RESOLUTION AGREEMENT

Among the Arizona Department of Education,
The U.S. Department of Education's Office for Civil Rights (Denver), and
The U.S. Department of Justice's Civil Rights Division

OCR Case Number 08-06-4006
DOJ Case Number 169-8-81

BACKGROUND

1. The U.S. Department of Education, Office for Civil Rights (OCR) and the U.S. Department of Justice (DOJ) investigated a complaint filed against the Arizona Department of Education (ADE) regarding ADE's policies and procedures by which English Language Learner (ELL) students are reclassified as Fluent English Proficient (FEP) and, thereby, exited from the English language acquisition services program. The complainant alleges that thousands of ELL students are being reclassified as "proficient" in English when test results indicate that they are not in fact proficient in English. The complainant alleges that this is due to ADE's use of the Stanford English Language Proficiency (SELP) test and, later, the Arizona English Language Learner Assessment (AZELLA) (along with their respective scoring procedures), as the sole reclassification criterion for ELL students statewide.

   ADE acknowledges that OCR received the aforementioned complaint regarding the AZELLA tests but refutes the allegations in the complaint. Moreover, ADE's participation in this Agreement does not constitute an admission to the allegations of this complaint or any other complaint or matter being investigated by OCR or DOJ involving ADE.

2. Title VI of the Civil Rights Act of 1964, 42 U.S.C. § 2000d and its implementing regulation at 34 C.F.R. pt. 100, taken together, provide, in relevant part, that no program or activity receiving Federal financial assistance from the United States may discriminate on the basis of national origin. ADE's Title VI obligations are enforced by OCR and DOJ.

3. The Equal Educational Opportunities Act (EEOA), 20 U.S.C. § 1703(f), requires that state and local educational agencies take appropriate action to overcome language barriers that impede equal participation by students in their instructional programs. The EEOA is enforced by DOJ.

4. OCR and DOJ found that ADE is not in compliance with Title VI regarding the issues investigated in this matter, and DOJ found that ADE is not in compliance with the EEOA. Specifically, OCR and DOJ found that due to ADE's mandated use of the composite scores on the AZELLA Form AZ-1 in the 2006-07, 2007-08, and 2008-09 school years, and the AZELLA Form AZ-2 since the 2009-10 school year for identification and reclassification purposes, ELL students (a) were not and are not being identified; (b) were and are being exited from the ELL program prematurely without yet obtaining proficiency in each of the four language skill areas (oral, comprehension, reading, and writing); and (c) were reclassified but are unable to participate meaningfully in the regular education environment and need further English language supports. Based on OCR's and DOJ's review, OCR and DOJ found that ADE's reliance on only the AZELLA's composite proficient score as its ELL identification and exit criteria is not a valid measure of English language proficiency and readiness to exit ELL services.

ADE disagrees with OCR's and DOJ's findings of noncompliance regarding the AZELLA tests and affirmatively asserts that the tests are valid measures to identify and reclassify ELL students. However, ADE enters into this Agreement to voluntarily resolve this matter. ADE's participation in this Agreement does not constitute an admission to OCR's and DOJ's findings regarding the AZELLA tests, nor does it constitute an admission to OCR's and DOJ's findings or determinations regarding any other matter involving ADE.

5. The purpose of this Resolution Agreement in the short term is to implement an interim plan beginning in the 2012-13 school year to offer and provide intervention services to qualifying students. The purpose of this Resolution Agreement in the long term is to institute a long-term plan that will include the implementation of a new, valid, and reliable English language proficiency assessment beginning in the 2012-13 school year for the purposes of identifying and exiting/reclassifying ELL students.

JURISDICTION

6. ADE is subject to Title VI because it receives Federal financial assistance from the U.S. Department of Education.

7. ADE is subject to the EEOA because ADE is a state educational agency as defined by 20 U.S.C. § 1720.

8. The parties to this Agreement are OCR, DOJ, and ADE. In light of this Agreement, the parties have determined that OCR Case Number 08-06-4006 and DOJ's investigation of this complaint can be resolved through the terms of this Agreement.

9. OCR agrees to refrain from initiating an enforcement action and DOJ agrees to refrain from initiating litigation regarding areas covered in the Agreement except as expressly discussed in this Agreement.

REMEDIAL ACTION

10. OCR and DOJ have determined that the AZELLA, in both forms AZ-1 and AZ-2, have failed to identify students in need of English language acquisition services and prematurely reclassified students as English proficient. ADE does not concur with OCR's and DOJ's determination; however, ADE is willing to implement the following measures to voluntarily resolve this matter.

INTERIM PLAN

11. Identification: For students identified as Initially Fluent English Proficient (IFEP) by the AZELLA-1 (AZ-1) and AZELLA-2 (AZ-2) in SY 2007-08, SY 2008-09, SY 2009-10, SY 2010-11, and SY 2011-12, ADE agrees to issue a directive to Arizona local education agencies (LEAs) within five (5) days of receipt of approval from OCR and DOJ on the draft directive as provided in Paragraph 19 to review their IFEPs' most recent AZ-1 or AZ-2 reading and writing sub-scale scores and their most recent AIMS Reading and Writing scores,[1] and to do the following for IFEP students who lack a "proficient"[2] score in (i)

[1] If a student lacks an AIMS Reading or AIMS Writing score, follow paragraph 15.

[2] As used in the Interim Plan portion of this Agreement, "proficient" shall equate to: (1) a score of proficient in the AZ-1, AZ-2, or AZ-3 reading and writing sub-scale scores, and (2) meets the standard on AIMS Reading or AIMS Writing as approved by the Arizona State Board of Education.

both AZ-2 reading and AIMS Reading, and/or (ii) both AZ-2 writing and AIMS Writing before December 15, 2012. If a student was in grades K-11 at the time of the IFEP decision, the LEA will review additional data, including, but not limited to, student grades, other standardized tests, teacher recommendations, and parental input, to determine whether to offer the student's parent the option of enrolling their child in "intervention services" targeted and appropriate to the student's (a) AZ-1 or AZ-2 reading score and AIMS reading score, and/or (b) AZ-1 or AZ-2 writing score and AIMS writing score if the student lacks proficiency in either or both reading and writing. The "intervention services" shall consist of English language support services targeted to the student's needs in the domains of reading and/or writing by a teacher qualified to deliver such services. LEAs may offer these intervention services via: (a) placement in an SEI class focused on reading and/or writing for no more than two hours daily; (b) a service plan similar to an Individual Language Learner Plan where interventions are delivered in a mainstream content class for no more than two hours daily; (c) another class focused on reading and/or writing instruction; or (d) before or after school (i) as a supplemental intervention service to those provided during the school day, or (ii) as the only intervention service if the parent so elects in writing. The LEA may consider the student's grades in core content classes when determining the type of intervention services to offer.[3] If the parent elects for their child to receive "intervention services," the LEA will begin providing the requested intervention services.

12. For students identified as IFEP on AZ-2 in SY2012-13 who are in grades 2-12, the LEA will offer the student's parent the option of enrolling their child in "intervention services" (subject to the terms in paragraph 11 above) in the current or upcoming school year that is targeted and appropriate to the student's AZ-2 score in (a) reading and/or (b) writing, as applicable, if the student lacks proficient scores in either or both reading and writing.

13. If a student is not proficient in the AZ-2 reading and/or writing domains in grades K-1 for SY 2012-13, the LEA will make a team determination with the student's parent and school personnel familiar with the student and second language acquisition as to whether the LEA will offer the student's parent the option of enrolling their child in "intervention services" (subject to the terms in paragraph 11 above) that are targeted and appropriate to the student's reading and/or writing needs.

14. Reclassification: For students reclassified as Fluent English Proficient (FEP) in SY 2007-08, SY 2008-09, SY 2009-10, SY 2010-11, SY 2011-12, and SY 2012-13[4] who were exited from ELL services and are enrolled in mainstream classes in Arizona LEAs, ADE agrees to issue a directive to Arizona LEAs within five (5) days of receipt of approval from OCR and DOJ on the draft directive as provided in Paragraph 19 that they must do the following before December 15, 2012:

    a. review (as part of the FEP1 and FEP2 monitoring process for those reclassified in SY 2010-11 and SY 2011-12) the student's most recent AZ-1 or AZ-2 sub-test scores and AIMS scores to determine if the student was proficient in reading and writing, as appropriate to his or her grade level; and

[3] For example, a student with an intermediate 2010 AZ-2 reading score, a partially proficient AIMS 2011 reading score, and passing grades in all subjects but social studies may receive English language support services targeted to achieving reading proficiency in his social studies class. A student with the same scores who is struggling in all subjects may receive a period targeted exclusively on reading comprehension skills using core content materials.

[4] This includes students reclassified under AZ-2 in SY 2012-13 before AZ-3 becomes operational.

b. regarding only those students not proficient in: (i) both AZ-1 or AZ-2 reading and AIMS Reading or (ii) AZ-1 or AZ-2 writing and AIMS Writing, review additional data, including, but not limited to, student grades, other standardized tests, teacher recommendations, and parental input, to determine if the LEA should offer the parent the option to enroll their child in "intervention services" (subject to the terms in paragraph 11 above) targeted to the student's lack of proficiency in reading and/or writing if the student lacks proficiency in either or both reading and writing; and

c. if the parent elects for their child to receive "intervention services," the LEA will begin providing the requested intervention services.

15. If, for any of the above requirements in paragraphs 11-14, a student lacks an AIMS Reading or Writing score, the LEA will review the student's core content grades within the last calendar year to determine if other data requires review to determine if the student needs "intervention services" (subject to the terms of paragraph 11 above) targeted to their lack of proficiency in reading and/or writing.

16. If pursuant to paragraphs 11-15 above, a parent elects for their child to receive "intervention services," the LEA (with input from a teacher knowledgeable about second language acquisition) shall develop a plan for each such student that documents the "intervention services" (subject to the terms of paragraph 11 above) provided and whether they prove effective over time based on formative assessments and other performance data. The LEA shall administer to each such student the AIMS Reading and/or Writing test, as applicable, and provide intervention services until the student achieves proficiency in reading and/or writing (whichever the student lacks) on AIMS.

17. ADE will remind LEAs in writing that if a classroom teacher believes that a student is ELL, the student can be referred for an ELL assessment regardless of the above procedures in paragraphs 11-16.

18. ADE agrees to direct Arizona LEAs to report to ADE the number of students who are enrolled in intervention services pursuant to paragraphs 11-17, separately reporting the subset of those students whose parents elected to enroll them only in before- or after-school intervention services. The LEAs will report this information disaggregated by school and grade level, and will indicate whether the students were originally identified as IFEP or reclassified as FEP. This information will then be provided to OCR and DOJ by January 31, 2013 for SY 2012-13 and by November 1, 2013 for SY 2013-14.

19. Within ten (10) days of signing this Agreement, ADE will provide to OCR and DOJ a draft written directive(s) to Arizona LEAs as provided in Paragraphs 11 and 14. OCR and DOJ will promptly approve the draft directive(s), or recommend revisions to the language contained therein before ADE is required to issue the written directive(s) to Arizona LEAs pursuant to Paragraph 32. Within thirty (30) days of ADE sending the directive, ADE will incorporate information regarding the Interim Plan and the directive in to (a) ADE's Office of English Language Acquisition Services (OELAS) training to LEAs, and (b) OELAS ELL Monitoring process. ADE will provide OCR and DOJ with documentation of the incorporation of this information in to the OELAS training and monitoring processes within ninety (90) days of ADE sending the directive.

LONG TERM PLAN

20. By **July 31, 2013,** ADE will adopt new, valid, reliable, and objective criteria for LEAs to use to ensure that each ELL student has obtained English language proficiency in reading, writing, listening, and speaking and in the composite of comprehension (reading and listening) and the composite of oral (listening and speaking) before: (a) being identified as IFEP; (b) being reclassified as FEP or RFEP; or (c) being required to exit English language acquisition services. ADE will require that LEAs administer a valid and reliable English language proficiency (ELP) assessment according to the Standards for Educational and Psychological Testing (AERA, APA, and NCME, 1999) and the publisher's instructions as a required criterion for identifying ELL students, determining their English language proficiency, and exiting ELL students from English language acquisition services when they become proficient in English.

21. To implement paragraph 20, ADE has chosen to create and administer a new version of AZELLA referred to as the "AZ-3" by the start of SY 2013-14 and a Kindergarten ELP screener ("K-screener") by the fall of 2012. ADE will demonstrate that the AZ-3 and the K-screener: [5]

    a. yield valid and reliable information regarding a student's English proficiency in each language domain of reading, writing, speaking, and listening and composites of oral and comprehension, as further explained in paragraphs 22-26 below;

    b. are aligned to the ADE revised English language proficiency (ELP) standards that:
        i. correspond to the Common Core State Standards (CCSS), and to ensure such alignment, ADE will follow the recommendations in The *Application of the Common Core State Standards for English Language Learners* published at http://www.corestandards.org/assets/application-for-english-learners.pdf);
        ii. include performance levels and achievement-level descriptors that identify a "proficient" level of English in each domain of listening, speaking, reading, and writing across grade spans K, 1-2, 3-5, 6-8, and 9-12; and
        iii. equate a "proficient" composite score and "proficient" subscores for listening, speaking, reading, and writing on the AZ-3 and K-screener with "proficient" performance levels on the revised ELP standards across domains for each grade span.

22. To comply with paragraphs 20 and 21 above, ADE may use a conjunctive model or a composite score of "proficient" on a valid and reliable ELP test as long as ADE can demonstrate that such a score (i) requires sufficient proficiency in each of the language domains (reading, writing, speaking, and listening) and the composites of oral and comprehension, as appropriate to grade level to permit students to effectively participate in grade-level content instruction and assessments in English, and (ii) overall is a valid and reliable measure of the student's proficiency in English that enables his/her meaningful and equal participation in the educational program without ELL services, consistent with the purpose for which the assessment is used. ADE will ensure that all weighting and cutscore decisions are fully supported by valid and reliable scientific research performed by ADE's, Pearson's, and WestEd's psychometricians throughout the AZ-3 development. If ADE uses a compensatory

[5] The terms of paragraph 21 apply to the K-screener except to the extent that it provides only a composite score, which includes all four domains.

model with a composite score of "proficient," then for purposes of this Agreement, a composite score of "proficient" shall require, inter alia:

a. For grades 9-12, at least 60% of the composite score shall be comprised of the reading and writing domains scores, with at least 29% of the composite score comprised of the reading score;
b. For grades 6-8, at least 60% of the composite score shall be comprised of the reading and writing domains scores, with at least 30% of the composite score comprised of the reading score;
c. For grades 3-5, at least 55% of the composite score shall be comprised of the reading and writing domains scores, with at least 28% of the composite score comprised of the reading score;
d. For grades 1-2, at least 52% of the composite score shall be comprised of the reading and writing domains scores, with at least 25% of the composite score comprised of the reading score;
e. For end-of-year Kindergarten AZ-3 test, at least 43% of the composite score will be comprised of the reading and writing domains scores, with at least 20% of the composite score comprised of the reading score; and
f. In addition to a "proficient" composite score, ADE will require students to achieve a minimum score for both the reading and writing subdomains ("the minimal achievement levels") that is externally validated by the Spring 2013 AIMS Reading and Writing data used during the standard setting procedures set forth in Paragraph 23(f)(i) and Exhibit A.

23. ADE must submit evidence demonstrating that the AZ-3 will yield valid and reliable assessment scores for both initial identification and exit from the ELL program. Toward that end, ADE will provide reports to OCR and DOJ throughout the development process of the AZ-3, that update OCR and DOJ on the status of AZ-3's development and contain at least the following information by the dates specified:[6]

    a. By **September 27, 2011:** a test blueprint that is: (i) developed based on the revised ELP standards, consistent with paragraph 21 above; (ii) includes no more than 25% of its items from AZ-1 and AZ-2 provided those items have been re-aligned to ADE's revised ELP standards that correspond to CCSS standards; and (iii) identifies the number of items that are (1) from AZ-1, (2) from AZ-2, or (3) newly developed. When ADE revises its test blueprint, ADE will ensure that it complies with this long-term plan and submit a copy to OCR and DOJ within five (5) days of revising the blueprint with an explanation of the revisions.

[6] If the reports and/or information required by paragraph 23 become available prior to the dates specified herein, ADE agrees to produce them to OCR and DOJ and not to delay their production until the dates specified herein. For any item required by paragraphs 23(a)-(d) for the end-of-the-year Kindergarten AZ-3 test that was not provided to OCR and DOJ by the date of this Agreement, ADE will report the data to OCR and DOJ by January 31, 2013.

b. By **November 1, 2011:**

i. the revised ELP standards that correspond to the CCSS standards in mathematics, English language arts and literacy in history/social studies, science, and technical subjects, and otherwise meet the requirements of paragraph 21(b) above;
ii. a plan showing that at least two times more test items were written as needed for use in AZ-3's operational Forms A and B with different levels of difficulty, including more difficult items than AZ-1 and AZ-2 have to improve AZ-3's discrimination power at the higher end of the proficiency distribution;
iii. several cognitive labs and bias and sensitivity reviews for each of the grade spans were conducted to identify items with unnecessary complex linguistics and format structures and possible cultural/linguistic biases;
iv. a plan showing that the item field test included non-ELL students to permit analyses such as differential item functioning (DIF), multiple group confirmatory analyses comparing structural differences between ELL and non-ELL students, and other comparisons; and
v. a plan showing that the item field test included a large enough sample of ELL students to permit analyses such as DIF by different background variables such as gender, length of time in Arizona and cultural and linguistic backgrounds, including an analysis of how Spanish-speaking students perform relative to non-Spanish-speaking.

c. **By March 30, 2012, the results of the item-level field testing, including evidence demonstrating that:**

i. the item field test included a large enough sample of ELL students to permit analyses within ELL subgroups such as DIF by different background variables such as gender, cultural and linguistic backgrounds, length of time in Arizona, and an analysis of how Spanish-speaking ELL students perform relative to non-Spanish-speaking ELL students;
ii. at least two times more test items were field tested as needed for use in AZ-3's operational Forms A and B with different levels of difficulty, including more difficult items than AZ-1 and AZ-2 have to improve AZ-3's discrimination power at the higher end of the proficiency distribution;
iii. the results of the item field test analyses involving non-ELLs and ELLs, including DIF analyses by at least the variables in paragraph 23(c)(i), multiple group confirmatory analyses comparing structural differences between ELL and non-ELL students, and other comparisons;
iv. interrater reliability indices in the scoring of extended constructed response items using at least the Kappa coefficient and/or intra-class correlation instead of or in addition to the percent of exact agreement and within-points agreements for the items field test; and
v. the achievement level descriptors based on the revised ELP standards that meet the requirements of paragraphs 21(b) and 23(b)(i) above.

d. **By June 1, 2012, evidence demonstrating that:**

i. How the results of the item-level field test were used to determine the items in the Spring 2013 test and to refine and improve the items for future field testing;
ii. analyses of distractors were conducted for the multiple-choice items and items with disproportional distracters were revised using standard distractor analyses with point biserial;
iii. AZ-3 items have undergone an examination of the IRT fit statistics for any indication of OUTFIT or INFIT and items were revised or removed accordingly;
iv. item-level analyses such as indices of internal validity were conducted and items with a low item discrimination of 0.3 or less were removed or revised, provided the overall score is a latent composite score;
v. when items from the field test versions were removed for any reason, a balance of items across categories of the revised ELP standards was ensured; and
vi. sufficient evidence, prior to the creation of the operational forms, that more difficult items were added to AZ-3 and that they establish adequate discrimination power particularly at the higher end of the distribution, including evidence that the negative skewness of the distribution does not cause ceiling effect that could result in inadequate initial identification of ELLs or the premature exit of ELLs, and that AZ-3 items reflecting insufficient levels of difficulty and discrimination power or disproportional distractors were replaced.

e. **By November 2, 2012 for item i and by December 31, 2012 for item ii:**

i. sufficient evidence of the alignment of AZ-3 with Arizona's ELP standards and correspondence with CCSS standards, using at least Norman Webb's alignment methodology; and
ii. information on the validity of the scoring rubric and training of the scorers for open-ended items.

f. **By July 10, 2013:**

i. sufficient evidence of validated cut scores exists for composite and domain subscores, particularly at the "proficient" level, and of validated minimal achievement levels for AZ-3 reading and writing, using valid external criteria; this evidence shall include the results of standard setting, including final cut scores, minimal achievement levels, and weighting, and a report providing sufficient evidence that the cutscores for AZ-3's composite scores and four domain subscores in grades 3-12 were set using the scientifically-accepted procedures for Bookmarking as overseen by Pearson and WestEd. These procedures will include the standard setting judges' evaluation of Spring 2013 AZ-3 and Spring 2013 AIMS Reading, Writing, and Math impact data used during round 4 of standard setting according to the procedures set forth in Exhibit A. The judges will convene in May 2013 for rounds 1 through 3 of standard setting and reconvene in June 2013 for round 4 when they will consider the impact data prescribed by Exhibit A and establish final cut scores.
    1. By May 15, 2013, ADE will provide a standard setting report that includes, *inter alia*, the number of judges and their qualifications, all of the judges' cut

scores and minimal achievement levels from each round of standard setting, and all directions and training provided to the judges.

2. By June 30, 2013, ADE will report all impact data generated by the decision consistency and logistics regression analyses required by Exhibit A;
3. By July 10, 2013, ADE will report any other impact data presented to the judges in round 4 and the judges' final cutscores after consideration of these impact data.

ii. reviewing the content for test items that have a correlation of less than 0.30 with the ELP test's overall score, provided the overall score is a latent composite score, to revise or replace them on a future form;

iii. sufficient evidence of the whole test reliability and the reliability of all domains, including using the parallel form reliability (AZ-3 Form A and Form B) approaches with a sample size of approximately 1000 students in grades 2, 4, 6, and 9, in addition to the internal consistency approach;

iv. data on the individual item's contribution to the overall test's internal consistency were checked (i.e., the alpha coefficient was examined when an item was removed) and items that did not contribute much to the overall reliability of the test (i.e., those that were not correlated with the total test score) assuming uni-dimensionality of the ELP domains will be reviewed to revise or remove them on a future form;

v. interrater reliability information in the scoring of extended constructed response items using at least the Kappa coefficient and intra-class correlation instead of or in addition to the percent of exact agreement and within-points agreements for the Spring test; and

vi. any changes to the achievement level descriptors of the AZ-3 ELP standards recommended by the judges in standard setting were adopted by ADE.

24. ADE will take the steps identified in paragraphs 21-23 above to develop the new K-screener by **Fall 2012**, and an end-of-year Kindergarten AZ-3 test by the **Spring of 2013**. ADE will use its existing data regarding the administration of the AZ-1 and AZ-2 Kindergarten tests to inform its development of both Kindergarten tests, and to the extent it uses AZ-1 and AZ-2 items or items from the ALL, ADE will ensure that these items are aligned to ADE's revised ELP standards required by paragraph 21 and 23 above. As set forth in this paragraph, ADE will provide evidence that the K-screener and end-of-year Kindergarten AZ-3 test validly and reliably measure English language proficiency in each language domain of reading, writing, speaking, and listening and in the composites of oral and comprehension, and such evidence will reflect all of the requirements of paragraphs 21-23 above regarding item development, pilot and field testing, validating cutscores, validating at the item level, and ensuring reliability. ADE will report to OCR and DOJ for the K-screener:

a. By September 20, 2012: (i) its analysis of the available new K-screener result data, comparing the numbers and percentages of tested K students identified as ELL (broken down by proficiency level) and as IFEP on the new K-screener vs. the AZ-2 in SY2011-12 for the same time period; and (ii) a preliminary test-development report, including, but not limited to, the item field test methods, item field test results, and how the results were used to create the final operational form;

b. By November 2, 2012, the draft alignment study and the standard setting report;

c. By January 31, 2013, the final alignment study;

d. By July 1, 2013, the end-of-year Kindergarten AZ-3 results of all students identified as IFEP and ELL by the K-screener and how these results informed the K-screener's development and final scoring; and

e. By October 15, 2013: (i) the consequential validity study being conducted by the National Center of Educational Outcomes; and (ii) concurrent external validity study using a valid and reliable K screener on a sample size of approximately 1,200 students who take both AZ's K-screener and the other K screener in the summer of 2013, using a stratified sampling that includes the major background variables, such as cultural and linguistic backgrounds, Spanish-speaking ELL students relative to non-Spanish-speaking ELL students, and socio-economic status.

25. By September 15, 2013, September 15, 2014, and September 15, 2015, ADE will report to OCR and DOJ the number of students by grade who took the AZ-3 and/or K-screener in SY2012-13, SY2013-14 and SY2014-15 and the percentage of such students who are classified as (a) IFEP, (b) ELL, (c) RFEP, (d) FEP1 who score overall proficient v. not-proficient, and (e) FEP2 who score overall proficient v. not-proficient. Separately for IFEP, RFEP, FEP1 who score overall proficient, and FEP2 who score overall proficient, ADE also will report the number and percentage of these students by grade who did not score proficient on (i) reading, (ii) writing, (iii) speaking, and (iv) listening.

26. ADE will provide evidence from the SY2012-13 administration of the K-screener and from the AZ-3 test for grades K-12 by January 2014, demonstrating that these tests validly and reliably measure English language proficiency in each language domain of reading, writing, speaking, and listening and in the composites of oral and comprehension. Part of this evidence will include the AZ-3 technical manual and sufficient evidence on the criterion-related validity of the AZ-3 in terms of predictive validity determined through analyses of the Spring 2013 AIMS Reading, Writing, and Math scores for (a) RFEP, (b) FEP1, and (c) IFEP students identified by AZ-3, as compared to the AIMS scores of never-ELLs, and (d) RFEPs who were proficient in reading, (e) RFEPs who were not proficient in reading , (f) IFEPs who were proficient in reading, (g) IFEPs who were not proficient in reading, (h) RFEPs who were proficient in writing, (i) RFEPS who were not proficient in writing, (j) IFEPs who were proficient in writing, and (k) IFEPs who were not proficient in writing.

27. In the event ADE, OCR, and DOJ do not agree on the final cutscores for the AZ-3 and K-screener by the start of the 2013-14 school year, the parties understand that (a) ADE and LEAs will nonetheless need to determine who will be treated as an ELL, FEP, and IFEP for purposes of placing students in classes for SY 2013-14 and may do so, and (b) OCR and DOJ reserve all rights regarding their enforcement of this Agreement and federal law, including challenging the validity of the SY 2013-14 student placements.

28. Reclassification: By July 31, 2013, ADE agrees to direct Arizona LEAs to reclassify ELL students based on the AZ-3 test to be administered in Spring 2013.

29. By July 15, 2013, ADE will provide to OCR and DOJ a draft written directive to Arizona LEAs as provided in Paragraph 28. OCR and DOJ will promptly approve the draft directive, or recommend revisions to the language contained therein before ADE is required to issue the written directive to

Arizona LEAs pursuant to Paragraph 28. Within thirty (30) days of ADE sending the directive, ADE will incorporate information regarding reclassification of ELL students based on AZ-3 in to: (a) OELAS training to LEAs, and (b) OELAS ELL Monitoring process. ADE will provide OCR and DOJ with documentation of the incorporation of this information in to the OELAS training and monitoring processes within sixty (60) days of ADE sending the directive.

NOTICE

30. ADE will require that Arizona LEAs communicate with parents of students who qualify under paragraphs 11-17 about the OCR and DOJ approved interim policies and procedures required by the Interim Plan provisions in paragraphs 11 through 17 above within fifteen (15) days of their adoption. ADE will notify LEAs that communication with these parents must be in a language they understand and must state that the purpose of the notification is to ensure that their children's educational needs are being met.

31. ADE will also require that Arizona LEAs communicate with parents of students about the OCR and DOJ approved Long-Term Plan policies and procedures within thirty (30) days of their adoption and notify LEAs that communication with these parents must be in a language they understand and must state that the purpose of the communication is to ensure that their children's educational needs are being met.

32. ADE will monitor the implementation of the corrective action required by paragraphs 11-18 and 30 through its OELAS monitoring process. The deadline for Arizona LEAs to complete this corrective action on the interim procedures will be December 15, 2012. Within five (5) days of OCR's and DOJ's approval of the draft written notice required by paragraphs 19 and 30, ADE will formally issue the notice to all Arizona LEAs and provide OCR and DOJ with written confirmation that the written notice has been issued.

33. ADE will monitor the implementation of the corrective action required by paragraphs 28 and 31 through its OELAS monitoring process. The deadline for Arizona LEAs to complete this corrective action on the long-term procedures will be listed in the notice and will be no later than August 2013 prior to the beginning of the 2013-14 school year. Within five (5) days of OCR's and DOJ's approval of the draft written notice required by paragraphs 28 and 29, ADE will formally issue the notice to all Arizona LEAs and provide OCR and DOJ with written confirmation that the written notice has been issued.

REPORTING

34. By August 15, 2013, and annually thereafter during each school year that this Agreement remains in effect, ADE will submit electronic files of all English Language Proficiency test scores (by each of the domain areas, and overall score) for each student who took the AZ-3 or K-screener, including each student's unique identifier, school district, school, grade, ELL status (ELL, RFEP, FEP by year, IFEP), language background (to the extent ADE receives and maintains this information), purpose of the test, and test date.

35. ADE will make AIMS data available to OCR and DOJ upon request, including AIMS scale scores in each of the performance levels of Reading, Writing, and Math. The following information will also be

available upon request: each student's unique identifier, school district, school grade, ELL status (for current ELLs, FEPs by year, RFEPs, IFEPs, and never-ELL/English-Only), language background, and test date.

36. To ensure compliance with this Agreement, OCR and DOJ may require additional monitoring reports or the ability to inspect data or other information maintained by ADE as determined necessary by OCR and DOJ.

ENFORCEMENT OF THIS AGREEMENT

37. If, at any time, ADE desires to modify any portion of this Agreement because of changed conditions making performance impossible or impractical or for any other reason, it will promptly notify OCR and DOJ in writing, setting forth the facts and circumstances thought to justify modification of this Agreement and the substance of the proposed modification. Until OCR and DOJ notify ADE in writing that they have agreed to the proposed modification, the proposed modification will not take effect. Any modifications must receive the prior written approval of OCR and DOJ. This approval shall not be unreasonably withheld or delayed.

38. It is a violation of this Agreement for ADE to fail to comply in a timely manner with any of its requirements without obtaining sufficient advance written agreement with OCR and DOJ for an extension of the relevant time frame imposed by the Agreement.

39. If OCR and DOJ believe that ADE has failed to comply in a timely manner with any requirement of this Agreement without obtaining sufficient advance written permission from OCR and DOJ regarding a modification of the relevant terms under the terms set forth above, OCR and DOJ will so notify ADE in writing and they will attempt to resolve the issue or issues in good faith. If OCR and DOJ are unable to reach a satisfactory resolution of the issue or issues raised within thirty (30) days of the date they provide notice to ADE, OCR and DOJ may take steps to initiate an enforcement action. This may include administrative proceedings through the Department of Education or as a referral to the U.S. Department of Justice, or the DOJ taking appropriate action to enforce the terms of this Agreement and to enforce Title VI, its implementing regulations and/or the EEOA.

40. Failure by OCR or DOJ to enforce this entire Agreement or any provision of it with regard to any deadline or any other provision shall not be construed as a waiver of OCR's or DOJ's rights to enforce other deadlines and provisions of this Agreement, or of ADE's obligation to comply with Title VI, its implementing regulations, or the EEOA.

41. This Agreement constitutes the entire Agreement between the parties on the matters raised herein, and no other statement, promise, or agreement, either written or oral, made by either party or agents of either party, that is not contained in this written Agreement, shall be enforceable.

42. This Agreement does not purport to remedy any existing or potential violations of Title VI, its implementing regulations, the EEOA, or any other federal law other than those addressed in this Agreement. The Agreement does not affect ADE's continuing responsibility to comply with these laws and regulations. For example, OCR and DOJ are investigating a complaint against ADE regarding its English Language Development program. This Agreement does not resolve that complaint or any

other regarding ADE's compliance with Title VI, its implementing regulation, the EEOA, or any other federal law.

ADE's participation in this Agreement in no way constitutes an admission of noncompliance with Title VI or the EEOA regarding this matter. In addition, ADE's participation in this Agreement in no way constitutes an admission of noncompliance regarding OCR/DOJ's investigation of the English Language Development Program, OCR/DOJ's investigation of the methods for identifying potential ELL students (involving the Home Language Survey and supplemental teacher referral process), or any other matter involving ADE that is under investigation by OCR and/or DOJ or for which OCR and/or DOJ have received a complaint.

IMPLEMENTATION OF THIS AGREEMENT

43. This Agreement will remain in effect until OCR and DOJ determine that ADE has fully complied with all of its provisions.

    With respect to the Interim Plan set forth in paragraphs 11-19, 30, and 32 of this Agreement, OCR and DOJ will consider these requirements to have been met and will close out the monitoring of the Interim Plan when ADE submits the written directive(s) to Arizona LEAs consistent with the above requirements, incorporates the notice within its ELL training and monitoring processes, submits all reports to OCR and DOJ required by paragraph 18, and completes the monitoring required by paragraph 32 for the SY2012-13 and SY2013-14. It is expected that monitoring of the Interim Plan part of the Agreement will be closed no later than the end of January 2014.

    OCR and DOJ will monitor ADE's implementation of the Long Term Plan part of the Agreement through its submission of the evidence required by the Long Term Plan. It is expected that monitoring of the Long-Term Plan part of the Agreement will be closed no later than December 31, 2015. If after ADE submits the September 15, 2015 report required by paragraph 25, OCR and/or DOJ determine that the AZ-3 is prematurely exiting ELLs or misidentifying ELLs as IFEP, OCR and/or DOJ will inform ADE in writing of the ELL students at issue and provide evidence supporting its determination to ADE. The parties will have ninety (90) days from the date of this written notice to ADE to attempt to reach agreement on the cut scores for the disputed group of students. If the parties cannot reach agreement within those ninety (90) days, then: (a) OCR and/or DOJ may initiate a Title VI and/or EEOA investigation of the AZ-3; (b) DOJ may initiate litigation to enforce this Agreement, the EEOA, and/or Title VI; and/or (c) OCR can initiate administrative proceedings to enforce this Agreement and/or Title VI.

44. During monitoring of this Agreement, OCR and DOJ will visit ADE and LEAs, interview staff, and students, and request such additional reports or data as are necessary for OCR and DOJ to determine whether ADE has complied with the terms of this Agreement and the provisions of Title VI and the EEOA that are applicable to this complaint. OCR and DOJ will provide sufficient notice to ADE at least five (5) business days prior to any visitations with or interviews of ADE staff. OCR and DOJ will provide ADE with sufficient time to comply with any additional requests for reports or data and any reports or data requested will be relevant to the Agreement.

45. The person signing for ADE represents that he or she is authorized to bind ADE to this Agreement.

46. The effective date of this Agreement is the date of the last signature below.

For ADE:

8/30/2012

John Huppenthal, Date
Superintendent of Public Instruction

For the U.S. Department of Education:

8/31/2012

J. Aaron Romine, Director Date
Denver Enforcement Office
U.S. Department of Education,
Office for Civil Rights, Region VIII

For the U.S. Department of Justice:

8/31/12

Anurima Bhargava, Section Chief Date
U.S. Department of Justice,
Civil Rights Division
Educational Opportunities Section
Washington D.C.

EXHIBIT A

As part of the Bookmark procedures to set the cut scores for AZ-3 composite scores and for each domain score and minimum achievement levels in the reading and writing domains, the following reports based on Spring 2013 AZ-3 data and Spring 2013 AIMS Reading, Writing, and Math data will be prepared by Pearson and WestEd and presented to all of the judges from rounds 1 through 3 as impact data during round 4 of the standard setting process.

1. Decision Consistency Analysis: Using scores for ELLs[7] with both (i) AZ-3 scores and (ii) AIMS Reading, Writing, and Math scores[8], Pearson and WestEd will examine the joint accuracy of performance level decisions from the two tests. A two by two table relating Meets the Standard performance on AIMS Reading, Writing and Math and sequential cut scores on AZ-3 will be constructed. The scale score for the Meets the Standard performance level on each AIMS test is fixed and each possible scale score on AZ-3 will be mapped as a potential cut score for English language proficiency. The report will provide the results of the decision consistency analysis, including information regarding the AZ-3 scale score that maximizes the sum of the percents of students in the cells labeled "Agreement" in the chart below separately for each AIMS subscale (i.e., Reading, Writing, and Math):

AIMS Meets the Standard	AZELLA 3 Cut Score	
	Below	**Above**
No	Agreement	
Yes		Agreement

2. Logistics Regression: Using scores and achievement levels for ELLs with both (i) AZ-3 scores and achievement levels and (ii) AIMS Reading, Writing, and Math scores and achievement levels,[9] Pearson and WestEd will conduct a logistics regression analysis to determine the minimal achievement levels for AZ-3 reading and AZ-3 writing that students in each grade span will need in addition to a "proficient" composite cut score in order to be designated as IFEP or RFEP. The logistic regression analysis will estimate the likelihood of ELLs "meeting or exceeding the standard" on AIMS Reading at varying scale scores for AZ-3 reading when they have an overall proficient AZ-3 score. The logistic regression analysis also will estimate the likelihood of ELLs "meeting or exceeding the standard" on AIMS Writing at varying scale scores for AZ-3 writing when they have an overall proficient AZ-3 score. The report will provide the results of the logistic regression analysis, including the AZ-3 reading and writing scores for which students have a probability of equal to or greater than 50-50 of Meeting the Standard on AIMS Reading and AIMS Writing respectively.

[7] This does not include FEP1 and FEP2 students during their two-year monitoring.
[8] An ELL need not have all three AIMS scores; provided an ELL has an AZ-3 score and at least one AIMS score, his/her AZ-3 and AIMS scores will be used in the decision consistency and logistics regression analyses.
[9] Same standard as footnote 2.