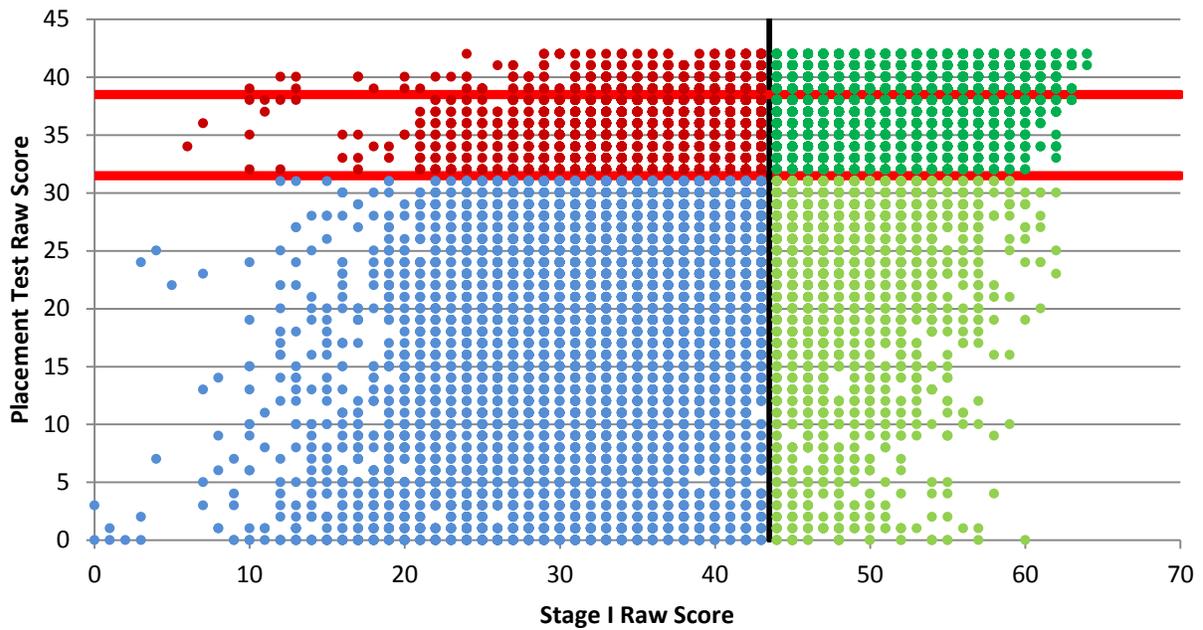


Figure D1. Proficient Cut 2013 Scatterplot. Scatterplot of kindergarten scores on the Placement Test in fall of school year 2012-2013 and on the Spring 2013 AZELLA Stage I Test with Proficient cuts noted.

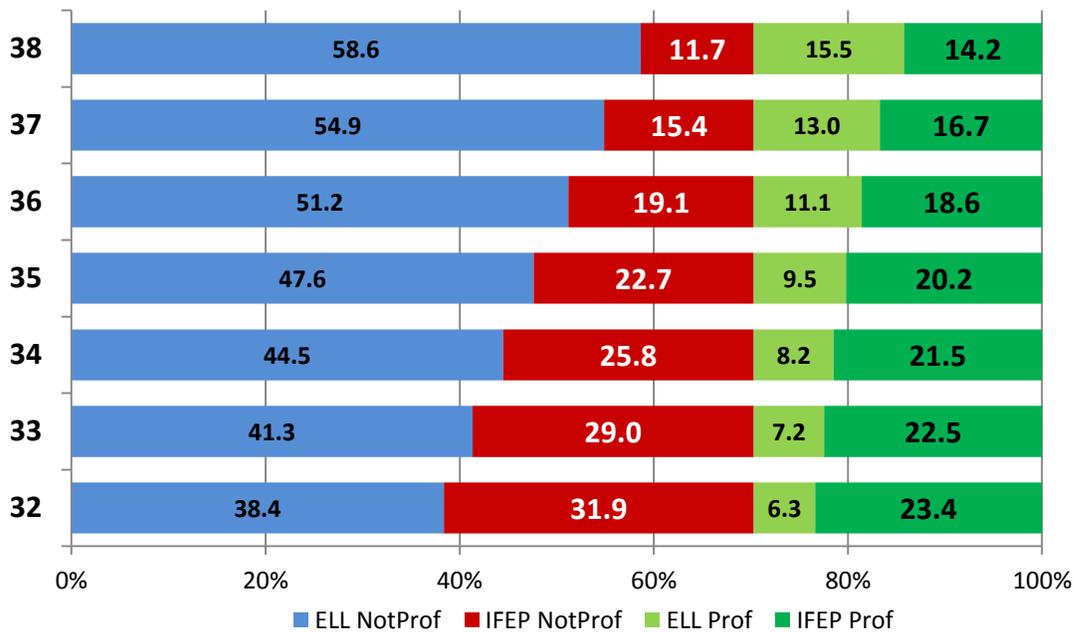


Figure D2. Proficient Cut 2013 Bar Graph. Percentage of kindergarten students in each performance level on the Spring 2013 AZELLA Stage I Test by their performance on their Placement Test in the fall of school year 2012-2013 for selected Placement Test raw score points.

Proficient Cut – 2014 Data

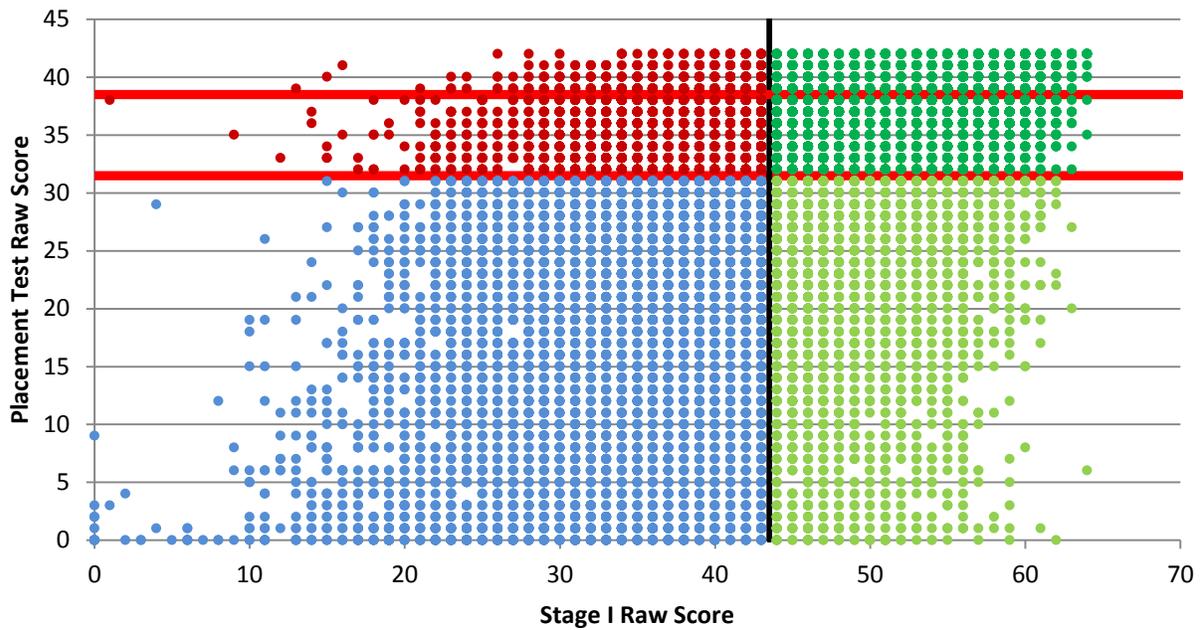


Figure D3. Proficient Cut 2014 Scatterplot. Scatterplot of kindergarten scores on the Placement Test in the fall of school year 2013-2014 and on the Spring 2014 AZELLA Stage I Test with Proficient cuts noted.

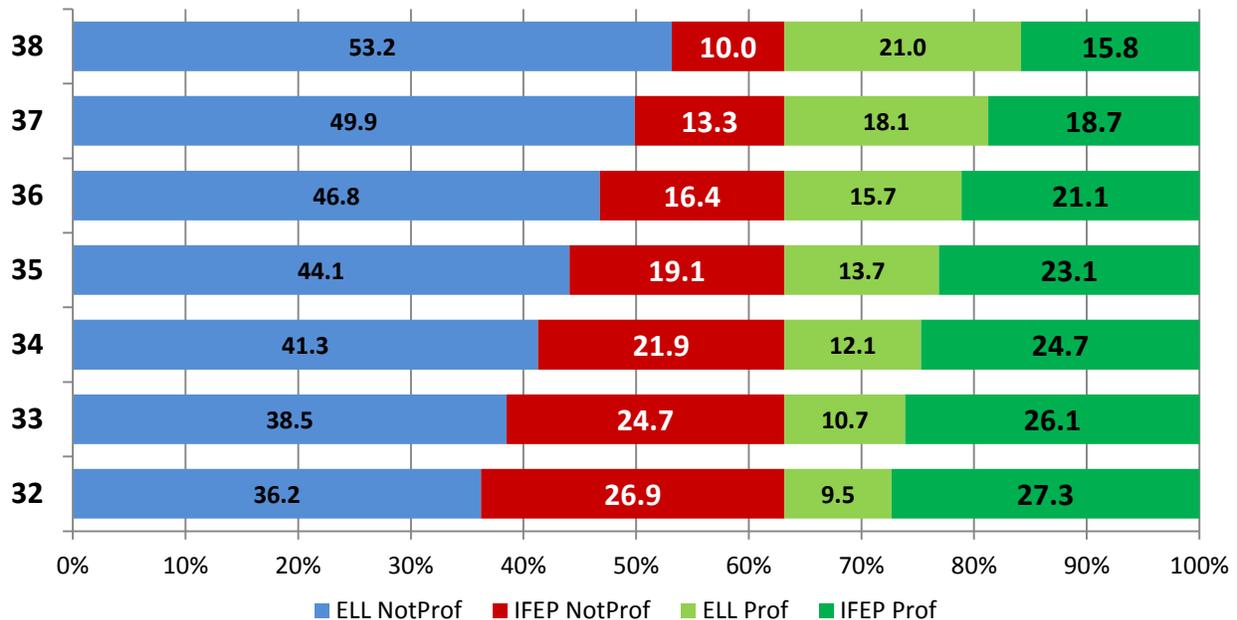


Figure D4. Proficient Cut 2014 Bar Graph. Percentage of kindergarten students in each performance level on the Spring 2014 AZELLA Stage I Test by their performance on their Placement Test in the fall of school year 2013-2014 for selected Placement Test raw score points.

Basic Cut – 2013 Data

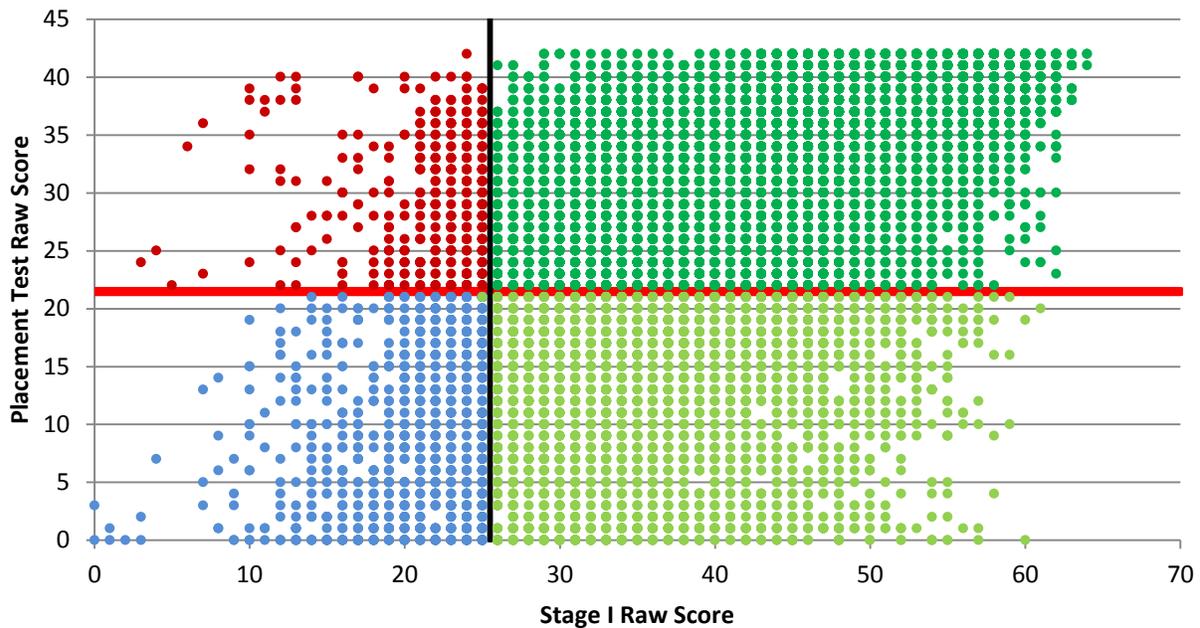


Figure D5. Basic Cut 2013 Scatterplot. Scatterplot of kindergarten scores on the Placement Test in the fall of school year 2012-2013 and on the Spring 2013 AZELLA Stage I Test with *Basic/Intermediate* and *Basic* cuts, respectively, noted.

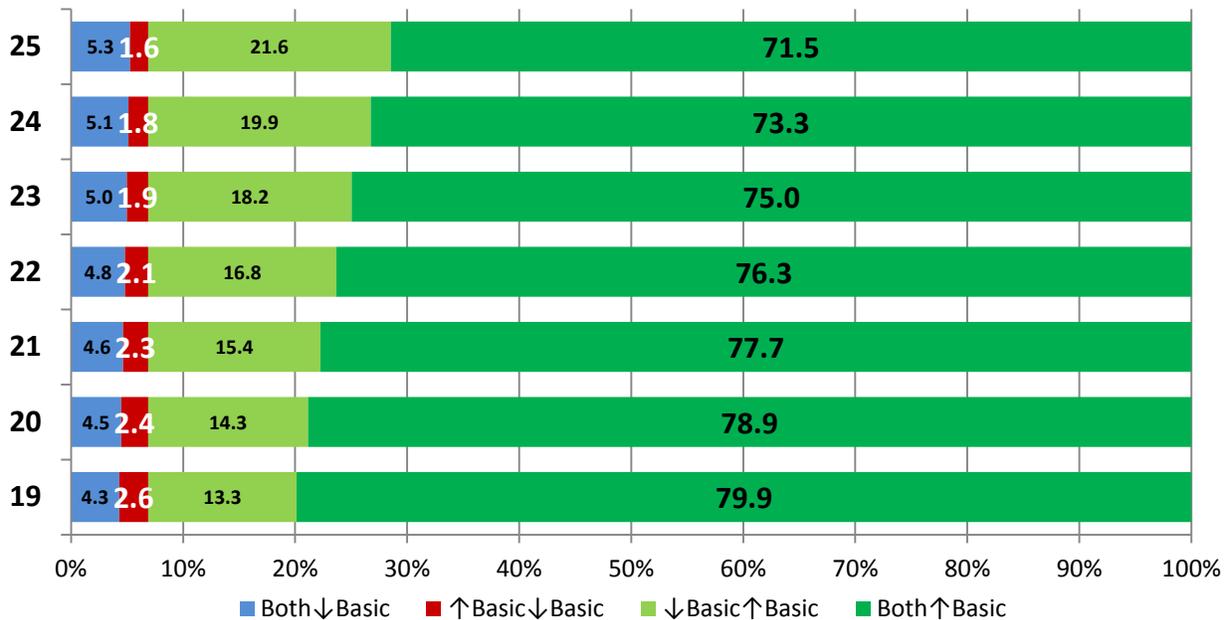


Figure D6. Basic Cut 2013 Bar Graph. Percentage of kindergarten students in each performance level on the Spring 2013 AZELLA Stage I Test by their performance on their Placement Test in the fall of school year 2012-2013 for selected Placement Test raw score points.

Basic Cut – 2014 Data

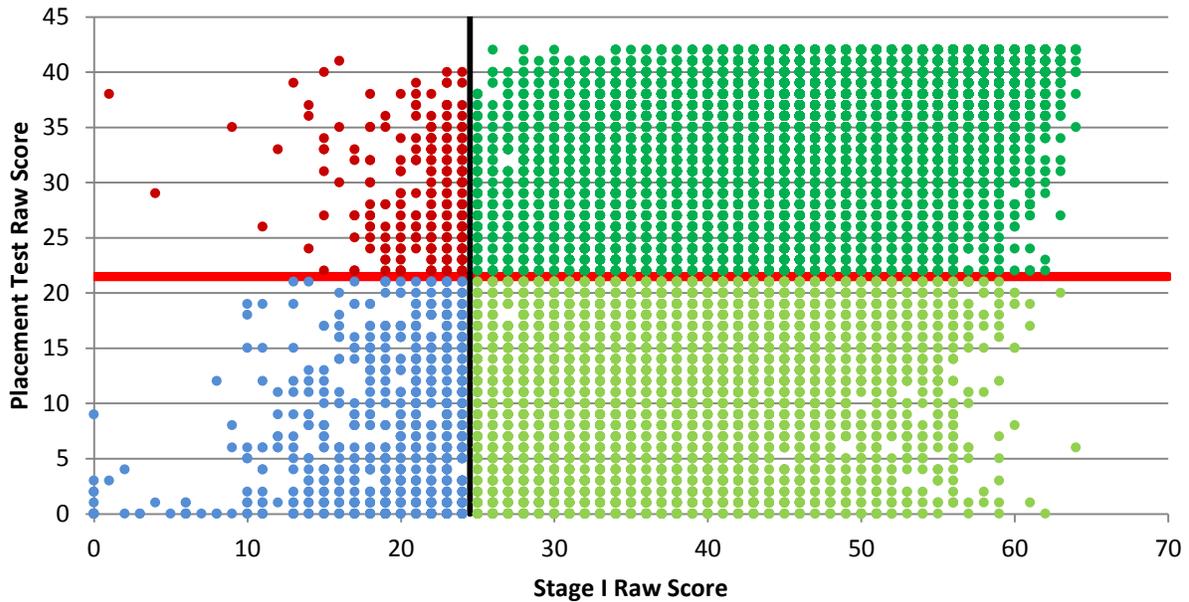


Figure D7. Basic Cut 2014 Scatterplot. Scatterplot of kindergarten scores on the Placement Test in the fall of school year 2013-2014 and on the Spring 2014 AZELLA Stage I Test with *Basic/Intermediate* and *Basic* cuts, respectively, noted.

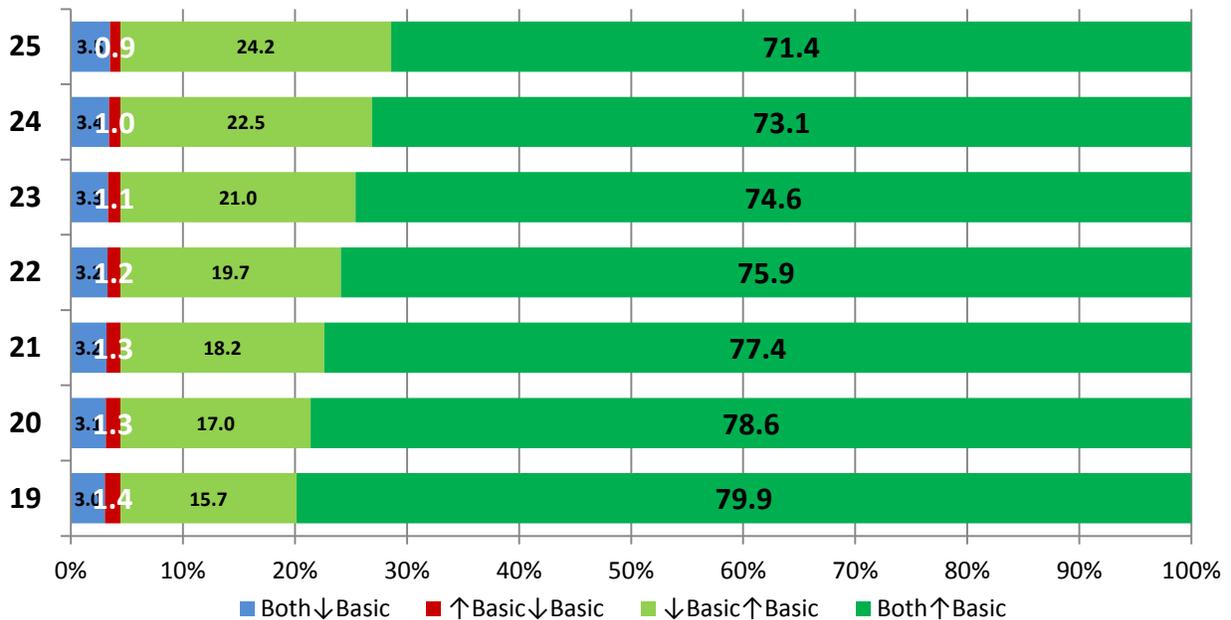


Figure D8. Basic Cut 2013 Bar Graph . Percentage of kindergarten students in each performance level on the Spring 2014 AZELLA Stage I Test by their performance on their Placement Test in the fall of school year 2013-2014 for selected Placement Test raw score points.

Appendix E

Page Number Recording Sheet



**ARIZONA ENGLISH LANGUAGE LEARNER ASSESSMENT
CUT-SCORE REVIEW
PAGE NUMBER RECORDING SHEET
KINDERGARTEN PLACEMENT TEST**



Panelist Number _____

Please enter the page number that corresponds to your bookmark for the Proficient Cut.

Proficient Cut Page Number

Round 1 _____

Round 2 _____

Please enter the page number that corresponds to your bookmark for the Basic Cut.

Basic Cut Page Number

Round 1 _____

Round 2 _____

Appendix F

Standard Setting Evaluation Forms with Responses



Arizona English Language Learner Assessment (AZELLA)
STANDARD SETTING
DECISION MAKING FACTOR SURVEY
KINDERGARTEN PLACEMENT TEST



Directions: Please respond to each statement by placing an “X” in the box corresponding to your opinion.

How much did each of the following factors influence your decisions on the cut score recommendations for the Arizona English Language Learner Assessment Kindergarten Placement Test?		Not at All	Somewhat	Moderately	Strongly	Very Strongly	No Response
1	Your experience in education	0	0	0	2	6	0
2	Prior to this standard setting meeting, your perceptions about students in each of the three performance levels	0	0	2	3	3	0
3	Your prior knowledge about current placement cut scores	0	1	3	3	1	0
4	The orientation on the Kindergarten Placement Test	0	1	1	2	4	0
5	Your perception of the high stakes versus low stakes context of the Kindergarten Placement Test	0	1	1	4	2	0
6	Your thinking about students in each performance level with whom you have had experience	0	0	0	3	5	0
7	Your understanding of the performance level descriptors	0	0	1	3	4	0
8	Frequency of page numbers presented after Round 1	0	0	3	2	3	0
9	Median page numbers presented after Round 1	0	1	1	3	3	0
10	The impact data presented	0	0	1	3	4	0
11	Your interactions with your fellow panelists after Round 1	0	1	3	2	2	0

Directions: Please respond to each statement by placing an “X” next to the category that best describes your school.

1. In general, my school/educational institution **mostly** serves students in the following socioeconomic status (choose one):

 5 Lower 2 Lower/Middle 1 Middle 0 Upper Middle 0 Upper

2. My educational institution is a charter school (choose one): 0 Yes 8 No



ARIZONA ENGLISH LANGUAGE LERNER ASSESSMENT (AZELLA)
STANDARD SETTING
FINAL STANDARD SETTING EVALUATION FORM
KINDERGARTEN PLACEMENT TEST



Directions: Please respond to each statement by placing an “X” in the box corresponding to your opinion. If you have any additional comments, please write them in the space provided at the end of this form.

NOTE: SD=Strongly Disagree; D=Disagree; A=Agree; SA=Strongly Agree; NR = No Response

	Statement	SD	D	A	SA	NR
1	The workshop was well organized.	0	0	0	8	0
2	The training materials were helpful.	0	0	0	8	0
3	I had a good understanding of what the test was intended to measure.	0	0	0	8	0
4	I had a good understanding of Performance Level Descriptors.	0	0	0	8	0
5	After the first round of ratings, I felt comfortable with the standard setting procedure.	0	0	0	8	0
6	I found the feedback on the frequency of ratings after Round 1 useful.	0	0	0	8	0
7	I found the feedback on median cut scores after Round 1 useful.	0	0	1	7	0
8	I found the feedback on the percentage of the students tested that would be classified at each performance level (Impact Data) useful.	0	0	0	8	0
9	Discussion after each Round was open and honest.	0	0	0	8	0
10	I believe that my opinions were considered and valued by my group.	0	0	0	8	0
11	I am confident that my Round 2 ratings for “ <i>Proficient</i> ” reflect the knowledge, skills, and abilities described in the performance level descriptors.	0	0	0	8	0
12	I am confident that my Round 2 ratings for “ <i>Basic/Intermediate</i> ” reflect the knowledge, skills, and abilities described in the performance level descriptors.	0	0	0	8	0
13	I would defend the standards recommended by our committee.	0	0	0	8	0
14	Overall, I valued the workshop as a professional development experience.	0	0	0	8	0

Please feel free to add comments on any of your responses above, make suggestions to improve future standard settings, and/ or tell us what you liked and did not like about this workshop.
 Thank you.

The Comments Made on the Evaluation Form by the Panelists

I enjoyed the experience and knowledge I gained today and do not have any other suggestions to improve any part of the day.
You both did an amazing job of delivering information in an understandable (way) and made it relevant and useable to the task at hand. Thank you.
Great day! Well-organized, information was presented in a meaningful way, both ladies were very pleasant.
It is so nice to feel my opinion is valued at the State level. Thank you for inviting me. I truly appreciate being a part of this process.
Such a wonderful experience. Thanks for everything and it was a true pleasure to add input and a learning experience that I feel confident about. 😊
This experience was a total eye opener! I enjoyed the process and having an opinion towards raising the bar for our students.
Thanks!