

This article was downloaded by: [University of Arizona]

On: 23 April 2014, At: 19:06

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Bilingual Research Journal: The Journal of the National Association for Bilingual Education

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/ubrj20>

Bilingual Two-Way Immersion Programs Benefit Academic Achievement

Viorica Marian^a, Anthony Shook^a & Scott R. Schroeder^a

^a Northwestern University

Published online: 05 Sep 2013.

To cite this article: Viorica Marian, Anthony Shook & Scott R. Schroeder (2013) Bilingual Two-Way Immersion Programs Benefit Academic Achievement, *Bilingual Research Journal: The Journal of the National Association for Bilingual Education*, 36:2, 167-186, DOI: [10.1080/15235882.2013.818075](https://doi.org/10.1080/15235882.2013.818075)

To link to this article: <http://dx.doi.org/10.1080/15235882.2013.818075>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Bilingual Two-Way Immersion Programs Benefit Academic Achievement

Viorica Marian, Anthony Shook, and Scott R. Schroeder
Northwestern University

The effects of bilingual education on reading and math achievement were examined by comparing test scores across different elementary school programs. Results revealed that bilingual Two-Way Immersion (TWI) programs benefited both minority-language and majority-language students. Minority-language students in TWI programs outperformed their peers in Transitional Programs of Instruction, while majority-language students in Two-Way Immersion outperformed their peers in Mainstream monolingual classrooms. Bilingual Two-Way Immersion programs may enhance reading and math skills in both minority-language and majority-language elementary school children.

INTRODUCTION

Bilingual education policy in the United States is a subject of intense debate (Baker, 2011; Wiley & Wright, 2004). While those who oppose bilingual education continue to cite early evidence that educational instruction in a minority-language student's native language is detrimental to academic success (Rossell, 1990; Rossell & Baker, 1996), there is increasing evidence that providing at least some instruction in the native language actually benefits academic performance (Greene, 1998; Rolstad, Mahoney, & Glass, 2005; Slavin & Cheung, 2005; Willig, 1985). In the present study, we investigated the effectiveness of bilingual education by examining whether a bilingual two-way immersion program could benefit academic performance in minority-language and majority-language elementary school students.

Minority-language students often show a disadvantage in academic performance compared to their majority-language peers (Kindler, 2002; Ruiz-de-Velasco & Fix, 2000). Thus, two crucial issues for policy makers are whether bilingual education can improve minority-language students' academic performance, and, if so, which bilingual education programs result in the largest improvements. Recently, Goldenberg (2008) addressed these issues by discussing the results of two large-scale reviews conducted by the Center for Research on Education, Diversity, and Excellence (CREDE; Genesee, Lindholm-Leary, Saunders, & Christian, 2005) and the National Literacy Panel (NLP; August & Shanahan, 2006) respectively. Consistent with earlier findings (Greene, 1998; Rolstad et al., 2005; Slavin & Cheung, 2005; Willig, 1985), CREDE

Viorica Marian, Anthony Shook, and Scott R. Schroeder are all affiliated with Northwestern University.

Address correspondence to Viorica Marian, Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston, IL 60208. E-mail: v-marian@northwestern.edu

and NLP in their large-scale reviews suggested that bilingual education programs (most commonly, transitional bilingual education programs) lead to better performance on English literacy measures than submersion, structured immersion, and ESL programs. In addition to these two large-scale reviews, there is also recent evidence from long-term, randomized studies that transitional bilingual education programs result in successful academic performance in English while further developing students' proficiency in their native language (Irby et al., 2010; Slavin, Madden, Calderón, Chamberlain, & Hennessy, 2011).

Moreover, while transitional bilingual education programs are known to be effective at educating minority-language students, dual-language two-way immersion programs may be even more effective (Lindholm-Leary, 2005; Lindholm-Leary & Howard, 2008; Thomas & Collier, 2002). For example, Collier and Thomas (2004) collected data from minority-language elementary students enrolled in a large school district in Texas and found that students in two-way immersion programs outperformed students in transitional bilingual education programs on both English and Spanish reading tests.¹

Because two-way immersion programs include not only minority-language students but also majority-language students (Christian, Howard, & Loeb, 2000), a complete evaluation of two-way immersion's effectiveness should take into account majority-language students' performance as well. Findings from studies assessing majority-language students' performance indicate that majority-language students in two-way immersion programs outperform their peers in mainstream classrooms (e.g., Genesee, 1983; Thomas & Collier, 2002), but the evidence remains limited. If indeed two-way immersion programs lead to improved academic performance in majority-language students, it would suggest that two-way immersion can be utilized as an enrichment tool for both minority-language speakers and majority-language speakers.

Despite the promising research on two-way immersion programs, current policies severely restrict bilingual education programs in certain states that have many minority-language students, such as Arizona and California (Gándara & Orfield, 2012; Mackinney & Ríos-Aguilar, 2012). The impact of these policies on students' performance has been analyzed recently, and the results suggest that restrictive bilingual education policies can lead to ineffectual academic programs (Gándara & Hopkins, 2010; Ríos-Aguilar, Canché, & Sabetghadam, 2012). Given the potential ineffectiveness of these current educational programs, along with the ongoing debate among educators and policymakers as to which programs best promote academic success, additional research is necessary.

Our aim in the current study was to examine whether two-way immersion programs benefit academic achievement in both minority- and majority-language students. We analyzed reading and math standardized test scores from third-, fourth-, and fifth-grade students enrolled in various educational programs in a single school district. We compared the test performance of minority-language students enrolled in a two-way immersion program to that of minority-language students enrolled in a transitional program of instruction. Comparisons also were made between majority-language students participating in the two-way immersion program and majority-language students in the English-only mainstream classroom. Based on previous evidence that two-way immersion is a highly effective program of instruction (Collier & Thomas, 2004; Genesee, 1983; Thomas & Collier, 2002), we predicted that minority-language students

¹One limitation of the Collier and Thomas (2004) study is that students were not randomly assigned to groups.

in the two-way immersion program would outperform their peers in the transitional program of instruction, and that majority-language students in the two-way immersion program would outperform their peers in a mainstream classroom.

METHODS

Participants

We obtained cross-sectional data from all 2,009 third-, fourth-, and fifth-grade public school students² enrolled in a school district in the Chicago-land area (see Note 3 for ethnic and financial information about the district).³ Of the 2,009 students, 157 were minority-language Spanish-speaking students. In the group of 157 minority-language Spanish-speaking students, 134 were third- ($N = 58$), fourth- ($N = 38$), or fifth- ($N = 38$) grade students enrolled in the two-way immersion program (abbreviated as TWI-S for two-way immersion native Spanish-speaking students), and 23 were third- ($N = 9$), fourth- ($N = 8$), or fifth- ($N = 6$) grade students enrolled in a transitional program of instruction/English as a second language (abbreviated as TPI). The TWI-S program was available to any student whose parents listed Spanish on a home-language survey and who qualified as a minority-language speaker as defined by performance on an English-language proficiency test (ACCESS; Assessing Comprehension and Communication in English State-to-State). Students who met these criteria were automatically enrolled in the TWI-S program; however, parents could waive this service if they preferred to place their child in the TPI program or a mainstream classroom. In the TWI-S group, 86.6% of students qualified for free or reduced price lunch and were therefore considered to be of low socioeconomic status (Caldas & Bankston, 1997; Sirin, 2005). In the TPI group, 47.8% of students qualified for free or reduced-price lunch and were considered to be of low socioeconomic status (SES). The TWI-S group had a significantly larger proportion of low SES students relative to the TPI group ($\chi^2 = 19.1, p < .001$); thus, if better performance in the TWI-S group is observed, it is unlikely to be attributable to increased social and economic resources. The TWI-S and TPI groups did not differ in gender distribution (51.5% females in the TWI-S group and 52.2% females in the TPI group; $\chi^2 = .004, p > .1$).

There were 1,852 majority-language students whose native language was English. Of these, 75 were third ($N = 37$), fourth ($N = 19$), or fifth ($N = 19$) graders enrolled in the two-way immersion program (abbreviated as TWI-E for two-way immersion native-English-speaking students), and 1,777 were third ($N = 574$), fourth ($N = 579$), or fifth ($N = 624$) graders enrolled in an English-only mainstream classroom setting (abbreviated as MC). Enrollment in the TWI-E program was determined through a lottery system. Information about the TWI-E program and the lottery process was provided at an orientation session held for parents of incoming kindergarten students. Parents entered their children into the lottery by completing an online form. On average, 17.4% of majority-language parents participate in the lottery each year. A subset of the students

²All students, including those with special needs, were counted in the sample.

³With respect to the ethnic and racial composition of the district, 42% of students are Caucasian, 39.1% are African American, 13.7% are Hispanic, 4.7% are Asian, 0.1% are American Indian, and 0.4% are multiracial. Regarding the financial and economic status of the district, the instructional expenditure per student in the district is \$7,468.

whose parents entered the lottery was randomly selected to enroll in the TWI program.⁴ English-native students who were not selected to enroll in the TWI-E program were placed in mainstream classrooms. The TWI-E group had a smaller proportion of low SES students (9.3%) than the MC group (37%; $\chi^2 = 24.0, p < .001$). Because the TWI-E group had fewer students in the low SES group, any performance advantages observed in the TWI-E group could be due in part to increased social and economic resources. Consequently, further analyses were done to match the TWI-E and MC group in SES (see Data Analysis Section). The TWI-E and MC groups did not differ in gender distribution (48% females in the TWI-E group and 49.1% females in the TPI group; $\chi^2 = .04, p > .1$).

Instructional Programs

All programs followed the same curriculum and differed only in the language of instruction. The minority-language students in the two-way immersion program (i.e., the TWI-S students) were initially taught reading and writing in their native language (i.e., in Spanish in kindergarten through second grade) and then in their second language (i.e., in English in third grade through fifth grade). From kindergarten to third grade, TWI-S students were taught math in Spanish; in fourth and fifth grades, math was taught in English. Social studies and science were taught in Spanish from kindergarten through fifth grade. (For math, social studies, and science, minority-language students in the two-way immersion program were integrated with majority-language students in the two-way immersion program.)

The minority-language students in the transitional program of instruction (i.e., the TPI students) were placed in a mainstream classroom along with majority-language students and were taught reading, writing, math, social studies, and science in English in all grades. The students in the transitional program of instruction also received as-needed pull-out ESL instruction.

The majority-language students in the two-way immersion program (i.e., the TWI-E students) were taught reading and writing in English from kindergarten through second grade, and then in Spanish from third grade through fifth grade. These students received math instruction in Spanish from kindergarten through third grade and in English in fourth and fifth grades. Social studies and science were taught in Spanish from kindergarten through fifth grade. (The majority-language students in the two-way immersion program were integrated with minority-language students in the two-way immersion program for math, social studies, and science.) The majority-language students in the mainstream classroom completed a similar curriculum as majority-language two-way immersion students, but English was used exclusively during all class instruction. The percentage of time each program used English versus Spanish in the core subjects is presented in Tables 1 and 2.

In all programs, the teachers met state-mandated standards. All teachers had completed a state-approved educational program, fulfilled a student teaching requirement, and passed the necessary Illinois Certification Testing System tests to obtain an Illinois Elementary teaching certificate. Teachers of two-way immersion classes also had a Transitional Bilingual certificate or bilingual approval or endorsement. The teachers in the district had an average of 13.1 years of teaching experience, all teachers had a bachelor's degree, and 57.7% also had a master's degree.

⁴All majority-language students were eligible to apply for the lottery; students did not have to meet any specific qualifications to be eligible.

TABLE 1
 Percentage of Time Minority-Language Students Were Exposed to English and Spanish in the Core Subjects (i.e., Reading, Writing, Math, Social Studies, and Science)

	<i>Two-Way Immersion Spanish Native</i>		<i>Transitional Program of Instruction</i>	
	<i>English</i>	<i>Spanish</i>	<i>English</i>	<i>Spanish</i>
Kindergarten	0%	100%	100%	0%
First Grade	0%	100%	100%	0%
Second Grade	0%	100%	100%	0%
Third Grade	40%	60%	100%	0%
Fourth Grade	60%	40%	100%	0%
Fifth Grade	60%	40%	100%	0%

TABLE 2
 Percentage of Time Majority-Language Students Were Exposed to English and Spanish in the Core Subjects (i.e., Reading, Writing, Math, Social Studies, and Science)

	<i>Two-Way Immersion English Native</i>		<i>Mainstream Classroom</i>	
	<i>English</i>	<i>Spanish</i>	<i>English</i>	<i>Spanish</i>
Kindergarten	40%	60%	100%	0%
First Grade	40%	60%	100%	0%
Second Grade	40%	60%	100%	0%
Third Grade	0%	100%	100%	0%
Fourth Grade	20%	80%	100%	0%
Fifth Grade	20%	80%	100%	0%

Measures

Performance in Reading and Math was measured using two state-mandated standardized tests, the State Measure of Annual Growth in English (taken by minority-language students enrolled in TWI-S and TPI programs) and the State Standards Achievement Test (taken by majority-language students enrolled in TWI-E and MC programs). The State Measure of Annual Growth had been administered since 1996, and the State Standards Achievement Test had been given since 1999. Both tests were based on state standards of academic performance. The measures of interest derived from these tests were the Scaled-Math and Scaled-Reading scores, which reflect a student’s overall performance based on a weighted average of several math and reading subtests. The State Measure of Annual Growth in English and the State Standards Achievement Test both consist of several measures of reading comprehension (e.g., short passages followed by multiple-choice questions) and math knowledge (e.g., single-sentence, multiple-choice word problems, and simple arithmetic) and tap into the same academic abilities. The State Measure of Annual Growth in English test was designed for English language learners and therefore

includes more simple language than that of the State Standards Achievement Test but addresses the same content.⁵ Questions on the State Measure of Annual Growth in English tended to contain higher-frequency words and less-complex sentence structures as compared to the State Standards Achievement Test. Sample math and reading items for both tests are provided in the Appendix.

The State Standards Achievement Test is regarded as a reliable and valid assessment of reading and math achievement. Reliability can be measured by calculating students' consistency in performance across the different items on the test (known as internal consistency). Cronbach's coefficient alphas were high for both reading and math in all three grades (reading in third grade = .91, fourth grade = .90, fifth grade = .91; math in third grade = .92, fourth grade = .92, fifth grade = .93), indicative of strong reliability. Analyses of construct validity and criterion-related validity can also be calculated. The term *construct validity* refers to the extent to which a test measures what it is intended to measure. This type of validity can be assessed through measures of dimensionality (i.e., how many different dimensions the tests contains). Divgi values for both subjects in all grades were greater than 3, meaning that the tests assess only a single dimension—math or reading—as they intend to assess. A component of construct validity called *internal construct* can be quantified as well. *Internal construct* refers to whether performance on subscales of the test correlates with performance on the whole test. If the subscales tap into the dimension of interest (i.e., reading or math), then the subscales will be highly correlated with the whole test. For both subjects in all grades, correlations were at least .85, suggesting high construct validity. Finally, a type of criterion-related validity called *concurrent validity* was considered. This type of validity refers to whether the test in question correlates with other tests that are known to be valid assessments of the constructs of interest. Performance on the State Standards Achievement Test correlated highly with the reading and math portions of Tenth Edition of the Stanford Achievement Test (higher than .85 in all grades), providing evidence for the test's concurrent validity. Taken together, these measures lend support to the notion that the State Standards Achievement Test is reliable and valid. Data regarding the reliability and validity of the State Measure of Annual Growth in English are not available but are likely similar to the State Standard Achievement test, given that the State Measure of Annual Growth was designed to address the same content as the State Standard Achievement Test but with simpler vocabulary.

Of the students who completed the State Standards Achievement Test, 82% of the third graders in the current sample met or exceeded state standards in reading, and 91% met or exceeded state standards in math. Among fourth graders, 81% met or exceeded state standards in reading, and 91% met or exceeded state standards in math. Among fifth graders, 76% met or exceeded state standards in reading, and 88% met or exceeded state standards in math. (These data are not available for the State Measure of Annual Growth in English.)

Data Analysis

The first set of analyses compared minority-language students in TWI-S and TPI groups on standardized test scores derived from the State Measure of Annual Growth in English. Specifically, TWI-S and TPI groups were compared at each grade level (third, fourth, and fifth graders) in

⁵Because the minority- and majority-language students were administered different tests, no direct statistical comparisons were made between these two tests and groups.

each subject (Reading and Math) using the Mann-Whitney U test. This test was chosen over the parametric *t*-test due to a discrepancy in sample sizes between groups. In all comparisons, the scores were independent of each other and came from two samples, confirming the appropriateness of the Mann-Whitney U test. Moreover, the groups' distributions were similar (as verified by nonsignificant Levene's tests). Because of sufficient sample sizes (i.e., either N_1 or $N_2 > 20$), asymptotic *p* values are provided for the Mann-Whitney U tests (Weinberg & Abramowitz, 2002). To supplement the Mann-Whitney U analyses, 95% confidence intervals and two types of effect size measurements are provided. Namely, the difference in mean ranks (Green & Salkind, 2008) and Pearson's *r* are used to index the size of the effects. In addition to comparing the TWI-S and TPI groups at each grade, differences from grade to grade within each program were examined by conducting Kruskal-Wallis tests with follow-up Tamhane T2 tests. It is important to note that comparisons across grades were done on cross-sectional data with no pretest and therefore should be interpreted with caution.

The second set of analyses compared the majority-language students in TWI-E and MC on standardized test scores derived from the State Standards Achievement Test. Comparisons were made at each grade level (third, fourth, and fifth graders) in each subject (Reading and Math) with Mann-Whitney U tests. The Mann-Whitney U test was used because of unequal sample sizes between groups. The scores for the Mann-Whitney U test were independent of each other and came from two different samples. In the third- and fourth-grade samples, variance distributions differed between the groups (as reflected in significant Levene's tests). The Mann-Whitney U test is sensitive to differing variances, but it is often less sensitive than its alternative, the *t*-test, and for that reason it is frequently used in cases of unequal variance (Cardinal, Pennicott, Sugathapala, Robbins, & Everitt, 2001; Delis et al., 1991; Sañudo-Peña, Patrick, Patrick, & Walker, 1996). Given the unequal sample sizes and unequal variances, the Mann-Whitney U test was deemed the most appropriate test for these analyses. For the Mann-Whitney U analyses, asymptotic *p* values are provided. In addition to the Mann-Whitney U results, 95% confidence intervals, differences in mean ranks, and Pearson's *r* effect sizes are included. Moreover, Kruskal-Wallis tests and follow-up Tamhane T2 tests were conducted to analyze differences between grades within each program.

We also conducted further comparisons between the TWI-E and MC groups in order to control for differences in SES. Because the MC group had more students with low SES, lower performance in this group could be attributable in part to the larger number of students who have limited economic or social resources. To address the potential confound of low SES in the MC group, we also conducted the analyses with the low SES students excluded from the two groups.

RESULTS

Test Performance in Minority-Language Students Enrolled in TWI-S or TPI

The Reading and Math scaled-scores (mean and 95% confidence intervals) for minority-language students enrolled in the TWI-S and TPI programs are presented in Figure 1 (Reading), Figure 2 (Math), and Table 3 (both Reading and Math).

In the sample of fifth graders, TWI-S students significantly outperformed their TPI peers in Math ($U = 39.5$, $p < .01$, difference in mean ranks = 13.9, $r = .38$) but not in Reading ($U = 70$,

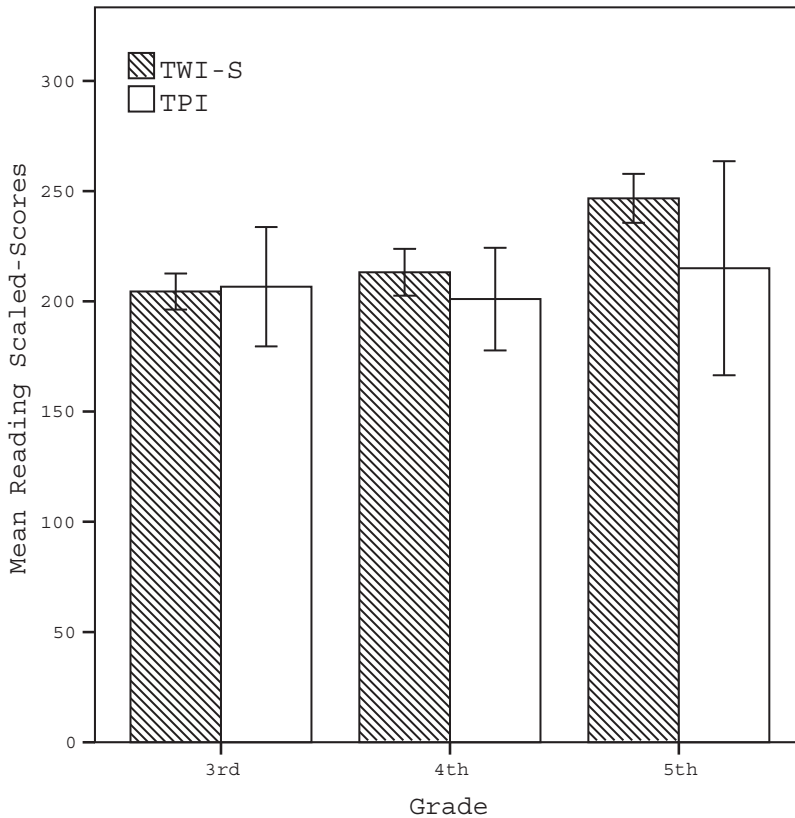


FIGURE 1 Reading scaled-scores for minority-language two-way immersion Spanish (TWI-S) and transitional program of instruction (TPI) students across Grades 3, 4, and 5. Error bars signify 95% confidence intervals.

$p > .1$, difference in mean ranks = 7.94, $r = .22$). Fourth graders did not show significant differences in Reading ($U = 102$, $p > .1$, difference in mean ranks = 4.7, $r = .13$) or Math ($U = 105$, $p > .1$, difference in mean ranks = 6.5, $r = .19$) and similar performance was also observed among the third graders (Reading $U = 205$, $p > .1$, difference in mean ranks = 2.7, $r = .05$; Math $U = 194$, $p > .1$, difference in mean ranks = 8.0, $r = .14$).

Notably, the TWI-S students showed significant improvements across grade levels. Specifically, both Reading and Math scaled-scores were better in TWI-S students in the higher grades relative to students in the lower grades (Reading: $\chi^2 = 31.0$, $df = 2$, $p < .001$, $\eta^2 = .24$; Math: $\chi^2 = 52.8$, $df = 2$, $p < .001$, $\eta^2 = .41$), indicating significant growth in spite of the fact that the standardized tests were grade and age adjusted. The TPI group did not show such improvements across grades (Reading: $\chi^2 = 0.42$, $df = 2$, $p > .05$, $\eta^2 = .02$; Math: $\chi^2 = 4.7$, $df = 2$, $p > .05$, $\eta^2 = .21$). For the TWI-S students, standardized Reading scores were significantly better among fifth graders relative to fourth and third graders (post hoc Tamhane T2, both $ps <$

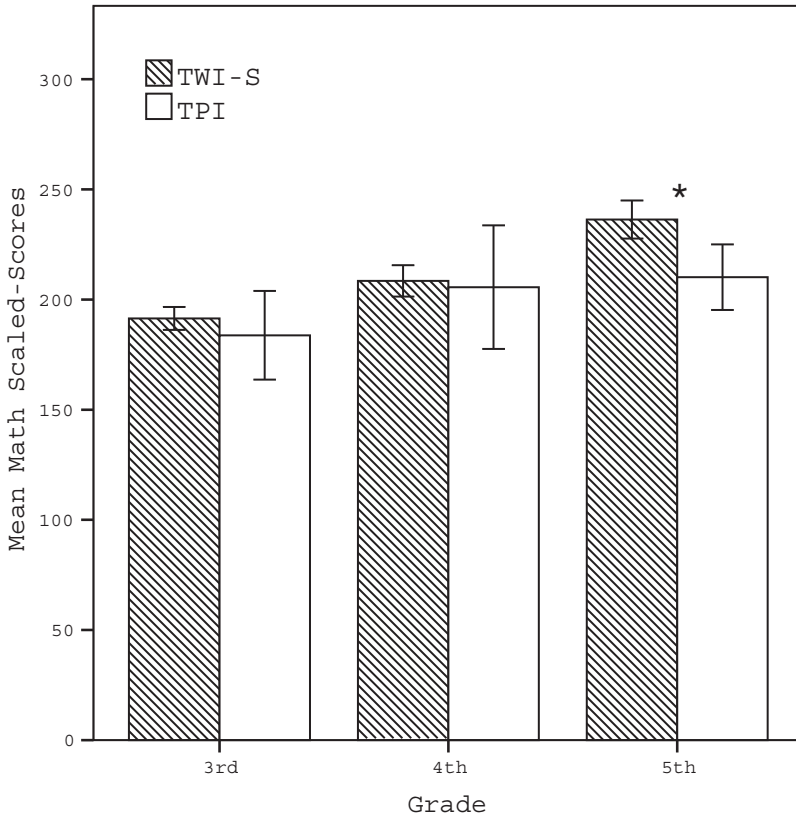


FIGURE 2 Math scaled-scores for minority-language two-way immersion Spanish (TWI-S) and transitional program of instruction (TPI) students across Grades 3, 4, and 5. *represents significance at $p < .05$. Error bars signify 95% confidence intervals.

TABLE 3
Reading and Math Scaled-Scores in Minority-Language Students

	<i>Two-Way Immersion Spanish Native</i>			<i>Transitional Program of Instruction</i>		
	<i>N</i>	<i>Reading Mean & CI</i>	<i>Math Mean & CI</i>	<i>N</i>	<i>Reading Mean & CI</i>	<i>Math Mean & CI</i>
Third Grade	58	205, 197–213	192, 187–197	9	207, 180–234	185, 162–208
Fourth Grade	38	213, 202–224	207, 210–214	8	201, 178–224	208, 175–241
Fifth Grade	38	248, 237–259	237, 228–246	6	215, 166–264	210, 195–225

.001). Standardized Math scores were significantly better among fifth graders relative to fourth and third graders and better among fourth graders relative to third graders (post hoc Tamhane T2, all $ps < .01$). These cross-sectional analyses suggest that the TWI program may lead to significant improvements across grades in minority-language students.

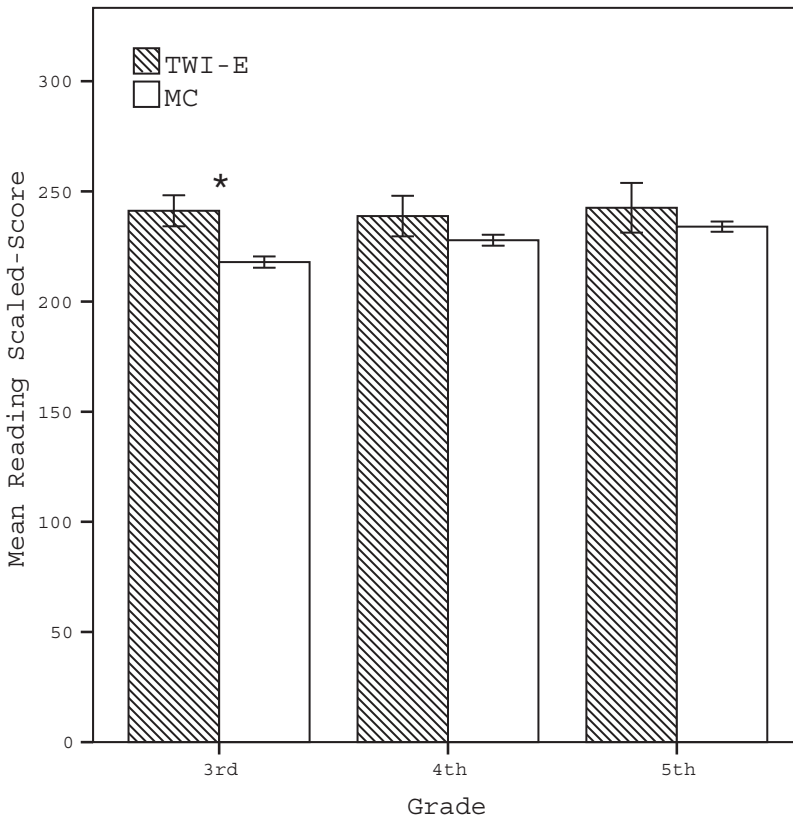


FIGURE 3 Reading scaled-scores for majority-language two-way immersion English (TWI-E) and mainstream classroom (MC) students across Grades 3, 4, and 5. *represents significance at $p < .05$. Error bars signify 95% confidence intervals.

Test Performance in Majority-Language Students Enrolled in TWI or MC

The Reading and Math scaled-scores (mean and 95% confidence intervals) for majority-language students enrolled in TWI-E or MC are presented in Figure 3 (Reading), Figure 4 (Math), and Table 4 (both Reading and Math).

Among fifth graders, TWI-E students outperformed MC students in Math ($U = 4046.5$, $p < .01$, difference in mean ranks = 100.5, $r = .09$), while no group differences were observed in Reading ($U = 5042.5$, $p > .1$, difference in mean ranks = 47.8, $r = .04$). Similarly, in fourth graders, the TWI-E group performed better than the MC group in Math ($U = 3442$, $p < .01$, difference in mean ranks = 111.8, $r = .11$) but not Reading ($U = 4248$, $p = .09$, difference in mean ranks = 68.1, $r = .07$). Among third graders, the TWI-E students outscored MC students in both Reading ($U = 5692$, $p < .001$, difference in mean ranks = 140.7, $r = .19$), and Math ($U = 5326$, $p < .001$, difference in mean ranks = 151.7, $r = .21$).

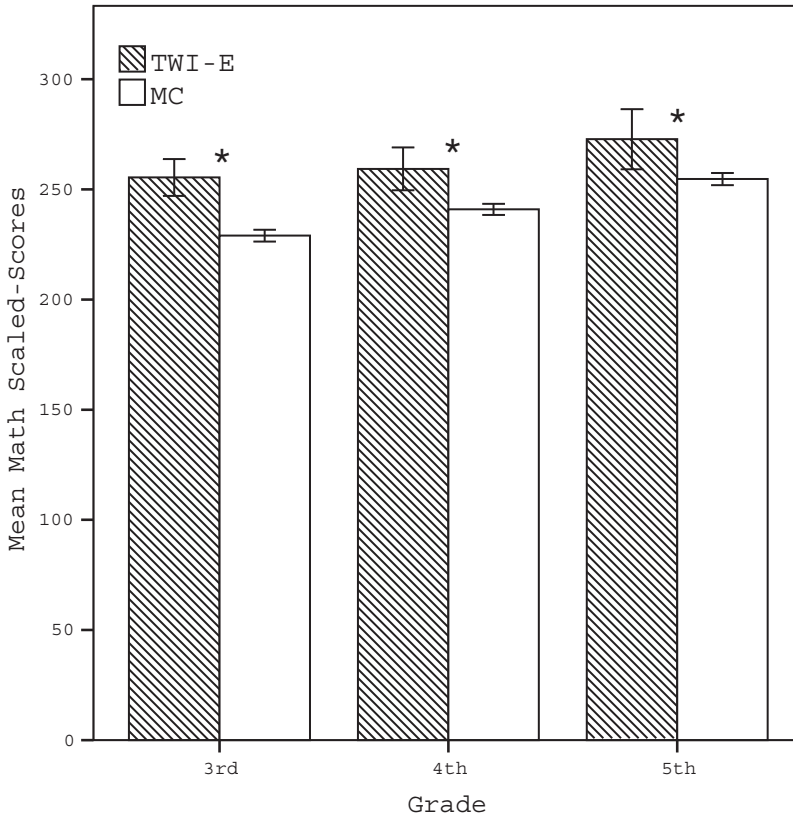


FIGURE 4 Math scaled-scores for majority-language two-way immersion English (TWI-E) and mainstream classroom (MC) students across Grades 3, 4, and 5. *represents significance at $p < .05$. Error bars signify 95% confidence intervals.

TABLE 4
Reading and Math Scaled-Scores in Majority-Language Students

	<i>Two-Way Immersion English Native</i>			<i>Mainstream Classroom</i>		
	<i>N</i>	<i>Reading Mean & CI</i>	<i>Math Mean & CI</i>	<i>N</i>	<i>Reading Mean & CI</i>	<i>Math Mean & CI</i>
Third Grade	37	241, 234–248	255, 247–264	574	218, 215–221	229, 226–232
Fourth Grade	19	239, 230–248	259, 250–269	579	228, 225–230	241, 238–243
Fifth Grade	19	243, 231–254	273, 259–286	624	234, 232–236	255, 252–257

Because the MC group had a larger proportion of low SES students than the TWI-E group, the comparisons were also done with all of the low SES students excluded from the analyses. The superior performance in the TWI-E group relative to the MC group persisted in both Reading ($U = 4879.5, p < .01$, difference in mean ranks = 57.5, $r = .14$) and Math ($U = 4440, p < .01$,

difference in mean ranks = 71.9, $r = .17$) for third grade students. The TWI-E advantage in Math among fourth- and fifth-grade students followed the same pattern but did not reach significance (fourth grade: $U = 2635$, $p > .1$, difference in mean ranks = 37.4, $r = .07$; fifth grade: $U = 2050.5$, $p > .1$, difference in mean ranks = 43.09, $r = .07$).

Analyses were also conducted to examine differences among grade levels in each group. The analyses were performed with all students included as well as with low SES students removed. Both analyses yielded similar results; therefore, we only report the analyses with all students included. In the TWI-E group, students in higher grades did not differ from students in lower grades in grade-adjusted Reading scores ($\chi^2 = 0.39$, $df = 2$, $p > .1$) or Math scores ($\chi^2 = 3.76$, $df = 2$, $p > .1$). Because the TWI-E students were already excelling in the early grades, the similar performance across grades likely reflects maintenance of high test performance. In the MC group, students in the higher grades had better grade-adjusted scores than students in the lower grades in Reading ($\chi^2 = 84.12$, $df = 2$, $p < .01$) and in Math ($\chi^2 = 148.71$, $df = 2$, $p < .01$). For both Reading and Math, fifth graders performed better than fourth and third graders, and fourth graders performed better than third graders (all $ps < .01$). As a result of the improved performance across grades, by fifth grade, the MC group performed similar to their TWI-E peers in Reading, but the MC group still had lower scores in Math.

DISCUSSION

The results of the present study suggest that bilingual two-way immersion education is beneficial for both minority- and majority-language elementary students. In the minority-language students, standardized reading and math scores in the two-way immersion (TWI-S) group increased across grades, with students in higher grades performing better than students in lower grades. Conversely, in the transitional program of instruction (TPI), standardized reading and math scores did not increase across grades, as students in the higher grades did not perform significantly better than students in the lower grades. Moreover, in the oldest students (i.e., the fifth graders), the TWI-S group outperformed the TPI group in math. In the majority-language students, the two-way immersion (TWI-E) group outperformed the mainstream classroom (MC) students in math in third, fourth, and fifth grade and in reading in third grade. These results are consistent with previous research (e.g., Collier & Thomas, 2004; Genesee, 1983; Lindholm-Leary, 2005; Lindholm-Leary & Howard, 2008; Thomas & Collier, 2002) and suggest that two-way immersion programs can benefit reading and math performance in elementary school children.

While both the majority-language and minority-language TWI students exhibited reading and math advantages over their non-TWI peers, these benefits manifested at different times in the two groups. The benefits were observed earlier in the majority-language TWI students (i.e., in the third graders) and later in the minority-language TWI students (i.e., in the fifth graders). Previous work suggests that it can take four to seven years for minority-language speakers to develop enough proficiency for successful academic performance (e.g., Hakuta, Butler, & Witt, 2000). The minority-language students may therefore have not had sufficient English-language proficiency in the early grades to perform successfully on the English-based tests, thereby delaying the benefits in minority-language students. In addition, the early emergence of academic advantages in the majority-language students may be due in part to their higher socioeconomic status. The majority-language TWI group had fewer low SES students (9.3%) than the minority-language

TWI group (86.6%), and may therefore have had the necessary social and economic resources and support to benefit immediately from bilingual education.

A further explanation for why the benefits emerged earlier in the majority-language TWI students is that by third grade, majority-language TWI students may have had more bilingual experience than the minority-language TWI students. From kindergarten to second grade, students enrolled in the TWI program were taught math, science, and social studies in Spanish and reading and writing in their native language. In other words, minority-language native-Spanish-speaking students received all of their instruction in Spanish until second grade, while majority-language native-English-speaking students received more balanced exposure to both of their languages. If degree of bilingual experience plays a role in developing the observed advantages, then majority-language TWI students may outperform their mainstream peers early on, while the minority-language students may not. However, as minority-language TWI students' curriculum shifts to more language-balanced education in Grades 3, 4, and 5, they begin to show improvements in math that are not found in the TPI students who undergo a less immersive bilingual experience. These results suggest that balanced-language instruction may promote academic achievement in both majority- and minority-language students.

As the minority-language TWI students begin to exhibit gains over their peers in the later grades, the early advantages observed in the majority-language TWI students may diminish to some degree. Indeed, among fifth-grade students, the majority-language TWI students showed an advantage only in math, and that advantage was not statistically significant when the low SES students were excluded. Nevertheless, the early benefits of two-way immersion education on reading and math performance should not be disregarded, as students are often tracked from early grades, and their performance in elementary school can determine whether they are placed in accelerated or remedial classrooms.

A possible explanation for TWI students' improved reading skills is that two-way immersion education may foster more direct attention to language use. That is, TWI students are more aware of their different languages, and they are encouraged to use both regularly. This intermixing of languages may result in increased metalinguistic knowledge, which has been shown to correlate with bilingualism (Bialystok, 1988; Campbell & Sais, 1995; Cromdal, 1999), and has also been shown to predict later reading skills (Dreher & Zenge, 1990). A second possible explanation is that the high transparency of Spanish orthography facilitated the acquisition of alphabetic principles in the TWI students.

In addition to enhancements in reading, the TWI groups also showed advantages on the math test. In fact, the strongest TWI advantages were observed in math. A potential explanation for increased math ability in TWI students comes from research on the nonlinguistic, cognitive benefits of bilingualism (for a review, see Bialystok, 2007). Bilingual children have been shown to exhibit increased executive functioning skills (Bialystok & Martin, 2004; Carlson & Meltzoff, 2008), which correlate with math performance (Blair & Razza, 2007; Bull, Espy, & Wiebe, 2008; Mazzocco & Kover, 2007; McClelland et al., 2007; Passolunghi & Seigel, 2001). For example, Bull and Scerif (2001) assessed third graders with executive functioning tasks (e.g., the Stroop task and the Wisconsin Card Sorting task) and with a mathematics test of addition and subtraction; through multiple regression analyses, the authors found that executive functioning ability reliably predicted math performance. Executive functioning may aid performance on math problems by allowing the student to hold the problem in working memory, to shift one's focus between different aspects of the problem and different approaches to the problem, and to suppress a tendency

to respond to salient but irrelevant elements of the problem (Blair & Razza, 2007). Moreover, executive functioning can help students focus on mathematical lessons during class instruction, which could result in enhanced learning of mathematical principles and ultimately better performance on standardized math tests. Thus, potential enhancements in executive functioning for TWI students could lead to their advantages in math (and to some extent in reading as well).

The improved math and reading performance observed in minority-language TWI students may also be due in part to these students being taught academic concepts in their native language. In the earlier grades, the minority-language TWI students were taking reading and math class in Spanish, their native and likely stronger language; this may have increased comprehension of the lessons and enabled the students to learn the material more effectively. In contrast, the TPI students were learning about reading and math in English, their nonnative language; this may have negatively impacted their ability to understand and subsequently learn the material. Thus, receiving instruction in the native language may account for the increased performance in TWI minority-language students, along with potential advantages in executive functioning and metalinguistic awareness.

While interpreting the reading and math enhancements observed in the TWI students, it is important to consider some limitations of the current study. Notably, students were not fully randomized to groups and were not tested before the start of the program. Therefore, there is a possibility that the superior performance of TWI students was driven by a selection bias, as students with higher social and economic resources may have been more likely to enroll in a TWI program and may have been advantaged from the outset. While potentially contributing to the results, a selection bias is not able to fully account for the TWI advantages. The minority-language TWI-S group actually had a higher proportion of low SES students than their TPI peers, and yet the TWI-S group still performed better than the TPI group. The majority-language TWI-E group had fewer low SES students than the MC group, but when comparisons were made with low SES students excluded, TWI advantages were still observed. Moreover, if a selection bias were driving the TWI advantages exclusively, then we would expect to see the same benefits in both math and reading measures, across all grades, but that was not the case. Thus, while a selection bias may have played a role in the results, it cannot fully account for the observed TWI benefits.

A second potential limitation is that minority-language students took the standardized test in their nonnative language. Therefore, the test may have been testing language ability more so than conceptual understanding of the material, thus negatively impacting internal validity. Indeed, even math tests, which ostensibly have a reduced language processing load, can impose high demands on language-comprehension skills (Abedi & Lord, 2001). However, it is important to reiterate that comparisons were made between two different minority-language groups (and not between minority-language groups and majority-language groups), and thus students who took the test in their nonnative language were only compared to other students who also took the test in their nonnative language. Moreover, it should be noted that although the test may have had reduced internal validity, this testing situation is common in educational settings, where students often have to take tests in their nonnative language (thereby increasing external validity).

A third limitation to consider is the small sample size in some of the groups, particularly in the TPI groups. Because performance by the TPI students was based on fewer data points, there is a possibility that the data are not fully reflective of the program's degree of effectiveness. The reduced sample sizes also limit the ability to detect significant differences between groups; yet

the sample sizes were adequate for statistically significant differences to emerge, as the TWI-S group reliably outperformed the TPI group in the oldest grade. Nevertheless, follow-up studies with larger samples are necessary to verify the current findings.

In future work, researchers should also consider the effects of two-way immersion education on other aspects of cognitive functioning (apart from reading and math test performance), as two-way immersion may have far-reaching benefits. For example, two-way immersion may improve components of the executive function system, such as selective attention, inhibitory control, and task switching. Indeed, practice focusing on one language, suppressing the other language, and shifting between two languages has been shown to increase executive function ability in bilingual children and may have the same effect in two-way immersion students (e.g., Bialystok & Viswanathan, 2009). To assess the effects of two-way immersion on executive function, in addition to reading and math standardized tests, reliable and valid measures of attention, inhibition, and switching, such as the Attentional Network Task, the Stroop task, and the Simon task, should also be administered. If two-way immersion benefits executive function, it could bring about improvements on several other components of cognitive performance, given that executive function is a domain-general process that is highly involved in a wide range of cognitive activities (Biederman et al., 2004; Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008). For example, as noted earlier, improved executive function may have contributed to the enhancements seen in reading and math, by helping students pay attention to the material taught during reading and math class. Furthermore, increased executive function skills could also lead to enhanced social reasoning, memory encoding and retrieval, and second- and third-language learning, among others (Bartolotti & Marian, 2012; Carlson, Moses, & Breton, 2002; Raj & Bell, 2010). If advantages in the aforementioned cognitive processes are observed on experimental tasks, further studies could be conducted to examine whether these benefits translate into noticeable real-world behaviors. To that end, researchers can administer more ecologically valid experimental tests along with subjective measures of academic and social functioning like peer, teacher, and parent evaluations of the students. Additionally, researchers should consider using longitudinal designs to determine whether these hypothesized effects of two-way immersion extend into later grades.

A second understudied topic to consider in future research is how learning academic concepts in one or both languages affects the ability to transfer knowledge to a nonacademic context. For example, if certain social studies or science concepts are taught to majority-language students only in Spanish, will these students remember them as well when they are out of the classroom and in an English-speaking context? According to studies on language-dependent memory (e.g., Marian & Neisser, 2000), remembering information is more difficult when the language context at retrieval is different from the language context at encoding (Marian & Kaushanskaya, 2007), so there is a possibility that in some cases, transferring knowledge may be less effective in two-way immersion students relative to mainstream classroom students (Marian & Fausey, 2006). On the other hand, since many of the concepts are likely taught in one language and later revisited in the other language, cases of mismatching linguistic contexts may be infrequent. Moreover, in situations where a concept is taught in both English and Spanish, memory for these concepts may be especially strong, as encoding in two different contexts may lead to deeper encoding and more retrieval routes. Thus, two-way immersion education may affect performance in and out of the classroom in many ways, and researchers should aim to understand these effects.

In closing, the results of the current study help to advance the debate surrounding the effectiveness of bilingual education programs. Our results are aligned with previous research (Collier &

Thomas, 2004; Genessee, 1983; Lindholm-Leary, 2005; Lindholm-Leary & Howard, 2008; Thomas & Collier, 2002) and suggest that bilingual two-way immersion programs provide an effective instructional approach for both minority-language and majority-language elementary school students. We found that minority-language and majority-language students enrolled in a TWI program show improved math and reading performance on standardized tests in English. Finally, beyond the direct focus of the current study, it is necessary to underscore another benefit of the TWI program. In addition to the potential improvements in academic performance, TWI students also stand to gain proficiency in both languages of instruction. This ability to communicate in two languages and interact with a larger proportion of the population is likely an asset for these students as they enter an increasingly globalized world. We conclude that two-way immersion models are beneficial in multiple ways and should be seriously considered when designing and implementing educational programs.

ACKNOWLEDGMENTS

The authors thank Dr. Steve Zecker for statistical advice, Dr. Margarita Kaushanskaya and Dr. Daphne Sajous-Brady for help with data entry and organization, and the Northwestern University Bilingualism and Psycholinguistics Research Group for helpful suggestions on the manuscript. We appreciate the generosity of the school district that shared the data to be included in the present analyses. The project was funded in part by Grant NICHD 1R01HD059858 to VM.

REFERENCES

- Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education, 14*(3), 219–234.
- August, D., & Shanahan, T. (Eds.) (2006). *Developing literacy in second-language learners*. Mahwah, NJ: Lawrence Erlbaum.
- Baker, C. (2011). *Foundations of bilingual education and bilingualism* (5th ed.). Clevedon, England: Multilingual Matters.
- Bartolotti, J., & Marian, V. (2012). Language learning and control in monolinguals and bilinguals. *Cognitive Science, 36*(6), 1129–1147.
- Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Developmental Psychology, 24*, 560–567.
- Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *International Journal of Bilingual Education and Bilingualism, 10*(3), 210–223.
- Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science, 7*(3), 325–339.
- Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition, 112*, 494–500.
- Biederman, J., Monuteaux, M. C., Doyle, A. E., Seidman, L. J., Wilens, T. E., Ferrero, F., . . . Faraone, S. V. (2004). Impact of executive function deficits and attention-deficit/hyperactivity disorder (ADHD) on academic outcomes in children. *Journal of Consulting and Clinical Psychology, 72*, 757–766.
- Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. E. (2008). Executive functions and school readiness: Impact, moderation and mediation in the Head Start REDI program. *Development and Psychopathology, 20*, 821–843.
- Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*(2), 647–663.

- Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, 33(3), 205–228.
- Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology*, 19(3), 273–293.
- Caldas, S. J., & Bankston C. III. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research*, 90, 269–277.
- Campbell, R., & Sais, E. (1995). Accelerated metalinguistic (phonological) awareness in bilingual children. *British Journal of Developmental Psychology*, 13, 61–68.
- Cardinal, R. N., Pennicott, D. R., Sugathapala, C. L., Robbins, T. W., & Everitt, B. J. (2001). Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science*, 292(5526), 2499–2501.
- Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, 11(2), 282–298.
- Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development*, 11, 73–92.
- Christian, D., Howard, E. R., & Loeb, M. I. (2000). Bilingualism for all: Two-way immersion education in the United States. *Theory Into Practice*, 39(4), 258–266.
- Collier, V. P., & Thomas, W. P. (2004). The astounding effectiveness of dual language education for all. *NABE Journal of Research and Practice*, 2(1), 1–20.
- Cromdal, J. (1999). Childhood bilingualism and metalinguistic skills: Analysis and control in young Swedish-English bilinguals. *Applied Psycholinguistics*, 20, 1–20.
- Delis, D. C., Massman, P., Butters, N., Salmon, D. P., Cermak, L. S., & Kramer, J. H. (1991). Profiles of demented and amnesic patients on the California Verbal Learning Test: Implications for the assessment of memory disorders. *Journal of Consulting and Clinical Psychology*, 1, 19–26.
- Dreher, M. J., & Zenge, S. D. (1990). Using metalinguistic awareness in first grade to predict reading achievement in third and fifth grades. *Journal of Educational Research*, 84(1), 13–21.
- Gándara, P., & Hopkins, M. (2010). *Forbidden language: English learners and restrictive language policies*. Multicultural Education Series. New York, NY: Teachers College Press.
- Gándara, P., & Orfield, G. (2012). Why Arizona matters: The historical, legal, and political contexts of Arizona's instructional policies and U.S. linguistic hegemony. *Language Policy*, 11(1), 7–19.
- Genesee, F. (1983). Bilingual education of majority-language children: The immersion experiments in review. *Applied Psycholinguistics*, 4(1), 1–46.
- Genesee, F., Lindholm-Leary, K., Saunders, W., & Christian, D. (2005). English language learners in U.S. schools: An overview of research findings. *Journal of Education for Students Placed At Risk*, 10(4), 363–385.
- Goldenberg, C. (2008). Teaching English language learners: What the research does—and does not—say. *American Educator*, 32, 8–23, 42–44.
- Green, S. B., & Salkind, N. J. (2008). *Using SPSS for Windows and Macintosh* (5th edition). Upper Saddle River, NJ: Prentice Hall.
- Greene, J. P. (1998). A meta-analysis of the Rossell and Baker review of bilingual education research. *Bilingual Research Journal*, 21(2–3), 103–122.
- Hakuta, K., Butler, Y. G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* University of California Linguistic Minority Research Institute Policy Report 2000-1. Santa Barbara, CA: University of California-Santa Barbara.
- Irby, B. J., Tong, F., Lara-Alecio, R., Mathes, P. G., Acosta, S., & Guerrero, C. (2010). Quality of instruction, language of instruction, and Spanish-speaking English language learners' performance on a state reading achievement test. *TABE*, 12(1), 1–42.
- Kindler, A. (2002). *Survey of the states' limited English proficient students and available educational programs and services: 2000–2001 summary report*. Washington, DC: National Clearinghouse for English Language Acquisition.
- Lindholm-Leary, K. J. (2005). The rich promise of two-way immersion. *Educational Leadership*, 62(4), 56–59.
- Lindholm-Leary, K. J., & Howard, E. R. (2008). Language development and academic achievement in two-way immersion programs. In T. W. Fortune & D. J. Tedick (Eds.), *Pathways to multilingualism: Evolving perspectives on immersion education* (pp. 177–200). Oxford, England: Blackwell.

- Mackinney, E., & Ríos-Aguilar, C. (2012). Negotiating between restrictive language policies and complex teaching conditions: A case study of Arizona's teachers of English learners. *Bilingual Research Journal*, 35(3), 350–367.
- Marian, V., & Fausey, C. M. (2006). Language-dependent memory in bilingual learning. *Applied Cognitive Psychology*, 20, 1–23.
- Marian, V., & Kaushanskaya, M. (2007). Language context guides memory content. *Psychonomic Bulletin & Review*, 14(5), 925–933.
- Marian, V., & Neisser, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General*, 129(3), 361–368.
- Mazzocco, M. M., & Kover, S. T. (2007). A longitudinal assessment of executive function skills and their association with math performance. *Child Neuropsychology*, 13(1), 18–45.
- McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, 43(4), 947–959.
- Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology*, 80(1), 44–57.
- Raj, V., & Bell, M. A. (2010). Cognitive processes supporting episodic memory formation in childhood: The role of source memory, binding, and executive functioning. *Developmental Review*, 30, 384–402.
- Ríos-Aguilar, C., Canché, M. S. G., & Sabetghadam, S. (2012). Evaluating the impact of restrictive language policies: The Arizona 4-hour English language development block. *Language Policy*, 11(1), 47–80.
- Rolstad, K., Mahoney, K., & Glass, G. V. (2005). The big picture: A meta-analysis of program effectiveness research on English language learners. *Educational Policy*, 19(4), 572–594.
- Rossell, C. H. (1990). The effectiveness of educational alternatives for limited-English-proficient children. In G. Imhoff (Ed.), *Learning in two languages* (pp. 71–121). New Brunswick, NJ: Transaction Publishers.
- Rossell, C. H., & Baker, K. (1996). The educational effectiveness of bilingual education. *Research in the Teaching of English*, 30(1), 7–74.
- Ruiz-de-Velasco, J., & Fix, M. (2000). *Overlooked & underserved: Immigrant students in U.S. secondary schools*. Washington, DC: The Urban Institute.
- Sañudo-Peña, M. C., Patrick, S. L., Patrick, R. L., & Walker, J. M. (1996). Effects of intranigral cannabinoids on rotational behavior in rats: Interactions with the dopaminergic system. *Neuroscience Letters*, 206, 21–24.
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417–453.
- Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction. *Review of Educational Research*, 75(2), 247–284.
- Slavin, R. E., Madden, N. A., Calderón, M., Chamberlain, A., & Hennessy, M. (2011). Reading and language outcomes of a multiyear randomized evaluation of transitional bilingual education. *Education Evaluation and Policy Analysis*, 33(1), 47–58.
- Thomas, W. P., & Collier, V. P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement*. Santa Cruz, CA: University of California at Santa Cruz, Center for Research on Education, Diversity, and Excellence.
- Weinberg, S. L., & Abramowitz, S. K. (2002). *Data analysis for the behavioral sciences using SPSS*. Cambridge, England: Cambridge University Press.
- Willig, A. C. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of Educational Research*, 55(3), 269–317.
- Wiley, T. G., & Wright, W. E. (2004). Against the undertow: Language-minority education policy and politics in the “age of accountability.” *Educational Policy*, 18(1), 142–168.

APPENDIX

1. Sample Third Grade Items from the *State Standards Achievement Test*

A. Reading

Would it be fun to be a fish? They are, after all, quite different from us. Fish have no ears as we do. Their bodies are covered with thin, flat plates called scales. The only sounds they know are what they feel using certain scales along their sides. These are special scales called lateral lines. We get oxygen from the air by using our lungs. Fish get oxygen from the water by using the gills on the sides of their heads. We can play in water and on land, but fish must stay in the water all the time. Fish never get hold or cold. They are called cold-blooded because they are always the same temperature as the water around them. That means they have no need for hot soup, or cold lemonade, or cozy blankets, or cool sandals. All in all, it's probably more fun being us.

This story mainly tells —

- how fish are different from people
- how many kinds of fish there are
- where fish can be found
- how fish swim

Which question does the article answer?

- Do fish have teeth?
- How can I catch a fish?
- Do fish sleep?
- What does cold-blooded mean?

B. Math

John bought 2 notebooks.

Each notebook costs \$1.80.

John gave the clerk \$5.00 to pay for the notebooks.

How much change should John receive?

- \$1.40
- \$2.40
- \$3.20
- \$3.60

2. Sample Third Grade Items from the *State Measure of Annual Growth in English*

A. Reading

Marie and Robert were going to the store. They left Marie's house and walked outside. The streets were not too busy until they got close to the store. Market Street was always busy. It was the last street they needed to cross. Robert and Marie checked for the "Walk" signal. Then they looked both ways for cars. When they were sure it was safe, they crossed Market Street quarter. After they left the store, Robert said, "Ask your dad if you can go to the park tomorrow. I'm going to call Jake when I get home

and see if he can go too.” “Okay,” said Marie. “I think that sounds like a good idea. We usually have a good time when the three of us do things together.”

Where were Robert and Marie going?

- a. To the movies
- b. To the library
- c. To the museum
- d. To the park
- e. To the store

How did Robert and Marie know it was safe to cross the street?

- a. Because a police officer told them
- b. Because the “Don’t Walk” sign was on
- c. Because they looked both ways
- d. Because they were crossing together
- e. Because they looked for the “Walk” sign

3. Math

John bought 2 notebooks.

Each notebook costs \$1.80.

John gave \$5.00 to pay for the notebooks.

How much change should John get?

- a. \$1.40
- b. \$2.40
- c. \$3.20
- d. \$3.60