Exploring Scores from the MI-LMU MLA (Chinese Language) Reading Test and the AAPPL Interpretive Reading Test

Objectives of the Report

As part of the Mandarin Institute-Loyola Marymount University (MI-LMU) STARTALK Infrastructure Grant around literacy in Mandarin Chinese, ACTFL prepared this report on how consistent the MI-LMU reading assessment is with the ACTFL Proficiency Guidelines. The tasks and scores of the MI-LMU reading assessment were compared with the tasks and scores of the Interpretive Reading segment of the ACTFL Assessment of Performance toward Proficiency in Languages (AAPPL). The following specific objectives form the focus for this report:

1. Examine the Mandarin Institute-Loyola Marymount University (MI-LMU) MLA (Chinese Language) Reading assessment tasks for the Grades 3, 4, and 5 MLA (Chinese Language) reading tests, and analyze how the tasks at each grade level are aligned with the ACTFL Proficiency Guidelines (targeting the Novice and Intermediate ranges).

2. Examine the MI-LMU assessment tasks for Grades 3, 4, and 5 MLA (Chinese Language) reading tests and analyze how the tasks at each grade level align with the Assessment of Performance toward Proficiency in Languages (AAPPL) Interpretive Reading tasks for Novice and Intermediate ranges.

3. Examine the AAPPL Interpretive Reading scores and scores from the Grades 3, 4, and 5 MLA (Chinese Language) Reading Tests, to provide an indication of how well the MI-LMU assessment also indicates the reading proficiency range of the learners.

To accomplish these objectives, this report will examine the MI-LMU MLA (Chinese Language) Reading Test, or MI-LMU, score data and the AAPPL Interpretive Reading scores for a sample of Grade 3, 4, and 5 students. The reporting of the data will be as follows:

The students who were assessed through the MI-LMU instrument and the AAPPL measure are currently in grades 3, 4, or 5, enrolled in immersion programs in Mandarin. The students are in 9 schools located in 5 states and the District of Columbia. For each grade level, students in four different classrooms were tested. Their teachers are all part of the STARTALK Infrastructure Grant project through MI-LMU.

First, for each grade level, student scores and score distributions for the AAPPL Interpretive Reading test will be reported. Second, a brief description of each grade's MI-LMU reading test will be provided as well as descriptive statistics for each grade level's MI-LMU data, along with the distribution of scores. It should be noted that not all students took, or had scores reported for, both tests, so the sample numbers will be different for the reporting of these two tests. Third, employing data from students who completed both the AAPPL and MI-LMU tests, the AAPPL proficiency rating (N1, N2, etc.) attained by each student will be used to group each student's corresponding MI-LMU reading test score to determine how well the MI-LMU assessment score for each student matches their performance on the AAPPL test. It should be noted that even

though some inferential data analysis techniques were employed to examine these data, this particular analysis is largely exploratory and observational due to a variety of factors, such as small sample size, unequal sample sizes due to the manner the data were separated, and a high degree of variability in the MI-LMU scores when the data were grouped and analyzed in this way. Still, it is informative to see how the data provide insights into the proficiency development of these students, especially considering what can be gained by comparing the Grade 3, 4, and 5 data. Lastly, overall comparisons will be put forth as to the differences between the MI-LMU tests and the AAPPL test.

Third Grade

Distribution of Grade 3 AAPPL Interpretive Reading Scores

According to the American Council on the Teaching of Foreign Languages website, the ACTFL Assessment of Performance toward Proficiency in Languages (AAPPL) is a performance assessment of standards-based language learning across the three modes of communication (Interpersonal, Presentational, and Interpretive) as defined by the National Standards for Foreign Language Learning. The AAPPL assesses Interpersonal Listening/Speaking, Presentational Writing, Interpretive Reading, and Interpretive Listening, and ratings are assigned according to the ACTFL Performance Descriptors for Language Learners (ACTFL. (2012). *ACTFL performance descriptors for language learners*, Alexandria, VA: Author). This report will focus on the AAPPL Interpretive Reading scores for samples of the immersion program learners in Grades 3, 4, and 5. Being a computerized test underpinned by the World Readiness Standards for Learning Language and the ACTFL Proficiency Guidelines, AAPPL tests the Interpretive Reading mode by providing specific tasks for the students that are immediately related to their lives as students and young people. Because the ACTFL Reading Guidelines stress factors such as visual/extra-linguistic support and a student's background knowledge while reading, AAPPL Interpretive Reading questions often contain pictures that aid the test taker, while receiving both the test questions and some of the accompanying contextual support in English since the test is not evaluating how well the test-taker can read the directions. According to the design of the AAPPL (https://www.languagetesting.com/pub/media/wysiwyg/manuals/AAPPLTopics2017.pdf) , texts used for learners in grade 6 and below include topics familiar to the learner, such as those dealing with one's school floorplan, chores, and hobbies for the Novice range or topics such as newspaper headlines, goals for one's PE class, and a letter from a teacher about the week's activities for the Intermediate range, and will involve a variety of tasks (making a poster, linking headlines with news story, etc.) and situations that are close to the students' experience (school election, restaurants, teaching visiting Chinese about our school, etc.). In other words, the content is highly focused and personalized to the student's more immediate environment.

N = 139
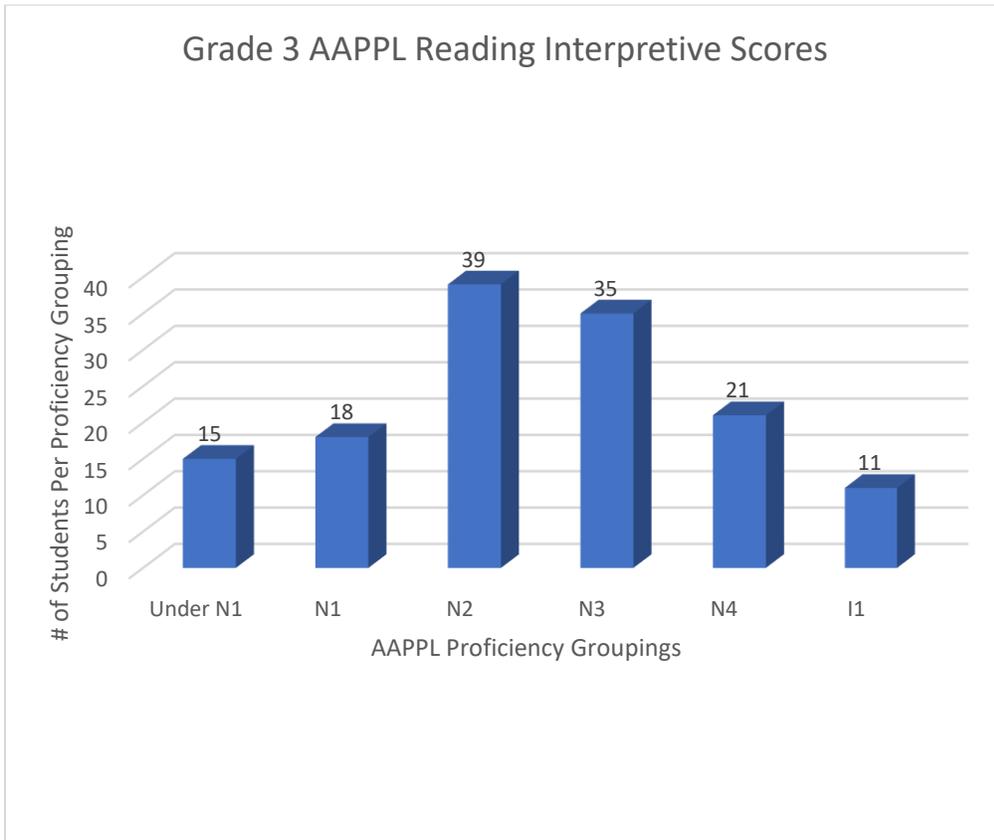
The graph above depicts the distribution of the Grade 3 students in the immersion schools that are participating in the MI-LMU STARTALK project (N=139) who took the AAPPL Interpretive Reading test October-December 2017. The distribution appears normally distributed, with 33 students (23.7%) scoring Below or in the Novice 1 level. This indicates that according to the AAPPL scoring guidance (see attached), roughly 24% of the students were at or approaching the level where they could understand individual words, are dependent upon visual support such as pictures to enhance meaning, and aided by repeated readings. With the largest number of 74 students (53.2%) scoring either in the Novice 2 (39=28%) or Novice 3 level (35=25%), learners here are demonstrating the ability to understand words, phrases, and the occasional simple sentence, aided by knowledge of the topic. The higher proficiency levels of the N4 and I1 groups are smaller and contain a total of 32 (23%) students, with the Novice 4 group containing 21 students (15%) and Intermediate 1, 11 students (7.9%). The N4 group is more likely to understand the main ideas of short paragraphs with familiar content, with visual cues and reading the passage more than once being helpful. I1 learners also can understand main ideas and are likely to understand supporting facts, with visual cues, topic familiarity, and reading the passage more than once being helpful.

The Grade 3 MI-LMU MLA Reading Test

The MI-LMU 3rd Grade MLA Reading Test is composed of 4 reading passages. Comprehension of each paragraph is assessed through a series of questions composed in multiple-choice (mc) format, with the choices that students can select sometimes being composed of words, and sometimes composed of phrases. There are also some word-level items in a fill-in-the-blank format. Each of the mc and fill in the blank questions is worth 1 point. Short-answer questions are also employed whereby the student responds to the question with short explanations in characters. Each short-answer question is evaluated via rubric, earning between 0-4 points. Maximum score on the test is 60 points.

Other pertinent characteristics of each of the Grade 3, 5, and 5 tests is that there is almost no support in the form of pictures, and except for the pinyin glossing noted below, there are no glossing aids. As well, all portions of the text including the test instructions are in Chinese characters (simplified). There is no English on the test at all.

The 4 texts of the test are paragraph-length and are from commercially available texts and include the following:

The 1st text, **The Luminescent Sun**, deals with the power of the sun and the benefits it provides nature and humans. It is composed of 3 paragraphs for a total of 277 characters. 9 words are glossed in the paragraph with pinyin romanization to describe how they are pronounced. There are 5 mc questions and 3 short answer questions.

The 2nd text, **Snowflake**, tells the story of Youyou's quest for spring among the hazards of winter snow. It is composed of 3 paragraphs for a total of 262 characters. 15 words are glossed in the paragraph with pinyin romanization to describe how they are pronounced. There are 4 mc questions, 2 fill in the blank questions, and 3 short answer questions.

The 3rd text, **The Squirrel**, describes attributes of squirrels. It is composed of 1 paragraph for a total of 216 characters. 11 words are glossed in the paragraph with pinyin romanization to describe how they are pronounced. There are 5 mc questions and 3 short answer questions.

The 4th text, **Cheng Yu* Stories: The Frog in the Well**, uses the story of a frog who thinks the sky is only so big, because this is all it can see from deep inside the well where it resides. In Chinese, use of the chengyu "frog in the well" describes someone who is narrow-minded and limited in their views. It is composed of 4 paragraphs for a total of 410 characters. 9 words are glossed in the paragraph with pinyin romanization to describe how they are pronounced, as are some of the words used in the test questions. There are 7 mc questions and 2 short answer questions.

In terms of assessment tasks, the test appears to be aligned with the **Novice to Intermediate Levels**. That is, while one could argue that multiple choice/fill in the blank items requiring identification of one word have the students working at the word identification level, some of the multiple-choice questions are written in phrase or sentence formats that require significant processing on the part of the students, possibly placing these items beyond the Intermediate level unless the vocabulary is high frequency for these immersion students. It is also important to

remember that these questions are based on understanding passages that are quite long (277, 262, 216, and 410 characters, respectively) with no pictorial support (except for one passage depicting a well with the Chinese word accompanying it) or glossing support (except for occasional pinyin), and deal with stories based on academic and cultural content, the level of detail of which is challenging for the student. Again, unless the texts are predictable and in a familiar format with vocabulary that is high frequency and familiar to these immersion students, these texts may be beyond the Intermediate level. Some of the questions also evaluate the students' ability to correctly identify the appropriate connective devices used in Chinese prose, or to identify antonyms from words encountered in text, or to identify the overall theme of the paragraph. Importantly, instructions given to the students for answering the short answer questions require the students to answer in sentences, complete sentences, or short answers. And while much of the test is based on identifying the content of the paragraphs, the short-answer prompts also evaluate argument-making, evidence gathering, and personalizing the content with questions such as "Why did…?", "What did the author say was…?" or asking the student to apply the story to their own life with questions such as, "(Like the character in the story) What do you do when…?"

*Chengyu are 4-character sayings or aphorisms derived from classical stories that permeate the Chinese language today.
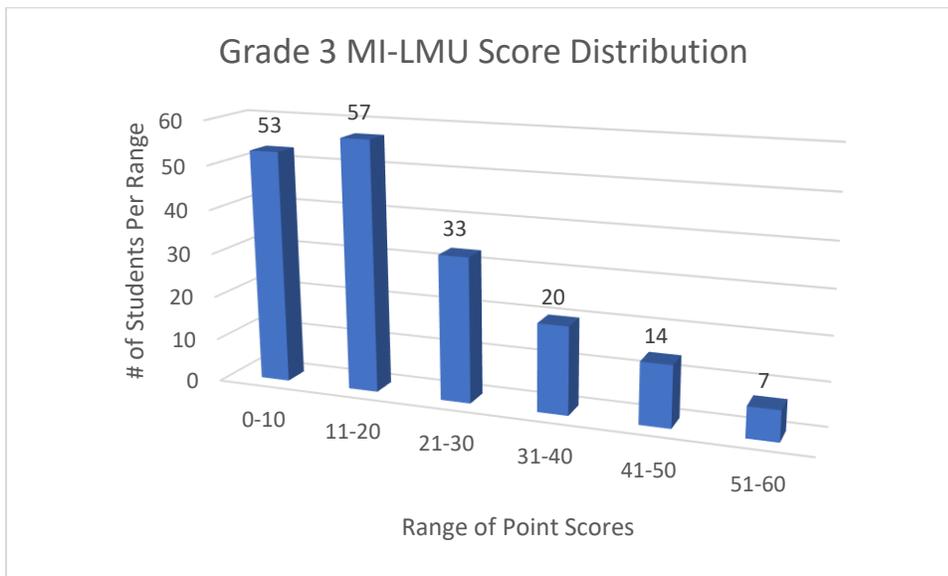
Grade 3 MI-LMU Test Data

Descriptive Statistics

N= 184

Mean = 20.25

SD = 13.86

Range of Scores = 0 to 57

Max Points: 60

Grade 3 MI-LMU Score Distribution

# of Students Per Range (y-axis: 0, 10, 20, 30, 40, 50, 60)

- 0-10: 53
- 11-20: 57
- 21-30: 33
- 31-40: 20
- 41-50: 14
- 51-60: 7

Range of Point Scores (x-axis)

The above graph displays the number of Grade 3 (N=184) students scoring within selected score ranges on the IM-LMU Chinese test where the maximum point score was 60 points and the mean was 20.25 points. 110 students (60%) scored below 20 points, and a total of 143 students (78%) scoring 30 points or less (50%) on the test, indicating that the test was challenging for the Grade 3 students. The distribution of scores shows progress towards increased proficiency, though for smaller numbers of students: fewer students demonstrate reading at the higher levels. Thus, we can see the challenge for Grade 3 students, many of whom are still challenged to read comfortably at the sentence level, and challenged even more to process larger units of extended text without rich pictorial support on passages of a more academic nature.

Comparison of Grade 3 MI-LMU Test Scores and AAPPL Scores

To determine how well the MI-LMU test scores relate to corresponding scores on the AAPPL Interpretive Reading test, Grade 3 students' AAPPL scores were matched with their scores on the MI-LMU reading test, with the six proficiency categories of the AAPPL test (i.e. Below N1, N1, N2, N3, N4, and I1) serving as grouping categories for the students' MI-LMU scores. In other words, the objective of the analysis was to determine how well the AAPPL groups seem to "fit" the MI-LMU test scores. Therefore, scoring data on the MI-LMU tests for the 15 students who scored Below N1 were analyzed together, the scoring data on the MI-LMU tests for the 18 students who scored N1 were analyzed together, etc. Group means and standard deviations for the scores in each group were as follows:

Proficiency Group	Below N1	N1	N2	N3	N4	I1
N =	15	18	39	35	21	11
Mean	9.26	14.13	15.32	25.34	25.31	36.00
SD	5.27	12.31	10.82	14.53	17.75	12.98
Score Range	0-26	3-48	0-45	3-53	3-53	17-57

An examination of the chart above indicates different scoring groupings, perhaps with the Below N1, N1, and N2 groups being the same, and the N3 and N4 groups being the same, with the possibility of the I(ntermediate) group being distinctive in its own right. A one-way ANOVA not surprisingly detected a difference between groups, ($[F (5, 133) = 8.94$, p =.0000), with post-hoc Bonferroni tests of all pairwise comparisons indicating no difference between the Below N1, N1, and N2 groups, and no difference between the N3, N4, and I1 groups. A significant difference was detected between the N2 and N3 group, but not, interestingly, between the N2 and N4 group, perhaps influenced by the higher variability in the N4 group. At least from these data, the MI-LMU reading tests seem to be grouping two levels of Novice proficiency, but the extreme range of the scores in each group limits the ability to make definitive conclusions.

Fourth Grade

Distribution of Grade 4 AAPPL Interpretive Reading Scores

N=167
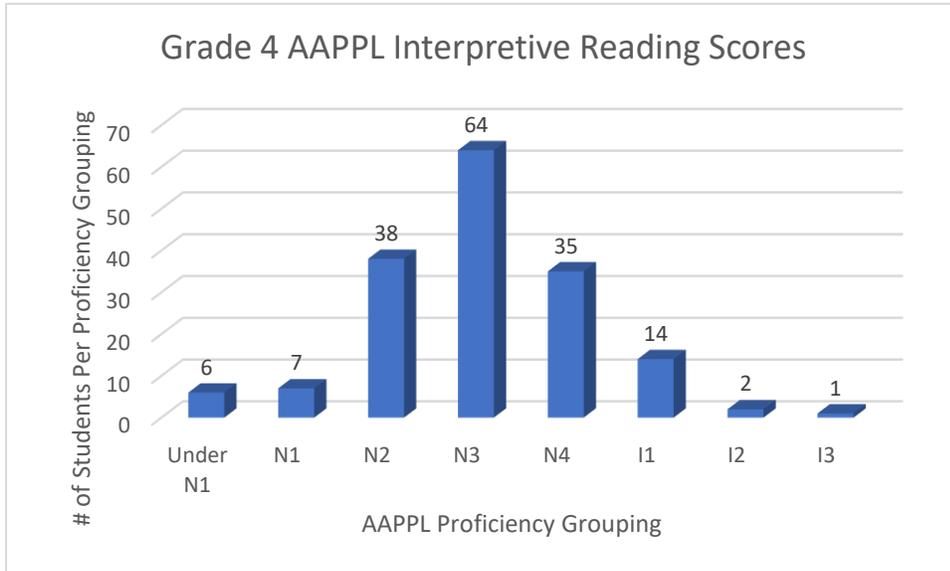
The graph above depicts the distribution of the Grade 4 students (N=167) who took the AAPPL Interpretive Reading test. The distribution appears somewhat normally distributed, with 13 students (7.7%) scoring Below or in the Novice 1 band. In this distribution, the largest group of 137 students (82%) scored either in the Novice 2 (38=22.7%), Novice 3 (64=38%), and Novice 4 (35 students=21%) ranges. 17 students (10%) comprising the Intermediate 1, 2, and 3 levels. In comparison to the Grade 3 students, the Grade 4 students show evidence of moving away from the Under N1 and N1 proficiency level, being solidly represented within the N2 and N4 levels, especially in N3, while also making inroads into the Intermediate level.

The Grade 4 MI-LMU MLA Reading Test

The MI-LMU 4th Grade Reading Test is composed of 4 texts. As with the 3rd Grade tests, comprehension of each text is assessed through a series of questions composed in multiple-choice (mc) format, with choices varying from word to phrase level; or fill-in-the-blank format. The expectations are increased, however, with the short-answer questions expecting longer answers, and the texts being paragraphs that are generally longer in length (509, 239, 316, and 433 characters respectively). Each of these mc/fill in the blank questions is worth 1 point, with the short-answer questions evaluated via rubric, earning between 0-4 points. Maximum score on the test is 76 points.

No sources for the texts used on the test were given, so it is unclear as to whether they are commercially available.

The 1st text, **Does an Ant Know How to Swim?**, deals with a youngster's experiment to determine whether or not an ant knows how to swim, while also teaching the joys of experimentation in discovering the world around us. It is composed of 8 paragraphs (some paragraphs one line long) for a total of 509 characters. No words in the text are glossed with pinyin romanization to describe how they are pronounced. There are 4 mc questions, 2 fill in the blank questions, and 4 short answer questions.

The 2nd text, **Killing the Chicken to Get the Eggs**, tells the story behind a 4-character chengyu, killing the chicken to get the eggs (doing something stupid), whereby the old women grows impatient that her special hen lays only one golden egg per day, and decides to get all the golden eggs at once by disemboweling the hen. It is composed of 3 paragraphs for a total of 239 characters. No words are glossed in the text with pinyin romanization to describe how they are pronounced. There are 6 mc questions, and 1 short answer question asking the students to write about what they would do if they had a hen that laid golden eggs.

The 3rd text, **Pandas our National Treasure**, describes attributes of pandas. It is composed of 1 paragraph for a total of 316 characters. No words/characters are glossed in the paragraph with pinyin romanization to describe how they are pronounced. There are 4 mc questions and 5 short answer questions.

The 4th text, **Eat More Vegetables**, explains the importance of eating fruits and vegetables and what happens to the human body if they are not eaten. It is composed of 4 paragraphs for a total of 433 characters. No characters/words are glossed in the paragraph with pinyin romanization to describe how they are pronounced. There are 7 mc questions and 4 short answer questions.

In terms of assessment tasks, the test appears to be aligned with the **Novice to Intermediate Level**, as the test format continues to embed responses by the students in sentence and phrase level formats that require significant processing on the part of the students. The length of the paragraphs increases for the 4th Graders, and the pinyin glosses enabling students to rapidly recover the pronunciation of words is discontinued. The content is again environmentally-based, with exploring another Chinese story that contains a moral and serves as a chengyu. It's also important to remember that these questions follow the reading of paragraphs that are quite long (277, 262, 216, and 410 characters, respectively) with no pictorial or glossing support, and deal with stories based on academic and cultural content, the level of detail of which is challenging for the student. To match the targeted Intermediate level, the vocabulary needs to be of high frequency and familiar to the students and the text needs to have predictable patterns of presentation. Besides testing an understanding of the paragraphs' content, tasks also evaluate the students' ability to correctly identify the appropriate connective devices used in Chinese prose, select alternative titles which would be appropriate for the paragraph, select other circumstances where the lessons learned in the text would fit, and to select the choice that best orders particular events in a story. Short answer questions also look for students' ability to express an opinion ("What kind of person is Mingming?"), justify an opinion (write a letter to a friend saying why eating vegetable is important), as well as to project the story's main character's experience onto their own life ("What are some of the discoveries you've made in life?" and "What would you do if you had a hen that laid golden eggs?").

4th Grade MI-LMU Test Data

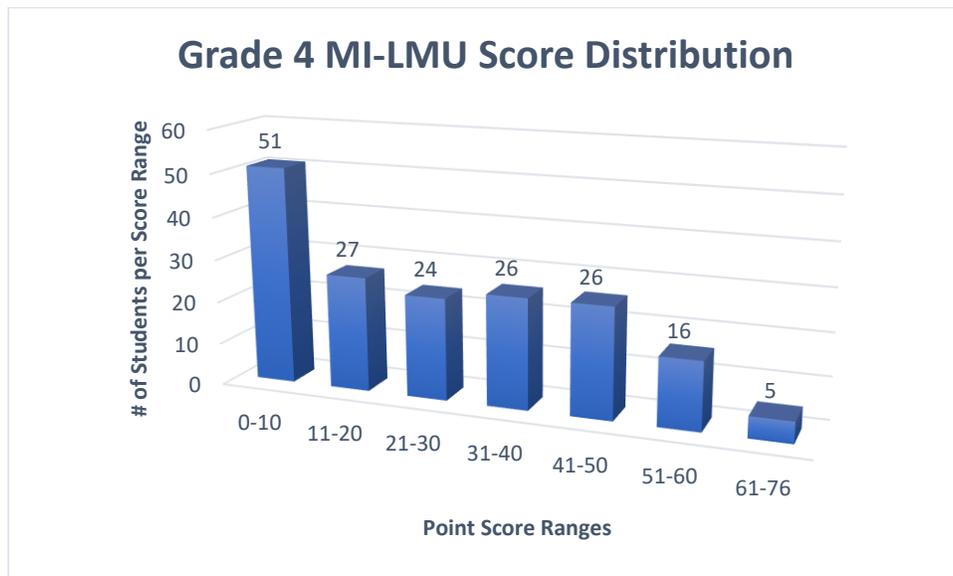**Descriptive Statistics: 4th Grade MI LMU Reading Test Data**

N = 175

Mean = 26.13

SD = 18.50

Range = 0 to 71

Max Score: 76

The above graph illustrates that for the Grade 4 students taking the IM-LMU Reading Tests (N=175), 91 students (52%) scored below the mean of 26.13. The distribution of scores differs from the 3rd graders', however, with more scores over the mean, and a more uniform clustering in the number of scores in the various point score ranges. For example, while 29% of the students only scored 10 points or below (13%), there are almost equal numbers of students whose scores fall into the ranges between 11 and 60 points (14%-65%) that include 103 of the students (59%), with 21 students (12%) scoring between 51 and 76 points (67-100%). Therefore, the distribution indicates that students within this sample are showing an increased level of reading proficiency compared to the third grade sample.

Comparison of Grade 4 MI-LMU Test Scores and AAPPL Scores

To determine how well the MI-LMU test scores relate to corresponding scores on the AAPPL Interpretive Reading test, students' AAPPL scores were matched with their scores on the MI-LMU reading test, with the six proficiency categories of the AAPPL test (i.e. Below N1, N1, N2, N3, N4, I1, and a collapsed group of I2-3) serving as grouping categories for the students' scores.

Proficiency Group	Below N1	N1	N2	N3	N4	I1	I2-3
N =	6	7	38	64	35	14	3
Mean	5.50	16.92	22.77	22.96	34.81	35.96	58.33
SD	5.95	13.17	14.71	17.64	19.88	15.06	6.11
Score Range	1-16	2-33	1-58	1-59	2-71	6-64	53-65

A look at the chart above, similar to the Grade 3 data, indicates that the Below N1 group not surprisingly has the lowest mean score, and smallest standard deviation. The mean scores between N1, N2 and N3 seem similar, with increasing variability. The N4 group seems to make a break from the N3 group, separated as it is by 11.85 points. The one way analysis of variance indicated a significant difference between group means ([$F_{(6, 160)}$ = 6.65, p =.0000), with post-hoc Bonferroni tests of all pairwise comparisons indicating no significant differences between the Below N1, N1, and N2 groups. A significant difference, however, was found between the N2/N3 score groupings and the N4 group, indicating that the N4 level of the AAPPL may be indicating a point in the MI-LMU data of detectable student progress, although there is still a large score range (2-71). The data also indicate consistent mean scores in the N4 and I1 categories, and a clear break away of the three scores in the I2-3 category, although admittedly, there are only three scores in the latter grouping.

Fifth Grade

Distribution of Grade 5 AAPPL Interpretive Reading Scores

N=88
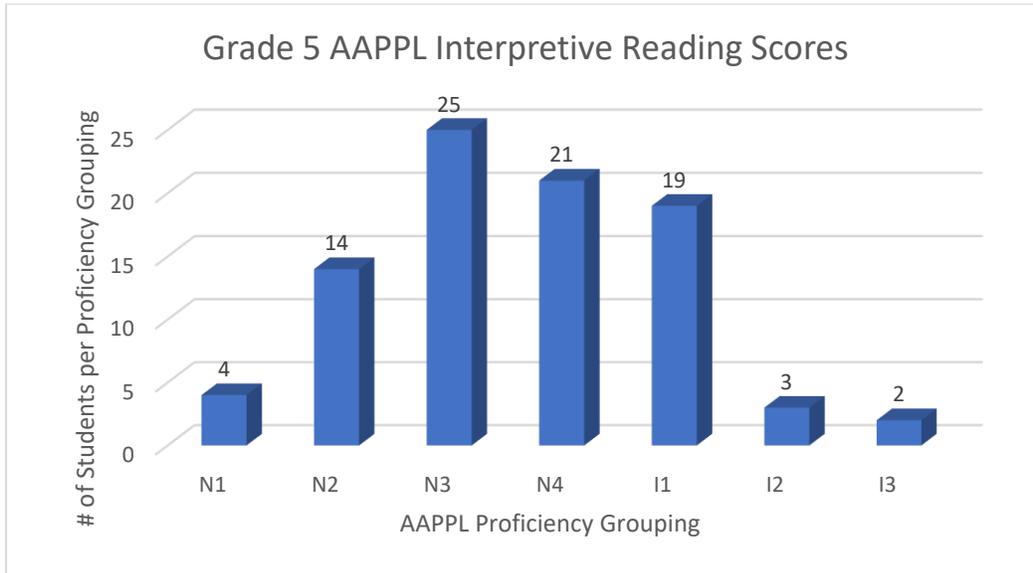
The graph above shows the distribution of the Grade 5 students (N=88) who took the AAPPL Interpretive Reading test. The distribution depicts students moving out of the N1 range (only 4 students (4.5%) scored at this level) – with no student scoring Below N1—and strong increases through the Novice and into the Intermediate levels. 14 students (16%) scored at Novice 2 with the largest group of 25 students (28%) scoring in the Novice 3 range. Noticeable is the lack of dramatic decline in the N4 and I1 levels, where 40 students (45%) scored at these levels. Noteworthy also is that with 3 students at I2 and 2 at I3, a total of 24 students (27%) reached the Intermediate range on the AAPPL Interpretive Reading Test.

The Grade 5 MI-LMU MLA Reading Test

The MI-LMU 5th Grade Reading Test is composed of 4 texts. Comprehension of each text is assessed through a series of questions composed in multiple-choice (mc) format, with choices varying from word to phrase level, or through the use of a fill-in-the-blank format. Each of these mc or fill in the blank questions is worth 1 point. Unlike the 3rd and 4th grade short answer questions, in the Grade 5 test two scores are given for these questions, one for comprehension and one for writing ability, each evaluated by rubric whereby the student earns a score of 0-4 for each (comprehension and writing) answer. Maximum score on the test is 108 points.

The 4 texts selected for the test are from commercially available texts.

The 1st text, **From Mexico to San Diego**, deals with one youngster's challenges in moving from Mexico to live in San Diego. It is composed of 5 paragraphs for a total of 474 characters. No words are glossed in the text with pinyin romanization to describe how they are pronounced. There are 5 mc questions and 3 short answer questions.

The 2nd text, **The Squirrel**, goes into a more detailed account of squirrels and their nesting habits and is composed of 3 paragraphs for a total of 322 characters. No words are glossed in the text with pinyin romanization to describe how they are pronounced. There are 4 mc questions, and 2 short answer questions.

The 3rd text, **The Moon is Always Wonderful**, discusses the importance of the moon in Chinese life, with a special emphasis on holidays where its presence is so central. It is composed of 6 paragraphs for a total of 484 words. No words are glossed in the text with pinyin romanization to describe how they are pronounced. There are 7 mc questions and 3 short answer questions.

The 4th text, **Covering One's Ears to Steal the Bell**, tells the story of a thief who wants to steal an ornate and beautiful bell from his neighbor, but the ringing of the bell as he tries to move it attracts too much attention. Since stuffing his ears with fabric prevents him from hearing the bell ring, he assumes no one else with hear it, either, making it safe now to steal. Thus, this four character chengyu has become a euphemism for self-deception. It is composed of 3 paragraphs for a total of 309 words. No words are glossed in the text with pinyin romanization to describe how they are pronounced. There are 4 mc questions and 3 short answer questions.

In terms of assessment tasks, the test appears to be aligned with the **Novice to Intermediate Levels**, as the test format continues to embed responses by the students in sentence and phrase level formats that require significant processing on the part of the students, with Paragraphs #2, #3, and #4 containing a high level of detail and, for #3 and #4, cultural knowledge that is tested in greater detail in the Grade 5 tests. The length of the paragraphs also increases to 474, 322, 484, and 309 characters respectively, with the pinyin glosses enabling students to rapidly recover the pronunciation of words also discontinued. Besides testing an understanding of the paragraphs' content, mc questions also evaluate the students' ability to correctly identify the appropriate connective devices used in Chinese prose, with the mc questions having generally longer choices for the students to evaluate. Short answer questions also look for students' ability to express an opinion ("What kind of person is Li Si?"), justify an opinion through a variety of "why?" questions, as well as questions about the nest construction of squirrels, and testing the students' abilities to compare and. contrast different holidays.

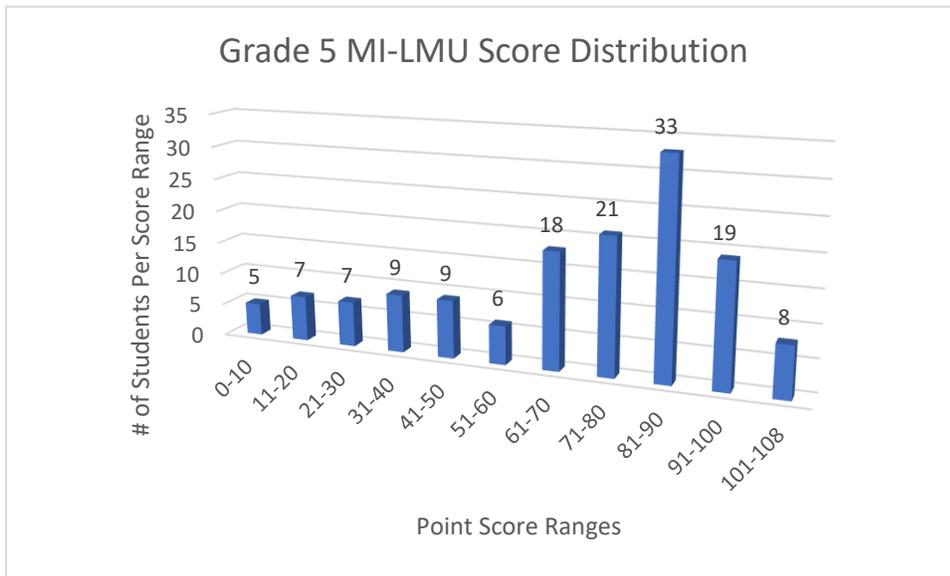**Descriptive Statistics: 5th Grade MI LMU Reading Test Data**

N = 142

Mean = 67.65

SD = 26.77

Max Score: 108

Range = 4 to 105

Grade 5 MI-LMU Score Distribution

Of the Grade 5 students taking the IM-LMU Reading Tests (N=142), 52 students (37%) scored below the mean of 67.65, with 90 students (63.4%) scoring above it. The graph above depicts a different distribution compared with scores of students in the 3rd & 4th grades. More Grade 5 students scored above the mean, and high numbers of students were in the upper ranges of scores: 27% between the 61-80 point range and 42% between the 81-100 point range.

Comparison of MI-LMU Test Scores and AAPPEL Scores

Proficiency Group	N1	N2	N3	N4	I1	I2	I3
N =	4	14	25	21	19	3	2
Mean	10.50	31.30	58.40	72.40	88.47	79.66	96.50
SD	9.19	19.45	28.51	23.20	15.50	25.73	6.36
Range of Scores	4-17	7-66	9-95	28-101	57-105	50-96	92-101

Despite the variability in the various distributions, the 5th grade MI-LMU mean scores appear to show the beginnings of more distinctive groupings via the AAPPL grouping categories, though there appears to be a noticeable clustering between the N4, I1, and I2 groups. To further explore the difference between group means, a one-way analysis of variance indicated a significant difference between group means ([$F_{(6, 77)} = 11.52$, $p = .0000$). Post-hoc Bonferroni tests of all pairwise comparisons indicated the N1 and N2 significantly different from the other groups, and the N3 group significantly different from the I1 group, most likely due to the smaller variability of the I1 scores. Again, some of the cell means are derived from extremely small sample sizes.

Summary Findings

In summary, it is important to stress that in both the AAPPL Interpretive Reading Scores and in the MI-LMU MAT Chinese Reading Test data, evidence is presented to indicate that learners are at increasing higher levels of reading proficiency as they progress through Grades 3, 4, and 5.
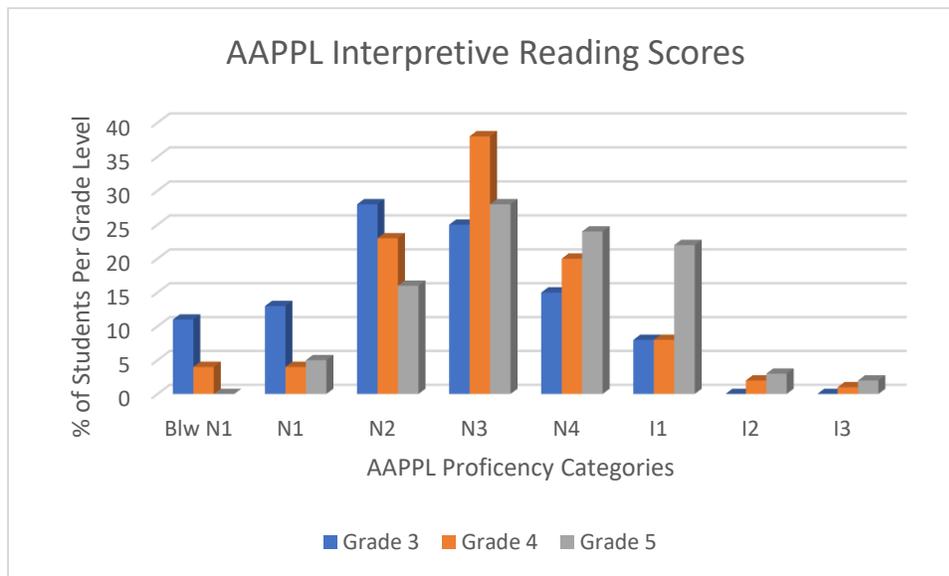

The chart above compares results from students in Grades 3, 4, and 5 in terms of the percentage of students who attained different proficiency levels on the AAPPL Interpretive Reading test. Grade 3 students are more prominent in achieving scores Below N1 and in the N1 range, and top off at the N2 range before demonstrating a decreasing though still impressive number of scores beginning with N3. Still, their scoring in the N4 and I1 range is notable.

The Grade 4 students show that their performances surpass the Below N1 and N1 range because the results show fewer scores in these two ranges as compared to the Grade 3 students. There is also a noticeable number of scores grouping in the N2, N3, N4 ranges. There is also evidence of a small number of Grade 4 student performances consistent with the Intermediate ranges.

Lastly, the Grade 5 students show that their performances exceed the Below N1 range of scores, with only 4% of their students scoring in the N1 level. With the 5th graders, the distinguishing characteristic appears to be not only their increase into the N2 range and performances falling solidly into the N3 and N4 proficiency ranges, but also the number of scores in the Intermediate range. As noted earlier, fully 27% of the 5th Graders who took the AAPPL Interpretive Reading test in this sample scored in the Intermediate range.

Turning to the MI-LMU MAT Chinese Reading Test data, comparing the three grades through comparing their performance on their respective tests also yields an interesting portrait of how the students demonstrate higher levels of performance in grade 5 compared to grade 3, and how their scores are reflecting trends in the AAPPL data.

Comparison of MI-LMU Scores by Grade Level
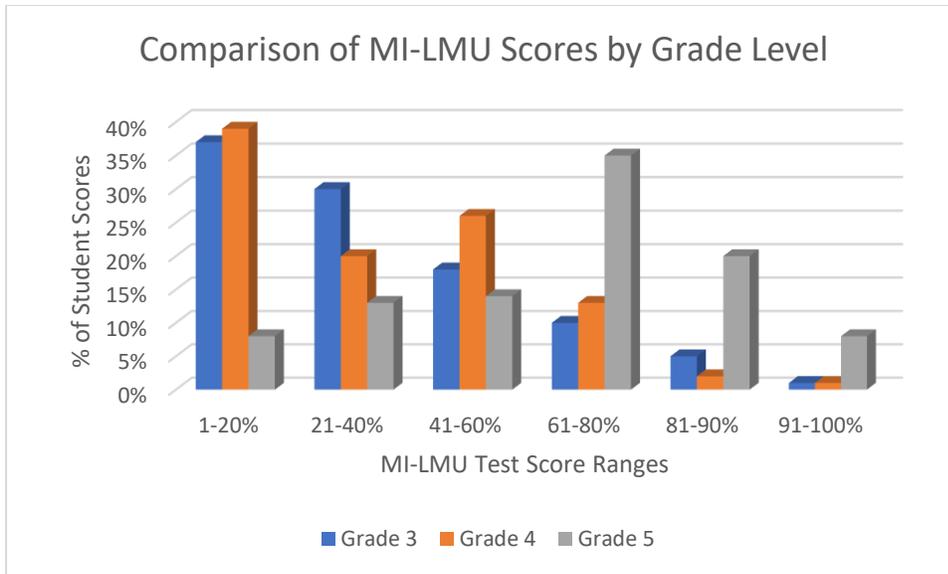
To illustrate this, the above graph depicts a comparison of how the Grade 3, 4, and 5 students performed on their MI-LMU MAT Chinese Reading tests. For this graph, the Grade 3, 4, and 5 test results have been converted to percentages, and divided into groups representing five ranges of percentage scores: 1-20%, 21-40%, 41-60%, 81-90%, and 91-100%. For each grade level, the percentage of student scores that fell within each range is depicted. It should be noted that while the tests for each of the grade level are different (as described in this report), it is useful to gain a general comparison of how students at different grade levels are performing on tests employing similar formats, but are longer with more challenging tasks, at each level.

Using this basis of comparison, we can observe that the test was challenging for the Grade 3 students, with 67% of the students showing a score of 40% or lower. The scores then decrease whereby 18% of the students scored between 41-60%, 10% scored between 61-80%, 5% scored between 81-90%, and 1% scoring between 91-100%. On the AAPPL scores, the 3rd grade students were more prominently grouped at and below the N3 range, although the MI-LMU data indicate that the rigors of reading longer texts with more complicated tasks was a challenging transition for many of these learners.

For the Grade 4 learners, the number of students who scored 40% or lower on their particular test decreases as compared to the Grade 3 learners, with more scores falling into the 41-60% range. A small group of students, 3%, scored between 80-100%. Given that the Grade 4 students were taking a test where the paragraphs were longer and more dense in terms of detail and cultural information, this can be taken as a sign of progress. Interestingly, there is a trend in the AAPPL data for 4th graders of them showing improvement in the N2-N4 ranges with their scores in the MI-LMU mid-ranges (21-80%) also improving.

Most striking, however, is the distribution of the 5th Grade scores. While there are clear indications of fewer students scoring in the lower test score bands, the percentage of Grade 5 students (63%) scoring between 61-100% on the test is striking. The distribution also shows consistent increases in scores for the Grade 5 students in the 1-60 point range. Thus, there is evidence of more students in Grade 5 adopting to the rigors of dealing with longer texts that

require them to perform tasks of greater complexity, possibly indicating that Grade 5 might be a point where students are beginning to move more uniformly into Intermediate level reading proficiency. As has already been noted, the 5th graders also showed impressive gains in the AAPPL scores not only in the N3 and N4 proficiency ranges, but most noticeably in the number of students in the Intermediate range. As noted earlier, fully 27% of the 5th Graders who took the AAPPL Interpretive Reading test in this sample scored in the Intermediate range.

Lastly, for students who took both the AAPPL Interpretive Reading test and the MI-LMU MAT, an exploratory analysis was conducted to determine the extent to which the AAPPL scores grouped the MI-LMU MAT scores. The results indicated that while there were scoring clusters of MI-LMU MAT scores that could be detected in the various AAPPL proficiency groupings, the extreme ranges of the MI-LMU MAT scores made meaningful interpretations of the grouping very difficult.

ACTFL AMERICAN COUNCIL ON THE
TEACHING OF FOREIGN LANGUAGES

INTERPRETIVE READING AND INTERPRETIVE LISTENING

Score	AAPPL Score Description	Strategy
N-1	Your AAPPL Interpretive Reading/Listening score of N-1 means that you understand individual words. You need to read/hear something more than once. You need visual cues such as pictures in order to understand what you read/hear.	Try to understand by looking for clues such as pictures, titles, words you recognize, words that are similar to words in your own language.
N-2	Your AAPPL Interpretive Reading/Listening score of N-2 means that you understand words and phrases. You may need to read/hear something more than once. You need visual cues such as pictures, your own knowledge of a topic, and words that are similar to English in order to understand what you read/hear.	Try to understand more by looking/listening for clues such as pictures, titles, words or parts of words that you recognize, words that are similar to words in your own language. Think about what you already know about a topic for additional clues about the topic. Then go back and reread/re-listen and see what more you can understand.
N-3	Your AAPPL Interpretive Reading/Listening score of N-3 means that you understand words, phrases and an occasional simple sentence. You may need to read/hear something more than once. You may need visual cues such as pictures, your own knowledge of a topic, familiarity with parts of words or words that are similar to English in order to understand what you read/hear.	Use what the passage says about the topic to understand the main idea. Try to use different approaches to understanding such as looking/listening for clues such as pictures, titles, words or parts of words that you recognize, words that are similar to words in your own language. Think about what you already know about a topic for additional clues about the topic. Then go back and reread/re-listen and see what more you can understand.
N-4	Your AAPPL Interpretive Reading/Listening score of N-4 means that you likely understand the main idea in short passages on very familiar topics but rarely the supporting facts. You usually need to read/hear the passage more than once. You may need visual cues such as pictures, your own knowledge of a topic, familiarity with parts of words or words that are similar to English in order to understand what you read/hear.	Use what the passage says about the topic to understand the main idea and look for some supporting facts. Try to use different approaches to understanding such as looking/listening for clues such as pictures, titles, words or parts of words that you recognize, words that are similar to words in your own language. Use what you already know about a topic for additional clues about the topic. Then go back and reread/re-listen and see what more you can understand.

Score	AAPPL Score Description	Strategy
I-1	Your AAPPL Interpretive Reading/Listening score of I-1 means that you likely understand the main idea and possibly some supporting facts in short passages on very familiar topics. You usually need to read/hear the passage more than once. You usually need visual cues, context clues, and prior knowledge to help you understand what you read/hear.	Use what the passage says about the topic to understand the main idea and look for some supporting facts. Try to use different approaches to understanding such as looking/listening for clues such as words or parts of words that you recognize, words that are similar to words in your own language, word order and form. Use what you already know about a topic for additional clues about what the author wants you to understand.
I-2	Your AAPPL Interpretive Reading/Listening score of I-2 means that you understand main ideas and supporting facts in short passages on familiar topics. You may need to read/hear the passage more than once. You may need visual cues, context clues, and prior knowledge to help you understand what you read/hear.	Practice reading/listening to longer passages or simple stories. To deepen your understanding start to look at how the passage is organized. Compare what you read or hear to what you already know.
I-3	Your AAPPL Interpretive Reading/Listening score of I-3 means that you understand main ideas and supporting facts in short passages on familiar topics. You may need to read/hear complex passages more than once. Context clues or prior knowledge may help you understand what you read/hear.	Frequently read/listen to longer passages or simple stories. To deepen your understanding start to look at how the passage is organized. Compare what you read or hear to what you already know and to what you expected to find out.
I-4	Your AAPPL Interpretive Reading/Listening score of I-4 means that you fully understand main ideas and supporting facts in short passages, simple narratives and descriptive passages on familiar topics. You may need to read/hear complex passages more than once. Context clues or prior knowledge may help you understand what you read/hear.	Frequently read/listen to longer passages, stories, news reports. To deepen your understanding start to look at how the passage is organized. Pause frequently to check your understanding. Compare what you read or hear to what you already know and to what you expected to find out. See what the details add up to? Any new insight?
I-5	Your AAPPL Interpretive Reading/Listening score of I-5 means that you fully understand and with ease main ideas and supporting facts in short passages, simple narratives and descriptive passages on familiar topics. In addition, you can understand some more complex passages on less familiar topics. You may need to read/hear complex passages more than once. Your knowledge of the language, your ability to tell the difference between a main idea and a detail, and your ability to use context clues and inferencing may help you understand what you read/hear.	Frequently read/listen to longer passages, stories, news reports on lots of different topics. Use the organization of the text to help you read and listen more efficiently. Pause frequently to check your understanding. Compare what you read or hear to what you already know and to what you expected to find out. See what the details add up to? Any new insight?

Score	AAPPL Score Description	Strategy
A	Your AAPPL Interpretive Reading/Listening score of A means that you are able to understand conventional narrative and descriptive texts that contain a clear underlying structure across major time frames, though comprehension may be uneven. Main ideas and supporting details are comprehensible with familiar situational subject matter containing high-frequency vocabulary and structures. Some new subject matter in areas of general interest will be generally understood.	Continue developing your skills by reading text or listening to passages that contain extended narration and more complex language on topics beyond the personal and general. Read articles, and other written works, or listen to spoken language about community or world issues. Develop strategies to understand others' opinions and support of those opinions as well as the abstract treatment of topics and discussion of a hypothetical nature.