THE ACADEMIC EFFECTS OF AN ELEMENTARY CHINESE IMMERSION PROGRAM

Sean R. Hill

A thesis submitted in partial fulfillment of
the requirements for the degree of
Master of Science

Department of Psychology

Central Michigan University
Mount Pleasant, Michigan
December 2014

Apprendre une autre langue, c'est comme le commencement d'une autre vie.

Michel Bouthot

This is dedicated to my family for
all of their support and patience
throughout this project and to the
leadership at Farwell Area Schools
that dared to think outside the box.

ACKNOWLEDGEMENTS

I wish to thank the members of my thesis committee: Dr. Hajime Otani, Dr. Sandra Morgan, and Dr. Roger Van Horn. These faculty members provided valuable direction when this project began, reviewed the draft of the document, and made many contributions to the final product. Throughout the process, these faculty members drew on their experiences of working with many graduate students as professors, dissertation advisors, and dissertation committee members. I would like to express additional gratitude to my advisor, Dr. Hajime Otani, for his intense effort on my behalf with the completion of this thesis and my program.

Arigatou gozaimashita.

ABSTRACT

THE ACADEMIC EFFECTS OF AN
ELEMENTARY CHINESE IMMERSION PROGRAM

by Sean R. Hill

The present study investigated whether a foreign language immersion experience would benefit academic performance in mathematics and English language arts. Classical education in the United States included instruction in ancient Greek and Latin; it was believed a mastery of these difficult subjects would create a mental toughness that would transfer across different disciplines. Although the learning theorist, Edward Thorndike, disputed the popular maxim, recent evidence suggests cognitive and academic benefits associated with bilingualism. Even though language immersion programs have grown in recent years, there is a paucity of published research about the benefits of such programs on learning math and other subject areas. Additionally, the research that is available was often derived from comparisons of non-equal groups in varied educational settings.

The current study examined the effect of a partial Chinese immersion program on academic performance in mathematics and English language arts in an impoverished elementary school located in rural Michigan. The research involving character-based language immersion education in the United States is practically non-existent. From the 2010-2011 school year to the end of 2013-2014, kindergarten, first, second, and third grade students spent half their instructional day in Mandarin Chinese. The school district believed the language instruction would help improve student achievement on standardized state tests and complete a mandated foreign language graduation requirement. This study examined the effect of the immersion experience utilizing four assessment instruments, the Brigance Screen for Kindergarten Readiness, the Scholastic Reading Inventory (SRI), STAR Early Literacy, and STAR Math,

between the immersion students and the students in the traditional classroom. In addition, the third grade Michigan Educational Assessment Program (MEAP) scores were compared.

The results showed that the partial Chinese immersion program did not jeopardize academic achievement in reading English or math. Further, immersion students at the third grade showed higher math performance relative to traditional peers. Due to the multitude of uncontrolled confounding variables, it is difficult to ascertain the reasons for the results but this study suggests that the immersion experience, using Chinese language as the target language, is not detrimental to academic achievement and may even produce a beneficial effect on mathematics achievement.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

CHAPTER I

INTRODUCTION

Can learning another language help students excel in other content areas, such as mathematics, science, and English Language Arts? However, spending time learning another language means less time devoted to learning mathematics. It does not seem to make much sense to spend less time when one would want to increase academic achievement in one particular subject area. This type of debate is not new; it has waxed and waned in the American education system. Benjamin Franklin, for example, proposed a utilitarian system where students learned material that would be beneficial to their future careers. In particular, he discouraged learning Latin and Greek in favor of English Language Arts as well as public speaking and recommended that vocational education be included as part of the curriculum (Ornstein & Levine, 2000). However, the National Education Association in 1892 still maintained much support for Latin and Greek as part of a college preparatory program and as main part of high school education.

One of the early psychologists who passed through such a classical education program, Edward Thorndike, reexamined the notion that mastering difficult coursework would foster a sort of mental toughness. Thorndike applied scientific principles and found that there was no evidence supporting the popular notion that learning in one discipline would transfer to learning in another discipline (Bruer, 1993). After the publication of his findings, the trend to disregard the classical education idea has begun to grow.

Foreign language education, in the context of learning Latin and ancient Greek, was not necessarily intent on building bilingualism and fluency. In contrast, the modern approach is to foster functional bilingualism, and this new approach is beginning to show cognitive benefits of

learning second language (L2) in public schools. Stewart (2005) describes three types of foreign language exposure in the elementary classroom: immersion, FLES, and FLEX. Of these, the immersion program offers the highest level of exposure to the target second language. In this program, classes are conducted in the target language, and each exposure period lasts for at least half a day. Some schools do not introduce English (native language or L1) for the entire first year, whereas others maintain a dual-language immersion where half a day is spent in one language and the other half is spent in another language. A FLES program, short for foreign language in elementary schools, is similar to traditional high school foreign language courses, where the language is taught as a specific subject. The exposure to the language is more limited in this traditional program. Lastly, in a FLEX or foreign language exploratory program, classroom exposure to the target language is limited and instead, much of the focus is on learning about the cultural practices of the foreign country rather than explicitly learning the language.

The basic assumption of these programs is that learning a second language produces a secondary benefit of enhancing cognitive performance. McLaughlin (1984), however, contends that although a child with two languages seems to have a clear linguistic advantage over a monolingual child in regard to systems of language, there remained a question as to whether increased sensitivity to languages would imbue benefits in general cognitive functioning. He concluded that it is difficult to show such an advantage scientifically because it is difficult to remove all the confounding variables in real life. Cognitive benefits of second language learning may not be clear; however, an interesting side effect of early foreign language education, unrelated to academic content, has been documented. Lambert, Tucker and d'Anglejan (1973) reported that after four years in the French language program, students had developed a balanced and less ethnocentric view toward French people and culture, indicating the usefulness of the

language exposure in promoting global awareness and understanding. Therefore, foreign language instruction in the early grades may potentially result in language proficiency, increased achievement in other content areas, and positive attitudes toward the target language and the culture.

Although Thorndike examined the issue of whether learning Latin or ancient Greek would lead to improved proficiency in learning other subject areas, he did not investigate whether bilingualism would lead to increased academic achievement and cognitive functioning. The goal of the present study, therefore, will be to investigate whether practical foreign language coursework at a young age would produce any measurable benefits in academic achievement. Stewart (2005) summarized three beneficial effects of foreign language programs in the elementary setting: (1) increased cognitive skills involving creativity and problem solving, (2) higher achievement in academic areas, and (3) higher scores on academic achievement tests.

For instance, Martin-Rhee and Bialystok (2008) found a cognitive advantage of bilinguals to monolinguals in controlled attention. Kempert, Saalbach, and Hardy (2011) found that highly proficient bilinguals showed highly efficient executive control processes. However, even among bilinguals who were not highly proficient, there were benefits to executive control. These researchers remarked that keeping two languages apart while performing various tasks may lead to development in executive control, a hypothesis supported by a finding that even in infants, being raised in bilingual homes show increased proficiency in executive control (Kovács & Mehler, 2009).

Lazaruk (2007) examined the findings from research on early language programs in Canada and concluded that students showed academic and cognitive benefits of these programs, including increased flexibility and creativity in problem solving, metalinguistic understanding,

and vocabulary control. In addition, learning another language increases students' understanding of their own native language. For example, Cunningham and Graham (2000) found that achievement scores in recognizing low-frequency English vocabulary words were higher among elementary students who had studied Spanish, relative to students in traditional classrooms. Further, there was a positive correlation between the length of time the students studied a foreign language and performance on the Scholastic Aptitude Test (SAT), particularly on the English vocabulary knowledge subtest. Another frequently reported benefit is increased phonological processing skills by bilingual students who learned one alphabet-based language to another, such as between French and English (Comeau, Cormier, Grandmaison & Lacroix, 1999; Jared, Cormier, Levy & Wade-Woolley, 2011).

The difference between the traditional foreign language instruction and the immersion education programs is that in the former type, students spend much of the time learning grammatical structures and vocabulary for communication, whereas in the latter, students use the foreign language as the method of mastering the entire school curriculum. In other words, in the immersion programs, the use of the language itself is teaching the language (Met, 1993, & Knell et al, 2007). Students learn how the language works in a communicative environment where the language has meaning, closely simulating the experience of students mastering their native language (Genesee, 1995). In the immersion programs, language is learned as students learn subject areas like science and art. Immersion education is on the rise in the United States as reported by the Center for Applied Linguistics (2011): in 1971 there were only three schools with immersion programs in the United States and by 2011, the number had grown to 448. Additionally, the variety of immersion languages offered has grown. Between 1995 and 2006, immersion instruction doubled from nine languages to eighteen (Lenker & Rhodes, 2007).

Most foreign language programs focus on students' foreign language proficiency and achievement whereas immersion programs are also responsible for demonstrating student proficiency in the curriculum areas taught in the target language. However, Chamot and El-Dinary (1999) found that the majority of the research in immersion education focused on program outcomes in language proficiency and an academic content area, English Language Arts, in particular. Immersion educators typically must teach mathematics, science, or social studies in the target language; however, Miller's (2010) survey of the field found little published research about learning math content through immersion.

Furthermore, Barwell (2010) concluded that the majority of the research involved in teaching the content areas through a L2 has focused on how the students learn the L2 through the content rather than how the students learn the content through the L2. Moreover, his survey of the field on learning mathematics in L2 environments showed contradictory claims. He stated that there are complicated issues, which he referred to as tensions that arise when one teaches math in a L2, that can range from the linguistic meaning of questions to the formation of mathematical concepts. For example, he cited an elementary student's understanding of polygons and her difficulty in naming each type. It was unclear whether the student was simply confusing the number of sides of an octagon and a pentagon or did not connect the concept with the appropriate noun. Her naming of hexagon, "hepsada," added further confusion. Notwithstanding, Barwell suggested that immersion teachers could learn much from a review of the literature and that immersion programs were not necessarily harmful to academic success.

Nevertheless, the wide range of effects reported is troubling. Marsh, Hau, and Kong (2000) claimed that the popular assumption that immersion programs lead to improved academic performance is based on the comparisons between non-equal groups, such as comparing the

students in the immersion programs, who are more selective, with the students in the traditional programs, who are not selective. Such selectivity may stem from attrition by low performing students from the immersion programs or increased investment in children's education by the parents of immersion students. Unfortunately, ruling out confounding variables is difficult because maintaining rigorous scientific controls is difficult, if not impossible, in a dynamic public education system.

In summary, whereas Thorndike dismissed the notion that learning foreign languages, particularly Latin and ancient Greek, would facilitate learning in other subject areas, language instruction in traditional education focused primarily on rote memorization and grammatical structures, and not on communication. Recent research has shown that bilingualism may foster creativity, problem solving, executive control, attention, and academic achievement. However, research has produced conflicting results on the effect of foreign language immersion programs on academic achievement. The disparate results may have been based on the fact that much of the research has been conducted in the public education system where uncontrolled extraneous variables could contaminate the results. Accordingly, further research on this topic is warranted.

Academic Gains of Students in Immersion Programs

The standard argument against implementation of immersion programs is that the time devoted to learning the language detracts from time spent learning other subject areas, particularly the student's first language (L1). However, Holobow, Genesee, and Lambert (1991) reported that student achievement in L2 does not come at a cost to achievement in L1, and that immersion students' progress is at pace academically, cognitively, and socially with their peers from traditional classrooms. Genesee (2007) stated that little association has been found between the amount of contact hours immersion students spend in English (L1) and achievement

of these students in English language arts. In contrast, the time spent in the target language (L2) tends to be positively correlated with second language proficiency. Genesee and Jared (2008) reviewed the literature and found that students enrolled in early immersion programs achieved at the same level on English reading assessments as students in traditional programs. While there was a period of underperformance on English assessments early by the total immersion students, these students were indistinguishable from traditional students within two years after English instruction was introduced.

Barik and Swain (1978) evaluated students enrolled in a French immersion program in Ottawa, Canada. These students had been in the program since kindergarten and were found to be at the same level as or exceeding a comparative group of non-immersion students in multiple course assessments. However, the students were not randomly assigned to respective groups because parents elected to enroll the students in the immersion program. It is, therefore, plausible that parents who were more involved with their child's education were more likely to register the child in the immersion program. Nevertheless, the comparison groups were matched on the basis of SES, IQ, age and other attributes.

In this program, Kindergarten was a half day program, and the first grade was an entire day, and lessons were taught completely in French. Students were introduced to English Language Arts in the second grade for one hour per day (even though some schools in the district did not introduce English instruction until the third grade), and this pattern continued through fifth grade, resulting in 20% of lessons taught in English (an exception of one school began teaching math using English in the fifth grade). Afterwards, an equal amount of time was allocated for lessons taught using English and French. The results indicated that the immersion students showed no significant difference in IQ scores from Kindergarten through the third grade

relative to those in the traditional program; however, there was an advantage of 3 and 5 IQ points, respectively, in the fourth and fifth grade. The cohorts that began two years subsequent to the initial cohort showed advantage in IQ scores in Kindergarten but the difference disappeared after Kindergarten. In English Language Arts, the immersion students achieved higher scores or scores equal to those of the comparison group. For the majority of mathematics assessments, there was no difference between the immersion and non-immersion students.

Lambert, Tucker and d'Anglejan (1973) found that the immersion students scored equally high as the English and French control groups on the Standard Progressive Matrices and the Lorge-Thorndike test by the fifth grade. As for creativity, the immersion students scored higher than the controls on measures of divergent thinking. These investigators commented that immersion education is not detrimental to the development of intelligence and may enhance problem solving skills compared to the traditional classroom. Similarly, Jones (2005), who examined a total immersion Texas school that introduced 30 minutes of English during the second semester of second grade, found no differences between the immersion and non-immersion students in 5 of 12 analyses on the Cognitive Abilities Test. Jones further reported that immersion students outscored traditional students on all tests of reading and mathematics when adjusted for cognitive ability.

Downs-Reid (2000) examined how immersion students faired against traditional students on the Fifth Grade Minnesota Comprehensive Assessments (MCA) of reading, writing, and mathematics when the two groups are matched on ethnicity and for free and reduced lunch. In his investigation of four different immersion programs in urban and suburban districts, all the immersion schools outscored their intermediate service districts on the 1999 MCA. One immersion school outranked the comparison schools on every subtest while the other three

immersion schools outranked their comparison schools on two out of three subtests (albeit the difference was not tested for statistical significance). These results indicate that these immersion students were equal to or exceeded similarly matched students from traditional classrooms. In addition, he showed that IQ growth rates were higher among students in the immersion program relative to students from traditional programs.

Essama (2007) investigated a French immersion program in Maryland and described the academic achievement of its students. These students did not receive instruction in English until the second semester of fourth grade. As second graders, they completed the nationally normed Comprehensive Test of Basic Skills (CTBS). The results over six years showed that these students scored above average in all areas of the test: reading, math, language mechanics, math computation, and language. In particular, the majority of these students were above the 75th percentile in mathematics achievement. When tested on math and reading with the Maryland School Assessment (MSA), these students in the third, fourth, and fifth grades outscored the state average during four years, with the exception of the third grade reading falling behind the state average score in 2006. By the fifth grade, the immersion students outscored the state average by approximately 20% in both reading and math achievement. Furthermore, 86% of the fifth graders read at or above the grade level on the Measure of Academic Progress—Reading (MAP-R) and almost one third of them read at the high school level or above. While these data indicate that even before instruction is introduced in English, these students averaged much higher than students in traditional education, Essama noted an extensive program of external support specifically designed to prepare the immersion students for the achievement tests. The gains that were reported may be confounded by this additional support, and it is plausible that traditional students would have scored much higher with similar assistance.

In conclusion, research on full immersion programs since the 1970s suggests that learning a second language (L2) through immersion instruction does not hinder the development of students' first language (L1). Further, the early French immersion programs in Canada showed that at the beginning of the program, academic achievement of immersion students may lag behind traditional students; however, the difference can be erased within two years of beginning English instruction, with added advantages in creativity and problem solving. Similar research in the United States supports the findings of the Canadian programs. For example, immersion students in Spanish and French in both Minnesota and Maryland outscored the local district averages as well as the state averages, even when the immersion and traditional groups were matched on variables such as ethnicity and free and reduced lunch qualification. However, a possibility cannot be ruled out that the higher scores by the immersion students were based on extra preparation these students received prior to state testing or some other extraneous factors.

Academic Gains of Students in Partial Immersion Programs

Not all foreign language immersion programs use a full immersion approach. Some are partial (or dual) immersion programs, in which lessons are taught using both L1 and L2. The allocation between L1 and L2 can be 50/50; however, some programs use a model in which instruction using L1 is increased as students advance in years. The present study will examine a partial immersion program in which instruction is evenly divided between Mandarin Chinese (L2) and English (L1).

In contrast to the full immersion model, the results of partial immersion programs are mixed. Haj-Broussard (2005) showed little benefit of a partial immersion program for students in an urban elementary school. In regard to mathematics achievement, de Courcy and Burston (2000) showed that students in a French partial immersion program scored worse when they took

the test in French compared to English, perhaps reflecting lower proficiency in French language than students from French full immersion programs. However, other investigators reported benefits. Caldas and Boudreaux (1999) reported general gains for African-American students in a French immersion program in late elementary classrooms compared to students in traditional classrooms.

Turnball, Lapkin, and Hart (2001) examined a province-wide test, which was administered to all third and sixth graders in Ontario, Canada, in 1998-1999. Given a variety of immersion and traditional programs, students were compared based on the amount of English instruction they had received up through the third grade. The results showed that the scores by the immersion students were similar to those of the traditional students in reading, writing, and mathematics, and that the number of hours spent in English instruction in the immersion programs did not show positive correlation with performance in English reading and writing. However, the immersion students in the sixth grade wrote better in English than the students in the traditional program. Mathematics achievement was not correlated with the amount of English instruction, unless the students had had less than 750 total hours of English by the third grade. However, by the sixth grade, students in the varied immersion programs outscored students in the traditional classrooms in mathematics.

Bournot-Trites and Reeder (2001) examined the effect of immersion instruction in French in the intermediate school years on mathematics achievement among the students who previously were taught in French in elementary school. There were two conditions: a 50-50 partial immersion program where math was taught in English and an 80-20 immersion program where math was taught in French. The more intense immersion environment produced better student achievement. The authors stated that the enhanced achievement scores were the result of

increased effort it took to master academic concepts using a non-native language, which involved deeper levels of processing. They also posited that students with greater proficiency in L2 were better able to learn the mathematical concepts and transfer them to L1. However, there were other factors that influenced the results as well, such as parent involvement, use of two different textbooks, and teaching methodologies.

Thomas, Collier, and Abbot (1993) examined the effect of Spanish, French, and Japanese partial immersion programs in Virginia. After the first year of math instruction in the target language, the immersion students showed slightly lower scores than a comparable group of students in traditional classrooms. However, the difference was statistically non-significant, indicating that instruction in a foreign language was not a hindrance to their ability to achieve on a mathematics test in English. As for the students receiving one year of instruction in mathematics through Japanese, they outscored both the comparison group and the school district average. After two years in an immersion program, all students showed scores no different from those in the comparison group, despite the fact that the scores of both groups were significantly above the school district average. In regard to performance in English language arts, the immersion students outscored both the comparison group and the school district average.

Haj-Broussard (2005) examined a French immersion program in regard to the minority/ majority achievement gap. On the Louisiana Educational Academic Performance—Mathematics (LEAP) test, white students in immersion education outscored white students in traditional classrooms whereas African-American students in immersion education outscored African-American students in traditional classrooms, even though the difference did not reach a statistical significance. In terms of the achievement gap between white and African-American students, there was no difference in scores of white students in regular education and African-American

students in immersion education, whereas white students outscored African-American students in traditional education.

Holobow, Genesee, and Lambert (1991) examined a partial French immersion program in Cincinnati, OH, over the course of four years, which included minority and working-class students in both immersion and traditional classrooms. Initially, the pretest scores did not show a significant difference between the traditional and immersion students on school readiness. Testing after the first grade indicated that the immersion students' achievement gains were similar to the traditional classrooms in English Language Arts. In addition, low SES at-risk students benefitted as much from the immersion program as students from middle class backgrounds. Further, the results of mathematics achievement on the California Achievement Test (CAT) indicated that the immersion students outscored the students in traditional classroom. However, the authors cautioned that the students in the immersion classes received math instruction in both languages, even though the authors did not report the proportion of instruction in each language.

Holobow, Genesee, and Lambert (1991) also posited that immersion education may mitigate the effects of SES, a hypothesis supported by Kruk and Reynolds (2012) who examined the effect of a French immersion program on at-risk readers from the first to third grade. The results showed that students at-risk of reading difficulties in the immersion programs showed higher literacy growth rate in decoding and reading comprehension than students at-risk in traditional classrooms, suggesting that the benefits of increased language sensitivity and exposure may accumulate over time with at-risk readers.

Bae (2007) examined the effect of a Korean/English two-way immersion program and its effects on writing development in first and second graders. The author found that there was no

statistically significant difference in scores for grammar, content/coherence, and text length between the immersion and traditional classrooms. A possible confounding variable is that the students in the immersion program were given five hours of additional tutoring weekly using both Korean and English, which may have led to higher development in English writing skills.

Some schools showed a marked gain in academic achievement with the partial immersion model. For example, Arthur (2004) described an immersion school in Indiana, which expanded its program based on the achievement scores of the immersion students. Originally the program began as an optional program for kindergarteners and first graders, which continued until they reached the fifth grade. In this 50/50 program, math, science and social studies were taught in Spanish with elements of Spanish language arts being incorporated into the content, whereas the remainder of the day was allocated for English language arts. Over a period of six years, the students in the immersion program continually outscored the students in traditional classrooms in mathematics achievement. In regard to English Language Arts, the immersion students outscored the traditional students on five out of six years. Due to the increased academic achievement, the school mandated all students participate in the immersion program. The success also led the district to allow both middle and high school to provide the continuation of immersion experience.

In summary, partial/dual immersion programs have grown in popularity due to the fact that these programs devote instructional time to both L1 and L2. The reported effects of such programs are more equivocal than the reported effects of full immersion programs. Some researchers found academic benefits of these programs whereas other researchers did not. Further, the beneficial effects of these programs were reported from the programs that provided additional tutoring to the immersion students or taught mathematics using both L1 and L2. As

stated earlier, the efficacy of these programs is difficult to ascertain in the dynamic environment of public school systems.

Immersion Education as Detrimental to Other Subject Areas

It has been reported that students and parents experience both excitement and frustration with immersion education (Brzezinski, 2008), mainly because of the exclusive use of the target language (L2) in the classroom. Knell and colleagues (2007) described that many parents fear that immersion students would fall behind academically even though Canadian programs showed that immersion students do not necessarily encounter achievement gaps in their first language or in other content areas. These authors also stated that the L2 language proficiency is much higher among immersion students relative to students taking a L2 as a foreign language class. However, not all studies showed the positive effect of immersion education.

In particular, partial immersion programs present a complicated situation where limited mastery of L2 can negatively impact mastery of other subjects. For example, Lazaruk (2007) described how partial immersion programs, as opposed to full immersion programs, were created after parents requested to ensure that immersion students received enough English instruction to keep pace with their peers in traditional programs. While students in partial immersion programs caught up to their traditional classroom peers in English Language Arts sooner than students from full immersion programs, they did not surpass the achievement scores of either traditional or full immersion students at the end of the immersion program. Further, Bournot-Trites and Reeder (2001) showed lower academic achievement by partial immersion students relative to total immersion students in content areas taught in the second language. Swain (1996) found similar lags in achievement by partial immersion students and concluded that these students did

not master L2 well enough to comprehend core content using L2 in later grades. Likewise, she found that students who were less proficient in L2 showed lower development of L1.

Marsh, Hau, and Kong (2000) found a negative effect of an English immersion program on Chinese students' achievement, when English was used as L2 to teach math and other subject beginning at the secondary level. While the immersion students did excel in English and Chinese compared to their non-immersion peers, they showed lower scores in history, geography, and science. Knell and colleagues (2007) posited that the implementation of the late immersion might have resulted in lower proficiency in L2, which hindered the mastery of various subjects that required technical language skills. The achievement gap noted was likely to be based on the timing of the first immersion experience because Chinese speakers from Singapore who began English immersion in the elementary school did not show detrimental effects of immersion. Indeed, these Singapore students were among the highest in the world in academic achievement when compared internationally.

Gibbons (1989) underscored the importance of L2 proficiency for educational attainment in other content areas. Marsh, Hau, and Kong (2000) found that students who used more L2 over multiple years at middle school (a proxy for proficiency) experienced a positive gain on history achievement. Further, continued use of L2 across classes was also positively correlated with performance.

Cummins (1998) agreed that while students in early immersion programs typically catch up with traditional students on academic achievement within a year, there are areas that need additional help, such as spelling and grammar. He also cautioned that the early Canadian studies did not take into account the student attrition from the program due to behavioral or academic problems. Likewise, De Jong (2002) criticized the Canadian studies and their applicability to

current education in the United States, arguing that the United States does not necessarily have the same levels of school readiness, parental involvement, and English proficiency and literacy at home as Canada. Once again, many confounding variables are precluding a firm conclusion.

Lastly, the immersion experience is analogous to the experience non-English speaking immigrant students face when they enroll in traditional programs in the United States. These students must learn all course curriculum through L2 (English). Willig (1985) found that teaching students in their first language (e.g., Spanish) resulted in gains in knowledge of content areas, whereas teaching content solely in English (L2) was detrimental to academic performance. Students had a difficult time learning content in English while learning English. Therefore, it was more beneficial academically to utilize the student's first language as the language of instruction while students learned English than to teach the entire curriculum in English. It is possible that immersion experience is similar to the experience of these immigrant students such that using L2 to learn course materials is detrimental to academic achievement. Notwithstanding, Swain, Lapkin, Rowen, and Hart (1990) investigated French immersion programs in Canada where immigrant children whose L1 was not English had begun immersion in French at the fifth grade. These researchers reported that these immigrant students were at pace academically with their non immigrant English-speaking peers. Consequently, immersion experience may be qualitatively different from the experience by the immigrant students in the United States.

In conclusion, although partial/dual immersion programs were designed to ensure adequate instruction in English (L1) as well as the target language (L2), the effectiveness of these programs on academic achievement is unclear. The main problem is the inadequate mastery of L2 by these students, preventing them from comprehending difficult concepts that are

introduced in the secondary education curriculum. Further, these students also showed less development in their L1. Similarly, non-English speaking immigrant students in the United States showed that instruction using L2 is not effective, especially when there is little exposure to L2 at home or in the community. Lastly, many previous investigations did not take into consideration uncontrolled variables, such as student attrition, non-equivalent groups, and the conditions of the student body, making it difficult to arrive at a clear conclusion.

Effects of Language Similarity on Achievement

Most of the investigations presented up to this point have examined the effect of French or Spanish immersion programs on native English speakers. These languages are similar, sharing vocabularies through cognates and, above all, using an alphabet that maintains similar phonetics. Fewer schools in the United States use a character-based language like Japanese or Chinese in their immersion programs. Fortune and Egenberger (2010) found that there has been an increase in popularity of Chinese language immersion programs during the recent decade. For example, there were only 12 programs in the United States using Chinese as L2 in 2005, but by 2009, the number grew to 42. By 2013, Weise (2013) estimated that across the United States, there were 147 schools with a Chinese immersion program, and that more than 50% these programs opened their doors in the 2009-2010 academic year or later. Within a span of five years, Mandarin Chinese instruction surged from 4% in 2006 (Lenker & Rhodes, 2007) to 13% of all language instruction in the United States (Center for Applied Linguistics, 2011). Miller (2008) commented that while many parents might be familiar with Spanish, very few parents are familiar with Mandarin Chinese. Furthermore, there is very little research examining the immersion programs in the United States using non-Roman alphabet language (Thomas, Collier, & Abbot, 1993).

Studies conducted in China suggest that learning dissimilar languages can be mutually beneficial for proficiency in L1 and L2. Knell et al. (2007) investigated that effects of early English partial immersion programs in mainland China. The researchers found that instruction in English did not hinder Chinese students' development in character reading compared to non immersion students. Additionally, they noted that instruction in pinyin (the phonetic transcribing system of Mandarin Chinese using the Roman alphabet) prior to character learning facilitated English proficiency compared to students from Hong Kong who did not learn pinyin. While the phonological similarity between English and pinyin helped Chinese children decode English vocabulary, learning English can also facilitate learning Chinese. Further, Wang, Cheng, and Chen (2006) reported that Chinese children showed increased awareness of compound word structure in reading and writing Chinese characters after receiving instruction of the same concept in English. These studies showed how language processing skills transfer across dissimilar languages in such a way that knowledge and skills in the first language (Chinese) can help acquire a second language (English) while learning the second language (English) help learn the first language (Chinese). The benefits can be seen in lexicon, morphology, and phoneme awareness, as well as which way to read script.

Lastly, the nature of language may be another factor that contributes to mathematics achievement. Kempert, Saalbach, and Hardy (2011) stated that cross-culturally, performance on mathematics assessments is related to how different languages express number systems. For example, Chinese uses a base-10 number system that recycles the same codes for every ten numbers; that is, after counting up to ten, the number eleven becomes "ten one" and the number twelve becomes “ten two.” Such a system may be more conducive to mathematics instruction. The math advantage enjoyed by Chinese language speakers may be due, in part, to the time it

takes to pronounce each number in Chinese. For example, Stigler, Lee and Stevenson (1986) found that children in China were able to remember more digits on average than their American and Japanese counterparts. The length of syllables needed to pronounce numbers in English and Japanese may be more taxing to working memory than Chinese. Therefore, the Chinese language would allow for more resources to be diverted toward mental calculations, yielding higher mathematical achievement.

In summary, the very nature of language may influence the efficacy of the immersion programs. The majority of past research investigated Spanish or French immersion programs; however, the number of programs using Chinese showed considerable growth within the last decade, and therefore, these programs warrant investigation. In particular, it has been assumed that there are advantages in using Chinese language, such as an efficient number system that may enhance working memory capacity. Accordingly, the effects of these programs on student achievement are of interest to educators and psychologists alike.

Present Study

The current investigation seeks to add to the sparse literature of the effect of Mandarin Chinese partial immersion programs in the United States by examining elementary students from an impoverished area of rural Michigan, who participated in an immersion program for four school years (2010-2011, 2011-2012, 2012-2013, 2013-2014). The students in both the immersion program and regular classrooms were required to complete local assessments as well as a final summative examination (the Michigan Education Assessment Program or MEAP), and therefore, it is possible to compare the scores of these assessments between the two groups. In particular, the present study will examine the impact of the immersion experience on academic achievement in English Language Arts and mathematics. Based on the literature on partial

immersion programs, it is predicted that the first year students in the immersion program will underperform their peers from traditional classrooms in both subjects whereas the fourth year students in the immersion program will outperform their peers from traditional classrooms in mathematics and show similar performance to their peers from traditional classrooms in English language arts.

The aim of the current investigation is similar to a recent publication by Padilla et al. (2013), who examined L2 proficiency along with the academic achievement in English language arts and mathematics among students in a Mandarin Chinese (L2)/English (L1) immersion program in a California school. In this program, students spent between 80% and 50% of their instructional time in Mandarin over multiple years. The researchers combined data from multiple grade levels and compared the scores on California state assessments between the immersion students and the students from the traditional classrooms. Their goal was to determine how such a character-based language would affect native English speakers in regard to mastery of Mandarin Chinese as well as academic achievement in other content areas. The results showed that immersion experience in a character-based language like Mandarin Chinese does not handicap students on the California-mandated tests in English language arts, writing, math, or science after five years of immersion instruction. The results also indicated that early in the program, the immersion students showed lower scores in English language arts than the traditional students: when they were tested in the second grade, the average score of the immersion students was 7% lower than that of the traditional students in proficient and advanced rankings. However, by the fifth grade, 100% of the immersion students, relative to 96% of the traditional students, received a proficient or advanced score. In regard to writing, at the fourth grade, more students in the immersion program attained an advanced rating relative to the

traditional students, 59% to 43% respectively. The math scores were similar between the two groups at the second grade; however, by the fifth grade, 92% of the immersion students, relative to only 47% of traditional students, scored at the advanced level. Unfortunately, the validity of these results is questionable because the dropout rate was higher for the immersion program than for the traditional program: from the second to the fifth grade, the number of students in the immersion program declined from 84 to 13 (85% dropout rate) whereas the number of students in the traditional program declined from 301 to 81 (73% dropout rate). Although the academic results of the immersion students were encouraging, the number of students who completed the program is disconcerting.

There are similarities between the present study and that of Padilla et al. (2013); however, there are important differences. For example, the California school was based on a two-way immersion model, in which native Mandarin speakers attended classes alongside native English speakers. In the current study, the program did not include native Mandarin speakers. In addition, the economic conditions as well as the academic achievement level was much less favorable in the Michigan school than in the California school. The California school with more than 12,000 students was located in an upper middle class suburban area, whereas the Michigan school with approximately 1,400 students was located in a rural area, which had been suffering from generational poverty. Indeed, in the Michigan school, more than 60% of the students qualified for free-or-reduced breakfast and lunch. Furthermore, the California school also received a grant to help fund the immersion program.

More critically, the amount of English used for instruction was more limited in the Michigan school than in the California school. In the California school, the students in kindergarten and first grade received 80% of the instruction in Mandarin, and the second and

third grade students received 60% of the instruction in Mandarin. Only the fourth and fifth grade instructional time was evenly divided between English and Mandarin. In contrast, in the Michigan school, the students received approximately 50% percent of the instruction in English at each grade level, resulting in less exposure to Mandarin over the course of four years compared to the students the California school. Moreover, the students in the California school were placed into mixed classrooms, with kindergarteners and first grade students sharing a classroom, so as the second and third graders and the fourth and fifth graders. This was not the case in the Michigan school. Lastly, the students in the Michigan school were not assessed on language proficiency in Mandarin; assessments were only given in English language arts and mathematics. These differences make it difficult to directly compare the results of the two studies. However, there is dearth of research regarding the effect of partial immersion programs with character based L2 language such as Mandarin Chinese, and therefore, the present study will be useful in terms of showing the effect of such a program in a low-performing, less wealthy, rural school.

Research on early foreign language education suggests that there are benefits of bilingualism in cognitive and academic domains. To capitalize on these benefits, the number of foreign language immersion programs is on the rise in the United States. Schools are also interested in the benefits of these programs to student assessment outcomes because school funding is increasingly tied to the results of state assessment. There are several variations in the immersion model, with the full immersion model being most prevalent and researched. Investigations revealed that students in full immersion programs typically achieve L2 language proficiency at a higher level than their peers while showing comparable academic achievement to their traditional peers in other subject areas. In contrast, research on partial/dual immersion

programs produced mixed results. An additional consideration is the target language in the immersion program. Although most of the programs are based on Spanish or French, Chinese language programs are on the rise in the United States. The goal of the present investigation will be to examine the effect of Chinese partial immersion programs by conducting an outcome assessment of a program at an elementary school in rural Michigan.

CHAPTER II

METHOD

Description of the Program

Interests in a partial foreign language immersion program began in the 2009-2010 academic school year in response to the National Race to the Top education campaign. The Michigan Department of Education challenged school districts to re-imagine how to create better student outcomes without raising student funding. The school leadership determined that a dual language program would help student acquire a 21st century skill set and could also improve the district's academic achievement on the Michigan Educational Assessment Program (MEAP). In addition, preschool and elementary foreign language programs are nearly unheard of in rural central and northern Michigan. Accordingly, this program was created in hope of setting the district apart as a school of choice, a school aiming to prepare students for the 21st world economy. A secondary goal of creating this program was to realize increased school budget by attracting students from other districts to come to this district as well as preventing the flight of students out of the district to schools in other districts with better academic achievement. Furthermore, engaging the community in Chinese language and culture would contribute to a reduction in negative attitudes toward Chinese people, which would, in turn, help students prepare for a future with Michigan's top trading partners. Lastly, the leadership of the school district anticipated that the program would act as a catalyst for changing students' collective attitudes toward academic achievement by exposing students to different norms and expectations regarding academic achievement and behavior.

The school culture at the onset of the program and throughout the four years is characterized by transitory students, low academic achievement, and discipline problems at the

district level. Table 1 shows how the third graders in the local school performed on the Michigan Educational Assessment Program (MEAP) relative to the state average (State of Michigan, 2014).

Table 1. Percent of Students in Local School Identified as Proficient on MEAP based on the State Average

	Reading		Mathematics	
Academic Year	Local School	State Average	Local School	State Average
2009-2010	59	65	22	36
2010-2011	62	63	20	35
2011-2012	51	62	19	36
2012-2013	51	67	32	41
2013-2014	51	61	34	40

Further, low achievement continued through and after the high school, evidenced by low college success. On the Michigan Top to Bottom list, which ranks schools based on academic indicators, the elementary school ranked at the 11th percentile in 2010-2011, 7th in 2011-2012, 15th in 2012-2013, and 35th in 2013-2014 (State of Michigan, 2014). The gain on the list for the 2013-2014 school year indicates that academic achievement was greater relative to many other schools in Michigan; however, the school was outperformed by 65% of other schools in Michigan.

Initial discussions resulted in contracting a Chinese teacher through the Confucius Institute at a Michigan university due to the lack of availability of U.S. nationals certified as a Chinese language teacher in rural Michigan. In 2010-2011, the school began the program with a Chinese instructor teaching one group of children for half the instructional day during the afternoon. These children came to the Chinese teacher from two other Kindergarten teachers, reducing the class sizes of those other teachers during the instruction in Chinese. All other kindergarten students received instruction in Chinese as a FLES (Foreign Language in

Elementary School) program for 40 minutes daily. In addition, a group of preschool students met for Chinese language instruction for approximately 20 minutes daily. In the initial cohort of students, no students had had prior instruction in Chinese.

For the 2011-2012 school year, the same basic program model expanded to include two immersion groups that alternated between the morning and afternoon sessions with students being drawn from four different kindergarten teachers. In order to accommodate the two immersion classes, students in the preschool met for approximately 15 minutes of Chinese language instruction daily. In addition, students in the first grade who had not been in the immersion program during the kindergarten year were allowed to enter the immersion program in order to fill vacancies in the classroom with a first grade Chinese teacher. The first grade was the last opportunity for students to enter the immersion program unless they sought out extensive tutoring beyond school hours. The Chinese Kindergarten FLES program was discontinued, replaced by a Spanish FLES program in which students attended a class for approximately 30 minutes once a week. Students from each elementary grade level continued weekly Spanish instruction until the 2014-2015 school year.

For the 2012-2013 school year, a change was made in the program model due to budgetary concerns. Students no longer came to the Chinese classroom from different teachers, and instead, they all came from one regular teacher who would also serve as the mentor teacher. The two teachers exchanged students halfway through the instructional day. This arrangement continued in the 2013-2014 academic year. The only notable program change was with personnel: the single kindergarten and single first grade teachers left the district and were replaced by other instructors. Likewise, a third grade teacher resigned in December and another

instructor filled the vacancy in January. The program was discontinued at the end of the 2013-2014 school year. Table 2 shows the history of the program.

Table 2. Program History

Academic Year	Program History
2010-2011	• The initial cohort of Kindergarten students enrolled in the fall 2010. • One Chinese instructor taught one group of immersion students in the afternoon. • No students had prior Chinese lessons. • Non-immersion kindergarten students took 40 minutes of FLES Chinese instruction daily. • Pre-school students received 20 minutes of FLES Chinese instruction daily.
2011-2012	• The program was expanded to two alternating immersion groups alternated between morning and afternoon in kindergarten and the first grade. • Pre-school students received 15 min of Chinese instruction daily. • Immersion students came from all kindergarten and first grade classrooms. • Chinese FLES instruction was eliminated in kindergarten replaced by Spanish FLES.
2012-2013	• At each grade, immersion students came from one American teacher who was paired with one Chinese instructor.
2013-2014	• Kindergarten and first grade teachers replaced. • A third grade instructor was replaced in January.

School Demographics

The school district is small, rural district in Central Michigan which consists of three schools: an elementary, a middle, and a high school, all in close proximity of each other. In the 2009-2010 academic year, there were 1492 students in the district, and in the 2013-2014 academic year, the student count was 1431 (MI School Data, 2014). The surrounding community lacks significant business, industry, educational attainment, and cultural diversity. The area has a history of both generational and childhood poverty, and the per pupil funding is at the lowest level in the State of Michigan. At the inception of the language immersion program in 2010, 65% of the students qualified for the federal free and reduced lunch program. In the elementary school, economically disadvantaged students accounted for 74% of all students in 2012-2013 and in 2013-2014. Ethnic and racial diversity at the school is low; a range from 92% to 98% of students classified as white, non Hispanic between 2009-2010 and 2013-2014. Further, between 2009 and 2014, legal guardians of all students reported that the first language of these students was English. In addition, students with disabilities (physical, cognitive, or emotional) ranged between 12% and 19% of all elementary school students.

Participants

Students enrolled in the Chinese immersion program solely based on parent choice. Forms for enrollment were included in parent information packets during the Kindergarten roundup (the enrollment process), and additional seats were offered to parents prior to the beginning of the first grade as a way to protect against student attrition. Children of school faculty and staff received the highest priority in admission followed by in-district students and then out-of-district students. A lottery system was incorporated for open seats when the demand for admissions exceeded the class size. It was communicated to parents that once students were

selected to become part of a cohort they were expected to remain in the program. Nevertheless, there was attrition each year for each cohort. Some parents removed their children from the immersion program but kept them in the district in the traditional classroom, and other parents removed their children from the district or moved out of the district. There were students who were removed for discipline or behavior management issues after conferencing with the school principal. Table 3 shows the number of students in the Chinese language immersion program and the traditional program at the end of each academic year, even though it is difficult to ascertain the exact number due to transient students moving in and out of the district in the middle of a school year. Table 4 shows the summary of the entire dataset used in the present study, with gender, SES (students with free or reduced lunch), and special education (students with special education status) information. As described in the results section, not all students in the dataset were used in each analysis because not all students met the inclusion criteria and because students with special education status as well as missing data were excluded from the analysis. To make it clear, for each analysis, the number of students in each group with gender and SES information will be specified.

Table 3. Number of Students Enrolled in the Immersion Program and the Traditional Program. Traditional Students in the Parentheses

	Year			
Grade Level	2010-2011	2011-2012	2012-2013	2013-2014
Kindergarten	25 (69)	51 (76)	34 (49)	42 (54)
First Grade		45 (52)	44 (38)	36 (68)
Second Grade			37 (58)	42 (64)
Third Grade				35 (38)

Table 4. Summary of the Entire Dataset with Gender, SES, and Special Education Information.

	Immersion		Traditional	
Cohort	N	%	N	%
Kindergarten (2013-2014)				
Male (M)	21	49	34	62
Female (F)	22	51	21	38
Both (M + F)	43		55	
SES	28	65	41	75
Special Edu	4	9	6	11
First Grade (2012-2013, 2013-2014)				
Male (M)	12	32	42	51
Female (F)	25	68	41	49
Both (M + F)	37		83	
SES	19	51	62	75
Special Edu	3	8	12	15
Second Grade (2011-2012, 2012-2013, 2013-2014)				
Male (M)	16	42	43	54
Female (F)	22	58	37	46
Both (M + F)	38		80	
SES	26	68	61	76
Special Edu	2	5	10	13
Third Grade (2010-2011, 2011-2012, 2012-2013, 2013-2014)				
Male (M)	13	32	24	39
Female (F)	28	68	38	61
Both (M + F)	41		64	
SES	23	56	46	74
Special Edu	3	7	13	21

Note: SES = Students with free or reduced lunch; Special Edu = Students with special education status; Percent shows the percent of students in each group.

The seven Chinese language teachers were all Chinese nationals who had completed undergraduate degrees in China. They were enrolled in graduate degree programs in a Michigan university as part of a teaching contract leading to a K-12 licensure in Mandarin language arts or

a full elementary license. While the terms of the contract lasted three years, one teacher left after the second year and another left after one semester. All began as first year teachers in the United States in the immersion program, except for one who had taught Chinese in Michigan for two and a half years prior to coming to this program. Regular teachers were all U.S. citizens with varying years of experience as a teacher in and outside of the district.

Assessment

Students in the elementary school were given a barrage of local assessments between the first and the third grade. In addition, they were assessed in the third grade with a statewide test of English language arts and mathematics, the standardized assessment tool for Michigan students from third through ninth grade known as the Michigan Educational Assessment Program (MEAP). It is a high-stakes assessment because the results are made public, and each school is rated, in part, based on the results (Michigan Department of Education, 2013). In the 2013-2104 school year, the MEAP was administered in fall 2013 assessing the English language arts and mathematics that were taught the year before. As such, the third grade results essentially reflect the achievement of students in the second grade with positive and/or negative effects of the summer recess. The third grade class was the only group of immersion students to take the MEAP. All other immersion cohorts received local assessments only.

Students typically take more than ten different assessments from Kindergarten through the third grade. For the current study, four measures will be examined: the Brigance Screen for Kindergarten Readiness, the Scholastic Reading Inventory (SRI), STAR Early Literacy, and STAR Math. The Brigance screen is a pre-Kindergarten assessment that all students take. This is a quick, criterion-based measurement consists of 12 subtests including language, number, body knowledge, motor skills, recognition, and ability to follow directions, resulting in the

maximum score of 100 (Curriculum Associates, 2013 and Mantzicopoulos, 1999). Although Mantzicopoulos (1999) discussed serious problems with this instrument to identify at-risk preschoolers from the Head Start program, she indicated that this assessment tool has high internal consistency (Alpha =.87), test-retest reliability (r = .82), as well as inter-rater reliability (r = .97). In the present study, the scores of this instrument will be compared between the immersion group and the traditional group to examine the group difference before the immersion students entering the program.

The SRI was used by the school district because of its high reliability and validity for reading ability and comprehension (Knutson, 2006), which enable teachers to monitor students' progress in reading. The SRI requires students to read narrative and expository passages from various content areas, draw conclusions, and make comparisons and generalizations (Scholastic, 2007). It is a targeted level testing that adjusts to the students' performance on each question, referred to as computer-adaptive testing, with high internal consistency (Alpha = .89). The SRI Lexile score (a measure of reading difficulty based on normative word frequency and sentence complexity) correlates well ($r \geq .90$) with other well-known instruments (the Stanford Achievement Test Tenth Edition, the Metropolitan Achievement Test Eighth edition, the Stanford Diagnostic Reading Test Version 4.0, the TerraNova and the Mississippi Curriculum Test (Scholastic, 2007).

The SRI was not administered kindergarten or first grade students for reading assessment. Instead, the STAR Early Literacy assessment was utilized because it can be used with students in the developmental stages of reading. It provides norm-referenced data as well as scaled scores for student performance for different types of literacy and print concepts, including sound-letter correlates and phonological awareness (Renaissance Learning, 2014). It is a computer-adaptive

test used by the school district in order to monitor student progress and to design interventions. This test consists of 27 test items with high internal consistency (overall Alpha =.85; within the grade levels Alpha ranging from .80 to .89) and test-retest reliability (overall *r* = .79; within grade levels Alpha ranging from .47 to .74). The validity of this instrument was demonstrated by moderate to high correlation (from the first to twelfth grade *r* = .52 to .77) with other literacy assessments as well as state assessments including Michigan (Renaissance Learning, 2014). Typically, the test lasted about 10 minutes and was administered three times per year per grade level.

The STAR Math assessment provides norm-referenced data as well as scaled scores for student performance for different types of mathematical operations (Renaissance Learning, 2014). It is a computer-adaptive test, administered to all students in the school district in order to monitor student progress in mathematics. This test consists of 34 test items with high internal consistency (overall Alpha =.97; within the grade levels Alpha ranging from .90 to .95) and test-retest reliability (overall *r* = .93; within grade levels Alpha ranging from .84 to .76). The validity of this instrument was demonstrated by moderate to high correlation (from the first to twelfth grade *r* = .55 to .80) with state assessments in many states including Michigan (Renaissance Learning, 2014). Typically, the test lasted about 20 minutes and was administered three times per year per grade level.

CHAPTER III

RESULTS

The first cohort enrolled at the beginning of the 2010-2011 school year, and each subsequent year, a new cohort entered the program until the end of the 2013-2014 school year. Accordingly, when the program ended at the end of the 2013-2014 year, there were four cohorts with varying numbers of years in the program: four years for the third graders who entered in the program at the beginning of the 2010-2011 school year; three years for the second graders who entered in the program at the beginning of the 2011-2010; two years for the first graders who entered in the program at the beginning of the 2012-2013 year; and one year for the Kindergarteners who entered in the program at the beginning of the 2013-2014 year. Each cohort was analyzed separately with grade level as a within-subject variable. Further, some students were excluded from the analysis because these students transferred into the district after the cohort had started and therefore, did not spend a sufficient amount of time in the program.

The following inclusion criterion was set for each cohort as well as for a comparison group of traditional students; three or four years in the program for the third graders, two or three years in the program for the second graders, one or two years in the program for the first graders, and one year in the program for the Kindergarteners. In addition, students with special education status as well as students with missing data were excluded from the analyses. Preliminary analyses showed that gender did not show a significant effect, and therefore, this variable was not included in the following analyses. Further, SES (free or reduced lunch) was coded as either 0 (no) or 1 (yes) and used as a covariate because this variable was significant in some analyses. For all the analyses, the significance level was set at $p < .05$ (two-tailed) unless otherwise specified.

Third Grade Cohort

For this cohort, the Brigance Screen for Kindergarten Readiness was not administered. The SRI and STAR Math were administered each year starting the first grade. Furthermore, this is the only cohort with scores on MEAP, which was administered at the beginning of the third grade (fall 2013).

MEAP Reading and Math. This assessment consisted of reading and math subtests, and therefore, for each subtest, a one-way analysis of covariance (ANCOVA) was conducted to compare the immersion and traditional groups with SES as a covariate. For both subtests, there were 38 students (27 females and 11 males) in the immersion group and 29 students (17 females and 12 males) in the traditional group. Further, the number of students with free or reduced lunch was 20 in the immersion group and 18 in the traditional group. The results showed that for the reading subtest, the group difference was not significant, $F(1, 64) = 0.81$, $MSE = 487.75$, $p = .37$, $\eta_p^2 = .01$, such that the scores were similar between the immersion group ($M_{\text{uncorrected}} = 330.50$, $SD = 20.85$) and the traditional group ($M_{\text{uncorrected}} = 325.00$, $SD = 23.04$). The effect of SES was not significant, $F(1, 64) = 0.63$, $MSE = 487.75$, $p = .43$, $\eta_p^2 = .01$. For the math subtest, the difference approached significance, $F(1, 64) = 3.13$, $MSE = 431.82$, $p = .08$, $\eta_p^2 = .05$. The immersion group showed a slightly higher mean ($M_{\text{uncorrected}} = 334.61$, $SD = 21.29$) than the traditional group ($M_{\text{uncorrected}} = 323.86$, $SD = 21.82$), even though the difference was not significant. The effect of SES was significant, $F(1, 64) = 5.69$, $MSE = 431.82$, $p = .02$, $\eta_p^2 = .08$.

Reading (SRI). Due to missing data, for this analysis, the number of students was 19 (13 females and 6 males) in the immersion group and 12 (8 females and 4 males) in the traditional group, with 8 students in the immersion group and 5 receiving in the traditional group receiving free or reduced lunch. A 2 (group: immersion and traditional) x 3 (grade level: one, two, and

three) ANCOVA was performed with SES as a covariate. Table 5 shows the mean SRI Lexile scores across group and grade level.

Table 5. Mean SRI Lexile Scores across Group and Grade Level for Third Graders

	Immersion (n = 19)		Traditional (n = 12)	
Grade Level	*M*	SD	*M*	*SD*
First	169.74	132.42	201.00	153.43
Second	437.00	153.94	620.25	94.89
Third	642.74	150.47	697.25	95.48

The results showed that the main effect for grade level was significant, $F(2, 56) = 159.71$, $MSE = 7783.76$, $p < .001$, $\eta^2 = .85$. LSD tests showed that all grade levels were significantly different from each other, indicating that both groups showed improvement in reading as they moved up from one grade to the next. The main effect of group was also significant, $F(1, 28) = 4.56$, $MSE = 38758.10$, $p = .04$, $\eta^2 = .14$, indicating that overall, the mean was lower for the immersion group ($M_{\text{uncorrected}} = 416.49$, $SD = 130.53$) than the traditional group ($M_{\text{uncorrected}} = 506.17$, $SD = 81.74$). However, the group x grade level interaction was significant $F(2, 56) = 6.31$, $MSE = 7783.76$, $p = .003$, $\eta^2 = .18$. The effect of SES was not significant, $F(1, 28) = 1.43$, $MSE = 38758.10$, $p = .24$, $\eta^2 = .05$.

To reveal the nature of the two-way interaction, independent samples *t* tests were conducted to compare the immersion and traditional groups at each grade level. Because SES was not significant in the omnibus test, the following analyses were uncorrected. Further, because the number of missing data was different at each grade level, the number of students was different from that of the omnibus test as well as for each grade level analysis. The results showed that there was no significant difference between the groups at the first, $t(30) = 1.06$, p

=.30 (n = 19 for the immersion group and n = 13 for the traditional group), second, $t(53) = 0.99$, $p = .33$ (n = 31 for the immersion and n = 24 for the traditional group), or third grade level, $t(59) = 1.11$, $p = .27$ (n = 36 for the immersion and n = 25 for the traditional group). The interaction was further analyzed by computing gain scores from the first to second grade and from the second to third grade. Independent samples t tests showed that the traditional group showed greater gain from the first to second grade than the immersion group, $t(30) = 2.37$, $p = .02$ (n = 19 for the immersion and n = 13 for the traditional group). The observed means were 267.27 (SD = 107.36) for the immersion group and 395.85 (SD = 198.52) for the traditional group. However, the immersion group showed greater gain from the second to third grade, $t(49) = 3.75$, $p < .001$ (n = 31 for the immersion and n = 20 for the traditional group). The observed means were 219.23 (SD = 130.05) for the immersion group and 90.15 (SD = 102.67) for the traditional group. These results showed that the immersion group was somewhat slower in the development of reading at the beginning than the traditional group; however, by the third grade the immersion group caught up with the traditional group.

Math (STAR Math). For this analysis, the number of students was 34 (25 females and 9 males) in the immersion group and 27 (16 females and 11 males) in the traditional group, with 17 students in the immersion group and 18 students in the traditional group receiving free or reduced lunch. Table 6 shows the means across group and grade level. A 2 (group: immersion and traditional) x 3 (grade level: one, two, and three) ANCOVA with SES as a covariate showed that the main effect of grade level was significant, $F(2, 116) = 157.76$, $MSE = 1490.78$, $p < .001$, $\eta^2 = .73$. LSD tests showed that all grade levels were significantly different from each other, indicating that both groups showed improvement from one grade level to another. The main effect of group was significant, $F(1, 58) = 4.05$, $MSE = 12521.56$, $p = .049$, $\eta^2 = .07$, such that

overall, the immersion group showed a higher mean ($M_{\text{uncorrected}} = 525.91$, $SD = 69.77$) than the traditional group ($M_{\text{uncorrected}} = 485.04$, $SD = 76.38$). Further, the group x grade level interaction was significant, $F(2, 116) = 6.31$, $MSE = 1490.78$, $p = .003$, $\eta^2 = .10$. The effect of SES covariate was significant, $F(1, 58) = 5.95$, $MSE = 12521.56$, $p = .02$, $\eta^2 = .10$.

Table 6. Mean STAR Math Scores across Group and Grade Level for Third Graders

	Immersion ($n = 34$)		Traditional ($n = 27$)	
Grade Level	*M*	*SD*	*M*	*SD*
First	414.09	61.24	404.07	65.53
Second	547.18	65.47	494.41	87.59
Third	616.47	77.17	556.63	89.33

The two-way interaction was analyzed further by a one-way ANCOVA performed at each grade level comparing the two groups. An ANCOVA was used instead of a *t* test because SES was significant in the omnibus test. Further, because the number of missing data was different at each grade level, the number of students was different from that of the omnibus test as well as between the first grade and the latter two grade levels. For the first grade, the group difference was not significant, $F(1, 59) = 0.08$, $MSE = 3743.81$, $p = .77$, $\eta^2 = .001$ ($n = 34$ for the immersion and $n = 28$ for the traditional group). The effect of SES approached significance, $F(1, 59) = 3.883$, $MSE = 3743.81$, $p = .05$, $\eta^2 = .06$. For the second grade, the group difference approached significance, $F(1, 63) = 2.91$, $MSE = 6032.19$, $p = .09$, $\eta^2 = .04$ ($n = 38$ for the immersion and $n = 28$ for the traditional group). The effect of SES approached significance, $F(1, 63) = 3.60$, $MSE = 6032.19$, $p = .06$, $\eta^2 = .05$. For the third grade, the group difference was significant, $F(1, 63) = 4.08$, $MSE = 6360.49$, $p = .048$, $\eta^2 = .06$ ($n = 38$ for the immersion and $n = 28$ for the traditional group). The observed mean was higher for the immersion group (*M*

$_{uncorrected}$ = 603.74, *SD* = 82.32) than for the traditional group ($M_{uncorrected}$ = 556.32, *SD* = 87.67). The effect of SES was significant, $F(1, 63) = 9.05$, $MSE = 6360.49$, $p = .004$, $\eta^2 = .13$.

Gain scores from the first to second and from the second to third grade were analyzed using ANCOVAs. From the first to second grade, the group difference was significant, $F(1, 59) = 9.30$, $MSE = 2666.54$, $p = .003$, $\eta^2 = .14$ ($n = 34$ students in the immersion group and $n = 28$ students in the traditional group). The observed means were 133.09 (*SD* = 47.74) for the immersion group and 92.04 (*SD* = 55.16) for the traditional group. The effect of SES was not significant, $F(1, 59) = 0.01$, $MSE = 2666.54$, $p = .94$, $\eta^2 = .00$. From the second to third grade, the group difference was not significant, $F(1, 62) = 0.04$, $MSE = 3517.30$, $p = .83$, $\eta^2 = .001$, with 38 students in the immersion group and 27 students in the traditional group. The observed means were 68.76 (*SD* = 67.44) for the immersion group and 62.22 (*SD* = 47.44) for the traditional group. The effect of SES was not significant, $F(1, 62) = 2.47$, $MSE = 3517.30$, $p = .12$, $\eta^2 = .04$.

Summary. The results from MEAP scores showed that there was no difference between the immersion and traditional group for both the reading and math subtests, even though for the math subtest, there was a trend showing that the scores were somewhat higher for the immersion group than for the traditional group. Similarly, SRI did not show a difference in reading between the two groups, even though there was evidence that gain from the first to second grade was slower for the immersion group than for the traditional group. However, from the second to third grade, gain was greater for the immersion group than for the traditional group, ultimately the immersion group catching up with the traditional group. Based on STAR MATH, there was evidence that the immersion group outperformed the traditional group at the third grade. At the first grade level, these groups were performing similarly, and therefore, the difference at the third

grade level is unlikely to be based on pre-existing advantage by the immersion group. Unfortunately, for this cohort, the Brigance Screen for Kindergarten Readiness was not administered. Accordingly, it was not possible to determine whether the difference was based on the difference in school readiness between the two groups.

Second Grade Cohort

For this cohort, the Brigance Kindergarten screen was administered prior to beginning Kindergarten. Further, for this group, the reading assessment was different between the first and second grade, with the STAR Early Literacy Assessment being administered for the first grade and SRI being administered for the second grade. These two tests were analyzed separately.

Brigance Kindergarten Screen. The scores were analyzed using a one-way ANCOVA with SES as a covariate. The number of students was 27 (15 females and 12 males) for the immersion group and 24 (10 females and 14 males) for the traditional group, with 16 students in the immersion group and 20 students in the traditional group receiving free or reduced lunch. The results showed that the group difference was significant, $F(1, 48) = 4.10$, $MSE = 294.33$, $p = .048$ $\eta^2 = .08$, showing a higher mean for the immersion group ($M_{\text{uncorrected}} = 64.52$, $SD = 17.54$) than for the traditional group ($M_{\text{uncorrected}} = 52.90$, $SD = 16.84$). The effect of SES was not significant, $F(1, 48) = 1.34$, $MSE = 294.33$, $p = .25$ $\eta^2 = .08$. The results indicated that even before entering Kindergarten, the immersion group showed higher school readiness than the traditional group. Because the group difference was significant, Brigance scores were used as an additional covariate for the following analyses.

Reading (STAR Early Literacy Assessment and SRI). For the first grade STAR Early Literacy Assessment, the number of students was 25 (16 females and 14 males) in the immersion group and 23 (22 females and 24 males) in the traditional group, with 19 students in the

immersion group and 38 students in the traditional group receiving free or reduced lunch. The results of a ANCOVA using SES and Brigance as covariates showed that the group difference was not significant, $F(1, 44) = 0.50$, $MSE = 3108.98$, $p = .49$, $\eta^2 = .01$, indicating that the means were similar between the immersion group ($M_{\text{uncorrected}} = 769.36$, $SD = 62.63$) and the traditional group ($M_{\text{uncorrected}} = 783.96$, $SD = 44.72$). Neither the effect of SES, $F(1, 44) = 0.04$, $MSE = 3108.98$, $p = .84$, $\eta^2 = .001$, nor Brigance, $F(1, 44) = 0.42$, $MSE = 3108.98$, $p = .52$, $\eta^2 = .01$, was significant. At the second grade, the analysis of SRI Lexile scores was based on 25 students (13 females and 12 males) in the immersion group and 22 students (9 females and 13 males) in the traditional group, with 16 students in the immersion group and 18 students in the traditional group receiving free or reduced lunch. The results showed that the group difference was not significant, $F(1, 43) = 0.00$, $MSE = 19139.22$, $p = 1.00$ $\eta^2 = .00$, such that there was no difference between the immersion group ($M_{\text{uncorrected}} = 449.16$, $SD = 143.24$) and the traditional group ($M_{\text{uncorrected}} = 417.50$, $SD = 139.26$). The effect of SES was not significant, $F(1, 43) = 0.04$, $MSE = 19139.22$, $p = .85$, $\eta^2 = .001$, whereas the effect of Brigance approached significance, $F(1, 43) = 3.66$, $MSE = 19139.22$, $p = .06$, $\eta^2 = .08$.

Math (STAR MATH). A 2 (group: immersion and traditional) x 2 (grade level: one and two) ANCOVA with SES and Brigance as covariates was conducted with 24 students (13 females and 11 males) in the immersion group and 21 students (8 females and 13 males) in the traditional group. The number of students with free or reduced lunch was 15 for the immersion group and 17 for the traditional group. Table 7 shows the means across group and grade level. The results showed that the main effect of grade level was significant, $F(1, 41) = 28.25$, $MSE = 1278.90$, $p < .001$ $\eta^2 = .41$, indicating that for both groups, there was improvement from the first ($M_{\text{uncorrected}} = 394.02$, $SD = 60.93$) to second grade ($M_{\text{uncorrected}} = 573.49$, $SD = 64.56$). Neither

the main effect of group, $F(1, 41) = 0.24$, $MSE = 6820.66$, $p = .63$ $\eta^2 = .01$, nor the group x grade level interaction, $F(1, 41) = 1.84$, $MSE = 1278.91$, $p = .18$, $\eta^2 = .04$, was significant. There was no effect of SES, $F(1, 41) = 0.01$, $MSE = 6820.66$, $p = .92$, $\eta^2 = .00$, or Brigance, $F(1, 41) = 1.65$, $MSE = 6820.66$, $p = .21$, $\eta^2 = .04$.

Summary. For the second grade cohort, there was no group difference in reading and math. Further, group did not interact with grade level, indicating that students in both groups developed similarly from the first to second grade.

Table 7. Mean STAR Math Scores across Group and Grade Level for Second Graders

	Immersion (n = 24)		Traditional (n = 21)	
Grade Level	*M*	*SD*	*M*	*SD*
First	389.38	65.24	399.33	56.72
Second	541.13	66.01	533.33	64.23

First Grade Cohort

For this cohort, the Brigance Kindergarten screen was administered for 12 students in the immersion group and 7 students in the traditional group. A small number of students with Brigance scores precluded the use of this variable as a covariate. However, an independent sample *t* test showed that the group difference was significant, $t(17) = 1.96$, $p = .03$, Cohen's d = 0.89 (one-tailed), with *a priori* expectation that the mean would be higher for the immersion group than the traditional group. As expected, the observed mean was higher for the immersion group ($M = 65.92$, $SD = 23.12$) than for the traditional group ($M = 41.86$, $SD = 30.22$). For this cohort, reading was assessed by the STAR Early Literacy Assessment, and math was assessed by STAR MATH. Unfortunately, only 7 students in the immersion group took STAR MATH at the kindergarten level, and therefore, only the first grade STAR MATH scores will be analyzed.

Reading (STAR Early Literacy Assessment). A 2 (group: immersion and traditional) x 2 (grade level: kindergarten and one) ANCOVA was performed with SES as a covariate. The number of students was 34 (23 females and 11 males) in the immersion group and 46 (23 females and 22 males) in the traditional group, with 17 students in the immersion group and 34 students in the traditional group receiving free or reduced lunch. The results showed that the main effect of grade level was significant, $F(1,77) = 68.39$, $MSE = 3068.50$, $p < .001$, $\eta^2 = .47$, indicating that there was improvement from Kindergarten to the first grade. The main effect of group was not significant, $F(1,77) = 0.06$, $MSE = 7453.88$, $p = .80$, $\eta^2 = .001$, showing similar means between the immersion group ($M_{\text{uncorrected}} = 714.69$, $SD = 40.76$) and the traditional group ($M_{\text{uncorrected}} = 714.29$, $SD = 72.29$). The group x grade level interaction approached significance, $F(1,77) = 3.46$, $MSE = 3068.50$, $p = .07$, $\eta^2 = .04$, showing a trend that at the Kindergarten level, the mean was somewhat lower for the immersion group than for the traditional group whereas at the first grade level, the mean was somewhat higher for the immersion group than for the traditional group. The effect of SES was not significant, $F(1,77) = 0.82$, $MSE = 7453.88$, $p = .37$, $\eta^2 = .001$.

Math (STAR MATH). A one-way ANCOVA was conducted with SES as a covariate. There were 35 students (24 females and 11 males) in the immersion group and 60 (29 females 31 males) in the traditional group. Further, the number of students with free or reduced lunch was 18 in the immersion group and 45 in the traditional group. The results showed that the group difference was not significant, $F(1,92) = 0.14$, $MSE = 4668.90$, $p = .71$, $\eta^2 = .002$, indicating that there was no difference between the immersion group ($M_{\text{uncorrected}} = 393.86$, $SD = 47.41$) and the traditional group ($M_{\text{uncorrected}} = 387.78$, $SD = 77.37$). The effect of SES was not significant, $F(1,92) = 0.01$, $MSE = 4668.90$, $p = .91$, $\eta^2 = .00$.

Summary. For this cohort, there was no difference between the immersion and traditional groups in reading and math. Similar to the second grade cohort, school readiness was higher for the immersion group than for the traditional group.

Kindergarten Cohort

For this cohort Brigance Kindergarten screen was administered. Further, reading was assessed by the STAR Early Literacy Assessment, and math was assessed by STAR MATH.

Brigance. A one-way ANCOVA with SES as a covariate was performed. The number of students was 36 (19 females and 17 males) in the immersion group and 35 (12 females and 21 males) in the traditional group with 24 students in the immersion group and 24 students in the traditional group receiving free or reduced lunch. The results showed that the main effect of group approached significance, $F(1, 68) = 3.64$, $MSE = 467.36$, $p = .06$, $\eta^2 = .05$, with a two-tailed test. However, a one-tailed test supported *a priori* expectation that school readiness would be higher for the immersion group than for the traditional group ($p = .03$). The observed means were 63.44 ($SD = 22.93$) for the immersion group and 53.46 ($SD = 21.46$) for the traditional group. The effect of SES was significant, $F(1, 68) = 4.85$, $MSE = 467.36$, $p = .03$, $\eta^2 = .07$.

Reading (STAR Early Literacy). A one-way ANCOVA with SES as a covariate was performed with 41 students (22 females and 19 males) in the immersion group and 50 students (21 females and 29 males) in the traditional group with 26 students in the immersion group and 37 students in the traditional group receiving free or reduced lunch. The results showed that the group difference was not significant, $F(1, 88) = 0.30$, $MSE = 7381.43$, $p = .59$, $\eta^2 = .003$, showing that the means were similar between the immersion group ($M_{\text{uncorrected}} = 690.80$, $SD = 80.74$) and the traditional group ($M_{\text{uncorrected}} = 676.96$, $SD = 92.07$). The effect of SES approached significance, $F(1, 88) = 3.59$, $MSE = 7381.43$, $p = .06$, $\eta^2 = .04$.

Math (STAR MATH). A one-way ANCOVA was performed with SES as a covariate. The number of students was 42 (22 females and 20 males) in the immersion group and 14 (3 females and 11 males) in the traditional group with 27 in the immersion group and 10 in the traditional group receiving free or reduced lunch. The results showed that the group difference was not significant, $F(1, 53) = 0.97$, $MSE = 6315.38$, $p = .33$, $\eta^2 = .02$, indicating that the means were similar between the immersion group ($M_{\text{uncorrected}} = 290.62$, $SD = 83.55$) and the traditional group ($M_{\text{uncorrected}} = 313.14$, $SD = 65.28$). The effect of SES was not significant, $F(1, 53) = 1.09$, $MSE = 6315.38$, $p = .30$, $\eta^2 = .02$

Summary. For this cohort, neither reading nor math showed a group difference. However, as expected, the immersion group showed higher school readiness than the traditional group.

CHAPTER IV

DISCUSSION

Although foreign language immersion programs are on the rise in the United States, research on these programs' effects on student achievement in math and other content areas has been scarce, particularly when a character based language, such as Mandarin Chinese, is used as L2. Although some of the previous studies supported the claim that a prolonged immersion experience yields general academic benefits, the results were not consistent across the studies, particularly when partial/dual immersion programs, as opposed to full immersion programs, were examined. The present study was similar to that of Padilla et al. (2013) who examined a Chinese two-way partial immersion program in California and confirmed that such a program indeed produces benefits in English language arts and math, even though at the beginning (second grade), the students in the immersion program experienced slow development in English language arts. As mentioned in the Introduction section, there are a number of differences between the present study and that of Padilla et al. One of the main differences is that the present study investigated a small program in a rural Michigan, which lasted only for four years.

The present study examined achievement scores in English language reading (MEAP, SRI, and STAR Early Literacy Assessment) and math (STAR MATH) by four cohorts of immersion students relative to students in the traditional classrooms. There was no indication that the immersion experience was harmful to academic development in either area, even though immersion experience may have slowed development in reading at the beginning, as evidenced by lower gain scores in reading from the first to second grade by the third grade immersion cohort relative to the traditional group. However, from the second to third grade, the gain scores were greater for the immersion group than for the traditional group, off-setting the difference

overall. Further, for the other cohorts, there was no indication that the immersion experience was associated with slower development in reading. In addition to showing no harmful effect of the immersion experience, there was an indication that the immersion experience was beneficial to achievement in math, as evidence by higher scores on STAR MATH by the immersion group than the traditional group at the third grade. Also, there was a trend on MEAP math scores, showing that the immersion students were slightly ahead of traditional students in math achievement. Taken together, the results showed that the immersion experience did not jeopardize academic achievement of these students in reading English and math, and for math, the immersion experience may have been beneficial to the students.

Padilla et al. (2013) concluded that academic achievement was similar between immersion and traditional groups in English language arts and math. Further, they reported an early performance lag by the immersion students in English language arts, even though by the fourth grade, these students outperformed the traditional students. These researchers also reported that by the fourth and fifth grade, the immersion students outscored the traditional students in math, even though there was no difference between the two groups at the second and third grade. Overall, the pattern of the results these researchers reported was similar to the pattern observed in the present study.

How is it possible that a foreign language immersion experience did not detract from achievement in English language (L1)? One possibility is that children have a natural ability to absorb their native language, and therefore, reduced instructional hours at school may not interfere with their mastery of the native language. It is also possible that their parents were providing additional lessons at home. In support of such a speculation, for the third grade cohort, the SRI Lexile scores showed greater standard deviations in the immersion group relative

to the traditional group (see Table 6). The difference may be based on the fact that some students in the immersion group received extra help while others did not. However, the presence of extra help remains speculative because such a difference in standard deviations was not observed for the reading scores of the other cohorts. Nevertheless, with a national tendency to learn the native language, coupled with additional lessons at home, the immersion students may have had sufficient exposure to English to compensate for the reduced hours at school. An intriguing question is whether learning a foreign language had a facilitative effect on learning the native language. That is, is it possible that the immersion students did not need as much time as the traditional students in learning English because learning a foreign language made learning English more efficient? Based on the present results, it is impossible to answer this question. However, the possibility that learning one language creates a learning set for another language warrants further investigation.

Another question is regarding the advantage in math shown by the immersion students at the third grade. How was it possible that learning math in Chinese language promoted higher achievement (albeit the difference was small) relative to learning math in native English language? One possibility is that Chinese language with its efficient number system was more conducive in learning math than English language. Another possibility is that parents may have provided additional lessons at home, particularly because the parents may have been concerned about their children falling behind.

Another important question is about the role of school readiness. It was clear from the Kindergarten, first grade, and second grade cohorts that school readiness, measured by the Brigance Screen, was higher for the immersion group than for the traditional group. Unfortunately, the Brigance Screen was not administered for the third grade. Can the results be

explained by higher school readiness by the immersion students relative to the traditional students? It is difficult to determine from the present data; however, additional analyses using multiple regressions did not show performance on Brigance Screen to be a reliable predictor of reading and math scores.

Social psychologists might also speculate the role of the self-fulfilling prophecy (Rosenthal & Jacobson, 1968) because it is possible that there was an expectation by the teachers, parents, and administrators that the immersion students would be more successful than their traditional peers. Rosenthal and Jacobson conducted a classic study in which they created a false teacher expectation by informing the teachers that some of their students were "late bloomers." One year later, the results indicated that the late bloomer group showed greater gains in IQ (12 points) relative to the control group (8 points). Over the years, hundreds of follow-up studies have been conducted in both social and education psychology to confirm this effect. Based on the present data, there is no way to determine whether the self-fulfilling prophecy played a role in the lack of disadvantages as well as the presence of advantages by the immersion students. However, there was an interesting indication: the advantage in math by the third grade cohort was primarily based on early gain (from the first to second grade) than late gain (from the second to third grade), a pattern consistent with the results obtained by Rosenthal and Jacobson. Unfortunately, this pattern was not observed with the reading scores. Further, an extensive review by Jussim and Harber (2005) showed that the effect of the self-fulfilling prophecy is miniscule at best and that the effect does not accumulate over time. Nevertheless, positive expectation may have played a role along with a multitude of other confounding variables.

In conclusion, the present study showed that a partial immersion program using Chinese language did not jeopardize academic achievement in reading English language (L1) or math.

Further, there was an advantage in math by the immersion students at the third grade relative to the traditional students. Because a plethora of confounding variables was not controlled, it is difficult, if not impossible to pinpoint the reasons for the results. However, the results were intriguing and therefore warrant further investigation. In particular, future research should use measures that are more sensitive to academic growth such as DIBELS for reading, decoding, and comprehension as well as CBM for mathematical computation and reasoning skills.

REFERENCES

Arthur, G. (2004). Partial Spanish immersion program expands by offering academic excellence for all. *The ACIE Newsletter, 7(2).* Retrieved October 8, 2012, from http://www.carla.umn.edu/immersion/acie/vol7/Feb2004_Partial_Spanish.html.

Bae, J. (2007). Development of English skills need not suffer as a result of immersion: Grades 1 and 2 writing assessment in a Korean/English two-way immersion program. *Language Learning, 57(2),* 299-332.

Barik, H. C. & Swain, M. (1978). Evaluation of a French immersion program: The Ottawa study through grade five. *Canadian Journal of Behavioral Science, 10(3),* 192-201.

Barwell, R. (2010). Tensions in teaching mathematics through a second language. *The ACIE Newsletter, 13(3),* 6-9.

Bournot-Trites, M. & Reeder, K. (2001). Interdependence revisited: Mathematics achievement in an intensified French immersion program. *The Canadian Modern Language Review/La Revue canadienne des langues vivants, 58(1),* 27-43.

Bruer, J.T. (1993). *Schools for thought: A science of learning in the classroom*. Cambridge, MA: The MIT Press.

Brzezinski, J. (2008). Amazing grace: A parent reflects on her daughter's kindergarten year in Chinese immersion. *The ACIE Newsletter, 11(3),* 1, 14-15.

Center for Applied Linguistics. (2011). *Directory of foreign language immersion programs in U.S. schools*. Retrieved August 22, 2014, from http://webapp.cal.org/Immersion/.

Chamot, A.U. & El-Dinary, P.B. (1999). Children's learning strategies in language immersion classrooms. *The Modern Language Journal, 83(iii),* 319-338.

Comeau, L., Comerier, P., Grandmaison, E. & Lacroix, D. (1999). A longitudinal study of phonological processing skills in children learning to read in a second language. *Journal of Educational Psychology, 91(1),* 29-43.

Cummins, J. (1998). Immersion education for the millennium: What have we learned from 30 years of research on second language immersion? In M.R. Childs and R.M. Bostwick (Eds.) *Learning through two languages: Research and practice. Second Katoh Gakuen International Symposium on Immersion and Bilingual Education.* (pp.34-47). Katoh Gakuen, Japan.

Cunningham, T. H. & Graham, C. R. (2000). Increasing native English vocabulary recognition through Spanish immersion: Cognate transfer from foreign to first language. *Journal of Educational Psychology, 92(1),* 37-49.

Curriculum Associates. (2013). *Brigance Early Childhood Screen III (K & 1).* n.d: Curriculum Associates, LLC. Retrieved from http://www.casamples.com/downloads/14299s.pdf.

De Courcey, M. & Burston, M. (2000). Learning mathematics through French in Australia. *Language and Education, 14(2),* 75-95.

De Jong, E.J. (2002). Effective bilingual education: From theory to academic achievement in a two-way bilingual program. *Bilingual Research Journal, 26,* 1 – 20.

Downs-Reid, D. (2000). Using English achievement data to promote immersion education. *The ACIE Newsletter: The Bridge, 1-4.*

Essama, L. (2007). Total immersion programs: Assessment data demonstrate achievement in reading and math. *The ACIE Newsletter: The Bridge, 1-8.*

Genessee, F. (2007). Top ten most consistent findings from research on foreign language immersion. *The ACIE Newsletter, 10(3).* Retrieved October 8, 2012, from http://www.carla.umn.edu/immersion/acie/vol10/may2007_researchfindings.html.

Genesee, F. & Jared, D. (2008). Literacy development in early French immersion programs. *Canadian Psychology, 49(2),* 140-147.

Gibbons, J. (1989). The issue of the language of instruction in the lower forms of Hong Kong secondary schools. In C. Kennedy (Ed.), *Language planning and English language teaching* (pp. 124-134). New York: Prentice Hall.

Haj-Broussard, M. (2005). Comparison Contexts: African-American Students, Immersion and Achievement. *The ACIE Newsletter, 8(3).* Retrieved October 8, 2012, from http://www.carla.umn.edu/immersion/acie/vol8/feb2005_research_africanamerican.html.

Holobow N.E., Genesee, F. Lambert, W.E., Gastright, J. & Met, M. (1987). Effectiveness of partial French immersion for children from different social and ethnic backgrounds. *Applied Psycholinguistics, 8,* 137-152.

Holobow, N. E., Genesee, F., & Lambert, W. E. (1991). The effectiveness of a foreign language immersion program for children from different ethnic and social class backgrounds: Report 2. *Applied Psycholinguistics, 12*, 179-198.

Jared, D., Levy, B. A., Cormier, P., & Wade-Woolley. L. (2011). Early predictors of biliteracy development in children in French immersion: A 4-year longitudinal study. *Journal of Educational Psychology, 103(1),* 119-139.

Jones, C. T. (2005). Spanish immersion and the academic success of Alamo Heights students. *The ACIE Newsletter, 9(1).* Retrieved October 8, 2012, from http://www.carla.umn.edu/immersion/acie/vol9/nov2005_researchreport_spanishimmersion.html.

Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review, 9,* 131-155.

Fortune, T. & Egenberger, A. (2010). STARTALK Mandarin Elementary Immersion Summer Teacher Education Program (MEISTEP). *ACIE Newsletter—The Bridge, 13(3),* 1 – 12.

Knell, E., Haiyan, Q., Miao, P., Yanping, C., Siegel, L.S., Lin, Z., & Wei, Z. (2007). Early English immersion and literacy in Xi'an, China. *The Modern Language Journal, 91(iii),* 395-417.

Knutson, K. (2006). *Because you can't wait until spring: Using the SRI to improve reading performance*. New York: Scholastic Inc.

Kovács, A. M. & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences of the United States of America, 106,* 6556-6560. Doi:10.1073/pnas.0811323106.

Kruk, R. S., & Reynolds, K. (2012). French immersion experience and reading skill development in at-risk readers. *Journal of Child Language, 39,* 580-610. doi:10.1017/S0305000911000201

Lambert, W. E., Tucker, G.R., & d'Anglejan, A. (1973). Cognitive and attitudinal consequences of bilingual schooling: The St. Lambert project through grade five. *Journal of Educational Psychology, 65(2),* 141-159.

Lazaruk, W. (2007). Linguistic, academic and cognitive benefits of French immersion. *The Canadian Modern Language Review/ La Revue canadienne des langues vivantes, 63(5),* 605-628.

Lenker, A. & Rhodes, N. (2007). *Foreign Language Immersion Programs: Features and Trends Over 35 Years.* Washington, DC: Center for Applied Linguistics.

Mantzicopoulos, P. (1999). Reliability and validity estimates of the Brigance K&1 screen based on a sample of disadvantaged preschoolers. *Psychology in the Schools, 36*, 11-19.

Martin-Rhee, M. & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition, 11*, 81-93. doi:10.1017/S1366728907003227

Marsh, H.W., Hau, K.T. & Kong, C.K. (2000). Late immersion and language of instruction in Hong Kong high schools: Achievement growth in language and nonlanguage subjects. *Harvard Educational Review, 70(3)*, 302-346.

McLaughlin, B. (1984). Early bilingualism: Methodological and theoretical issues. In M. Paradis and Y. Lebrun (Eds.) *Early bilingualism and child development.* Lisse: Swets & Zeitlinger B.V.

Met, M. (1993). *Foreign language immersion programs*. Retrieved from ERIC Digest. (ED363141)

Michigan Department of Education. (2013). *MEAP Fall 2013 Test Administrator Manual Grade 3*. State of Michigan: State Administrative Board.

Miller, K. (2010). From the editor. *The ACIE Newsletter, 13(3)*, 2.

Ornstein, A. C. & Levine, D. U. (2000). *Foundations of Education. 7th ed*. Boston: Houghton Mifflin Co.

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(4)*, 661 - 679.

Renaissance Learning. (2014). *STAR Math*. Wisconsin Rapids, WI. Renaissance Learning Inc. Retrieved from http://doc.renlearn.com/KMNet/R004589585C0EE14.pdf.

Renaissance Learning. (2014). *The Science of STAR*. Wisconsin Rapids, WI: Renaissance Learning Inc. Retrieved from http://doc.renlearn.com/KMNet/R003957507GG2170.pdf.

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectations and student intellectual development*. New York: Holt.

Scholastic. (2007). *Scholastic Reading Inventory Technical Guide*. New York: Scholastic Inc. Retrieved from http://teacher.scholastic.com/products/sri_reading_assessment/pdfs/SRI_TechGuide.pdf.

State of Michigan. (2014). *Top-to-Bottom Rankings*. Retrieved September 13, 2014, from http://www.michigan.gov/mde/0,4615,7-140-22709_56562---,00.html.

State of Michigan. (2014). *2013-14 MEAP Trend*. Retrieved September 13, 2014, from https://www.mischooldata.org/DistrictSchoolProfiles/AssessmentResults/Meap/MeapPerformanceSummary.aspx.

Stewart, J. H. (2005). Foreign language study in elementary schools: Benefits and implications for achievement in reading and math. *Early Childhood Education Journal, 33(1),* 11-16.

Stigler, J. W., Lee, S. & Stevenson, H. W. (1986). Digit memory in Chinese and English: Evidence for a temporally limited store. *Cognition, 23,* 1-20.

Swain, M. (1996). Discovering successful second language teaching strategies and practices: from program evaluation to classroom experimentation. *Journal of Multilingual and Multicultural Development, 17*, 117-129.

Swain, M., Lapkin, S., Rowen, N. & Hart, D. (1990). The role of mother tongue literacy in third language learning. *Vox, 4*, 111-121.

Thomas, W. P., Collier, V. P., & Abbott, M. (1993). Academic achievement through Japanese, Spanish or French: The first two years of partial immersion. *The Modern Language Journal, 77(ii),* 170-179.

Turnball, M., Lapkin, S. & Hart, D. (2001). Grade 3 immersion students' performance in literacy and mathematics: Province-wide results from Ontario (1998-99). *The Canadian Modern Language Review/ La Revue canadienne des langues vivantes, 58(1),* 9-26.

Wang, M., Cheng, C., & Chen, S. (2006). Contribution of morphological awareness to Chinese-English biliteracy acquisition. *Journal of Educational Psychology, 98(3),* 542-553.

Weise, E. (2013). *Mandarin Immersion Schools in the United States in 2014.* Retrieved November 16, 2014, from http://miparentscouncil.org/2013/11/20/mandarin-immersion-schools-in-the-united-states-in-2014/.

Willig, A.C. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of Educational Research, 55,* 269-318.