Concurrent External Validity Study
of
Arizona's
AZELLA Kindergarten Placement Test

October 2013

Addendum

August 2015

Lietta Scott

Arizona Department of Education – Assessment Unit

Concurrent External Validity Study of Arizona's
AZELLA Kindergarten Placement Test
October 2013

       This concurrent external validity study of the *AZELLA Kindergarten Placement Test* was undertaken by the Arizona Department of Education (ADE) in response to the requirements in paragraph 24(e)(ii) of the Resolution Agreement among the ADE, the United States Department of Education's Office of Civil Rights (OCR) in Denver, and the United States Department of Justice's Civil Rights Division (DOJ) relating to OCR Case Number 08-06-4006 and DOJ Case Number 169-8-81 of August 31, 2012. It used the *preLAS* (CTB/McGraw-Hill, LLC), an off-the-shelf, pre-literacy, English proficiency screener test for students entering kindergarten, to compare student results to that of the newly developed *AZELLA Kindergarten Placement Test*. Both tests were each given during a six week window during July and August of 2013 to a sample of kindergarten students entering schools from around Arizona for the first time. The sample was limited to students who had a non-English response to any of the three Primary Home Language Other Than English (PHLOTE) survey questions. The goal of the study was to find out how well the *AZELLA Kindergarten Placement Test* correlated with a nationally accepted, previously validated, assessment.

       Currently, there is no globally or nationally accepted definition of language proficiency. Arizona, using the well-established, Modified Angoff standard setting procedure (Angoff, 1971; Plake & Cizek, 2012) set the cut-score for the *AZELLA Kindergarten Placement Test* during the summer of 2012. The standard setting panel, which consisted of 13 educators from around the state, made their decisions following Arizona Administrative Code (R7-2-306 G.1) which specifies that a student who is not ELL "has the English language skills necessary to succeed in the English language curricula." A complete description of this standard setting meeting, including a description of panelist qualifications as well as intermediate and final cut-score determinations for the whole group and for each of the panelists, is available online (ADE, 2013).

       The *preLAS* currently has five performance levels, the two highest of which are used to determine students who have sufficient English language skills for placement in a mainstream classroom. It was developed to assess the oral language proficiency of young children (ages 4 to 6) from homes where the first language spoken is not English. The *preLAS* was re-normed in 2000 using a sample of 963 students at nine sites from around the nation. These students spoke 25 different languages of which Spanish was the most common. Included in the sample were 251 students who only spoke English. In their re-norming process, CTB/McGraw-Hill defined language proficiency as students who have the "linguistic elements necessary for successful communication within the school environment" (De Avila, & Duncan, 2000, p. 2). These authors found that age, grade, and home language all played significant roles in "determining the English aptitude, [of both] oral and Pre-Literacy skills, of the students tested" (p. 11).

While the *AZELLA Kindergarten Placement Test* and the *preLAS* definitions of language proficiency are very similar in context, the cut-scores associated with each were developed in very different ways with what might be found to be very different results. Arizona chose to depend on a proven psychometric process informed by the expertise of state educators where CTB/McGraw-Hill set the cut-scores for the *preLAS* using only statistical methods. These differences might be observed in a difference in percentage of students scoring proficient on each of the tests. The goal of the *AZELLA Kindergarten Placement Test* is to appropriately select students in need of the additional support of an English language learner kindergarten classroom. In the interest of providing policy makers additional information upon which to determine future decisions for the *AZELLA Kindergarten Placement Test*, decision consistency analysis was performed to find the point(s) at which the *AZELLA Kindergarten Placement Test* cut-score maximizes the percentage of students who passed both tests while minimizing the percentage of students who scored less than proficient on both.

Background

In school year 2012-2013, ADE administered the newly developed *AZELLA Kindergarten Placement Test*, an English Language Proficiency Screener for entering kindergartners. The *AZELLA Kindergarten Placement Test* was administered to all incoming kindergarten students with any non-English response to the three question Primary Home Language Other Than English (PHLOTE) survey. This assessment was developed as part of a new system of assessments, the Arizona English Language Learner Assessment (AZELLA) aligned to Arizona's English Language Proficiency Standards. Arizona's English Language Proficiency Standards are designed to prepare English language learners for the instruction required by Arizona's College and Career Ready Standards in the mainstream classroom after reclassification.

The *AZELLA Kindergarten Placement Test* is a pre-literacy test aligned to the Pre-Emergent, Emergent, and Basic levels of the Stage I English Language Proficiency Standards focusing primarily on receptive and productive oral language skills. The *AZELLA Kindergarten Placement Test* is administered orally in a one-to-one situation by test administrators who have completed an ADE-developed training and have passed a qualifying exam. The test administration typically takes less than 20 minutes, contains 38 questions with a maximum of 42 points. Using established rubrics, the test administrator scores each placement test item while administering the test. The item scores and appropriate student demographic data are then entered into an online system hosted by the test vendor. The overall test results are available immediately upon completing the data entry. *AZELLA Kindergarten Placement Test* results fall into three performance levels: Pre-Emergent/Emergent; Basic/Intermediate; Proficient. Students who score in one of the levels below Proficient are placed into an English Language Learner program.

Based on the documentation supplied to Arizona's Request for Proposal, the *preLAS* was selected by a committee from among several off-the-shelf proficiency screener assessments as the one that 1) was most similar in administration and 2) contained items that were most similar to the *AZELLA Kindergarten Placement Test* while presenting strong evidence of validity and reliability for the intended target population. The *preLAS* is an "early childhood assessment of English language proficiency and pre-literacy skills and is part of the *LAS Links* product family. The assessment uses graphics and stories based on early childhood literature and kindergarten readiness skills. The assessment takes about 10 minutes to administer and is composed of game-like tests that address general and specific features of a child's language proficiency" (Haley, 2013).

The sample plan called for approximately 1200 students from around the state to be assessed with each test as closely in time as possible with the limitation that state law requires that all new PHLOTE students must be assessed for language service need within thirty days of their first day in school. Using a stratified sample plan of PHLOTE students which considered linguistic backgrounds (Spanish, Native American language, and other languages), socio-economic status (ranging from 0% to 100% free and/or reduced lunch students), school type (regular district, charter, traditional, and magnet), as well as proportionally by county[1], ADE selected 32 schools. These were selected based on the number and demographics of students assessed during July and August of 2012. The number of schools and students in the planned sample as well as the corresponding number for all schools with PHLOTE students in each county are presented in Table 1.

Administration and data collection

ADE worked with each sample school to schedule the days within the six week window that the maximum number of students would be available to be assessed. Considerations included a desire that about one-half of the students be assessed first with each test (the *preLAS* and the *AZELLA Kindergarten Placement Test*), and the school's intended Placement testing dates. A survey of the 32 schools resulted in dates ranging from July 22, 2013 through August 30, 2013, with some schools indicating that they were planning on testing students for the whole six weeks. Since it was only practical for ADE staff to assess up to 20 students per day with the *preLAS* and to limit the travel time especially for those schools away from the Phoenix Metro area, schools were assigned specific ADE *preLAS* days. These assigned days were determined in conjunction with and agreed upon by the school's administration and fell close to or in the middle of the school's intended *AZELLA Kindergarten Placement Test* assessment window.

From July 22, 2013 through August 30, 2013, on the school's assigned day(s) ADE employees trained in the administration of the *preLAS* gave the test to all PHLOTE kindergarten students available. Thirteen ADE employees, normally in pairs, traveled as far west as Lake Havasu City, as far south-west as Gadsden, as far south-east as Willcox, and as far north-east as Kayenta to administer the *preLAS*. These cities are 199, 200, 197, and 293 miles from Phoenix,

respectively. Generally, each student's *AZELLA Kindergarten Placement Test* was given to the student by the school's regular prequalified assessment administrator. The one exception was in Tuba City where their prequalified administrator had recently left the school. In this one instance one ADE staff member administered the *AZELLA Kindergarten Placement Test* and another administered the *preLAS* to the eight kindergarten PHLOTE students.

Table 1. Number of schools and PHLOTE students in the planned sample and in each county.

County	Schools in Sample	Students in Sample	Total Schools	Total Students
Apache			5	30
Cochise	1	29	15	293
Coconino	1	16	12	140
Gila			2	18
La Paz			3	29
Maricopa	17	861	483	12,261
Mohave	1	13	19	154
Navajo	1	39	12	121
Pima	6	143	126	2,038
Pinal	2	35	36	435
Santa Cruz	1	34	11	538
Yavapai	1	27	21	251
Yuma	1	69	49	1,276
Total	**32**	**1266**	**778**	**17,333**

Note: Graham and Greenlee counties were excluded from consideration for the sample because no kindergarten PHLOTE students in these counties were assessed during July and August of 2012. These counties combined only had two kindergarten students registered in English language services during the whole school year.

Student name, date of birth, age, gender, school, and student identifier were captured along with their responses to the *preLAS* questions on the supplied answer documents which were then returned to the test vendor for scoring and reporting. Scores for the *preLAS* were reported for 1,008 students. Students who were repeating kindergarten in the 2013-2014 school year and those for whom no *AZELLA Kindergarten Placement Test* was submitted were eliminated from the study. Of the 988 students with both a *preLAS* and *AZELLA Kindergarten Placement Test* score, most were age 5 (914), followed by age 4 (69), with very few age 6 (5). A slightly higher number of female students (546, 55.3%) were assessed than male (442, 44.7%). Table 2 presents the students' race and ethnic demographics based on data submitted by the schools along with the students' *AZELLA Kindergarten Placement Test* responses.

Table 2. Student demographics from the *AZELLA Kindergarten Placement Test*.

	Asian	Black	Hawaiian/ Pacific Islander	Native American	White	Total
Hispanic	1	5	1	5	545	837
Non-Hispanic	67	11	2	39	25	151

Note: The Hispanic and Non-Hispanic totals do not equal the sum of the races since the choice not to respond or to respond affirmatively to more than one race is allowed.

Statistical Methodology

To examine how the two tests function and to determine the degree of agreement between the determinations of proficiency by the two tests, both correlational and decision analyses were performed. The rationale for the appropriateness of the use of only raw scores for these analyses is explicated below.

Using the student responses from *AZELLA Kindergarten Placement Tests* given during July and August of 2012, a principle axis factoring[2] with Varimax rotation was performed on the covariance matrix using all students in the State for which valid overall proficiency levels (OPL) had been assigned ($N = 17748$). The unrotated analysis produced one main factor accounting for 44.80% of the observed variance. In addition, three minor factors (each accounting for less than 7% of the observed variance) were identified. When rotation analysis on the four factors was performed, the analysis failed to converge. However, when rotation analysis for three factors was performed, convergence was achieved (60 iterations were required). The rescaled rotated factor matrix for the resultant three factors is presented in Table 3. This table reveals that all items weigh most heavily on the first factor indicating that the *AZELLA Kindergarten Placement Test* is essentially a unidimensional assessment. Since none of the items weigh most heavily on either of the minor factors, Pearson product-moment correlation was performed only between the total raw scores for the two tests.

Additionally, since the dichotomous proficiency determination of the *AZELLA Kindergarten Placement Test* is of particular import, decision consistency analysis was performed. It was based on the work presented by Tom Fawcett (2006). In this analysis, the number of students at each raw score point on the *AZELLA Kindergarten Placement Test* are tabulated and classified by whether or not they achieved one of the two proficient levels on their preLAS test. Based on this information the number of True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) cases at that raw score point are computed. These values are then combined into an F-Score which when maximized identifies the *AZELLA Kindergarten Placement Test* raw score point that (based on this evaluative score for these two

Table 3. Fall 2012 *AZELLA Kindergarten Placement Test* rotated factor matrix.

Item	Factor		
	1	2	3
QS1	.381	.267	.022
QS2	.661	-.051	-.014
QS3	.716	.134	-.063
QS4	.639	.044	.026
QS5	.694	.105	-.055
QS6	.665	-.130	.029
QS7	.680	.192	-.210
QS8	.681	.262	-.234
QS9	.774	-.016	-.239
QS10	.789	-.025	-.229
QS11	.595	.192	-.028
QS12	.721	.210	-.094
QS13	.698	.209	-.083
QS14	.661	.338	-.081
QS15	.711	.103	-.060
QS16	.685	.052	-.108
QS17	.636	.135	-.101
QS18	.508	.050	-.029
QS19	.457	-.090	.026
QS20	.502	.318	.226
QS21	.506	.367	.128
QS22	.484	.386	.204
QS23	.639	.068	.110
QS24	.604	.080	.141
QS25	.354	-.135	.155
QS26	.557	.229	.226
QS27	.538	.319	.226
QS28	.527	.224	.215
QS29	.624	.044	.192
QS30	.608	-.002	.267
QS31	.576	-.014	.261
QS32	.570	-.010	.234
QS33	.373	-.008	.058
QS34	.411	-.031	.044
QS35	.344	-.065	.034
QS36	.503	-.164	.041
QS37	.596	-.111	-.023
QS38	.767	-.203	.063

tests) maximizes the True Positive and True Negative cases while minimizing the False Positive and False Negative cases. Fawcett formulated F-Score as:

$$\text{F-Score} = \frac{2}{\frac{1}{precision} + \frac{1}{recall}}$$

where precision is equal to the True Positive/(True Positive + False Positive) and recall is True Positive/(True Positive + False Negative). In this analysis, the raw cut scores at and around the established *AZELLA Kindergarten Placement Test* raw cut score of 32 out of a possible 42 points on the test is of particular interest.

Results

      The correlation coefficient was computed between the total raw scores of the *preLAS* and the *AZELLA Kindergarten Placement Test* using the 988 students having both tests. This coefficient was .861, which is significant at the $p < .001$ level. Figure 1 presents the scatterplot of the students' raw scores on the *preLAS* versus the *AZELLA Kindergarten Placement Test* which confirms the relatively linear relationship found via the correlation analysis.
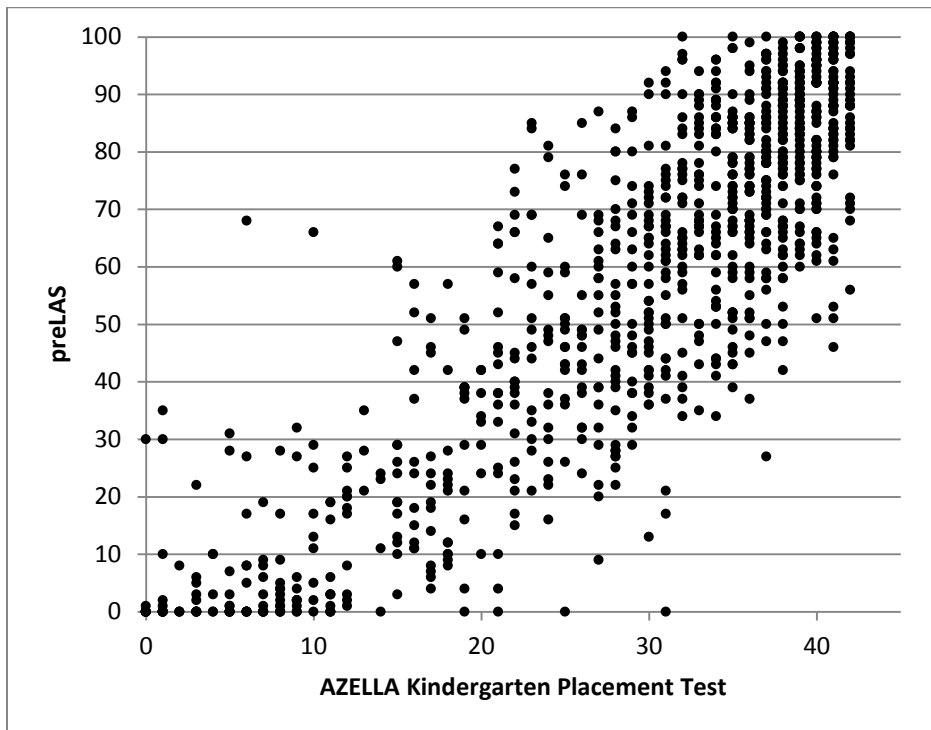

Figure 1. Raw Scores for the *preLAS* and *AZELLA Kindergarten Placement Test*.

      Table 4 presents the F-Score for each *AZELLA Kindergarten Placement Test* raw score point as well as the number of students at each of those raw score points and how many scored non-proficient or proficient on their *preLAS* assessment. Of particular interest are the F-Scores at and around the *AZELLA Kindergarten Placement Test* cut score of 32. This analysis indicates that the F-Score at the *AZELLA Kindergarten Placement Test* proficient cut score is .661 where

Table 4. Decision Consistency F-Score and Data

Raw Score	Number of Students	Did Not Pass *preLAS*	Passed *preLAS*	F-Score
0	11	11	0	0.444
1	12	12	0	0.448
2	4	4	0	0.452
3	10	10	0	0.454
4	5	5	0	0.457
5	11	11	0	0.459
6	11	11	0	0.463
7	11	11	0	0.468
8	14	14	0	0.472
9	10	10	0	0.478
10	9	9	0	0.482
11	11	11	0	0.485
12	11	11	0	0.490
13	3	3	0	0.495
14	5	5	0	0.496
15	14	14	0	0.498
16	11	11	0	0.504
17	13	13	0	0.509
18	13	13	0	0.516
19	12	12	0	0.522
20	7	7	0	0.528
21	17	17	0	0.531
22	19	19	0	0.540
23	16	14	2	0.550
24	16	16	0	0.554
25	17	17	0	0.563
26	17	16	1	0.573
27	21	20	1	0.581
28	29	28	1	0.592
29	22	20	2	0.609
30	36	34	2	0.619
31	35	32	3	0.641
32	**33**	**25**	**8**	**0.661**
33	34	24	10	0.668
34	39	21	18	0.672
35	50	35	15	0.658
36	49	32	17	0.663

37	58	35	23	0.660
38	63	32	31	0.646
39	62	20	42	0.603
40	60	26	34	0.494
41	54	10	44	0.390
42	33	5	28	0.178
Totals	**988**	**706**	**282**	

Note: The *AZELLA Kindergarten Placement Test* proficient cut-score for school years 2012-2013 and 2013-2014 is set at 32 raw score points.

the maximum F-Score is slightly higher at .672 at raw score point 34. Figure 2 displays this information graphically. As can be seen within this graph there is very little difference between the F-Score at the current cut-score of 32 (indicated by the vertical red line) and that at its maximum and there appears to be a leveling off or plateauing of F-Scores between the raw scores of 32 and 37.
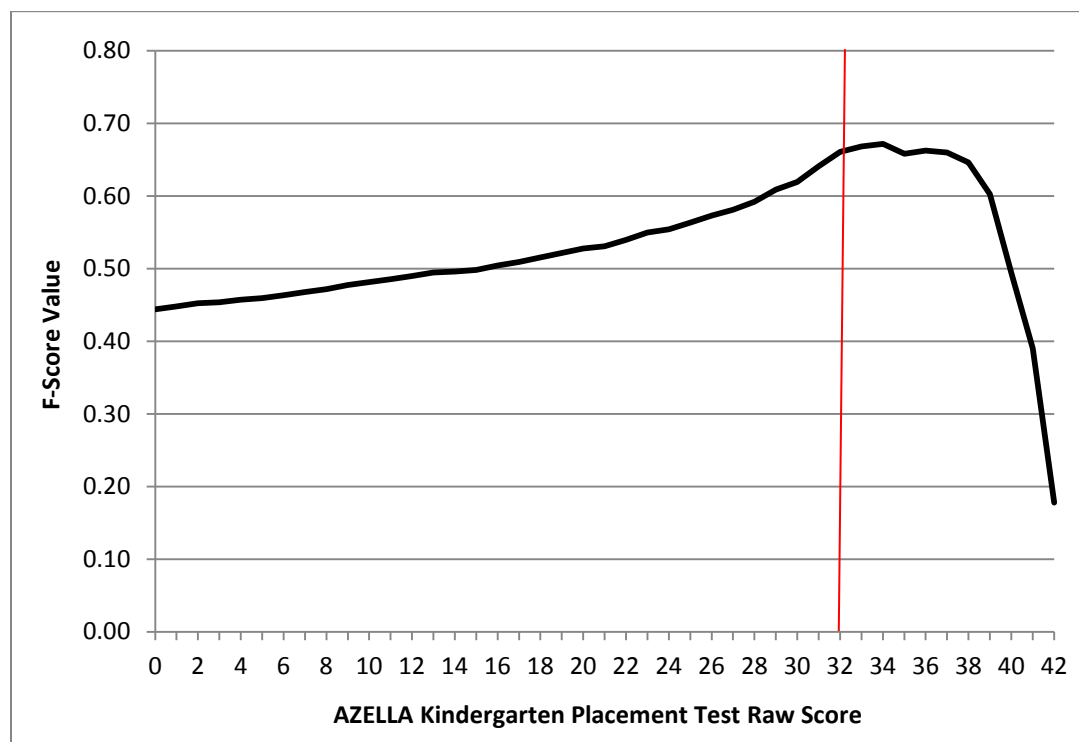

Figure 2. F-Scores at each of the *AZELLA Kindergarten Placement Test* raw scores.

Discussion and Conclusion

As with all policy decisions, multiple data points and points of view must be considered and in many situations there generally are multiple "right" answers. This study presents just one view of the *AZELLA Kindergarten Placement Test*, that which compares its results to that of the *preLAS*. Based on both the correlation and decision consistency analyses it is quite comparable.

10

The significant correlation coefficient of .861 was well over the industry standard of .80. Additionally, while the F-Score at the cut score from the decision consistency analysis of .661 was slightly lower than the maximum of .672, it could be considered to be within the plateau which is evident from a raw score of 32 to a raw score of 37. The minor differences in F-Score across this plateau might simply be due to sampling. While this study provides some indication that a change in the *AZELLA Kindergarten Placement Test* proficient cut-score *could* be made (to anywhere in the 32 to 37 raw point range), given the general consistency of results from these two assessments, based solely on this study, indications for the *need* for changes in either the *AZELLA Kindergarten Placement Test* or the test cut-score were not found.

Addendum to
Concurrent External Validity Study of Arizona's
AZELLA Kindergarten Placement Test
October 2013

*Analysis of Predictive Value of
the preLAS and the AZELLA Kindergarten Placement Test
Performance Levels on Student Performance
on the Spring 2014 AZELLA Stage I Reassessment.
August 24, 2015*

This addendum adds the examination of the *pre*LAS's and AZELLA Kindergarten Placement Test's ability to predict student performance on the Spring 2014 AZELLA Stage I Reassessment Test (Spring Test) to the original Concurrent External Validity Study between the two tests. The results of all students who had a score on the both of the placement tests (*N*=988) were matched by State student identifier number to their score on the Spring Test. Of the original sample, 941 (95.2%) were matched. The 4.8% that were found to not have a score on the Spring Test were split approximately equally across the three AZELLA Placement Test performance levels.

A series of regression analyses were conducted to predict the student performance level on the Spring test from the performance levels on the placement tests. For the AZELLA Placement Test and the Spring Test, the student's final proficiency level was numerically coded from 1 for the lowest level to the highest level (Proficient). This resulted in performance values of 1, 2, and 3 for the AZELLA Placement Test and 1 through 4 for the Spring Test. Since the *pre*LAS had two proficient levels, they were combined for this study as level 4. This resulted in the *pre*LAS also having performance values of 1 through 4.

The first multiple regression analysis evaluated how well the two tests together predicted performance on the Spring Test. The linear combination of the placement tests was significantly related to Spring Test performance level, $F(2, 938)=170.81$, $p<.001$. The sample multiple correlation coefficient was .52, indicating that approximately 27% of the variance of the Spring Test performance level in the sample can be accounted for by the linear combination of placement test performance levels. The use of both placement tests to predict student performance on the Spring Test had an effect size of .36. Both the AZELLA Placement Test and *pre*LAS performance levels displayed a significant ($p<.001$) correlation with the Spring Test performance level (.50 and .43, respectively). These correlations indicate the relative strength of the two placement tests as predictors. They both also have significant ($p<.001$) correlations with the Spring Test performance level after controlling for the performance level on the other placement test (.32 and .14, respectively).

The second set of regression analyses evaluated how well each of the two placement tests, individually, predicted performance on the Spring Test. The regression equation for predicting the Spring Test performance level from the *pre*LAS performance level is

Spring Test Performance = .279 preLAS Performance + 2.53

The 95% confidence interval for the slope, .241 to .317, does not contain the value of zero, and therefore *pre*LAS performance is significantly related to Spring Test performance. The standardized slope for predicting Spring Test performance is .426. The use of the *pre*LAS alone accounts for approximately 18% of the variance of the Spring Test performance level. The effect size for *pre*LAS, singly, predicting student performance on the Spring Test is .22.

The regression equation for predicting the Spring Test performance level from the AZELLA Placement Test performance level is

Spring Test Performance = .517 AZELLA Placement Performance + 1.96

The 95% confidence interval for the slope, .460 to .574, does not contain the value of zero, and therefore AZELLA Placement Test performance is also significantly related to Spring Test performance. The standardized slope for predicting Spring Test performance is .503. The use of the AZELLA Placement Test alone accounts for approximately 25% of the variance of the Spring Test performance level. The effect size for the AZELLA Placement Test, singly, predicting student performance on the Spring Test is .34.

The third set of regression analyses evaluated how well each of the two placement tests predicted performance on the Spring Test over and above the other. The AZELLA Placement Test performance level accounted for a significant proportion of the variance in Spring Test performance after controlling for the effects of the *pre*LAS performance level, R^2 change = .09, $F(1, 938) = 109.08$, $p < .001$. The *pre*LAS performance level also accounted for a significant proportion of the variance in Spring Test performance after controlling for the effects of the AZELLA Placement Test performance level, R^2 change = .01, $F(1, 938) = 17.49$, $p < .001$.

While all of these analyses should be used to inform the use of placement tests for incoming PHLOTE Kindergarten students within Arizona, the third set is perhaps the most informative. In these analyses, while both the AZELLA Placement Test and the *pre*LAS Placement Test significantly account for variance over and above the other, the amount of change observed with the *pre*LAS (.01) is much lower than that with the AZELLA Placement Test (.09). This indicates that the AZELLA Placement Test accounts for more of the variance in the Spring Test performance level than the *pre*LAS Placement Test and that the *pre*LAS adds very little predictive value over and above that observed by the AZELLA Placement Test.

Another way to look at this data is to examine the percentages of students who scored at each of the performance levels on the two placement test as compared to the percentage of

students at each performance level on the Spring Test. Table A1 presents the number and percentage of students at each of the *pre*LAS performance levels by Spring Test performance level. As can be seen in this table, approximately 63% of the students who scored *Proficient* on the *pre*LAS also scored Proficient on the Spring Test.

Table A2 presents the number and percentage of students at each of the AZELLA Placement Test performance levels by Spring Test performance level. As can be seen in this table, approximately 58% of the students who scored *Proficient* on the AZELLA Placement Test also scored Proficient on the Spring Test. This percentage is similar to that found when the same students took the *pre*LAS.

For the AZELLA Placement Test, however, there have been changes to both the *Basic/Intermediate* and the *Proficient* cut scores that will be implemented starting in school year 2014-2015. Based on an Evidenced Based Standard Setting Review held May 30, 2014, the AZELLA Placement Test *Basic/Intermediate* cut score will be set 1 raw score point lower and the *Proficient* cut score will be set 3 raw score point higher than previously. Because of this change, student scores on AZELLA Placement Test were recoded and the number and percentage of students at each of the new AZELLA Placement Test performance level by Spring Test performance level are presented in Table A3. As can be seen within this table, with the new performance level cut scores the percentage of students who score *Proficient* on both the AZELLA Placement Test and the Spring Test during school year 2014-2015 (62%) can be expected to be even more similar to that found with the *pre*LAS.

Table A1. Number (and Percentage) of Students at each *pre*LAS Performance Level by Spring Test Performance Level.

preLAS Proficiency	Stage I Overall Proficiency				Total
	Pre-Emergent/ Emergent	Basic	Intermediate	Proficient	
Beginning	40 (9)	107 (25)	190 (45)	88 (21)	425
Early Intermediate	4 (3)	16 (12)	53 (40)	58 (44)	131
Intermediate	0 (0)	7 (6)	52 (44)	59 (50)	118
Proficient (Levels 4-5)	0 (0)	9 (3)	90 (34)	168 (63)	267
Total	44 (5)	139 (15)	385 (41)	373 (40)	941

Note: Percentages are based on the number of students at each of the *pre*LAS proficiency levels.

Table A2. Number (and Percentage) of Students at each AZELLA Placement Test **Old** Performance Level by Spring Test Performance Level.

Placement Test Proficiency	Stage I Overall Proficiency				Total
	Pre-Emergent/ Emergent	Basic	Intermediate	Proficient	
Pre-Emergent/ Emergent/	32 (15)	70 (33)	89 (42)	23 (11)	214
Basic/ Intermediate	10 (5)	39 (18)	133 (52)	55 (25)	217
Proficient	2 (0)	30 (6)	183 (36)	295 (58)	510
Total	44 (5)	139 (15)	385 (41)	373 (40)	941

Note: Percentages are based on the number of students at each of the old AZELLA Placement Test proficiency levels.

Table A3. Number (and Percentage) of Students at each AZELLA Placement Test **New** Performance Level by Spring Test Performance Level.

Placement Test Proficiency	Stage I Overall Proficiency				Total
	Pre-Emergent/ Emergent	Basic	Intermediate	Proficient	
Pre-Emergent/ Emergent/	32 (15)	70 (33)	89 (42)	23 (11)	214
Basic/ Intermediate	11 (3)	49 (15)	164 (51)	96 (30)	320
Proficient	1 (0)	20 (5)	132 (32)	254 (62)	407
Total	44 (5)	139 (15)	385 (41)	373 (40)	941

Note: Percentages are based on the number of students at each of the new AZELLA Placement Test proficiency levels.

[1.] The three counties considered for the sample but not included each had a total of no more than 40 ELL Kindergarten students assessed by August 30, 2012 for the 2012-2013 school year.

[2.] Principal axis factoring was chosen for this analysis over principal components factoring because some research indicates that it is more sensitive to identifying minor factors especially when those factors are highly correlated (Crawford, Green, Levy, Lo, Scott, Svetina, & Thompson, 2010).

References

Arizona Department of Education (2013). *Arizona English Language Learner Assessment (AZELLA), standard setting report: Kindergarten Placement Test.* Phoenix, AZ: Arizona Department of Education. Available at: http://www.azed.gov/standards-development-assessment/files/2013/09/az-kpt-standard-setting-report-final_092413.pdf.

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 on Education.

Crawford, A. V., Green, S. B., Levy, R., Lo, W. J., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Measurement, 70*(6), 885-901.

De Avila, E. A., & Duncan, S. E. (2000). preLAS2000 English & Spanish, technical notes. Monterey, CA: CTB/McGraw-Hill.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*, 861-874.

Haley, E. (2013). Personal communication in response to Solicitation Number: ADED13-00002799. CTB/McGraw-Hill, LLC. Monterey, CA.

Plake, B. S., & Cizek, G. J. (2012). Variations on a theme: The modified Angoff, extended Angoff, and Yes/No standard setting methods. In G. J. Cizek (Ed.), *Setting performance standards: foundations, methods, and innovations.* (2nd ed., pp. 181-199). New York, NY: Routledge.