Analysis of Accommodations for Limited English Proficient Students

As required in the amended Quality Education Accountability Act, 79-760, the Nebraska Department of Education (NDE) has built Nebraska State Accountability assessments to measure reading achievement (NeSA-R) in grades 3, 4, 5, 6, 7, 8 and 11 for the purposes of accountability (Update: Standards, Assessment, and Accountability (SAA-6), Sept. 2010). With the help of their partners, Data Recognition Corporation (DRC) out of Maple Grove, Minnesota, and Computer Assisted Learning (CAL) from the University of Kansas in Lawrence, Kansas, NDE has involved Nebraska teachers in the design and development of the first state test in reading. In the spring of 2010 the NeSA-R was administered in both online and paper/pencil modes to approximately 147,000 students including Limited English Proficient (LEP) students. Many LEP students were provided with appropriate accommodations to ensure inclusion in the state assessment system. NDE reported the results of the NeSA-R assessments to Nebraska educators and the public through the State of the Schools Report (SOSR) on the NDE website. Students classified as redesignated English fluent two years or less were included in the ELL subgroup for reporting purposes.

Additionally states receiving Title 1 funding are required to administer annually an assessment of English language proficiency (ELP) measuring progress and attainment of benchmarks for English Language Learners (ELLs). Title III of the No Child Left Behind Act (NCLB) specifies that these annual assessments of English proficiency measure the domains of reading, writing, listening, and speaking (section 3113(a)(1)). Along with eighteen other states in a consortium, Nebraska developed reliable and valid ELP tests, aligned with state ELP

standards, in order to accurately assess students' levels of English proficiency. The tests called the *English Language Development Assessments* (*ELDA*: AIR, 2005) was administered during the months of February and March of each year to students in Kindergarten through grade 12. The students are tested in clusters with cluster 1 being Kindergarten, cluster 2 being grades 1-2, cluster 3 being grades 3-5, cluster 4 being grades 6-8, and cluster 5 being 9-12. The same form of the ELDA assessment is administered to each LEP student in the cluster.

Testing accommodations are adjustments to testing procedures, testing materials, or the testing process to allow the student fair and meaningful participation in an assessment (Acosta, Rivera, Shafer, and Willner, 2008). Effective accommodations for ELLs address the unique linguistic and socio-cultural needs of the student and enable them to demonstrate knowledge of the content without changing the test construct. Guidance for accommodations is provided in the document "Nebraska State Accountability (NeSA) Approved Accommodations."

It is evident that there is a need to determine whether achievement on the NeSA-R assessment was effected by accommodations provided for ELL students or not. According to the 2009-2010 SOSR, 99.2% of the ELL subgroup participated in the NeSA-R assessment, but only 33% of the ELL students attained "Meets the Standards" or "Exceeds the Standards" while 69% of the total population of students scored at the proficiency or above levels.

The Present Study

This non-experimental study investigated the effect of accommodations on the overall student reading achievement as measured by the state reading assessment for each ELDA reading level by grade level and ELDA cluster level. This research also compared reading results for students currently in the LEP program and those students redesignated as English fluent two years or less. This research was designed to answer the following questions:

1. Is there a difference in reading scores between LEP students receiving accommodations, LEP students without accommodations, and those redesignated as English fluent less than two years?
2. Is there a difference between LEP students receiving accommodations, LEP students without accommodations, and those redesignated as English fluent less than two years for grades 3, 4, 5, 6, 7, 8, and 11?
3. For each cluster, is there a difference between NeSA-R scores for students receiving LEP accommodations and those without LEP accommodations based on ELDA reading levels?
4. For each cluster, is there a difference between NeSA-R scores for students with different ELDA reading fluency levels?

Method

Participants

NeSA-R results for 9445 students were analyzed to address the first and second question about scale score differences based on student language proficiency status and ELL test accommodation participation. The total group of 9445 students in grades 3, 4, 5, 6, 7, 8,

and 11 included 1269 students redesignated as English fluent two years or less and 8176 students receiving LEP support. Initial analysis compared scale scores for LEP students with accommodations (ACC_ELP = 1), LEP students without accommodations (ACC_ELP =2), and previous LEP students redesignated as English fluent for two years or less (ACC_ELP=3). Follow-up analyses within each grade level (grades 3, 4, 5, 6, 7, 8, and 11) compared reading scores for these same categories. No comparison between grade levels was conducted. Table 1 shows the statistics for the ELL students included in the first analysis.

NeSA-R assessment results for 7862 LEP students with scores from the ELDA and NeSA-R were analyzed to address the third question concerning the impact of ELL accommodation participation. The redesignated students (1269) and LEP students without ELDA results (314) were deleted from this analysis. Follow-up comparisons for ELL students with accommodations students without accommodations were conducted for each ELDA cluster (Cluster 3 – grades 3, 4, and 5; Cluster 4 – grades 6, 7, and 8; Cluster 5 – grade 11). Table 7 shows the statistics for the LEP students included in the second analysis.

Measures and Procedures

The *Nebraska State Accountability assessment for reading* (NeSA-R) developed by Nebraska Educators was administered to students in grades 3, 4, 5, 6, 7, 8, and 11 during March and April 2010. The number of multiple choice items to measure reading achievement varied across the grades with 45 items for grade 3 and 4, 48 items for grade 5, 6, and 7, and 50 items for grade 8 and 11. The raw score from each grade level test was converted to a scale score between 0 and 200 based on a grade level conversion chart located at: http://www.education.ne.gov/Assessment/NeSA.Scoring.htm. The student's reading scale

score from the NeSA-R assessment was included in the model as the dependent variable. Scale scores were converted to three reading performance levels – *Below the Standards* (0-84), *Meets the Standards* (85-134), and *Exceeds the Standards* (135-200) by a cut score process in June after the test administration. The reading performance levels were not included in the model.

The NeSA reading assessments for grades 3, 4, 5, 6, 7, 8, and 11 were administered to all public students enrolled in Nebraska schools. Each student either completed an online or paper/pencil assessment during March-April 2010. Student responses were scored and returned to NDE for reporting and analysis. Some of the ELL students were removed from analysis due to missing NeSA-R scores for the following reasons – recently arrived LEP (332), no longer enrolled - moved (244), 1% alternate assessment (18), student absent during testing (7), other reasons (8), or illness (4). Since the missing scores were either random (moved, illness, absent, or other) or not required to be tested (recently arrived or alternate), the data for these students were deleted from the study (listwise deletion).

The *English Language Development Assessment* (ELDA; AIR, 2005) (ELDA) is a series of tests designed to measure ELL students' level of English language proficiency (ELP) separated into listening, speaking, reading, and writing subscales at each of four grade clusters: K-2, 3-5, 6-8, and 9-12. For this study, the measure of a student's ELP in the reading domain was determined from a single test form at each cluster. The assessment was developed by the American Institute of Research (AIR) under the direction of the Council of Chief State School Officers (CCSSO) and 18 state education departments. The tests measured the reading fluency domain in each grade with fifty multiple-choice items. The reading raw scores were converted

to reading performance levels based on conversion tables specific for each grade level and test form. The performance levels are pre-functional (1), beginning (2), intermediate (3), advanced (4), and fully English proficient (5). The student's reading performance level from the ELDA was included in the model as an independent variable.

During February-March 2010, a single form of the *ELDA* was administered to LEP students in the same cluster according to standardized test conditions defined in the test administration manual. For the three clusters included in this analysis, students were allowed sixty minutes to complete the ELDA reading assessment. Student answers were submitted to Measurement, Inc. for scoring, and the ELDA results were returned to NDE for reporting and analysis. Since the missing data seemed to be completely at random, the data for students with incomplete profiles or missing subscale scores were deleted from the analysis. Only the student's ELDA reading performance level (1, 2, 3, 4, or 5) along with variables for grade level (3, 4, 5, 6, 7, 8 or 11), English Language Fluency Status (1, 2, or 3), Accommodations (1 or 2), and ELDA cluster level (3, 4, or 5) were included in the analysis.

Categorical indicators for providing LEP accommodations (*ACCOMMODATIONS_NESA*) and English language fluency status (*ENGLISH_PROFICIENCY_CODE)* were included in the model as independent variables. Grade level (*GRADE_LEVEL*) and cluster level (*Cluster*) were included in follow-up analysis for splitting the data into additional groups for comparing with-in group variance, but no between groups analysis was conducted by grade level or cluster level.

A General Linear Modal (GLM-Univarate) procedure with PASW Statistics 17.0 software (SPSS) was used to perform an analysis on unbalanced data. Since the analysis involved non-experimental research GLM Method 3 was selected to compare unequal sample sizes reflecting

the importance of cells and thus unequal priority (Tabachnick & Fidell, 2007, pp 217-218). Two separate analyses were conducted. One compared the effect on NeSA-Reading scale scores by English Language Fluency Status and accommodations overall followed by split analysis for the seven tested grade levels. A second analysis compared the effect of NeSA-Reading scale scores by ELDA reading performance level and accommodation participation for the three ELDA assessment clusters on NeSA reading scale scores. All ELL students who completed the NeSA-R assessment were included in the first analysis, whereas the second analysis included ELL students with scores on both the NeSA-R and ELDA assessments.

Three assumptions underlying one-way ANOVA analysis, namely normality of dependent variable, homogeneity of variance, and independence of dependent variable for each cell in the design are addressed below.

1. Normality of dependent variable – refer to table 1, table 4, and table 7 for statistics on skewness and kurtosis for distribution of the dependent variable in each cell of the design along with sample sizes. Since the sample sizes for each cell were large, the Central Limit Theorem states that with sufficiently large sample sizes ($n > 30$) the sampling distributions of means are normally distributed regardless of the distributions of variables (Gravetter and Wallnau, 2007, pp 201). Based on these statistics, the NeSA-R scale scores are reasonably normally distributed for all cells.
2. Homogeneity of Variance – refer to table 1, table 4 and table 7 for variance statistics. This assumption is assessed with F_{max} in conjunction with sample-size ratios. With larger sample size discrepancy, an F_{max} as small as 3 is associated with inflated Type I error if the larger variance is associated with the smaller cell size (Milligan, Wong, and

Thompson, 1987). The largest F_{max} ratio for the first analysis, 1.277 (1167.604/914.46), was below the threshold for a violation of homogeneity of variance. Calculations for sample size discrepancy for the second analysis produced similar results with largest F_{max} ratios ranging from 1.137 to 1.776. Calculations for sample size discrepancy for the third analysis produced similar results with largest F_{max} ratios ranging from 1.012 to 2.325. Since all F_{max} ratios were less than 3, homogeneity of variance holds for all comparisons.

3. Independence of dependent variable – refer to table 1, table 4 and table 7 for sample sizes. Scale scores are a one-time unique measure of student achievement on the NeSA-R assessment. Since each student's scale score is independent of the scale score for other students and no repeated measures of the dependent variable are included in the analysis, the assumption for independence is met (Gravetter and Wallnau, 2007, pp 248).

Results

Comparison of Students Designated as English Language Fluent (redesignated two or less years) and LEP Students with or without Accommodations on NeSA-R Scale Scores: A one-way analysis of variance (GLM-Univariate) was performed to evaluate the relationship between English Language Fluency Status and student accommodation participation on NeSA-R scale score. The independent variable (fluency status and accommodations) included three levels: LEP students with accommodations (*ACC_ELP* = 1), LEP students without accommodations (*ACC_ELP* =2), and previous LEP students redesignated as English fluent for two years or less (*ACC_ELP*=3). Means and standard deviations for NeSA-R scale scores (*Readscale*) are

presented for the three levels in Table 1. The reading scale scores varied significantly with fluency status and accommodation participation, as summarized in table 2, with $F(2,9442) = 237.78$, $p < .001$. The strength of relationship between fluency status/accommodations and reading scale score, as assessed by partial η^2, was .048 accounting for 4.8% of the variance in the dependent variable, a small effect. Although there was a significant relationship, it did not account for a large percentage of the variance in the reading scale scores.

Since significant difference between groups was found, follow-up tests were conducted to evaluate pairwise differences among the means. The post hoc comparisons were conducted with Tukey HSD, a test that does assume equal variances among the three groups. The pairwise tests indicated a significant difference in the means between the three groups ($p < .001$), with a mean difference of 7.30 between LEP students with accommodations (64.75) and LEP students without accommodations (72.05), a mean difference of 23.21 between LEP students with accommodations (64.75) and students redesignated as English fluent (87.96), and a mean difference of 15.91 between LEP students without accommodations (72.05) and students redesignated as English fluent (87.96). The redesignated group had the highest mean scale score in comparison to students receiving LEP support while LEP students with accommodations had the lowest mean scale score. The 95% confidence intervals for pairwise differences, as well as the means and standard deviations for the three groups, are reported in table 3.

For the next analysis, the ELL student data was split between the seven grade levels to compare the reading scale scores for the subgroups within a grade. Table 4 presents grade level reading scale score means, standard deviations, variance and sample size for each group.

The sample size for each cell ranges from low of 96 (11th grade LEP students with accommodations) to a high of 1141 (3rd grade LEP students with accommodations). A one-way ANOVA was conducted for each grade level. To control for Type I error across the seven simple main effects, the alpha level for each was set at .007 (.05 / 7 = .007). The reading scale scores varied significantly between the three groups, as summarized in table 5, with $F(2,2161) = 50.397$, $p < .001$, partial $\eta^2 = .045$ for third grade, $F(2,1843) = 42.741$, $p < .001$, partial $\eta^2 = .044$ for fourth grade, $F(2,1530) = 87.007$, $p < .001$, partial $\eta^2 = .102$ for fifth grade , $F(2,1313) = 43.544$, $p < .001$, partial $\eta^2 = .062$ for sixth grade, $F(2,1037) = 60.304$, $p < .001$, partial $\eta^2 = .104$ for seventh grade, $F(2,876) = 59.457$, $p < .001$, partial $\eta^2 = .120$ for eighth grade, and $F(2,664) = 33.744$, $p < .001$, partial $\eta^2 = .092$ for eleventh grade. The strength of relationship between fluency status/accommodations status and reading scale score, as assessed by partial η^2, varied from a low of .044 (4th grade) to a high of .120 (8th grade) accounting for 4.8% and 12% of the variance in the dependent variable, respectively.

Follow-up tests were conducted to evaluate pairwise differences among the means in each of the seven grades. The post hoc comparisons were conducted with Tukey HSD, a test that assumes equal variances among the three groups. For each grade level, the mean reading scale scores for LEP students without accommodations and the redesignated students were significantly higher than the mean reading scale score for LEP students receiving accommodations ($p < .001$), except grade 11 where the mean scores showed no difference between LEP students with accommodations and LEP students without accommodations ($p = .042$). Students redesignated as fluent two or less years scored significantly higher than the LEP students without accommodations at all grade levels, as well. Table 6 presents the pairwise

differences for each grade. Since the primary purpose of the study was to determine whether accommodation participation had an effect on reading scale scores, the analysis continues with the LEP students split into two groups, students with accommodations and students without accommodations groups.

Comparison of LEP Students' Scale Scores for Reading by Accommodation Participation and ELDA Fluency Level: For this analysis, only LEP students who had a reading scale score for NeSA-R and an ELDA reading fluency level determination were included. Since students in a cluster took the same assessment, the LEP student data was split by cluster to compare the reading scale scores for each ELDA reading fluency level and each level of accommodation participation. A total of 7862 LEP students were included in the analysis with 4818 students in cluster 3, 2567 students in cluster 4, and only 477 students in cluster 5. A 5 x 2 way analysis of variance (GLM-Univariate) was performed to evaluate the relationship between student participation with accommodations on NeSA-R (1 = Yes and 2 = No) and the ELDA Reading Level (1 = Pre-functional, 2 = Beginner, 3 = Intermediate, 4 = Advanced, 5 = Fully English Proficient). Means and standard deviations for NeSA-R scale scores (*Readscale*) are presented for the fluency levels and accommodation participation in Table 7.

The ANOVA indicated no significant interaction between reading levels and accommodations for cluster 4 and 5 with $F(4,2557) = 2.22$, $p = .065$, partial $\eta^2 = .003$, and $F(4,467) = 1.192$, $p = .313$, partial $\eta^2 = .010$,respectively. A significant interaction was indicated for cluster 3 $F(4,4808) = 7.202$, $p < .001$, partial $\eta^2 = .006$, but the partial Eta squared was so small that the interaction can be safely dropped from the analysis. The main effects for accommodations within cluster 3, 4, and 5 were not significant with $F(1,4808) = .146$, $p = .702$,

partial η^2 = .000, F(1,2557) = 1.579, p = .209, partial η^2 = .001, and F(1,467) = 3.471, p = .063, partial η^2 = .007, respectively. The results from analysis of the main effects are presented in Table 8. The primary purpose of the study was to determine whether accommodation participation had an effect on reading scale scores. The results indicate that accommodations had no significant effect on the student's scale score for reading for the three cluster groups. Table 9 presents information on comparison of students with accommodations and students without accommodations.

ANOVA indicated significant main effects for ELDA reading levels in cluster 3, 4, and 5, F(4,808) = 422.319, $p < .001$, partial η^2 = .260, F(4,2557) = 356.805, $p < .001$, partial η^2 = .358, and F(4,467) = 48.948, p = .295, partial η^2 = .295, respectively. The strength of relationship between the ELDA reading fluency level and the reading scale score, as assessed by partial η^2, varied from a low of .260 (cluster 3) to a high of .358 (cluster 4) with .295 (cluster 5) in the middle. The ELDA reading fluency level accounted for 26%, 35.8%, and 29.5% of variance in the dependent variable for clusters 3, 4, and 5, respectively. Since there were five ELDA reading levels, a follow-up analysis to the significant main effect was conducted for each cluster. The follow-up analysis among the five reading levels consisted of the Turkey HSD procedure across all pairwise comparisons, with alpha set at .01 (.05 / 5 = .01) to control for Type I error. The results of this analysis indicated that students with higher reading fluency did significantly better on the NeSA-R assessment than students with lower reading fluency ($p < .001$). For all three clusters, students with a 5 reading level scored significantly higher than students at lower levels, while students with a 4 reading level scored significantly higher than students at lower levels except for cluster 5 where no difference was found between level 4 and level 3 students.

The level 3 students at all clusters scored higher on average than level 2 or level 1 students. Students at level 2 had significantly higher scores than students at level 1, as well. The highest mean difference was between reading level 1 (pre-functional) and level 5 (fully English proficient) with a significant difference of 52.27 for cluster 3, 70.062 for cluster 2, and 68.070 for cluster 5. The student's ELDA reading level was a significant indicator for the student's NeSA Reading Scale Score at all three clusters. Results for the pairwise comparisons are presented in Table 10.

Findings and Discussion

The analysis of the data resulted in significant finding for some of research hypotheses, but not all. It is recommended that research reports include a measure of effect size in addition to the measure of significance. Most of the effect sizes were small. The findings included:

1. Is there a difference in reading scores between LEP students receiving accommodations, LEP students without accommodations, and those redesignated as English fluent less than two years?

    There was a significant difference between the mean scores for the three groups. Initially the analysis of data for all ELL students who took the NeSA reading assessment resulted in a overall significant difference between LEP students with accommodations, LEP students without accommodations and students redesignated as English fluent, but the effect size was less than .05, a small effect. Further analysis found that LEP students with accommodations scored significantly lower than both LEP students without accommodations and students redesignated as English fluent, while LEP students without

accommodations scored significantly lower than students redesignated as English fluent.

2. Is there a difference between LEP students receiving accommodations, LEP students without accommodations, and those redesignated as English fluent less than two years for grades 3, 4, 5, 6, 7, 8, and 11?

   There was a significant difference between the mean scores for the three groups within each grade level. Analysis of results for LEP students with accommodation, LEP students without accommodation, and redesignated English fluent for each of the seven grade levels resulted in a similar outcome. There were significant differences between the three groups, but each difference had a small effect ranging from .044 to .120. The effect size of .12 indicates that only 12% of the variability in reading scale scores was accounted for by accommodations.

3. For each cluster, is there a difference between NeSA-R scores for students receiving LEP accommodations and those without LEP accommodations based on ELDA reading levels?

   There was no significant difference between the mean scores for the two LEP groups. The second analysis of LEP students with ELDA results started with a comparison of LEP student with accommodations and LEP without accommodations for each cluster. The students were grouped into three clusters, cluster 3 for grades 3, 4, and 5, cluster 4 for grades 6, 7, and 8, and

cluster 5 for grade 11. The main effect of accommodations was not significant for any of the clusters.

4. For each cluster, is there a difference between NeSA-R scores for students with different ELDA reading fluency levels?

The second analysis showed a significant difference between students at the five ELDA fluency levels within each cluster. Follow-up analysis found that student's scores in cluster 3 were significantly different from each level of fluency with students at the higher fluency levels scoring significantly higher than students at lower fluency levels. Likewise the results found that students in cluster 4 with higher fluency levels scored higher on the reading assessment than students at lower fluency levels. The only exception to this pattern was students in cluster 5 where no difference was found between students with fluency level of 4 and 3. The analysis showed that reading fluency level accounted for 26% of the variability in cluster 3, 36% of the variability in cluster 4, and 30% of the variability in cluster 5.

Conclusion

The analysis showed a significant difference between students with accommodations and students without accommodations does exist but they are extremely small. The effect of a student's level of fluency on NeSA-R scale scores was significant and yielded a larger effect size between level of fluency and the students reading performance scores.

References

Acosta, B., Rivera, C., and Shafer Willner, L. (2008). *Best Practices in State Assessment Policies for Accommodating English Language Learners: A Delphi Study.* Arlington, VA: The George Washington University Center for Equity and Excellence in Education.

American Institute for Research (AIR). (2005). *English language Proficiency Standards and Test and Item Specifications, K-12.* Washington, DC: Council of Chief State School Officers. Retrieved from http://www.ccsso.org/projects/ELDA/research_studies

Gravetter, F. and Wallnau, L. (2006). *Statistics for Behavioral Sciences.* Belmont, CA: Thomas Higher Education.

Milligan, G., Wong, D., and Thompson, P. (1987). *Robustness Properties of Nonorthogoal Avalysis of Variance.* Psychological Bulletin. 101 (3)

Nebraska Department of Education. (2010). *Update: Standards, Assessment, and Accountability (SAA-6).* Lincoln, NE: Statewide Assessment Office. Retrieved from http://www.education.ne.gov/Assessment/documents/SAAUpdate6.pdf

Nebraska Department of Education. (2010). 2009-2010 *State of the Schools Report A Report on Nebraska Public Schools.* Lincoln, NE: Nebraska Department of Education. Retrieved from http://reportcard.education.ne.gov/Main/Home.aspx

Nebraska Department of Education. (2011). *Nebraska State Accountability (NESA) Approved Accommodations.* Lincoln, NE: Statewide Assessment Office. Retrieved from http://www.education.ne.gov/assessment/documents/NeSA_Accommodations.pdf

PAWS Statistics 17.0. (2009). Somers, NY: IBM Corporation

Tabachnick, B. and Fidell,L. (2007). Using Multivariate Statistics. Boston, MA: Pearson and Allyn & Bacon.

U.S. Congress. (2002). *No Child Left Behind Act.* Public Law 107-110. Washington, DC: Government Printing Office. Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/indes.html

Table 1: Descriptive Statistics for Analysis of Reading Scale Scores for Three ELL Groups

NeSA – R Scores	N	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Total	9445	71.45	33.448	1118.751	.347	.025	.207	.050
LEP Students with Accommodations	3539	64.75	30.240	914.460	.200	.041	.012	.082
LEP Students without Accommodations	4637	72.05	34.170	1167.604	.379	.036	.149	.072
Redesignated	1269	87.96	33.352	1112.372	.291	.069	.368	.137

Table 2

Tests of Between-Subjects Effects for NeSA Reading Scale Scores by Fluency Status / Accommodations

Dependent Variable:SCALE_SCORE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	506630.196[a]	2	253315.098	237.781	.000	.048	475.561	1.000
Intercept	3.928E7	1	3.928E7	36868.535	.000	.796	36868.535	1.000
ACC_ELP	506630.196	2	253315.098	237.781	.000	.048	475.561	1.000
Error	1.006E7	9442	1065.331					
Total	5.879E7	9445						
Corrected Total	1.057E7	9444						

a. R Squared = .048 (Adjusted R Squared = .048)

b. Computed using alpha = .05

Table 3

95% Confidence Intervals of Pairwise differences in Mean Differenced in Reading Scale Scores For Three Groups

Dependent Variable:SCALE_SCORE

Fluency Status/ Accommodations Group	Mean	Mean Difference	Standard Error	Significance	98.3% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
LEP with Accommodations (1) (3-1)	64.75	---	---	---	---	---
LEP without Accommodations (2) (2-1)	72.05	7.30*	.729	.000	5.31	9.30
Redesignated (3) (3-2)	87.96	15.91*	1.034	.000	13.08	18.74
Redesignated (3) (3-1)	87.96	23.21*	1.068	.000	20.29	26.14

Based on estimated marginal means

*. The mean difference is significant at the .017 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Table 4: Descriptive Statistics for Analysis of Reading Scale Scores for Three ELL Groups by Grade (3, 4, 5, 6, 7, 8, and 11)

NeSA – R Scores	N	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
3rd Grade								
LEP Students with Accommodations (3rd)	1141	73.19	25.251	637.634	.428	.072	.210	.145
LEP Students without Accommodations (3rd)	795	82.78	31.073	965.519	.607	.087	.799	.173
Redesignated (3rd)	228	90.37	29.580	874.975	.554	.161	.784	.321
4th Grade								
LEP Students with Accommodations (4th)	874	70.54	30.416	925.160	.281	.083	.246	.165
LEP Students without Accommodations (4th)	807	79.33	34.465	1187.853	.222	.086	.033	.172
Redesignated (4th)	165	94.48	34.242	1172.519	.453	.189	.350	.376
5th Grade								
LEP Students with Accommodations (5th)	637	57.58	30.658	939.882	.167	.097	-.361	.193
LEP Students without Accommodations (5th)	701	69.63	34.180	1168.242	.325	.092	.200	.184
Redesignated (5th)	195	92.13	32.396	1049.525	.169	.174	.607	.346
6th Grade								
LEP Students with Accommodations (6th)	415	57.65	33.679	1134.278	.300	.120	-.341	.239
LEP Students without Accommodations (6th)	675	67.60	34.115	1163.840	.363	.094	-.114	.188
Redesignated (6th)	226	84.39	38.043	1447.271	.332	.162	-.108	.322
7th Grade								
LEP Students with Accommodations (7th)	181	51.23	28.905	835.510	.535	.181	.201	.359
LEP Students without Accommodations (7th)	698	70.89	33.746	1138.780	.512	.093	.183	.185
Redesignated (7th)	161	90.20	32.823	1077.351	.067	.191	-.169	.380
8th Grade								
LEP Students with Accommodations (8th)	195	48.82	25.293	639.760	.395	.174	.256	.346
LEP Students without Accommodations (8th)	537	68.96	33.711	1136.438	.457	.105	.185	.210
Redesignated (8th)	147	85.70	30.417	925.170	.564	.200	1.016	.397
11th Grade								
LEP Students with Accommodations (11th)	96	47.88	29.859	891.542	.624	.246	.281	.488
LEP Students without Accommodations (11th)	424	55.01	30.863	952.551	.589	.119	.375	.237
Redesignated (11th)	147	76.69	31.845	1014.104	.073	.200	.487	.397

Table 5
Tests of Between-Subjects Effects for Reading Scale Scores for the Three Subgroups by Grade Level
Dependent Variable:SCALE_SCORE

GRADE_LEVEL	Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
03	Corrected Model	78925.684[a]	2	39462.842	50.397	.000	.045
	Intercept	9307375.598	1	9307375.598	11886.245	.000	.846
	ACC_ELP	78925.684	2	39462.842	50.397	.000	.045
	Error	1692144.018	2161	783.037			
	Total	1.511E7	2164				
	Corrected Total	1771069.702	2163				
04	Corrected Model	90787.277[b]	2	45393.639	42.741	.000	.044
	Intercept	7070882.668	1	7070882.668	6657.737	.000	.783
	ACC_ELP	90787.277	2	45393.639	42.741	.000	.044
	Error	1957367.312	1843	1062.055			
	Total	1.286E7	1846				
	Corrected Total	2048154.589	1845				
05	Corrected Model	184152.542[c]	2	92076.271	87.007	.000	.102
	Intercept	5921421.189	1	5921421.189	5595.416	.000	.785
	ACC_ELP	184152.542	2	92076.271	87.007	.000	.102
	Error	1619142.348	1530	1058.263			
	Total	8784854.000	1533				
	Corrected Total	1803294.890	1532				
06	Corrected Model	104773.478[d]	2	52386.739	43.544	.000	.062
	Intercept	5284758.246	1	5284758.246	4392.659	.000	.770
	ACC_ELP	104773.478	2	52386.739	43.544	.000	.062
	Error	1579655.276	1313	1203.089			
	Total	7652676.000	1316				
	Corrected Total	1684428.754	1315				
07	Corrected Model	129854.548[e]	2	64927.274	60.304	.000	.104
	Intercept	3423334.557	1	3423334.557	3179.584	.000	.754
	ACC_ELP	129854.548	2	64927.274	60.304	.000	.104
	Error	1116497.674	1037	1076.661			
	Total	6409351.000	1040				
	Corrected Total	1246352.222	1039				
08	Corrected Model	117871.069[f]	2	58935.534	59.457	.000	.120
	Intercept	3001539.061	1	3001539.061	3028.090	.000	.776
	ACC_ELP	117871.069	2	58935.534	59.457	.000	.120
	Error	868319.111	876	991.232			
	Total	4965996.000	879				
	Corrected Total	986190.180	878				
11	Corrected Model	64609.907[g]	2	32304.954	33.744	.000	.092
	Intercept	1647227.371	1	1647227.371	1720.600	.000	.722
	ACC_ELP	64609.907	2	32304.954	33.744	.000	.092
	Error	635684.666	664	957.356			
	Total	3003515.000	667				
	Corrected Total	700294.573	666				

a. R Squared = .045 (Adjusted R Squared = .044)
b. R Squared = .044 (Adjusted R Squared = .043)
c. R Squared = .102 (Adjusted R Squared = .101)
d. R Squared = .062 (Adjusted R Squared = .061)
e. R Squared = .104 (Adjusted R Squared = .102)
f. R Squared = .120 (Adjusted R Squared = .118)
g. R Squared = .092 (Adjusted R Squared = .090)

Table 6
95% Confidence Intervals of Pairwise Differences in Means for the Reading Scale Scores by Fluency/Accommodation Level

Dependent Variable:SCALE_SCORE

Fluency Status/ Accommodations Group	Mean	Mean Difference	Standard Error	Significance	98.3% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
Third Grade						
LEP with Accommodations (1)	73.188	---	---	---	---	---
LEP without Accommodations(2) (2-1)	82.785	-9.596*	1.293	.000	-13.596	-5.597
Redesignated (3)(3-2)	90.373	-7.588*	2.102	.000	-14.092	-1.084
Redesignated (3)(3-1)	90.373	-17.184*	2.030	.000	-23.465	-10.904
Fourth Grade						
LEP with Accommodations (1)	70.543	---	---	---	---	---
LEP without Accommodations(2) (2-1)	79.326	-8.782*	1.591	.000	-13.706	-3.859
Redesignated (3)(3-2)	94.479	-15.153*	2.784	.000	-23.770	-6.536
Redesignated (3)(3-1)	94.479	-23.935*	2.766	.000	-32.496	-15.375
Fifth Grade						
LEP with Accommodations (1)	57.579	---	---	---	---	---
LEP without Accommodations(2) (2-1)	69.631	-12.051*	1.781	.000	-17.564	-6.539
Redesignated (3)(3-2)	92.128	-22.498*	2.634	.000	-30.651	-14.345
Redesignated (3)(3-1)	92.128	-34.549*	2.662	.000	-42.791	-26.307
Sixth Grade						
LEP with Accommodations (1)	57.646	---	---	---	---	---
LEP without Accommodations(2) (2-1)	67.597	-9.951*	2.164	.000	-16.651	-3.252
Redesignated (3)(3-2)	84.394	-16.797*	2.666	.000	-25.051	-8.543
Redesignated (3)(3-1)	84.394	-26.748*	2.867	.000	-35.627	-17.869
Seventh Grade						
LEP with Accommodations (1)	51.227	---	---	---	---	---
LEP without Accommodations(2) (2-1)	70.891	-19.665*	2.737	.000	-28.144	-11.185
Redesignated (3)(3-2)	90.205	-19.314*	2.869	.000	-28.202	-10.426
Redesignated (3)(3-1)	90.205	-38.978*	3.555	.000	-49.991	-27.966
Eighth Grade						
LEP with Accommodations (1)	48.815	---	---	---	---	---
LEP without Accommodations(2) (2-1)	68.955	-20.140*	2.632	.000	-28.299	-11.981
Redesignated (3)(3-2)	85.701	-16.745*	2.931	.000	-25.829	-7.662
Redesignated (3)(3-1)	85.701	-36.885*	3.439	.000	-47.545	-26.226

Eleventh Grade						
LEP with Accommodations (1)	47.875	---	---	---	---	---
LEP without Accommodations(2) (2-1)	55.012	-7.137	3.497	.042	-17.987	3.713
Redesignated (3)(3-2)	76.694	-21.682*	2.962	.000	-30.870	-12.494
Redesignated (3)(3-1)	76.694	-28.819*	4.060	.000	-41.416	-16.222

Based on estimated marginal means

*. The mean difference is significant at the .002 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Table 7
Descriptive Statistics for Analysis of Reading Scale Scores for Clusters by Reading Level and Accommodations

NeSA – R Scores	N	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Total – Cluster 3	4818	72.61	31.151	970.363	.288	.035	.315	.071
With Accommodations	2603	68.45	28.945	837.788	.189	.048	.205	.096
Without Accommodations	2215	77.50	32.900	1082.385	.280	.052	.233	.104
Reading Level = 1	599	46.14	24.339	592.363	.150	.100	-.165	.199
Reading Level = 2	1099	58.18	24.375	594.127	.205	.074	.969	.147
Reading Level = 3	829	69.05	24.003	576.161	.102	.085	.269	.170
Reading Level = 4	1420	80.77	27.998	783.871	.177	.065	.438	.130
Reading Level = 5	871	99.12	28.617	818.935	.286	.083	.422	.166
Total – Cluster 4	2567	63.70	32.507	1056.682	.407	.048	.075	.097
With Accommodations	775	54.41	30.993	960.563	.426	.088	-.044	.175
Without Accommodations	1792	67.72	32.332	1045.344	.409	.058	.117	.116
Reading Level = 1	191	24.06	19.225	369.591	.788	.176	-.061	.350
Reading Level = 2	612	41.82	23.423	548.650	1.023	.099	2.582	.197
Reading Level = 3	503	56.97	21.564	464.993	.384	.109	.810	.217
Reading Level = 4	774	73.97	23.341	544.819	.256	.088	.526	.176
Reading Level = 5	487	97.39	28.738	825.855	.255	.111	-.041	.221
Total – Cluster 5	477	53.42	30.011	900.668	.504	.112	.246	.223
With Accommodations	90	47.79	29.772	886.348	.648	.254	.385	.503
Without Accommodations	387	54.73	29.953	897.199	.480	.124	.264	.247
Reading Level = 1	88	25.16	18.586	345.423	.707	.257	.427	.508
Reading Level = 2	72	38.32	17.504	306.389	.000	.283	.010	.559
Reading Level = 3	105	49.13	21.751	473.117	1.000	.236	3.143	.467
Reading Level = 4	145	64.20	24.282	589.592	-.214	.201	.699	.400
Reading Level = 5	67	90.13	26.690	712.330	.358	.293	-.270	.578

Table 8
Tests of Between-Subjects Effects for NeSA Reading Scale Scores by ELDA Reading Level/ Accommodations

Dependent Variable:SCALE_SCORE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Cluster 3						
Corrected Model	1.386E6	9	153956.742	225.086	.000	.296
Intercept	1.902E7	1	1.902E7	27813.716	.000	.853
ReadLevel	1155448.845	4	288862.211	422.319	.000	.260
ACCOMMODATIONS_NESA	100.087	1	100.087	.146	.702	.000
ReadLevel * ACCOMMODATIONS_NESA	19705.411	4	4926.353	7.202	.000	.006
Error	3288626.414	4808	683.991			
Total	3.008E7	4818				
Corrected Total	4674237.093	4817				
Cluster 4						
Corrected Model	1.255E6	9	139459.500	244.864	.000	.463
Intercept	5441675.518	1	5441675.518	9554.535	.000	.789
ReadLevel	812857.588	4	203214.397	356.805	.000	.358
ACCOMMODATIONS_NESA	899.400	1	899.400	1.579	.209	.001
ReadLevel * ACCOMMODATIONS_NESA	5056.386	4	1264.096	2.220	.065	.003
Error	1456309.925	2557	569.538			
Total	1.313E7	2567				
Corrected Total	2711445.423	2566				
Cluster 5						
Corrected Model	200493.976[c]	9	22277.108	45.584	.000	.468
Intercept	628523.289	1	628523.289	1286.107	.000	.734
ReadLevel	95684.158	4	23921.040	48.948	.000	.295
ACCOMMODATIONS_NESA	1696.358	1	1696.358	3.471	.063	.007
ReadLevel * ACCOMMODATIONS_NESA	2330.459	4	582.615	1.192	.313	.010
Error	228224.003	467	488.702			
Total	1789788.000	477				
Corrected Total	428717.979	476				

a. R Squared = .296 (Adjusted R Squared = .295)

b. R Squared = .463 (Adjusted R Squared = .461)

c. R Squared = .468 (Adjusted R Squared = .457)

Table 9
95% Confidence Intervals of Pairwise Differences in Means for the Reading Scale Scores by Fluency/Accommodation Level

Dependent Variable:SCALE_SCORE

Fluency Status/ Accommodations Group	Mean	Mean Difference	Standard Error	Significance	98.3% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
Cluster = 3						
With Accommodations (1)	69.988	---	---	---	---	---
Without Accommodations (2) (2-1)	70.310	.322	.841	.702	-1.846	2.490
Cluster = 4						
With Accommodations (1)	57.603	---	---	---	---	---
Without Accommodations (2) (2-1)	59.104	1.500	1.194	.209	-1.577	4.578
Cluster = 5						
With Accommodations (1)	57.821	---	---	---	---	---
Without Accommodations (2) (2-1)	52.110	-5.711	3.065	.063	-13.639	2.217

Table 10
Pairwise Differences in NeSA Reading Scale Scores by ELDA Fluency Level

Dependent Variable:SCALE_SCORE

Fluency Status/ Accommodations Group	Mean	Mean Difference				Significance for Mean Difference			
Cluster = 3		Reading Level One (1)	Reading Level Two (2)	Reading Level Three (3)	Reading Level Four (4)	Reading Level One (1)	Reading Level Two (2)	Reading Level Three (3)	Reading Level Four (4)
Reading Level One (1)	45.568								
Reading Level Two (2)	57.562	11.994*				.000			
Reading Level Three (3)	69.005	23.437*	11.442*			.000	.000		
Reading Level Four (4)	80.768	35.200*	23.206*	11.764*		.000	.000	.000	
Reading Level Five (5)	97.841	52.273*	40.279*	28.837*	17.073*	.000	.000	.000	.000
Cluster = 4									
Reading Level One (1)	24.223								
Reading Level Two (2)	41.943	17.720*				.000			
Reading Level Three (3)	56.923	32.700*	14.980*			.000	.000		
Reading Level Four (4)	74.393	50.171*	32.450*	11.764*		.000	.000	.000	
Reading Level Five (5)	94.285	70.062*	52.342*	28.837*	19.892*	.000	.000	.000	.000
Cluster = 5									
Reading Level One (1)	26.139								
Reading Level Two (2)	39.037	12.897*				.001			
Reading Level Three (3)	52.776	26.636*	-3.739*			.000	.001		
Reading Level Four (4)	62.667	36.527*	23.630*	-9.891		.000	.000	.014	
Reading Level Five (5)	94.210	68.070*	55.173*	41.434*	31.543*	.000	.000	.000	.000

Based on estimated marginal means

*. The mean difference is significant at the .01 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).