



# Learner Performance in Mandarin Immersion and High School World Language Programs: A Comparison

Xiaoqiu Xu

*Pearson Knowledge Technologies*

Amado M. Padilla

*Stanford University*

Duarte M. Silva

*Stanford University*

**Abstract:** *This study compared the Mandarin performance of elementary immersion program students and high school world language program students in the same school district. A cross-sectional design was employed to gather information on Mandarin proficiency of fourth and fifth graders and Level 4 and Level 5 (AP Chinese) high school students who took the Mandarin Standards-Based Measurement of Proficiency assessment at the end of the school year. Results indicated that immersion students slightly outperformed the comparison high school group in reading but lagged slightly behind in writing and speaking skills, for which assessment tasks required higher levels of cognitive awareness. Findings also showed that while only a few nonheritage speakers in the high school world language program continued to Level 5 (AP), most nonheritage speakers remained in the immersion program for the full duration and performed as well, or nearly as well, as the heritage speakers when exiting the program. Further studies are needed, but the findings provide a strong rationale for the role of Mandarin immersion programs in assisting students to develop higher levels of linguistic proficiency in Mandarin.*

**Key words:** *AP Chinese, Mandarin, immersion program, high school Mandarin, proficiency assessment*

---

*Xiaoqiu Xu (PhD, Stanford University) is Test Development Manager at Pearson Knowledge Technologies, Menlo Park, CA.*

*Amado M. Padilla (PhD, University of New Mexico) is Professor and Chair of Developmental and Psychological Sciences, Stanford Graduate School of Education, Stanford, CA.*

*Duarte M. Silva, (EdD, University of San Francisco) is Executive Director of the California World Language Project, Stanford Graduate School of Education, Stanford, CA.*

## Introduction

In the past two decades, there has been a dramatic increase in the study of Chinese<sup>1</sup> in the United States: The number of Chinese language programs in the United States for learners of all ages, from elementary school through adulthood, tripled between 1995 and 2005 and has continued to expand during the last 10 years (Dobuzinskis, 2011; Neely, 2011; Rogers, 2012). In 2012, approximately 125 schools, mostly at the elementary level, offered Mandarin immersion programs, with the largest number of such programs on the West Coast (Worthen, 2012).

However, data on learners' developing mastery of either Mandarin or English for children in such programs are limited. Lindholm-Leary (2011) reported students' ratings of their own Chinese oral language proficiency, but external assessment data validating students' self-reports were not provided. Chang (2011) examined the impact of Mandarin instruction on learners' academic, attitudinal, and cognitive development. While Chang stated that parents and teachers believed that students' academic work was enhanced thanks to instruction in Mandarin, no quantitative data were offered to show that instruction actually led to increased academic or cognitive development.

Another large-scale study on students' language performance in dual-immersion programs was conducted by the Second Language Teaching and Research Center at the University of Utah (Rubio, 2014). All available third to fifth graders enrolled in Chinese, Spanish, and French dual-immersion programs were tested on their language proficiency using the ACTFL Assessment of Performance Toward Proficiency in Languages (AAPPL). Preliminary results showed that more than 75% of Mandarin dual-immersion students performed at or above the stated "target" at each grade level in terms of speaking, which for third graders was Novice Mid and for fourth and fifth graders was Novice High. However, data showed that students faced greater challenges on the reading portion of the AAPPL: At the end of fourth grade, 78% of the students were

rated below the stated targeted level—Novice High—in Mandarin (Rubio, 2014, n.p.). Unfortunately, data on reading performance were not provided for the third and fifth graders in this study, nor was writing proficiency measured for any of the Mandarin immersion programs.

Recently, a longitudinal study of students in a Mandarin/English dual-immersion program (Padilla, Fan, Xu, & Silva, 2013) followed students from kindergarten, when they entered the immersion program, through fifth grade, when they exited that program. Padilla et al. reported data for all second- through fifth-grade students on mandated California standards tests for English language arts, writing (fourth grade), math, and science (fifth grade): Data showed that while the nonimmersion second- and third-grade students from the same school, who were taught in English, had higher scores on the English language arts and math tests than the immersion students in the same grades, Mandarin immersion students in the upper grades scored higher on these tests than their nonimmersion peers. In addition, Padilla et al. reported data on immersion students' oral/listening, reading, and writing performance in Mandarin using program-created assessment measures. As expected, results showed that as students moved from one grade level to the next, they acquired increasing mastery in oral/listening, reading, and writing in Mandarin. Interestingly, a comparison of Mandarin heritage students with nonheritage students in the immersion program initially favored heritage students; however, this advantage was not statistically significant in the later grades. Finally, in fifth grade, students' progress toward proficiency in speaking, reading, and writing was measured using the Mandarin Standards-Based Measurement of Proficiency (STAMP) 4Se test. These data were initially reported in Padilla et al. and showed that students exiting the Mandarin dual-immersion program had acquired an intermediate level of competency in speaking, reading, and writing. Upon completion of this study and prompted by parents of exiting

fifth-grade immersion students, a new research question emerged: How does the proficiency of students who exit from a dual-language immersion program compare with that of high school students who study the same language but as part of their high school world languages curriculum? Since a search of the existing literature did not yield answers to this question, this current follow-up study was undertaken.

## Methods

### *Context*

Both the elementary Mandarin immersion school and the high schools were located in a suburban school district of approximately 12,500 students in an upper-/middle-income community in Northern California. The district is known for its high academic performance, with a graduation rate of 98%, and most high school graduates continue on to higher education.

### *Participants*

#### **Elementary Immersion Students**

The elementary Mandarin immersion program initially started in fall 2008 with 20 kindergarten and 20 first-grade students divided into two K/1 classrooms of 20 students each. In subsequent years, 20 new kindergarten students were added each year. Children whose home language was Mandarin were given an oral interview to assess their age-appropriate proficiency in Mandarin and, based on these results, were then selected as the native Mandarin speakers. Simultaneously, the names of non-Mandarin-speaking children were entered into a lottery from which the other half of each new class of students was selected. A total of 48 Mandarin immersion students participated in the study,<sup>2</sup> including 18 fourth graders (all from the 2014 school year,<sup>3</sup> including nine heritage learners) and 30 Mandarin immersion fifth graders (14 students from the 2013 school year and 16 students from the 2014 school year, including 11 heritage learners).

#### **Elementary Immersion Teachers**

The six immersion teachers were all native Mandarin speakers who completed most of their education in China and Taiwan and who received a multiple subject credential in bilingual education from universities in California with a specialization in Mandarin.

#### **High School Mandarin Students**

Mandarin was first offered in one high school beginning in fall 2006 and was made available at the other high school in fall 2007. Students chose to enroll in Mandarin as part of their world language elective. Nonheritage students were placed into a Level 1 class; heritage students or learners who had received prior Mandarin instruction were placed into an appropriate level based on a teacher-developed assessment. A total of 119 high school students participated in the study. The 95 Level 4 students (48 students from the 2009 school year and 47 students from the 2010 school year, including 52 heritage learners) had completed four years of instruction or its equivalent. The 24 Level 5 AP students (all from the 2010 school year,<sup>4</sup> including 19 heritage learners) had also had four years of classroom instruction, plus two consecutive summer intensive Mandarin programs (5.5 hours per day for four weeks in each summer), which was approximately equivalent to two additional semesters of Mandarin instruction (Xu, Padilla, & Silva, 2012).

#### **High School Mandarin Teachers**

The first Mandarin teacher was a native speaker of Japanese who had studied Chinese linguistics at both the undergraduate and graduate level. He served as the primary consultant when the Mandarin immersion program was being planned and was instrumental in the design of the curriculum. He also served as the cooperating teacher for the other two teachers during the student teaching phase of their teacher preparation program. These two high school teachers were native Mandarin speakers who had completed most of their education in China or Taiwan and who received a single subject

credential in world language education from universities in California with a specialization in Mandarin.

### *Instruction*

Both the Mandarin immersion and world language programs used simplified characters and pinyin in instruction, emphasized correct form and stroke order for writing in Chinese, and offered explicit language instruction in reading and writing Chinese characters. The Mandarin immersion program followed an 80–20% model of Mandarin/English instruction in the K/1 class, 60–40% in the second-/third-grade class, and 50–50% in the fourth/fifth grades. It had several innovative features. First, following the mixed-grade philosophy of the school, mixed-grade classes (K/1, 2/3, 4/5) facilitated collaboration between older and younger children. Younger students in a mixed-grade classroom had the older students who could serve as linguistic and content guides. Second, the program adopted an “open education” approach that sought to help each child acquire the tools of learning—reading, language, mathematics—in a way that would allow the child to experience the joy of learning. Third, parent involvement was an important and crucial element of the program. Parents were actively involved in the program by assisting teachers in teaching subjects, organizing and leading activities, or maintaining creative and productive classroom environments.

In the high school program, because the two native speaker teachers had trained under the direction of the program director, all three teachers had similar philosophies of instruction and similar pedagogical approaches. Specifically, all three teachers used an immersive, standards-based communicative approach to instruction; thus, learners received instruction mostly in Mandarin, even at the lower levels. English was used only to communicate classroom rules and explain complicated tasks. To motivate students to maximize their use of oral language, a system for rewarding

oral language participation was used. Students also spent at least one hour per week in the language lab writing in Mandarin and completing other computer-assisted language learning tasks.

### *Assessments*

The STAMP4S and the STAMP4Se were selected as the measurements of the study because they are based on the same standards and performance guidelines and can provide direct comparisons between the cohort groups. On each version of the assessment, learners must demonstrate sustained performance at a specific language level prior to advancing to the next level (Avant Assessment, 2014b, 2014c). The STAMP4S is designed for adolescent and adult learners. High school students who were completing the fourth and fifth (AP) levels of instruction in Mandarin were assessed using the STAMP4S in speaking, reading, and writing. The STAMP4Se is designed for younger learners and includes comparable, but age-appropriate, content and item prompts. Fourth- and fifth-grade students in the Mandarin immersion program were assessed using the STAMP4Se on the same skills. Both the STAMP4S and the STAMP4Se are computerized (online) and standardized assessments. During this study, external researchers and teachers administered the STAMP4S and the STAMP4Se to the students in their own classrooms toward the end of the semester with all of the results reported directly to Avant Assessment. The rationale for proctoring the exam in the students’ own classrooms was to lower the learners’ affective filter and have the assessment parallel to the greatest extent possible a normal learning experience.

In terms of the levels measured in the STAMP tests, the Avant Assessment literature states that STAMP levels are “related to” and “defined by” the proficiency levels and sublevels that are described in the ACTFL Proficiency Guidelines (e.g., Novice High, Intermediate Mid; ACTFL, 2012). Some STAMP results are reported using

those designations; however, they are not equivalent to official ACTFL ratings (Avant Assessment, 2014a). It should be pointed out, however, that the scales for assessing students' writing on the STAMP4S differ slightly from those used to assess writing samples on the STAMP4Se: The STAMP4S Level 5 for writing is described as Intermediate Mid for adolescent and adult learners but is categorized as Intermediate Mid or Intermediate High for younger students. Similarly, the STAMP4S Level 6 corresponds to Intermediate High for adolescent and adult learners but Advanced Low for elementary students. Therefore, in the analysis of the writing and speaking sections, Levels 5 and 6 were combined and marked as "Intermediate Low and above."

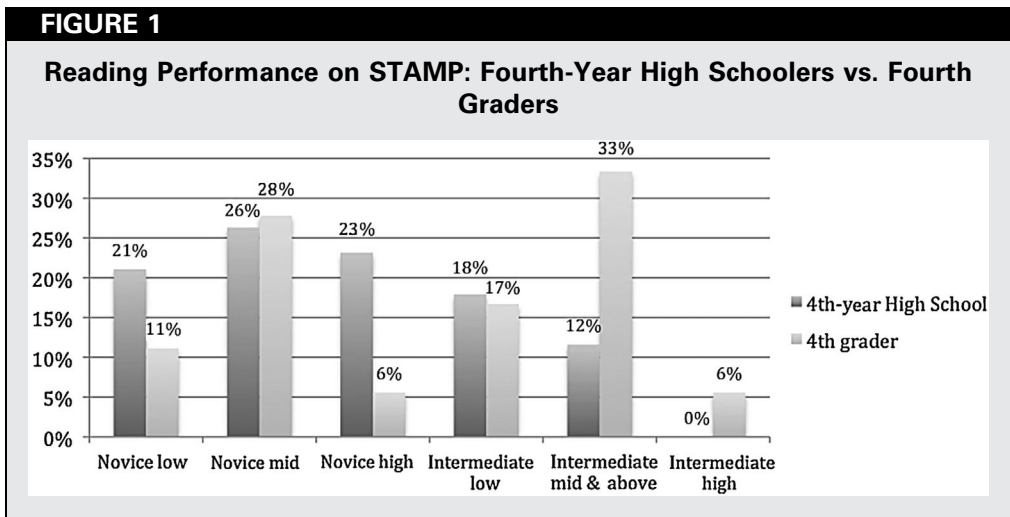
**Results**

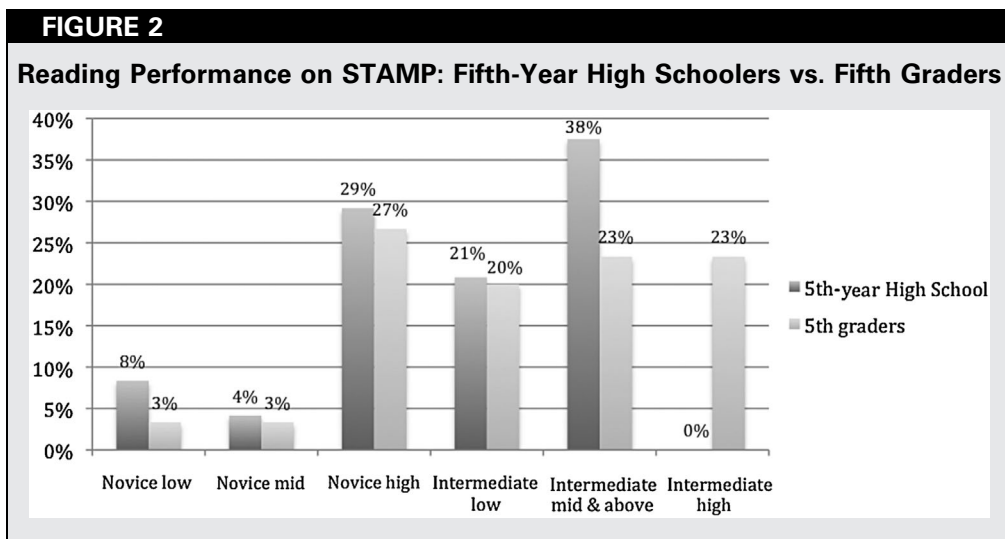
*Reading Performance*

Figure 1 shows that, by the end of the fourth year of Mandarin instruction, 70% of high school students were rated at the Novice Low, Novice Mid, and Novice High levels, and the remaining 30% were all rated either Intermediate Low or Intermediate Mid. No fourth-year high school students achieved ratings of Intermediate High. In comparison, at the end of fourth grade, 44% of immersion students were rated at the Novice level but

56% were at the Intermediate level, including 6% who were rated Intermediate High. The independent-sample t test between the fourth-year high school students and fourth graders favored the immersion students, and the difference showed a distinct trend toward statistical significance ( $t [116] = 1.84, p = 0.07$ ). Nonetheless, the fourth graders showed a trend in the direction of higher reading scores on the STAMP reading proficiency than the exiting fourth-year high school students.

Data comparing the reading scores for high school students who had completed five years of Mandarin instruction in the world language program and fifth-grade immersion students who were exiting the elementary Mandarin immersion program are presented in Figure 2. By the end of the fifth year of Mandarin (AP) instruction, 41% of the high school students were rated at a Novice level in reading and 59% were rated Intermediate Low or Intermediate Mid. At the time of the study, the district, in collaboration with a STARTALK program (About STARTALK, n.d.), offered the opportunity for students to participate in a two-consecutive-summers program, the equivalent of an additional year of study, and become eligible to enroll in the fifth-year AP Mandarin course in their senior year. No fifth-year AP student was rated Intermediate High in reading. However,





33% of the fifth graders in the immersion program were rated in the Novice range, and 67% were rated at the Intermediate level. Of the elementary students who were rated at the Intermediate level, 23% were rated Intermediate High. The independent-sample *t* test between the fifth-year high school students and fifth graders failed to attain statistical significance ( $t [52] = 1.49, p > 0.1$ ). Interestingly, though, the descriptive analysis still showed that Mandarin immersion students achieved overall higher proficiency in reading in Mandarin, as 23% of the immersion students reached Intermediate High while none of the high school students reached that level.

**Writing Performance**

As shown in Figure 3, by the end of the fourth year of Mandarin instruction, 47% of high school students were rated at the Novice level in writing and 53% were at rated at the Intermediate level, with ratings predominantly of Intermediate Mid and Intermediate High.<sup>5</sup> In comparison, at the end of fourth grade, only 22% of the immersion students scored at the Intermediate level, all receiving ratings of Intermediate Low, while the majority (85%) of the fourth graders were rated in the Novice range. The independent-sample *t* test between the

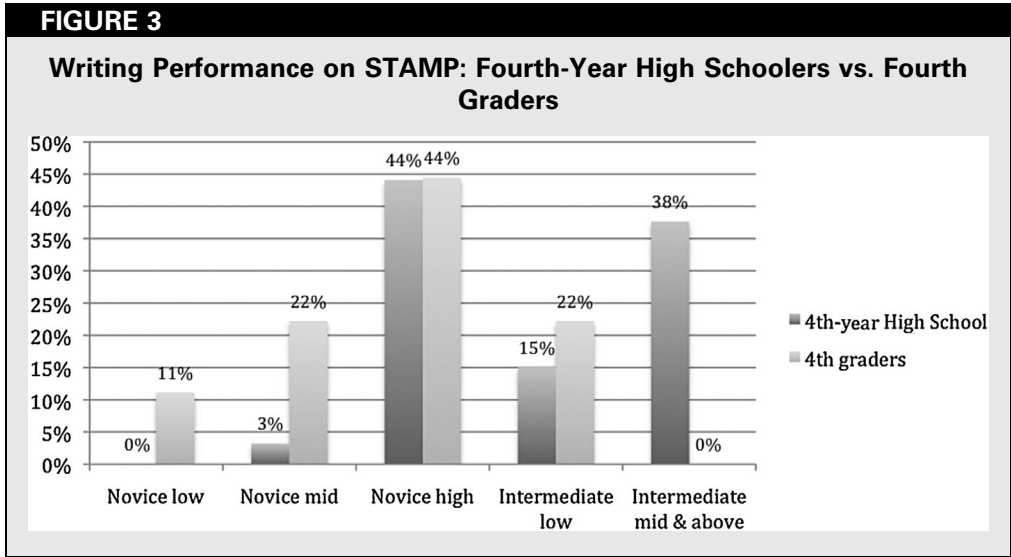
fourth-year high school students and fourth graders demonstrated significant differences in their ratings for writing ( $t [109] = 4.33, p < 0.01$ ).

As Figure 4 depicts, by the end of the fifth year of study, only 4% of the AP high school students were rated in the Novice range; the remaining 96% were all rated at the Intermediate level. In contrast, 39% of the fifth graders were still rated at the Novice level. The independent-sample *t* test between the fifth-year high school students and the immersion students demonstrated a significant difference in writing ( $t [50] = 5.01, p < 0.01$ ).

**Speaking Performance**

As shown in Figure 5, fourth-year high school Mandarin students consistently performed about one level higher than the fourth-grade immersion students. The independent-sample *t* test between the fourth-year high school students and fourth graders was statistically significant ( $t [95] = 2.20, p < 0.05$ ).

When comparing the speaking of fifth-year high school AP students with that of the fifth graders who were exiting the Mandarin immersion program, the researchers found an even more obvious performance gap. Figure 6 shows that most of the high school students (71%) scored at an



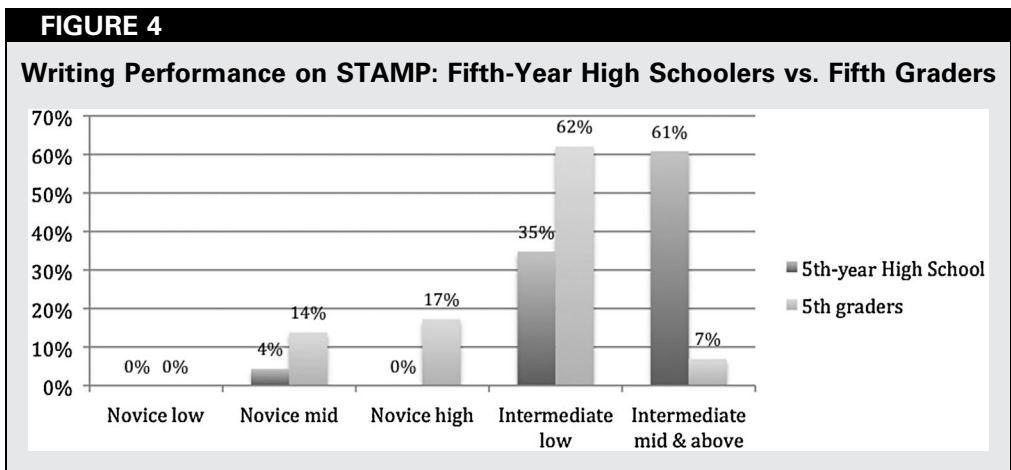
Intermediate Mid level or above while most of the immersion students (64%) scored one level lower, receiving ratings of Intermediate Low. The independent-sample *t* test between the fifth-year high school students and fifth graders demonstrated a significant difference ( $t [43] = 3.47, p < 0.01$ ).

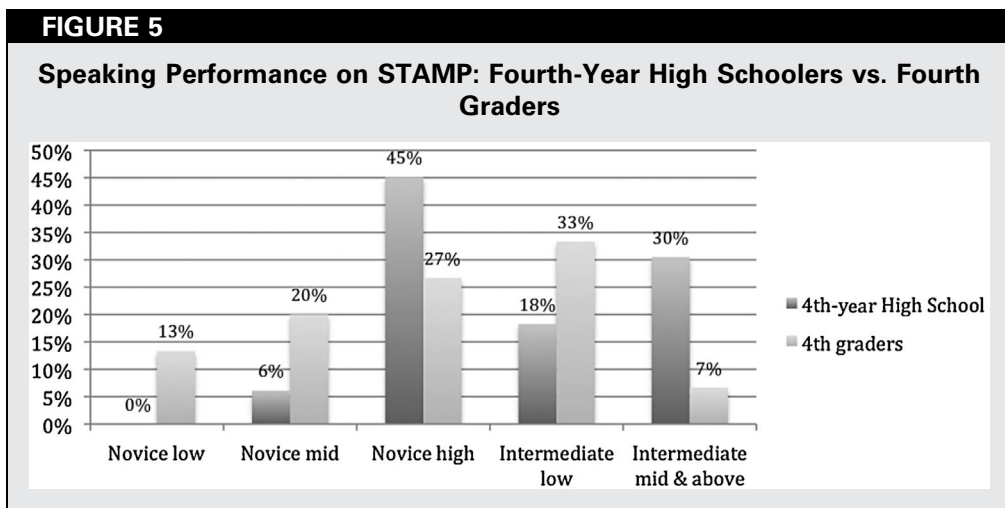
*Overall Comparison and Heritage vs. Nonheritage Performances*

To reveal the overall comparison of the Mandarin language proficiency between the immersion and high school students, Tables 1 and 2 display the medians of stu-

dents' performance on the STAMP4S and STAMP4Se. Results also include the comparison of medians of heritage and nonheritage student test scores in the two programs. Finally, the total number of participants from both programs as well as the numbers of heritage and nonheritage students at each targeted program level are also provided.

Median scores in the tables facilitate comparisons between the fourth-year high school students and fourth-grade immersion learners and between the fifth-year (AP) high school students and fifth graders



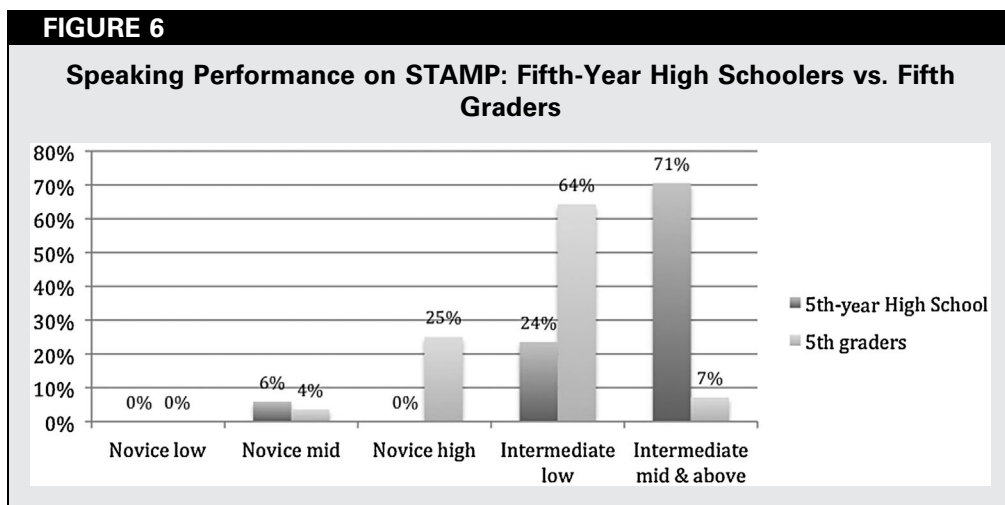


at the conclusion of the elementary immersion program. The tables also provide data on heritage and nonheritage learners for each grade/year of study.

Inspection of Tables 1 and 2 reveals that, in the high school, at the end of the fourth-year level of Mandarin instruction, there was a balanced distribution of heritage ( $n = 52$ ) and nonheritage students ( $n = 43$ ). However, by the end of the fifth year of Mandarin AP instruction, only a few nonheritage students remained ( $n = 5$ ), whereas the majority were heritage speakers ( $n = 19$ ). In contrast, in the immersion program, by the end of fifth grade, there were slightly more nonheritage speakers ( $n = 14$ ) than heritage speakers

( $n = 11$ ). Data showed that the medians of the heritage and nonheritage speakers in the Mandarin immersion program were close, even nearly identical, to each other, indicating that these elementary-level, immersion nonheritage speakers performed as well as their elementary heritage counterparts across all three language skills as assessed by the STAMP4Se.

Together, Tables 1 and 2 show that very few high school nonheritage speakers advanced into the AP Level 5 Mandarin class, whereas in the immersion program, nonheritage speakers remained in the program for the full duration from kindergarten through fifth grade and by the end of that





**TABLE 1**

**Fourth-Year Comparison, Including Heritage and Nonheritage Students**

All	Reading	Writing	Speaking
High School	3	4	3.5
High School ( <i>n</i> )	95	93*	82
Immersion	4	3	3
Immersion ( <i>n</i> )	18	18	18
Heritage	Reading	Writing	Speaking
High School	3	4	4
High School ( <i>n</i> )	52	51	46
Immersion	4	3	3
Immersion ( <i>n</i> )	9	9	9
Nonheritage	Reading	Writing	Speaking
High School	2	3.3	3.3
High School ( <i>n</i> )	43	42	36
Immersion	2	3	3.5
Immersion ( <i>n</i> )	9	9	9

\* A mismatch of the numbers for reading, writing, and speaking can be observed in the tables due to technical issues encountered in the tests. A few tests were stopped in the middle of the writing or speaking sections due to an Internet connection problem. A few other tests were rated “unratable” in the speaking section due to headset problems (line noises or silence).

program had skills that were comparable, or nearly comparable, to those of the elementary heritage speakers.

**Discussion**

Chinese has often been categorized as one of the most challenging languages to learn as a world language. The Foreign Service Institute of the U.S. Department of State has categorized Mandarin Chinese as one of just a handful of languages assigned to Category III languages, which are exceptionally difficult for native English speakers to learn (Defense Language Institute Foreign Language Center, 1995). Teachers, researchers, and parents of students who are enrolled in or have exited a Mandarin immersion program have often asked about the proficiency of exiting immersion program students in comparison with high school students who have studied the same language but as part of a

world languages program in high school. This may be an even more salient question to ask when the language is as challenging to learn as Chinese, which typically requires many years of study. This investigation has addressed this concern: Evidence shows that both groups of students attained a relatively high level of performance in Mandarin as measured by the STAMP assessments and provides strong support for the level of proficiency that can be gained in a Mandarin immersion program when compared to that of students in a high school Mandarin program who have completed fourth-year Mandarin and/or AP Chinese. When the high schoolers exited the world languages program at the conclusion of Mandarin 5 (AP level), the medians for their performances in reading, writing, and speaking were respectively at Intermediate Low, Intermediate Mid, and Intermediate High levels. In contrast, the elementary immersion

**TABLE 2****Fifth-Year Comparison, Including Heritage and Nonheritage Students**

All	Reading	Writing	Speaking
High School	4	5	5
High School ( <i>n</i> )	24	23	17
Immersion	4	4	4
Immersion ( <i>n</i> )	30	29	28
Heritage	Reading	Writing	Speaking
High School	4	5	5
High School ( <i>n</i> )	19	18	12
Immersion	4	4	4
Immersion ( <i>n</i> )	11	11	9
Nonheritage	Reading	Writing	Speaking
High School	2	4.3	4
High School ( <i>n</i> )	5	5	5
Immersion	4.5	4	4
Immersion ( <i>n</i> )	14	14	14

students, who were on average at least six years younger than the high school students, attained proficiency levels that were nearly the same as the comparison high school student group in reading, writing, and speaking.

The distribution and the *t* tests of the STAMP levels in both programs reveal that the immersion students performed slightly better than the high school students in the interpretive mode, reading, perhaps because the STAMP reading items were all multiple-choice questions that measured test-takers' reading comprehension in reference to age-appropriate content. However, the immersion students lagged slightly behind the high school students in the presentational modes of writing and speaking. This could possibly be related to the nature of the test items presented in the STAMP assessment instruments: The writing items were all open-ended questions that required the writer to develop a topic theme and to more fully elaborate with a detailed written answer to the prompt. Differences in students' scores, then, across the two instructional programs may be due to differences between the younger and the older students' level of cognitive

development and skills (Skehan, 1998). Interestingly, a closer examination of the data on immersion students reveals that they made a tremendous jump in writing proficiency from fourth grade to fifth grade (15% vs. 61% attained Intermediate status).

As in the writing section, the speaking section of the STAMP test also required that students respond to open-ended questions, which also may have benefited the older test-takers who are more likely to have higher levels of cognitive ability as well as greater facility and experience structuring and articulating their thoughts in any language, including in the second language.

Another noticeable finding was that only a few nonheritage speakers in the high school world languages program proceeded to the AP level: Among the participants in this study, only 5 out of 24 were nonheritage speakers. In contrast, the fifth-grade immersion classes included a greater number of nonheritage speakers (see Table 2) than heritage students. Interestingly, even when taking into account scores from a greater percentage of nonheritage speakers, the immersion students still

performed at higher levels than the high school students in STAMP reading proficiency. It is thus encouraging that nonheritage speakers remained in the immersion program for the full duration and performed as well as their heritage speaker counterparts in all Mandarin skills when exiting the program.

This finding supports the critical period hypothesis (Lightbown & Spada, 2006), which claims that in order for a second language learner to achieve a native-like proficiency, he or she should ideally learn the language before the age of puberty. Considering the fact that Mandarin is undoubtedly a challenging language to acquire and if the goal of learning Mandarin is to achieve native-like proficiency, it is essential that the acquisition of Mandarin begin as early as possible and that learners engage in longer and continuous learning sequences. The findings of the current study clearly provide a strong rationale for the role of Mandarin immersion programs in helping students to attain high—and very desirable—levels of proficiency in Mandarin. Since students exiting the Mandarin immersion program at the end of fourth and fifth grade were performing at Intermediate levels of proficiency on the STAMP4Se, it is likely that they could attain Advanced levels of linguistic performance with continued study during their middle and high school years.

Finally, it is important to note several limitations. First, a relatively small sample size limits this study's generalizability. Second, all participants in this study were from the same upper-/middle-class suburban school district, which is well known for the excellence of its schools (<http://www.usnews.com/education/best-high-schools/California>). Therefore, the overall performance of the high school and immersion students when exiting the programs may not be representative of all Mandarin programs in the United States. Third, the STAMP assessment has limited item types and only measures overall language skills in reading, speaking, and writing. It would be interesting to assess other linguistic sub-

skills, such as pronunciation, tone, or fluency, and compare differences between the high school and immersion students. In addition, toward the higher scales of the STAMP speaking and reading assessments, the different measurement scales for immersion and high school students are not perfectly aligned with each other, which may cause a biased comparison unless data are carefully reported. Future research with a larger sample size, a randomized research design, and more robust and detailed measurements is called for.

## Conclusion

This study reveals timely and useful findings that compare the performance in reading, writing, and speaking of students enrolled in an elementary Mandarin immersion program with that of fourth- and fifth-year students in a high school world languages program. Results demonstrated that the immersion students, at an early stage of their schooling, can attain high levels of performance in Mandarin that are nearly comparable to those attained by high school students in advanced-level classes. This study may thus provide a reference for setting both Chinese immersion program goals and student benchmarks at different levels of instruction.

Of equal importance, the study suggests that participation in a Mandarin immersion program during the elementary grades positions students to achieve extremely high levels of Mandarin proficiency at the conclusion of sustained, K–12 sequences of instruction. Thus, the data also provide a guide to language educators for planning and implementing well-articulated K–5, middle, and high school world language programs so as to ensure that students who wish to continue with the study of Mandarin beyond elementary school are offered sustained opportunities to do so. The resulting level of proficiency will no doubt equip these learners with a level of Mandarin proficiency that will be a valuable asset to themselves and to the nation.

## Acknowledgments

The authors wish to acknowledge the support of the Foreign Language Assistance Program funded by the U.S. Department of Education and the STARTALK project funded by the National Security Agency, via the National Foreign Language Center, which supported the instructional programs researched for this study. In addition, the authors are grateful for our partnership district, the instructors who taught in the two programs studied, and the students who earnestly engaged in the learning of Mandarin and demonstrated what they had learned. Finally, we wish to thank Lorraine Fan for her support with data analysis and Anne Nerenz and the anonymous reviewers whose comments and suggestions strengthen our article.

## Notes

1. In China, there are two major common Chinese languages: Mandarin and Cantonese. As Mandarin is the standardized language used in China, it has become increasingly popular worldwide. In this study, the terms Chinese and Mandarin are used interchangeably.
2. Two fourth graders and five fifth graders in the immersion program missed the test due to sickness or traveling during the time of assessment.
3. The STAMP4Se was administered to only fifth graders in 2013. Based on that experience, we were comfortable with delivering the test to fourth graders; therefore, we administered the test to both fourth and fifth graders in 2013.
4. As the world languages program was established in 2006, there were no Level 5 AP students in 2009.
5. As mentioned in the Methods section, as the scales of STAMP4S and STAMP4Se differ slightly on Levels 5 and 6, all performances at or above Intermediate Mid were put together and reported as “Intermediate Mid and above.”

## References

- About STARTALK. (n.d.). About STARTALK. Retrieved January 7, 2015, from <https://startalk.umd.edu/about>
- ACTFL. (2012). ACTFL proficiency guidelines 2012. Alexandria, VA: Author
- Avant Assessment. (2014a). STAMP4S score interpretation guide. Retrieved November 13, 2014, from [http://avantassessment.com/docs/stamp\\_scoreinterpretation.pdf](http://avantassessment.com/docs/stamp_scoreinterpretation.pdf)
- Avant Assessment. (2014b). STAMP4S benchmarks and rubric guide. Retrieved November 13, 2014, from [http://avantassessment.com/docs/stamp4s\\_benchmarkandrubicguide.pdf](http://avantassessment.com/docs/stamp4s_benchmarkandrubicguide.pdf)
- Avant Assessment. (2014c). STAMP4Se benchmarks and rubric guide. Retrieved November 13, 2014, from [http://avantassessment.com/docs/stamp4se\\_benchmarkandrubicguide.pdf](http://avantassessment.com/docs/stamp4se_benchmarkandrubicguide.pdf)
- Chang, W. D. (2011). Perceptions on the impact of K-5 Mandarin foreign language instruction on students' academic, attitudinal, and cognitive development (Unpublished doctoral dissertation). California State University, Northridge
- Defense Language Institute Foreign Language Center. (1995). Guidelines, policies and procedures for DoD command language programs. Retrieved November 13, 2014, from <http://www.ut.ngb.army.mil/clp/clpm/regs/musthave/DLIFLC%20Pam%20350-9.pdf>
- Dobuzinskis, A. (2011). More U.S. students learning Chinese as school language programs expand. Retrieved April 20, 2012, from [http://www.huffingtonpost.com/2011/04/21/more-us-students-learning\\_n\\_852093.html](http://www.huffingtonpost.com/2011/04/21/more-us-students-learning_n_852093.html)
- Lightbown, P., & Spada, N. M. (2006). *How languages are learned*, 3rd ed. Oxford: Oxford University Press
- Lindholm-Leary, K. (2011). Student outcomes in Chinese two-way immersion programs: Language proficiency, academic achievement, and student attitudes. In D. Tedick, D. Christian, & T. Fortune (Eds.), *Immersion education: Practices, policies, possibilities* (pp. 81–103). Avon, UK: Multilingual Matters
- Neely, B. (2011). More Americans learn their ABCs in Chinese. Retrieved April 20, 2012, from <http://www.npr.org/2011/01/19/133031008/american-interest-in-learning-chinese-skyrockets>

Padilla, A. M., Fan, L., Xu, X., & Silva, D. (2013). A Mandarin/English two-way immersion program: Language proficiency and academic achievement. *Foreign Language Annals*, 46, 661–679. doi: 10.1111/flan.12060

Rogers, M. (2012). Demand high for dual immersion programs in Utah. *The Salt Lake Tribune*. Retrieved November 1, 2014, from <http://www.sltrib.com/sltrib/news/..language-programs-Chinese-dual.html.csp>

Rubio, F. (2014, October). Overview of performance assessment results in the Utah dual language immersion program. Paper presented at the Fifth International Conference on Language Immersion Education, Salt Lake City, UT

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press

Worthen, B. (2012). More area schools embrace Chinese immersion method. *The Wall Street Journal*, August 22

Xu, X., Padilla, A. M., & Silva, D., (2012). Time factor in Mandarin language learning: Four week intensive vs. regular high school semester. *The Language Learning Journal*, 11, 26–39. doi:10.1080/09571736.2012.677054

*Submitted November 14, 2014*

*Accepted December 3, 2014*