Evaluation of 2007 Extended Learning Opportunities Summer Adventures in Learning (ELO SAIL) Program: Implementation and Outcomes

Office of Shared Accountability

February 2009

Helen Wang, Ph.D.

Montgomery County Public Schools

OFFICE OF SHARED ACCOUNTABILITY

Dr. Stacy L. Scott, Associate Superintendent
850 Hungerford Drive
Rockville, Maryland 20850
301-279-3925

Dr. Jerry D. Weast
Superintendent of Schools

Dr. Frieda K. Lacey
Deputy Superintendent of Schools

Table of Contents

Executive Summary vi
Program goal and evaluation question. vi
Methodology. vi
Findings. vi
Recommendations for Future Program Implementation. viii

Introduction 1
2007 ELO SAIL. 1
Program structure. 1
Student eligibility and recruitment. 1
Curriculum and instruction. 1
Staffing and training. 2
Literature Review 2
Summer academic loss. 2
Evaluation designs and findings. 2
Evaluation of ELO SAIL. 3
Evaluation Questions. 4

Methodology. 5
Implementation Evaluation. 5
Stakeholder surveys. 5
Instruments. 5
Data collection 5
Analysis 5
Classroom observations 5
Protocols 5
Selection of classes. 5
Inter-rater agreement 6
Data collection 6
Analysis. 6
Outcome Evaluation. 6
Design. 6
Student sample. 7
Outcome measures and data collection. 7
Analysis. 8

Results . . . 9
Implementation Results . . . 9
Results from stakeholder surveys . . . 9
To what extent was ELO SAIL perceived as effective by the administrators with respect to administrative, student-related, and instruction-related activities? . . . 9
To what extent was ELO SAIL implementation perceived as effective by classroom teachers with respect to curriculum, instruction, planning, and resources? . . . 11
What were perceived as ELO SAIL strengths and weaknesses by the administrators and classroom teachers? . . . 13
Results from classroom observations . . . 15
Reading instructional blocks in Grades 1 and 2 . . . 15
Mathematics instructional blocks in Grades 3 and 4 . . . 19
Summary of classroom observations . . . 23
To what extent were the mastery objectives of ELO SAIL lessons communicated to students? . . . 24
To what extent were the components of ELO SAIL curricula applied in the classroom? . . . 24
To what extent was ELO SAIL instruction differentiated to meet student needs? . . . 24
To what extent was a rich learning environment facilitated in ELO SAIL classrooms to support student engagement? . . . 24
Academic outcomes . . . 25
Was ELO SAIL effective in maintaining/improving the text reading and comprehension levels of full-program participants relative to that of nonprogram participants? . . . 25
Was ELO SAIL effective in maintaining/improving the performance of full-program participants in number relations and computation relative to that of nonprogram participants? . . . 25
Did the academic benefit of ELO SAIL vary by student subgroups with respect to race/ethnicity, gender, and receipt of ESOL, FARMS, and special education services? . . . 26
Grade 1 fall 2007 text reading and comprehension levels by student subgroups . . . 27
Grade 2 fall 2007 text reading and comprehension levels by student subgroups . . . 29
Grade 4 fall 2007 mathematics scores by student subgroups . . . 29
Conclusion . . . 33
Findings . . . 33
Administrator/principal surveys . . . 33
Classroom teacher surveys . . . 33
Fidelity of curriculum implementation . . . 33
Academic outcomes . . . 34
Recommendations for Future Program Implementation . . . 35
Class size . . . 35

Curriculum 35
Instruction. 35
Student recruitment 35
Strengths and Limitations of the Evaluation. 36
Strengths. 36
Limitations 36

References 37

Appendixes. 39

Appendix A: List of Title I Schools Included in 2007 ELO SAIL Program Evaluation. 39

Appendix B: Inter-rater Reliability for 2007 ELO SAIL Classroom Observation. 40

Appendix C: Item Analyses for Fall 2007 ELO SAIL Mathematics Test. 41

Appendix D: Administrator and Teacher Respondents' Background. 42

Appendix E: Demographic Characteristics of Students Included in 2007 ELO SAIL Outcome Analyses by Grade. 46

Appendix F: Descriptive Statistics of Scores in Reading and Mathematics Assessments 47

Appendix G: Adjusted Mean Differences and Effect Sizes for Reading and Mathematics Assessments, by Student Subgroups. 48

List of Tables

Table 1	Administrators' Perceptions of 2007 ELO SAIL Implementation	10
Table 2	Teachers' Perceptions of 2007 ELO SAIL Implementation	14
Table 3	General Statistics for Reading Block Observations, by Grade	16
Table 4	Number and Percentage of Volunteers and Helping Staff in the Reading Block, by Grade	16
Table 5	Number and Percentage of Mastery Objectives Presented in the Reading Block, by Grade	16
Table 6	Number and Percentage of Book Levels Used for the Whole-Group Reading Instruction, by Grade	17
Table 7	Number and Percentage of Activities Observed during Whole-Group Instruction in the Reading Block, by Grade	17
Table 8	Number and Percentage of Activities Observed during Student Independent Practice in the Reading Block, by Grade	18
Table 9	Number and Percentage of Activities Observed during Small-Group Instruction in the Reading Block, by Grade	19
Table 10	General Statistics for Mathematics Block Observations, by Grade	20
Table 11	Number and Percentage of Observed Volunteers and Helping Staff in the Mathematics Block, by Grade	20
Table 12	Number and Percentage of Mastery Objectives Presented in the Mathematics Block, by Grade	20
Table 13	Number and Percentage of Observed Lessons Taught in the Mathematics Block, by Grade	21
Table 14	Number and Percentage of Observed Warm-Up and Closure in the Mathematics Block, by Grade	21
Table 15	Number and Percentage of Activities Observed during Focus Lesson in the Mathematics Block, by Grade	22
Table 16	Number and Percentage of Activities Observed during Student Independent Practice in the Mathematics Block, by Grade	23
Table 17	Number and Percentage of Activities Observed during Small-Group Instruction for the Mathematics Block, by Grade	23
Table 18	Adjusted Mean Differences and Effect Sizes for Grades 1 and 2 Fall 2007 Text Reading and Comprehension Levels	25
Table 19	Adjusted Mean Differences and Effect Sizes for Grade 4 Fall 2007 Mathematics Scores	26

List of Figures

Figure 1 Adjusted means for fall 2007 text reading and comprehension levels, by FARMS and ELO SAIL participation 26

Figure 2 Adjusted means for fall 2007 text reading and comprehension levels, by ESOL and ELO SAIL participation. 27

Figure 3 Adjusted means for fall 2007 text reading and comprehension levels, by special education and ELO SAIL participation. 28

Figure 4 Adjusted means for fall 2007 text reading and comprehension levels, by gender and ELO SAIL participation. 28

Figure 5 Adjusted means for fall 2007 text reading and comprehension levels, by race/ethnicity and ELO SAIL participation. 29

Figure 6 Adjusted means for Grade 4 fall 2007 mathematics scores, by FARMS and ELO SAIL participation 30

Figure 7 Adjusted means for Grade 4 fall 2007 mathematics scores, by ESOL and ELO SAIL participation. 30

Figure 8 Adjusted means for Grade 4 fall 2007 mathematics scores, by special education and ELO SAIL participation. 31

Figure 9 Adjusted means for Grade 4 fall 2007 mathematics scores, by gender and ELO SAIL participation. 31

Figure 10 Adjusted means for Grade 4 fall 2007 mathematics scores, by race/ethnicity and ELO SAIL participation. 32

Executive Summary

The Office of Shared Accountability (OSA) conducted an evaluation of the 2007 Extended Learning Opportunities Summer Adventures in Learning (ELO SAIL) program in Montgomery County Public Schools (MCPS). The ELO SAIL is a schoolwide Title I program, designed to provide supplemental academic support to students from low-income communities. More than 5,500 students attended the 2007 program in 22 schools that received Title I funds. This evaluation report summarizes the implementation and academic outcomes of the program.

Program goal and evaluation question

The primary goal of ELO SAIL is to provide opportunities to acquire and preview concepts and skills in reading and mathematics to alleviate summer academic loss and promote continued learning by students (MCPS, 2007). The following broad evaluation questions address whether the 2007 ELO SAIL was implemented as planned and accomplished its academic goal:

1. Was the program perceived to be effective by the administrators with respect to administrative, student-related, and instruction-related activities; and by the teachers with respect to curriculum, instruction, planning, and resources?
2. Were the ELO SAIL reading and mathematics curricula implemented with fidelity in classes, in terms of communication of mastery objectives, application of instructional component, differentiated instruction, and facilitation of rich learning environment?
3. Was the program effective in alleviating summer academic loss and promoting continued learning, as measured by local reading assessment and ELO SAIL mathematics assessment? Did the academic benefit of the program vary by student subgroups?

Methodology

A multimethod study was performed to address the evaluation questions. Administrator/principal and teacher surveys were administrated to assess stakeholders' experiences with the 2007 ELO SAIL and their perceptions of the program effectiveness. To examine the fidelity of ELO SAIL curriculum implementation, classroom observations were conducted for Grades 1 and 2 reading classes and Grades 3 and 4 mathematics classes. Student performance in Grades 1 and 2 reading and Grade 4 mathematics in fall 2007 was compared between full participants (attending 16 or more days) and same-grade nonparticipants (eligible but not enrolled), adjusted for student initial abilities and socio-demographic characteristics. Program-relevant and timely assessments were used for the outcome measures.

Findings

Administrator/principal surveys. The 32 administrators who responded viewed the 2007 ELO SAIL implementation positively in general. In particular, 85% or more of them received support and information for administration, program implementation, and procuring material and supplies. They were able to recruit qualified teachers and paraeducators and receive adequate support from multiple staff members. Most summer administrators had prior program communication with the principal, but there was less agreement about the effectiveness of

assigning responsibilities to summer administrators (60%). Also, 80% or more agreed that the procedures of student arrival and dismissal went well and the afternoon session enhanced the program effectiveness. Though more than 80% agreed that the administration of mathematics assessments worked well, more than 25% believed the administration of the running records during the program did not go smoothly. Further, more than 90% agreed the classroom observations were minimally disruptive to instruction; and between 80% and 100% also agreed that the program met the needs of students in each grade.

Classroom teacher surveys. Teachers also held a positive view of the ELO SAIL implementation, in general. Almost all of the 271 teachers perceived the program as successful in providing review and preview sessions to facilitate continued learning in reading. There was less agreement about the effectiveness of the mathematics curriculum among the teachers (about 75%). With respect to differentiated instruction, teachers responded more favorably to the reading curriculum (81%) than to the mathematics curriculum (60%). Nearly 75% agreed that sufficient time was provided for planning and setup, and 65% agreed that supplies were adequate. Among non-Reading First teachers, almost 80% agreed that the reading material was adequate. More than 90% agreed that the library media center resources were available, and the media centers were more often used for checking out books by teachers and students than for lesson planning.

Fidelity of curriculum implementation. Findings from classroom observations demonstrated that the mastery objectives of the ELO SAIL curricula were appropriately communicated to students through implementation of the instructional components, such as alignment with the mastery objectives of the lesson, connection to student prior knowledge, and modeling of effective reading strategies and mathematics problem solving.

In almost all of the Grade 1 and the Grade 2 reading classes, components of the instructional block, including whole group, small groups, independent practice, and lesson warm-up and reflection, were observed. The small-group instruction in the reading classes was well organized and involved most students. However, there was inconsistent implementation of the instructional block components in the mathematics classes. For instance, small-group instruction was observed in less than 50% of Grade 3 and slightly more than 60% of Grade 4 mathematics classes. Center activities and differentiated class assignments also were observed more frequently in Grades 1 and 2 reading than in Grades 3 and 4 mathematics classes.

Furthermore, teachers in both reading and mathematics classes supported student engagement through teacher modeling, student practices, discussions, various learning activities, and games. Opportunities for multiple problem solutions were observed in about three quarters of Grade 3 and one half of Grade 4 mathematics focus lessons.

Academic outcomes. The hypothesis about the impact of ELO SAIL on text reading and comprehension levels was supported by the results for Grade 1 but not for Grade 2. Specifically, Grade 1 full participants significantly outperformed their nonparticipating counterparts on fall text reading and comprehension levels. However, no significant association emerged between program participation and Grade 2 text reading and comprehension levels. The hypothesis about the impact of ELO SAIL on mathematics was supported by Grade 4 fall test scores; full participants significantly outperformed their nonparticipating counterparts. The between-group difference was large enough to be of practical significance to educators for Grade 4 mathematics but not for Grade 1 reading.

Also as anticipated, the positive impact of ELO SAIL was more likely to occur in certain socio-demographic subgroups. For Grade 1 text reading and comprehension levels, full ELO SAIL participation had a significant benefit for students receiving Free and Reduced-price Meals System (FARMS) and English for Speakers of Other Languages (ESOL) services, whereas it did not make a notable change for non-recipients of these services. As for Grade 4 mathematics, full participation was significantly beneficial for students receiving FARMS services but not for students who were not eligible to receive FARMS services. Grade 4 students benefited from full attendance in mathematics but with a greater benefit for ESOL students than for non-ESOL students.

In addition, ELO SAIL participation was associated with a significant reading benefit for students not receiving special education services in Grade 1; it had a greater mathematics benefit for recipients than for non-recipients of these services in Grade 4. Grade 1 African American and Hispanic students improved their text reading and comprehension levels significantly; the improvement for same-grade White students approached practical significance. As for Grade 4 mathematics, the program benefit was evident only for Asian American and Hispanic students.

In summary, the positive academic impacts associated with 2007 ELO SAIL remained evident in fall. The higher Grade 1 fall text reading and comprehension levels and the higher Grade 4 fall mathematics scores for full participants suggest that the program was effective in prevention of summer academic loss and even resulted in more academic gain in these grades. Findings from the disaggregated data suggest that students who were impacted by poverty and limited English language proficiency benefited more in Grade 1 fall text reading and comprehension levels and Grade 4 fall mathematics scores.

Recommendations for Future Program Implementation

Class size. Investigate why a smaller class was among the most common recommendations for program improvement proposed by the teachers.

Curriculum. Provide sufficient review lessons in reading for all grades, in order to reinforce academic skills already acquired and help students acquire grade-level concepts and skills. Update the mathematics curriculum to ensure that sufficient lessons and resources are provided to challenge students at different levels and accelerate student learning. Investigate and identify factors responsible for the lack of significant program impact on Grade 2 text reading and comprehension levels.

Instruction. Provide teachers with guidance and assistance to support differentiated instruction in Grade 3 and 4 mathematics classes. Provide sufficient instructional material and time for classroom setup and lesson planning. Encourage teachers to increase student-to-student and student-to-teacher discourse, student sharing of multiple problem-solving strategies, and mathematics journals. Encourage the use of summary activities by students, beyond exit cards. Further, provide support for teachers during the administration of mathematics assessments; and provide training to help teachers administer running records efficiently, if the evaluation is repeated in future years.

Student recruitment. Continue and expand the afternoon program, if possible, because it would enhance the effectiveness of ELO SAIL.

Evaluation of 2007 Extended Learning Opportunities Summer Adventures in Learning (ELO SAIL) Program: Implementation and Outcomes

Helen Wang, Ph.D.

Introduction

Extended Learning Opportunities Summer Adventures in Learning (ELO SAIL) is a schoolwide Title I program, designed to provide supplemental academic support to students from low-income communities in Montgomery County Public Schools (MCPS). The program is aligned with scientifically research-based practices, as required by the *No Child Left Behind Act of 2001* (NCLB). The primary goal of ELO SAIL is to alleviate summer academic loss and promote continued learning in students by providing opportunities to acquire and preview basic concepts and skills in reading and mathematics. Since 2002, the program has serviced several thousand elementary students each summer. More than 5,500 students attended the 2007 program at 22 ELO SAIL sites that received Title I funds (see a school list in Appendix A).

2007 ELO SAIL

Program structure

The 2007 ELO SAIL program lasted four weeks, from July 9 through August 3, with a two-week extension of mathematics instruction at the Broad Acres site. The program featured a three-hour instructional day of reading/language arts and mathematics. About one half of the 22 participating schools also provided afternoon arts and recreational activities with external funding. All participating students received free breakfast, lunch, and transportation. The expected student-teacher ratio was 17:1 for kindergarten, 19:1 for Grades 1 and 2, and 25:1 for Grades 3 to 5.

Student eligibility and recruitment

All incoming kindergarten through Grade 5 students in fall 2007 from the 22 Title I schools were eligible for the ELO SAIL summer program. According to the Division of Title I Programs, the data used to rank and identify Title I schools were based on the percentage of students who qualify for Free and Reduced-price Meals System (FARMS) services as of October 31, 2006. Title I schools were responsible for the recruitment of students, with significant support from the Division of Title I Programs. Strategies for recruiting students and maintaining attendance in previous years, including use of mailings, phone calls, incentives, and the availability of afternoon programs, were shared with ELO SAIL administrators and recruitment staff.

Curriculum and instruction

The ELO SAIL lessons were developed based on the MCPS reading and mathematics curricula. The reading curriculum focuses on foundational skills and text reading and comprehension. The mathematics curriculum places an emphasis on number relations and computation, with grade-appropriate content and levels of difficulty. Both reading and mathematics curricula consisted of

a review session of previously learned material and a preview session of academic work for the coming school year.

Staffing and training

During staff recruitment, preference was given to applicants who were highly qualified in early childhood, elementary, English for Speakers of Other Languages (ESOL), or special education within the Title I schools. All ELO SAIL teachers must meet the Maryland State Department of Education (MSDE) "highly qualified" standards in the subject areas. A training session was held for teachers new to the program or the grade level. A planning and preparation day was also scheduled for all ELO SAIL staff prior to the start of the program. In addition, each school was assigned one or two ESOL teachers, one special educator, and one library media specialist. Assistant principals were assigned to help with the administration of the summer program at each Title I school.

Literature Review

Summer academic loss

Research has shown that students from low-income communities suffer greater academic loss during the summer vacation than their counterparts from more wealthy communities (Cooper, Nye, Charlton, Lindsay, and Greathouse, 1996; Cooper, 2001; and Entwisle and Alexander 1992). For example, Cooper et al. (1996) reviewed 39 studies about summer loss and combined 13 of them for a meta-analysis. Their meta-analysis indicated that the summer loss equated to about one month on a grade-level equivalent scale. While the detrimental effect of summer break was more observable for mathematics than for reading, middle-class students appeared to gain in grade-level equivalent reading-recognition tests over summer while students from low-income families showed losses on them.

Evaluation designs and findings

To reduce the summer learning loss among low-income students, various public funds, including federal Title I funds, have been available for summer educational programs. An emerging body of evaluation research has provided evidence of summer programs as vehicles to remediate, reinforce, and accelerate learning, according to Cooper, Charlton, Valentine, and Muhlenbruck's synthesis of evaluation studies of summer learning programs (2000). The meta-analyses for 54 summer program evaluations included in the Cooper et al. study yielded small mean effect sizes in favor of program attendees, indicating that the benefits of the summer programs were large enough to be practically significant.

Studies included in this literature review mostly focused on short summer learning programs (3–6 weeks) targeting low-income and low-achieving students in public school systems. The main objective of these programs was to remediate and improve poor reading and mathematics skills among struggling students. Evaluation findings showed only sporadic effectiveness of the summer programs in certain subject matters (reading or mathematics) of certain grade levels. This is consistent with the small average effect size for remedial summer programs reported by Cooper et al. (2000) in their meta-analyses.

The sporadic effects might be understandable given the short length of the summer programs. Shortcomings associated with evaluation designs and measurements also may contribute to low, or lack of, effects found for the summer programs. Some evaluation studies (Joseph, 1999; Joseph, 2001; Portz, 2004; and Roberts and Nowakowski, 2004) simply looked at the pre- to posttest changes in the academic performance of participating students. These studies did not include a comparison group and significance tests.

Several evaluators, including Green (1998), Holdzkom (2002), Paris, Pearson, Cervetti, and Carpenter (2004), and Roderick, Jacob, and Bryk (2004) applied, more scientifically rigorous designs such as a nonrandomized comparison group pre- and posttest quasi-experimental design to assess the academic benefit of summer learning programs. These researchers incorporated a comparison group to assess the relative gain associated with the summer programs. For example, Roderick et al. (2004) found that scores increased by 4 to 6.5 months for Grade 6 and 8 students in reading and mathematics; and by 2 months in reading and more in mathematics for Grade 3 students. In the Paris et al. study (2004), the pre- to posttest improvement in reading for Grade K–3 students was modestly significant and the treatment group did not fall further behind in the follow-up test. However, the comparability of the treatment group and the comparison group was unknown for these studies. Further, these studies mostly adopted local or commercial standardized assessments or the local retention/promotion criteria for outcome measures. Due to the fact that the alignment of these assessments with the summer curricula was not clearly addressed by the evaluators, it is not readily apparent that an academic gain on these assessments was obtained from the summer programs.

A few recent summer programs (e.g., Borman, Overman, Fairchild, Boulay, and Kaplan, 2004; Borman and Dowling, 2006; and Chaplin and Capizzano, 2006) have used an experimental design by randomly assigning eligible students or program applicants to the treatment group (receiving the program) and the control group (not receiving the program). For example, Borman et al. (2004) evaluated a multiyear summer program in Baltimore using random assignment. Although their study showed a very slight and insignificant immediate treatment effect on the reading vocabulary and comprehension of kindergarteners, it indicated a potential cumulative program effect after the third summer program. Chaplin and Capizzano (2006) conducted an experimental study of a commercial summer learning program in the New York and Boston areas provided to low-income children in Grades 1 to 6. The result showed that the treatment group gained one month's worth of reading skills more than the control group.

Evaluation of ELO SAIL

Early evaluation reports of ELO SAIL in MCPS showed significant academic benefits on the MCPS Assessment Program for Primary Reading (MCPS AP-PR) and ELO SAIL mathematics assessment, after controlling for student socio-demographic characteristics. For example, first and second grade full participants in the 2002 program demonstrated a larger improvement in both reading and mathematics than nonparticipants, and the academic benefits were similar across all racial/ethnic groups (Sunmonu, Larson, Van Horn, Cooper-Martin, and Nielsen, 2002). An improvement in mathematics among Grade 4 full participants was observed for the 2003 program, and recipients of ESOL services, African American, Hispanic, and Asian American students in certain grades also showed academic benefits (Sunmonu, Curry-Corcoran, and Mordica, 2004).

Moreover, 10 Title I schools have exited school improvement or corrective action over the time period since ELO SAIL was implemented. There is no Title I school at any level of school improvement based on NCLB.

Evaluation Questions

The current study evaluated three areas related to ELO SAIL implementation and its impact on student academic performance in the 22 elementary schools receiving Title I funds. First, the evaluation addressed whether the 2007 ELO SAIL program was implemented effectively through the following questions:

1. To what extent was ELO SAIL perceived as effective by the administrators with respect to administrative, student-related, and instruction-related activities?
2. To what extent was ELO SAIL implementation perceived as effective by classroom teachers with respect to curriculum, instruction, planning, and resources?
3. What were perceived as ELO SAIL strengths and weaknesses by the administrators and classroom teachers?

Second, the evaluation assessed whether the essential components of the ELO SAIL reading and mathematics curricula were implemented with fidelity, guided by the following questions:

1. To what extent were the mastery objectives of ELO SAIL lessons communicated to students?
2. To what extent were the components of ELO SAIL curricula applied in the classroom?
3. To what extent was ELO SAIL instruction differentiated to meet student needs?
4. To what extent was a rich learning environment facilitated in ELO SAIL classrooms to support student engagement?

Finally, this evaluation assessed whether the program alleviated summer academic loss and promoted continued learning in students of selected grades. It was anticipated that full participants, with less summer loss or more academic gain due to ELO SAIL, would receive a higher average score in reading and mathematics than same-grade nonparticipants in fall 2007. The following questions were investigated:

1. Was ELO SAIL effective in maintaining/improving the text reading and comprehension levels of full program participants relative to that of nonprogram participants?
2. Was ELO SAIL effective in maintaining/improving the performance of full program participants in number relations and computation relative to that of nonprogram participants?
3. Did the academic benefit of ELO SAIL vary by student subgroup with respect to race/ethnicity, gender, and receipt of ESOL, FARMS, and special education services?

Methodology

Implementation Evaluation

Stakeholder surveys

Instruments. To address the effectiveness of the 2007 ELO SAIL implementation, the administrator/principal and teacher online surveys were administrated. With Likert-scale, multiple-response, and open-ended statements or questions, these surveys assessed stakeholders' experiences with the different components of the program and their perceptions of the effectiveness of the program.

Data collection. The survey Web links along with a password and instructions for completing the surveys were delivered via e-mail to all ELO SAIL administrators, principals, and classroom teachers during the last week of the program. Follow-up e-mails were sent as reminders to complete the surveys. A total of 32 ELO SAIL administrator/principal surveys, 224 non-Reading First classroom teacher surveys and 47 Reading First classroom teacher surveys were submitted. This yielded an overall response rate of 76% for the administrator/principal survey, 97% for the non-Reading First teacher survey, and 100% for the Reading First teacher survey. Tables in Appendix D describe respondents' background information. Among the 32 respondents who returned the administrator survey, there were 26 ELO SAIL administrators, four principals, and two in other positions. The average number of years of teaching experience was 9.6 (*SD*=8.7) and 7.4 (*SD*=6.6) among the 224 non-Reading First and 47 Reading First teachers, respectively.

Analysis. Percentages of respondents favoring the survey statements and questions were calculated based on the total number of valid responses, including N/A for each of the survey items. Open-ended responses were coded into categories addressing the same phenomena.

Classroom observations

Protocols. To evaluate fidelity of curriculum implementation, observations were conducted to gather information on the variety of teaching and learning activities in the ELO SAIL classrooms of selected grades. The reading and mathematics classroom observation protocols were developed by the Office of Shared Accountability (OSA) in collaboration with the Division of Title I Programs and the Office of Curriculum and Instructional Programs. Instructional specialists involved in ELO SAIL curriculum development and teacher training contributed to the measures included in the instruments as a way to strengthen the content validity of the instruments.

Selection of classes. Given that Grades 1, 2, and 4 were chosen for reading and mathematics outcome analyses, respectively, classroom observations were performed in these grades and extended to Grade 3 mathematics classes. One classroom per selected grade per school was randomly chosen for observation. Of the 22 participating Title I schools (see Appendix A), four were primary (K–2) only and one was upper-grade only (Grades 3–5) schools, in which there were no mathematics or reading classes of selected grades available for observation. Also, reading classes in four Title I Reading First schools were not observed because they used a different curriculum.

Inter-rater agreement. The observer team consisted of OSA staff and instructional specialists from the Division of Title I Programs who received training sessions in using the observation instruments. In order to calculate the inter-rater agreement on the instructional activities, nine reading and eight mathematics lessons were rated by two observers independently. The inter-rater agreement was assessed using weighted kappa with a higher coefficient indicating higher agreement (see kappa coefficients in Appendix B). The mean (0.67; *SD*=0.12) and median (0.65) of all kappa coefficients for paired reading classroom observations fall in the range of substantial inter-rater agreement. For paired mathematics classroom observations, the mean (0.57; *SD*=0.27) and median (0.61) of all kappa coefficients are at or close to the upper limit of the moderate inter-rater agreement. Two mathematics observations with a kappa coefficient below the lower boundary of the moderate range were not included in current analysis.

Data collection. Classroom observations took place during the second and third week of the program. Each observation covered a two-hour reading or one-hour mathematics instructional block. Teachers were informed of the purpose and selection process of the observation through e-mail with a pre-observation survey attached. The instructional activities were recorded through both structured categories and narrative comments. The observers would put a check mark in the structured items when activities took place and add comments in the next column as needed.

Analysis. The final analysis involved data from 16 Grade 1 reading, 16 Grade 2 reading, 15 Grade 3 mathematics, and 16 Grade 4 mathematics classroom observations. The observation data were summarized with descriptive statistics and content analysis.

Outcome Evaluation

Design

A nonrandomized comparison group pre- and posttest quasi-experimental design was employed to estimate the effect of ELO SAIL on academic performance in the current study. For the purpose of analysis, students were grouped by the number of days of ELO SAIL participation and enrollment status as follows: 1) "full participants" includes students who attended the summer program for 16 or more days, 2) "partial participants" includes students with 6 to 15 days of attendance; 3) "least participants" includes students with five or fewer days of attendance, and 4) "nonparticipants" includes eligible students who were not enrolled in the summer program.

Comparisons were made between each pair of these groups in preliminary analyses. There were no significant results for comparisons involving partial participants and the number of students in the least-participation group was insufficient for a comparison. Therefore, only comparisons between full participants and nonparticipants were reported. The results from comparisons were also disaggregated by student socio-demographic subgroups. Mathematics data collected from Grade 4 students at the Broad Acres site were analyzed and reported separately because there was a two-week extension of mathematics instruction at the site. The following diagram illustrates the nonrandomized comparison group pre- and posttest quasi-experimental design.

Group	Pretest (spring 2007)	Summer Program	Posttest (fall 2007)
Treatment group	O1	Full participation	O2
Comparison group	O1	Nonparticipation	O2

Student sample

Although the 2007 ELO SAIL program served K−5 students in Title I schools, only certain grades were involved in the outcome analysis. Grade 1 and 2 students were assessed on the mClass Reading 3D running records, because it was only administered in the primary grades. Grade 4 students were given ELO SAIL mathematics assessments, given that they were able to read the test by themselves.

The samples for the reading outcome analysis then contained Grade 1 and 2 students who had valid scores on spring and fall 2007 text reading and comprehension levels, as measured with the mClass Reading 3D. The sample for the mathematics outcome analysis contained Grade 4 students who had valid scores on spring and fall 2007 ELO SAIL mathematics tests. These samples were also broken down into student socio-demographic subgroups for the disaggregated analysis. To allow for reliable results, Native American students were not included due to the insufficient number of students.

Outcome measures and data collection

The mClass Reading 3D running records data obtained in spring and fall 2007 from Grade 1 and 2 students were used as the reading outcome measure. This assessment is designed to assess progress in text reading and comprehension among students of primary grades. This assessment is consistent with the content and purpose of the ELO SAIL reading curriculum. Teachers would ask the student to read aloud a book at an appropriate level of difficulty until he/she reached 90% or higher accuracy of word recognition and demonstrated comprehension orally, in writing, or both. Then the reading level of the student was determined.

The ELO SAIL mathematics assessments were used to gather data on students' academic retention and mastery of basic skills through the summer. The assessments, aligned with the content and purpose of the summer mathematics curriculum, were developed by MCPS mathematics content specialists. All Grade 4 students eligible for ELO SAIL were administrated the assessment at the end of the spring term and the beginning of the fall term in 2007.

There were 28 items covering number relations and computation in the mathematics assessments. Students would circle a choice from multiple responses and compute, estimate, or use their learned knowledge to explain their answers. For each question, the student received a score of 1 for a correct answer and a score of 0 for an incorrect or missing answer. The assessments were aligned with the review and preview sessions in the summer curriculum and included review and preview subtests. The review subtest covered mathematics concepts from the previous grade, while the preview subtest covered mathematics concepts that would be taught in fall. The ELO SAIL academic outcome in mathematics was assessed based on the combined score of review and preview subtests.

Item analyses were performed to examine the internal consistency of the 28 items in the spring 2007 mathematics test. The results, presented in Appendix C, show that the test has acceptable difficulty values, ranging from 0.28 to 0.93 for the review subtest and 0.05 to 0.62 for the preview subtest. This indicates that items in the review subtest are less difficult than those in the preview subtest. The discrimination index shows that the association of students' total performance with their performance on each item was reasonably strong for most of the items. The test has a reliability coefficient of 0.86 (Cronbach's Alpha), which indicates the high internal consistency of the test. In general, the results of the item analyses suggest the mathematics test is a reliable assessment. The fall and spring tests are identical.

Analysis

This evaluation used the analysis of covariance procedure (ANCOVA) in SPSS General Linear Model (GLM) to examine whether the full participants, on average, scored higher than the nonparticipants in the outcome measures. Text reading and comprehension levels and mathematics scores in fall 2007 were outcome measures, serving as the dependent variable. Program participation at two levels (full participation and nonparticipation) was the primary independent variable. Text reading and comprehension levels and mathematics scores in spring 2007 were included as measures of student initial abilities.

The ANCOVA adjusts for the confounding effects related to the nonequivalence of the comparison group through controlling students' initial abilities as well as student socio-demographics, such as gender; race/ethnicity; and participation in FARMS, ESOL, and special education programs. Propensity scores were also included to further improve the similarity of the two comparison groups (Dehejia and Wahba, 2002). The scores were computed from a set of specific predictors (student initial abilities and socio-demographic characteristics) of being either a full participant or nonparticipant through the logistic regression procedure. The scores were then divided into five categories and used as a categorical covariate in the analysis.

The analyses reveals average spring-to-fall changes for the full-participation and same-grade nonparticipation groups, and also estimates the adjusted mean difference between the two groups. In preliminary analyses, the interaction effect was also examined to determine whether the academic gain from ELO SAIL varied by student initial ability. The final analyses did not include interaction terms because they were not significant.

In order to avoid acceptance of a significant but trivial program effect, the effect size was computed to assess how meaningful the adjusted group mean difference is. Within the education research field, an effect size of 0.2 is considered small but acceptable; at least 0.5 is considered medium; and 0.8 or greater is considered large (Cohen, 1988; Datnow, Borman, Stringfield, Overman, and Castellano, 2003).

Results

Implementation Results

Results from stakeholder surveys

To what extent was ELO SAIL perceived as effective by the administrators with respect to administrative, student-related, and instruction-related activities?

The administrators perceived the implementation of the 2007 ELO SAIL program to be effective, in general.

Administrative activities. A majority of the administrators responded favorably to survey items regarding administrative activities (Table 1). About 90% were satisfied with the support they received in addressing administrative concerns. Nearly 85% agreed they were provided with information for program implementation and support in procuring materials and supplies. Among those who were involved in the process, 23 out of 24 reported they were able to recruit highly qualified and skilled teachers and paraeducators. Nearly 70% of the summer administrators had an opportunity to meet with the school principal for early planning. More than 60% responded favorably to the statement that assigning daily ELO SAIL administrative responsibilities to summer administrators was a smooth and efficient process.

Furthermore, the respondents were asked to identify all positions listed under the multiple-response item about staff support. A high percentage of the respondents received adequate support from multiple staff members, as follows:

- Paraeducators (100%)
- Secretarial and clerical staff (97%)
- ESOL teachers (91%)
- Media specialists (91%)
- Food service staff (88%)
- Special educators (78%)

Student-related activities. Among the administrators, 80% or more felt that the procedures of student arrival and dismissal went well (Table 1). Among those who had the afternoon component available at their ELO SAIL sites, 18 out of 22 (82%) agreed that the afternoon component enhanced program effectiveness in terms of student recruitment and outcome. With the multiple-response items, almost all the respondents cited one or more strategies used to recruit students and strategies to support daily attendance. Strategies for student recruitment reported by at least two thirds of respondents were as follows:

- Registration forms were distributed through backpacks (100%)
- Reminder notices were sent home frequently (93%)
- Follow-up registration forms were sent home (90%)
- Phone calls were made to families (77%)
- Information was distributed during kindergarten orientation (67%)

Table 1
Administrators' Perceptions of 2007 ELO SAIL Implementation (*N*=32)

Program implementation	Strongly agree		Agree		Disagree/ strongly disagree		Not applicable		Total
	n	%	*n*	%	*n*	%	*n*	%	*N*[a]
1. I was satisfied with the support I received in addressing administrative concerns.	20	62.5	9	28.1	3	9.4	0	0.0	32
2. I was provided with ELO SAIL information needed to implement the program in my school.	12	37.5	15	46.9	4	12.5	1	3.1	32
3. I was satisfied with the support I received in procuring materials and supplies.	17	53.1	9	28.1	3	9.4	3	9.4	32
4. I was able to recruit highly qualified and skilled teachers and paraeducators.	19	59.4	4	12.5	1	3.1	8	25.0	32
5. Administrators had an opportunity to meet with the school principal to plan for program implementation prior to the program start.	11	34.4	11	34.4	6	18.8	4	12.5	32
6. Assigning daily ELO SAIL administrative responsibilities to summer administrators was a smooth and efficient process.	13	40.6	7	21.9	9	28.1	3	9.4	32
7. The procedure established for student arrival at my ELO SAIL site worked well.	18	60.0	6	20.0	6	20.0	0	0.0	30
8. The procedure established for student dismissal at my ELO SAIL site worked well.	17	54.8	9	29.0	5	16.1	0	0.0	31
9. The afternoon component enhanced program effectiveness (in terms of student recruitment and outcome).	9	28.1	9	28.1	4	12.5	10	31.3	32
10. The administration of ELO SAIL mathematics assessments to Grade 4 at my site worked well.	8	25.0	12	37.5	4	12.5	8	25.0	32
11. The administration of running records to Grade 1 at my site worked well.	7	21.9	14	43.8	8	25.0	3	9.4	32
12. Classroom observations caused minimum interruptions to observed classes.	15	46.9	14	43.8	3	9.4	0	0.0	32

[a] Missing data are excluded.

Moreover, strategies for attendance reported by at least two thirds of respondents were as follows:

- Phone calls were made to families of absent students (100%)
- School staff encouraged students to attend the program regularly (94%)
- Parents were greeted at arrival time (91%)
- Weekly incentives were used (75%)
- Perfect attendance was recognized and announced (72%)

Instruction-related activities: More than 90% of the respondents agreed that the classroom observations caused minimum interruptions to the classes observed (Table 1). Among respondents for whom the items were applicable, 20 out of 24 (83%) agreed that the administration of ELO SAIL mathematics assessments to Grade 4 students in spring and fall worked well, but 8 out of 29 (28%) did not feel so about the on-site administration of running records to Grade 1 students. Other instruction-related activities surveyed include whether student learning needs were met and what student assessment data were available for instructional planning. With the multiple-response item, all 32 respondents indicated one or more grades in which they believed the learning needs of students were met. The percentages below were computed based on the total number of respondents, including those with no upper grades in their schools:

- Kindergarten students' needs were met (100%)
- Grade 1 students' needs were met (100%)
- Grade 2 students' needs were met (97%)
- Grade 3 students' needs were met (81%)
- Grade 4 students' needs were met (78%)
- Grade 5 students' needs were met (78%)

In addition, the administrators were asked to check as many types of student assessment data as were available. Of the 32 administrators, 28 reported at least one type, as follows:

- DIBELS-MClass Reading 3D (71%)
- Articulation cards (reading and mathematics) (71%)
- Reading fluency (54%)
- MAP-R (46%)
- MCPS AP (43%)
- Pre-K reading assessment (43%)
- Pre-K mathematics assessment (39%)
- Other assessments (18%)

To what extent was ELO SAIL implementation perceived as effective by classroom teachers with respect to curriculum, instruction, planning, and resources?

Despite the response variation across the survey items, teachers held a positive view of the program implementation, in general. Data obtained from the non-Reading First and Reading First teacher surveys are reported in an integrated way.

Curriculum and instruction. A great majority of teachers believed that the program was successful in helping students preview the reading material for the coming school year (95%) and review the reading material already taught (89%) (Table 2). More than 80% agreed that MCPS Quarter 1 Reading/Language Arts Guide or the ELO SAIL Reading First notebook provided differentiated instructional strategies to meet student learning needs in reading. There was less agreement about the effectiveness of the mathematics curriculum. About three quarters agreed that ELO SAIL helped students review (75%) and preview (74%) the mathematics material. There was also less agreement (60%) that ELO SAIL mathematics lessons provided differentiated instructional strategies to meet student learning needs.

Some teachers also provided open-ended comments on the ELO SAIL curriculum. These comments were not converted into quantitative data but were summarized by content. Non-Reading First teachers were concerned about a lack of review lessons available in reading or that the review session was not applicable for Grade 1. Reading First teachers felt the Reading First curriculum was either too easy or too difficult, depending on the grade, and reflected little of what was taught in class. A common concern about mathematics was the lack of challenging material and a smooth flow among the lessons or between the review and preview sessions.

Planning and resources. Nearly three quarters of teachers felt sufficient time was provided for lesson planning and classroom setup (Table 2). About 65% agreed that adequate supplies (e.g., paper, markers, pens, erasers) were available. In addition, almost 80% of the non-Reading First teachers agreed that there was enough reading material to meet their students' needs. The availability and full utilization of the library media center resources on the student day was rated most favorably (93%) among items addressing planning and resources. With the multiple-response item regarding how the library media center resources were used, 253 out of 271 teachers provided at least one valid answer. Teachers and students were more likely to borrow instructional and reading materials than use the center for lesson planning and literature appreciation, respectively, as presented below:

- Teachers borrowed instructional materials (81%).
- Teachers planned lessons with media specialist integrating information literacy skills with reading content standards (32%).
- Teachers planned lessons with media specialist integrating information literacy skills with mathematics content standards (5%).
- Students used the center to check out books for free/voluntary reading (88%).
- Students used the center for literature appreciation (63%).
- Other usage of the library media center (7%).

Out of 271 teachers, 235 reported using at least one type of student assessment data for instructional planning. The percentages of teachers using each assessment data are listed below:

- Articulation cards (reading and mathematics) (49%)
- DIBELS-MClass Reading 3D (46%)
- Reading fluency (21%)
- MAP-R (12%)
- MCPS AP (11%)
- Pre-K reading assessment (9%)

- Pre-K mathematics assessment (5%)
- Assessment data (unspecified) (10%)
- Specialist/ESOL/special education teachers (9%)
- Observation (7%)
- Other assessments (20%)

For each of the first eight assessments listed above, the percentage of teachers that reported using it was lower than the percentage of administrators who reported that it was available. This might be partially explained by the fact that some of these assessments were not applicable for some teachers.

What were perceived as ELO SAIL strengths and weaknesses by the administrators and classroom teachers?

Through open-ended responses, 78% of the administrators and 82% of the teachers commented on effective practices of 2007 ELO SAIL; also 75% and 65% of them, respectively, recommended areas of improvement for future programs.

Administrators. Of the 25 administrators who provided at least one positive comment on 2007 ELO SAIL, 14 stated that the program provided students with an opportunity for academic continuity. The same number of administrators stated that school staff/volunteers were qualified, skillful, and effective. Ten administrators cited structured routine and good coordination of the program. Other positive comments included valuable support provided by the administrative team (5) and effective incentive programs for attendance (4).

A total of 24 administrators offered at least one recommendation, including the following:

- Improving accurate routes and safety of transportation (9)
- Establishing a better routine and facilitating program logistics (8)
- Facilitating administrator transitions (5)
- Using effective strategies to improve attendance (5)
- Providing adequate staff especially during assessment periods (4)
- Increasing the afternoon program (3)

Table 2
Teachers' Perceptions of 2007 ELO SAIL Implementation (*N*=271)

	Program implementation	Strongly agree		Agree		Disagree/ strongly disagree		Not applicable		Total
		n	%	*n*	%	*n*	%	*n*	%	*N*[a]
Curriculum and instruction	1. ELO SAIL was successful in helping students preview the reading material for the coming school year.	69	25.8	185	69.3	6	2.2	7	2.6	267
	2. ELO SAIL was successful in helping students review previous academic work in reading.	59	21.9	181	67.3	16	5.9	13	4.8	269
	3. MCPS Quarter 1 Reading/Language Arts Guide provided differentiated instructional strategies; ELO SAIL Reading First notebook supported student learning needs.	48	17.9	168	62.7	26	9.7	26	9.7	268
	4. ELO SAIL was successful in helping students review previous academic work in mathematics.	47	17.5	154	57.5	42	15.7	25	9.3	268
	5. ELO SAIL was successful in helping students preview the mathematics material for the coming school year.	52	19.4	146	54.5	46	17.2	24	9.0	268
	6. ELO SAIL mathematics lessons provided differentiated instructional strategies to meet student learning needs.	32	11.9	129	48.1	79	29.5	28	10.4	268
Instruction planning and resources	7. Library media center resources were available and fully utilized on the students' day.	139	51.7	112	41.6	11	4.1	7	2.6	269
	8. Sufficient time was provided for setting up my classroom.	55	20.5	143	53.4	64	24.9	6	2.2	568
	9. Sufficient time was provided for lesson planning.	54	20.1	143	53.2	68	25.3	4	1.5	269
	10. Adequate supplies (e.g., paper, markers, pens, erasers) were available.	58	21.5	117	43.3	90	33.3	5	1.9	270
	11. There was enough reading material to meet the needs of my students.[b]	71	31.8	106	47.5	38	17.0	8	3.6	223

[a]Missing data are excluded
[b]This item is only included in the non-Reading First Teacher survey

Classroom teachers. Of the 222 teachers who indicated at least one positive aspect about 2007 ELO SAIL, 123 believed student learning continued through review and preview of the academic work. Teachers also reported that the program was well structured and logistically balanced (45), and that administration and staff provided valuable support and created a positive educational atmosphere (43). Other positive aspects commented on by the teachers included available instructional materials (27) and motivated students (24). Finally, several Reading First teachers indicated an effective ELO SAIL curriculum (7) and collaborative team work (3).

With respect to program improvement, 37 out of 176 teachers who offered at least one recommendation requested more time for lesson planning and classroom setup. Some teachers recommended a smaller class size (33) and adequate material supply (26). Some non-Reading First teachers highlighted concerns for improvement in program structure and routine (28) and program management and logistics (25). Reiterating their concerns about the curriculum, 30 teachers suggested the need for more variety and differentiation in the mathematics curriculum, 28 non-Reading First teachers asked for more variety in the reading curriculum and resources, and 18 non-Reading First teachers wanted improvements in ELO SAIL curriculum and instruction but did not specify subject. Finally, five Reading First teachers stated that the reading block and literacy centers needed to accommodate a variety of student abilities and needs.

Results from classroom observations

Reading instructional blocks in Grades 1 and 2

General information. Statistics regarding number of students in the class, length of instructional activities, and structure of the class are summarized in Table 3. Sixteen Grade 1 reading classes included in this analysis lasted an average of 112 minutes and contained an average of 15 students. Sixteen Grade 2 reading classes lasted an average of 116 minutes and contained an average of 16 students. The average time spent on the whole-group instruction, independent practice, and small-group instruction was more evenly distributed in Grade 1 (41, 51, and 51 minutes) than in Grade 2 (33, 69, and 64 minutes). Relative to Grade 1, the Grade 2 reading block spent more time on student independent practice and the small-group instruction than on the whole-group instruction. Small groups were observed in all classes, except for one Grade 1 class, with an average of 3 and 4 small groups in Grades 1 and 2, respectively. Sixty-five percent of Grade 1 and 83% of Grade 2 students were involved in small-group learning. Volunteers, helping staff, or both were present in every classroom (Table 4).

Mastery objectives and curriculum resources. In 13 out of 16 Grade 1 and 11 out of 16 Grade 2 classrooms, mastery objectives were presented both on the board and orally (Table 5). Some observers noted that teachers incorporated the mastery objectives in the morning message or asked students to read them aloud. Some also noted that Grade 2 teachers discussed with students the importance of mastering the objectives for a good reader. Books for whole-group instruction as recorded by the observers consisted of reading levels that reflected both the review and preview purposes of the ELO SAIL program (Table 6). But it is worth noting that more than one half of the observations for Grade 1 reading classes did not record the book level.

Table 3
General Statistics for Reading Block Observations, by Grade

	Grade 1 (*N*=16)				Grade 2 (*N*=16)			
General statistics	Min	Max	Mean	*SD*	Min	Max	Mean	*SD*
Minutes of observation	75	135	112	16.5	95	128	116	8.3
Minutes for whole-group instruction	23	60	41	13.6	20	67	33	12.1
Minutes for independent practices	15	81	51	18.9	42	98	69	17.5
Minutes for small-group instruction	0	97	51	24.4	30	92	64	17.6
Number of small groups in classroom	0	4	3	1.2	2	6	4	1.0
Number of students in classroom	10	18	15	1.9	11	22	16	3.0
Percentage of students in small groups	0	100	65	31	56	100	83	14

Note. *SD* = Standard Deviation

Table 4
Number and Percentage of Volunteers and Helping Staff in the Reading Block, by Grade

	Grade 1 (*N*=16)		Grade 2 (*N*=16)	
Volunteers and helping staff	*n*	*%*	*n*	*%*
Volunteers and helping staff	3	19	3	19
Volunteers only	12	75	12	75
Helping staff only	1	6	0	0
No volunteer or helping staff	0	0	1	6

Table 5
Number and Percentage of Mastery Objectives Presented in the Reading Block, by Grade

	Grade 1 (*N*=16)		Grade 2 (*N*=16)	
Mastery objectives	*n*	*%*	*n*	*%*
Presented on board and orally	13	81	11	69
Presented on board or orally	3	19	5	31

Whole-group instruction. Almost all Grade 1 and 2 whole-group lessons included read aloud, modeled reading, and/or shared reading strategies (Table 7), along with picture walk, think aloud, prediction, and other reading strategies, according to observers. This was consistent with the requirement of the ELO SAIL reading curriculum. As part of the whole-group instruction, library media center activity was also observed in a few classes. All whole-group lessons were aligned with the mastery objectives. They connected to student prior learning or knowledge,

promoted interactive discussion regarding the objectives, and checked student understanding through various strategies. Word work was observed in 12 Grade 1 and 9 Grade 2 classrooms. Developmentally appropriate writing activities took place in 11 and 7 classes of Grade 1 and Grade 2, respectively.

Table 6
Number and Percentage of Book Levels Used for the Whole-Group Reading Instruction, by Grade

	Grade 1 (*N*=16)		Grade 2 (*N*=16)	
Book level	*n*	%	*n*	%
G	0	0	1	6
H	4	25	3	19
I	2	13	4	25
J	0	0	2	13
K	1	6	1	6
L	0	0	1	6
M	0	0	1	6
No book used	0	0	1	6
Not recorded	9	56	2	13

Table 7
Number and Percentage of Activities Observed During Whole-Group Instruction in the Reading Block, by Grade

	Grade 1 (*N*= 16)		Grade 2 (*N*=16)	
Whole-group instruction	*n*	%	*n*	%
Read aloud, modeled reading, shared reading, or combination	16	100	15	94
Alignment with lesson objectives	16	100	16	100
Connection to prior learning or knowledge	15	94	16	100
Interactive discussion regarding the lesson objectives	16	100	16	100
Word work	12	75	9	56
Developmentally appropriate writing activities	11	69	7	44
Strategies for checking understanding	15	94	16	100
Media center activity[a]	4	25	3	19
Student sharing of learning in reflection time	5	31	6	38

[a]Media center activity would not occur if it was not scheduled for the lesson.

Independent practice. In almost all Grade 1 and 2 classrooms, independent practice integrated work assignments (Table 8). In almost all classrooms, students participated in various literacy center activities and had books of different levels available to them. Developmentally appropriate writing activities were observed in 14 out of 16 classrooms for both grades. In 12 Grade 1 and 10 Grade 2 classrooms, the teacher provided specific feedback on student work.

Table 8
Number and Percentage of Activities Observed during Student Independent Practice in the Reading Block, by Grade

	Grade 1 (*N*= 16)		Grade 2 (*N*=16)	
Independent practice	*n*	*%*	*n*	*%*
Work on assignments	15	94	16	100
Books available to students	15	94	16	100
Various literacy activities	15	94	15	94
Developmentally appropriate writing activities	14	88	14	88
Specific feedback on student work	12	75	10	63

Small-group instruction. Small-group instruction in both grades was used to reinforce the whole-group instruction and build foundational skills, according to the observers. Almost all the teachers provided books matched to student reading levels in the small groups (Table 9). Almost all the teachers introduced the book and purpose for reading the book, allowed students to read through the book independently, reinforced the reading strategies taught through prompting or probing, and provided word work opportunities. Small-group activities were connected to the whole-group lesson in 11 of 15 Grade 1 and 15 of 16 Grade 2 classrooms. Small groups in more than 50% to 87% of Grade 1 and 2 classrooms included instructional activities such as student written or oral responses to text reading, other developmentally appropriate writing, and student talk. Although specific feedback on student work was not observed, the observers noted that some teachers did collect student work for review and feedback after the observation period. Moreover, 10 Grade 1 and all 16 Grade 2 teachers did check student understanding through other strategies during small-group instruction. Finally, interactive writing only occurred in few classes for both grades.

Lesson reflection. Student reflection on the lesson was observed in only a few classrooms, 5 and 6 out of 16 in Grades 1 and 2, respectively (Table 7).

Table 9
Number and Percentage of Activities Observed during Small-Group Instruction in the Reading Block, by Grade

	Grade 1 (*N*= 15)[a]		Grade 2 (*N*= 16)	
Small-group instruction	*n*	%	*n*	%
Use of books with various levels	14	93	15	94
Connection to the whole-group lesson	11	73	15	94
Introduction of the book, purpose for reading, or both	15	100	15	94
Student reading through the text independently	15	100	15	94
Prompting or probing to direct attention to reading strategies	14	93	16	100
Word work	15	100	13	81
Written or oral responses to text reading	13	87	11	69
Other developmentally appropriate writing activities	11	73	9	56
Student talk	10	67	13	81
Strategies for checking understanding	10	67	16	100
Interactive writing	2	13	3	19
Specific feedback on student work	0	0	0	0

[a]One Grade 1 reading class did not have small-group instruction.

Mathematics instructional blocks in Grades 3 and 4

General information. Statistics regarding the number of students in the class, length of instructional activities, and structure of the class are summarized in Table 10. The 15 Grade 3 mathematics classes included in this analysis lasted an average of 62 minutes and contained an average of 19 students. The 16 Grade 4 mathematics classes lasted an average of 63 minutes and contained an average of 17 students. On average, Grade 3 classes spent more time on the focus lesson than on the small-group instruction (22 vs. 7 minutes), while Grade 4 spent a similar amount of time on both (17 vs. 15 minutes). Both grades spent an average of 26 minutes on independent practice. Small-group instruction was observed in only 7 out of 15 Grade 3 and 10 out of 16 Grade 4 classrooms, with 7 and 15 minutes on average for the two grades, respectively. Volunteers, helping staff, or both were present in 11 of Grade 3 and 13 of Grade 4 classrooms (Table 11).

Mastery objectives and lesson topics. In 11 out of 15 Grade 3 and 12 out of 16 Grade 4 classrooms, mastery objectives were presented both on the board/overhead and orally (Table 12). The ELO SAIL mathematics curriculum placed an emphasis on number relations and computation. While Grade 3 classes were more likely to involve lessons related to fact strategy ideas and computations, Grade 4 lessons focused more on place value, multiplication, and division (Table 13).

Table 10
General Statistics for Mathematics Block Observations, by Grade

	Grade 3 (*N*=15)				Grade 4 (*N*=16)			
General statistics	Min	Max	Mean	*SD*	Min	Max	Mean	*SD*
Minutes of observation	40	75	62	8.9	50	105	63	12.9
Minutes for focus lesson	10	34	22	7.9	5	30	17	7.3
Minutes for independent practices	10	40	26	8.1	15	40	26	7.6
Minutes for small groups	0	31	7	10.0	0	50	15	15.7
Number of students in classroom	15	27	19	3.4	12	23	17	2.9
Percentage of students in small groups	0	42	10	15.0	0	100	28	29.5

Note. *SD* = Standard Deviation

Table 11
Number and Percentage of Observed Volunteers and Helping Staff in the Mathematics Block, by Grade

	Grade 3 (*N*=15)		Grade 4 (*N*=16)	
Volunteers and helping staff	*n*	%	*n*	%
Volunteer and helping staff	1	7	3	19
Volunteer only	9	60	10	62
Helping staff only	1	7	0	0
No volunteer or helping staff	4	26	3	19

Table 12
Number and Percentage of Mastery Objectives Presented in the Mathematics Block, by Grade

	Grade 3 (*N*=15)		Grade 4 (*N*=16)	
Mastery objectives	*n*	%	*n*	%
Presented on board/overhead and orally	11	73	12	75
Presented on board/overhead or orally	3	20	3	19
Not presented	1	7	1	6

Lesson warm-up. Lesson warm-up took place in almost all 15 Grade 3 and 16 Grade 4 classrooms (Table 14). Warm-up activities in most classes were connected to prior learning (11 in Grade 3 and 13 in Grade 4), but also were aligned to the mastery objectives (12 in Grade 3 and 15 in Grade 4).

Table 13
Number and Percentage of Observed Lessons Taught in the Mathematics Block, by Grade

Lessons	Grade 3 (*N*=15)		Grade 4 (*N*=16)	
	n	%	*n*	%
Place value	1	7	3	19
Fact strategy ideas	2	13	0	0
Two ways to add	1	7	0	0
Addition cover up	2	13	0	0
Compare and subtract	1	7	0	0
Holders and blockers	2	13	0	0
Dividing dominoes	1	7	0	0
Round about	2	13	0	0
Choice time and review	1	7	0	0
Number sense (Add and Subtract)	0	0	1	6
Multiplication	0	0	6	38
Division	0	0	6	37
Missing	1	6	0	0

Table 14
Number and Percentage of Observed Warm-Up and Closure in the Mathematics Block, by Grade

Warm-up and closure	Grade 3 (*N*=15)		Grade 4 (*N*=16)	
	n	%	*n*	%
Warm-up	13	87	15	94
Warm-up connected to prior learning	11	73	13	81
Warm-up connected to mastery objectives	12	80	15	94
Student sharing of learning at closure	10	67	13	81

Note. Observers might miss start of class due to overlap of another class for observation.

Focus lesson. All focus lessons in Grade 3 were aligned with the mastery objectives (Table 15). Almost all the focus lessons included instructional activities such as direct instruction, interactive discussions regarding mathematics problem solving, and checking of student understanding with various strategies. Some observers commented that class discussions usually involved simple and direct teacher-to-students interaction. Also, according to observers' notes, activities and games were integrated in the focus lesson. In both grades, 13 or more focus lessons were connected to students' prior knowledge and allowed students to practice by themselves. Less frequently, the opportunity to solve mathematics problems with different strategies was observed in 11 out of 15 Grade 3 and in only 8 out of 16 Grade 4 classrooms.

Table 15
Number and Percentage of Activities Observed During Focus Lesson in the Mathematics Block, by Grade

	Grade 3 (*N*= 15)		Grade 4 (*N*=16)	
Focus lesson	*n*	*%*	*n*	*%*
Alignment with the mastery objectives	15	100	16	100
Connection to students' prior knowledge	13	87	13	81
Direct instruction	15	100	15	94
Interactive discussion regarding math problem solving	14	93	15	94
Guided student practice	13	87	15	94
Opportunity to solve math problems with different strategies	11	73	8	50
Strategies for checking understanding	15	100	16	100

Independent practice. Independent practice was often integrated in the focus lesson of the mathematics block. In almost all the classrooms of both grades, students practiced for a better understanding of the lesson and received specific feedback from their teachers (Table 16). Some observers noted that students were usually given uniform assignment for practice. In at least two thirds of the classes in both grades, students used manipulatives to solve mathematics problems and had access to mathematical tools. Students in 12 classrooms of both grades also had the opportunity to work cooperatively. Blackline masters or activity sheets were used in 8 out of 15 Grade 3 and 13 out of 16 Grade 4 classes. However, writing in mathematics journals and independent center activities only occurred in two or three classrooms of both grades.

Small-group instruction. As mentioned earlier, small groups were only included in 7 out of 15 Grade 3 and 10 out of 16 Grade 4 mathematics classes. Observers commented that small groups were used mostly to provide extra assistance to help students better understand concepts taught and complete the activities, worksheets, or games assigned. The following activities were observed in all or almost all the small groups (Table 17): the instruction was connected to the focus lesson, students used manipulatives or tools, teachers helped students with problem solving through prompting or probing, students practiced through guided instruction, students were encouraged to talk, and teachers checked student understanding. Problem solving with multiple solutions was observed in 4 out of 7 Grade 3 and 8 out of 10 Grade 4 classes that included small-group instruction.

Lesson closure. Student sharing of learning at closure was marked for 10 out of 15 Grade 3 and 13 out of 16 Grade 4 lessons (Table 14). According to observers, some teachers asked students to reflect on what they learned from the lesson. However, the use of exit cards was the only closure activity in many classrooms, in which students did not have the opportunity to wrap up the class by themselves.

Table 16
Number and Percentage of Activities Observed during Student Independent Practice in the Mathematics Block, by Grade

Independent practices	Grade 3 (*N*=15)		Grade 4 (*N*=16)	
	n	*%*	*n*	*%*
Practice to support understanding of the lesson	14	93	16	100
Student using manipulative relevant to lesson's objectives	12	80	11	69
Blackline masters or activity sheets	8	53	13	81
Writing in mathematics journals	3	20	3	19
Various independent center activities	2	13	3	19
Opportunity to work cooperatively	12	80	12	75
Mathematical tools accessible to students	12	80	13	81
Specific feedback on student work	14	93	15	94

Table 17
Number and Percentage of Activities Observed During Small-Group Instruction for the Mathematics Block, by Grade

Activity	Grade 3 (*N*=7)[a]		Grade 4 (*N*= 10)[b]	
	n	*%*	*n*	*%*
Connection to the focus lesson	6	86	9	90
Manipulative or tools used	7	100	10	100
Prompting or probing to help students with problem solving	7	100	9	90
Practice through guided instruction	6	86	10	100
Problem solving with multiple solutions encouraged	4	57	8	80
Evident student talk	7	100	9	90
Strategies for checking understanding	6	86	9	90

[a]Excluding eight Grade 3 mathematics classes without small-group instruction.
[b]Excluding six Grade 4 mathematics classes without small-group instruction.

Summary of classroom observations

This summary is organized to answer the evaluation questions addressing the fidelity of ELO SAIL curriculum implementation.

To what extent were the mastery objectives of ELO SAIL lessons communicated to students?

The observations of Grades 1 and 2 reading blocks and Grades 3 and 4 mathematics blocks confirm that the mastery objectives of the ELO SAIL curricula were appropriately communicated to students through implementation of the instructional components. The lessons were aligned with the mastery objectives that generally reflected the ELO SAIL curricula and also were connected to prior student knowledge. Teachers modeled effective reading strategies and mathematics problem solving. They also checked student understanding and provided specific feedback on student work. Students had opportunities to talk and practice individually and cooperatively with regard to the mastery objectives. Though interactive writing rarely occurred, other developmentally appropriate writing activities were observed throughout the reading block, especially during independent practice. However, few mathematics lessons included writing in mathematics journals.

To what extent were the components of ELO SAIL curricula applied in the classroom?

The components of the instructional blocks, such as whole-group instruction/focus lesson, small groups, student independent practice, lesson warm-up, and reflection/closure, were observed in both reading and mathematics classes. In particular, the reading small-group instruction was well organized and involved a great majority of students in the class. However, there was inconsistent implementation of the instructional block components. For instance, student reflection of the lesson was observed only in a few Grade 1 and 2 reading classes. More than 50% of Grade 3 and nearly 40% of Grade 4 mathematics classes integrated the focus lesson and student independent practice but did not include any small-group activities. Finally, exit cards were used at closure in many mathematics classes in which student sharing of learning was not actually observed.

To what extent was ELO SAIL instruction differentiated to meet student needs?

Differentiated instruction was better performed in Grades 1 and 2 reading blocks than in Grades 3 and 4 mathematics blocks. Books of various reading levels were used and instruction and student practices were managed in accordance with student reading levels in the small groups. Literacy center activities also varied to meet different needs of students. However, differentiation of instruction in the mathematics block was limited by a lack of small groups, absence of center activities, and uniform class assignments.

To what extent was a rich learning environment facilitated in ELO SAIL classrooms to support student engagement?

Observed teachers in both reading and mathematics classes made an effort to facilitate a rich learning environment to support student engagement. The lessons were filled with teacher modeling, student practices, discussions, activities, and games. Books of various levels, small groups, student practice, and literacy center activities made the reading block stimulating. Utilization of manipulatives plus games made the mathematics block fun; however, there was a lack of small groups, center activities, and differentiation in student practices in many mathematics classes. Furthermore, problem solving using multiple solutions and neither student-to-student nor student-to-teacher discourse were often observed in mathematics classes.

Academic outcomes

Socio-demographic characteristics of students by comparison groups are described in Appendix E. There were a relatively higher percentage of FARMS and ESOL recipients and Hispanic students in the full participation group than in the nonparticipation group across all grades selected for the outcome evaluation. Appendix F presents descriptive statistics regarding average spring-to-fall changes in Grades 1 and 2 text reading and comprehension levels and average changes in Grade 4 mathematics scores, comparing full-participants with their nonprogram counterparts. The result about reading levels as measured on-site for non-Reading First Grade 1 full program attendees are also included in Appendix F.

Results from outcome analyses were organized to answer evaluation questions below.

Was ELO SAIL effective in maintaining/improving the text reading and comprehension levels of full-program participants relative to that of nonprogram participants?

For Grade 1, the text reading and comprehension level in fall 2007 was 0.65 points higher for full participants than for nonparticipants on average (Table 18). This difference was statistically significant (F=14.76; P<.0001), but its practical significance was not confirmed by an effect size of 0.13. For Grade 2, the difference between full participants and nonparticipants was both statistically insignificant and practically negligible. These results suggest that ELO SAIL was beneficial for Grade 1 but not for Grade 2 students, with respect to fall text reading and comprehension levels.

Table 18
Adjusted Mean Differences and Effect Sizes for Grade 1 and 2
Fall 2007 Text Reading and Comprehension Levels

	Adjusted mean[a]		Treatment effect	
	Full participants	Nonparticipants	Adjusted mean difference	Effect size
Grade 1	8.21 (N=590)	7.56 (N=381)	0.65 (F=14.76; p<.00)	0.13
Grade 2	17.74 (N=608)	17.66 (N=433)	0.09 (F=0.01; p=.756)	0.01

[a] Student initial ability, propensity score, FARMS, ESOL, race/ethnicity, special education, and gender were controlled.

Was ELO SAIL effective in maintaining/improving the performance of full-program participants in number relations and computation relative to that of nonprogram participants?

The analysis reveals that the score of the fall mathematics test was 1.23 points higher for Grade 4 full participants than for same-grade nonparticipants on average (Table 19). This difference was statistically significant (F=19.94; p<.0001). The corresponding effect size of 0.24 confirms that the better performance of Grade 4 full participants was practically meaningful. For the Grade 4 at Broad Acres ELO SAIL site, full participants, on average, scored 1.65 points (F=3.71; p=.063) higher than nonparticipants on the fall mathematics test. This difference was not statistically significant, possibly due to the small sample size, but the effect size of 0.27 confirms a practical significance of this difference. These results suggest that Grade 4 students benefited from full ELO SAIL participation with respect to number relations and computation.

Table 19
Adjusted Mean Differences and Effect Sizes for Grade 4 Fall 2007 Mathematics Scores

Number relations and computations	Adjusted mean[a]		Treatment effect	
	Full participants	Nonparticipants	Adjusted mean difference	Effect size
Fall 2007 test	15.23 (*N*=421)	14.0 (*N*=372)	1.23 (*F*=19.94; *P*<.000)	0.24
Fall 2007 test (Broad Acres ES)	10.52 (*N*=33)	8.88 (*N*=8)	1.65 (*F*=0.81; *P*=.376)	0.27

[a] Student initial ability, propensity score, FARMS, ESOL, race/ethnicity, special education, and gender were controlled.

Did the academic benefit of ELO SAIL vary by student subgroups with respect to race/ethnicity, gender, and receipt of ESOL, FARMS, and special education services?

Differences between full participants and nonparticipants in Grades 1 and 2 fall text reading and comprehension levels and Grade 4 fall mathematics scores were assessed within student socio-demographic subgroups. The results in Figures 1 to 10 show that the benefits of ELO SAIL varied by student subgroups (Appendix G).

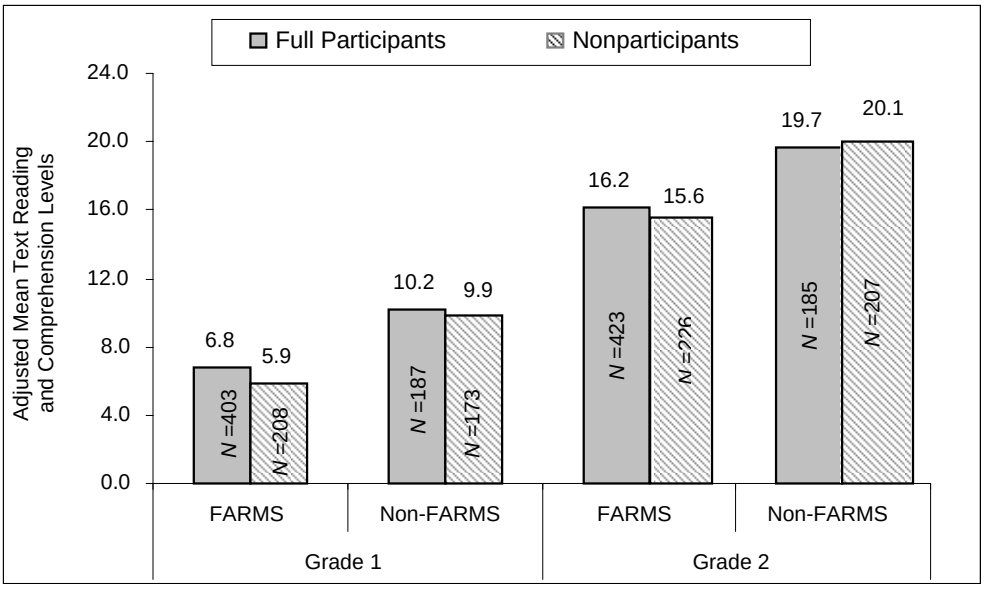

Figure 1. Adjusted means for fall 2007 text reading and comprehension levels, by FARMS and ELO SAIL participation.

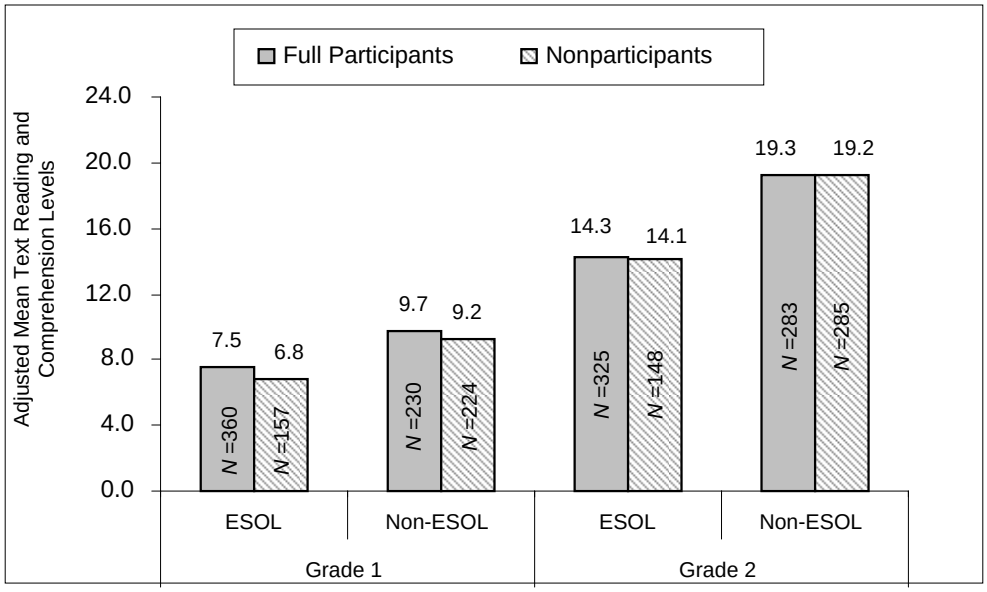

Figure 2. Adjusted means for fall 2007 text reading and comprehension levels, by ESOL and ELO SAIL participation.

Grade 1 fall 2007 text reading and comprehension levels by student subgroups

For FARMS and ESOL students, full participants scored 0.86 (F=26.24; p<.0001) points and 0.71 points (F=13.83; p<.0001) higher than nonparticipants respectively on fall text reading and comprehension levels (Figures 1 and 2). These statistically significant differences were also educationally meaningful, confirmed by the corresponding effect sizes of 0.22 and 0.19. For non-FARMS and non-ESOL students, however, the between-group differences were not significant. These results suggest that full participation in ELO SAIL was beneficial in reading for recipients of FARMS and ESOL services.

Among non-recipients of special education services, full participants scored 0.70 levels (F=15.24; p<.0001) higher than nonparticipants on fall text reading and comprehension levels (Figures 3 and 4), with an effect size of 0.14; the between-group difference for recipients of special education service was not observed. Looking into gender groups, the differences of 0.49 (F=3.74; p<.05) for female and 0.83 (F=13.84; p<00) for male students were both significant, with a greater effect size for males (0.17) than for females (0.09).

Moreover, the effect of ELO SAIL on fall text reading and comprehension levels varied by racial/ethnic groups (Figure 5). Full program attendance was significantly beneficial for African American and Hispanic students, with between-group differences of 0.89 (F=8.89; p<.005) and 0.51 (F=5.91; p<.05). The corresponding effect sizes were 0.18 and 0.13. An insignificant but approaching small effect size of 0.16 for White students may be due to the relatively small sample. There was no difference observed between full participants and nonparticipants for Asian American students.

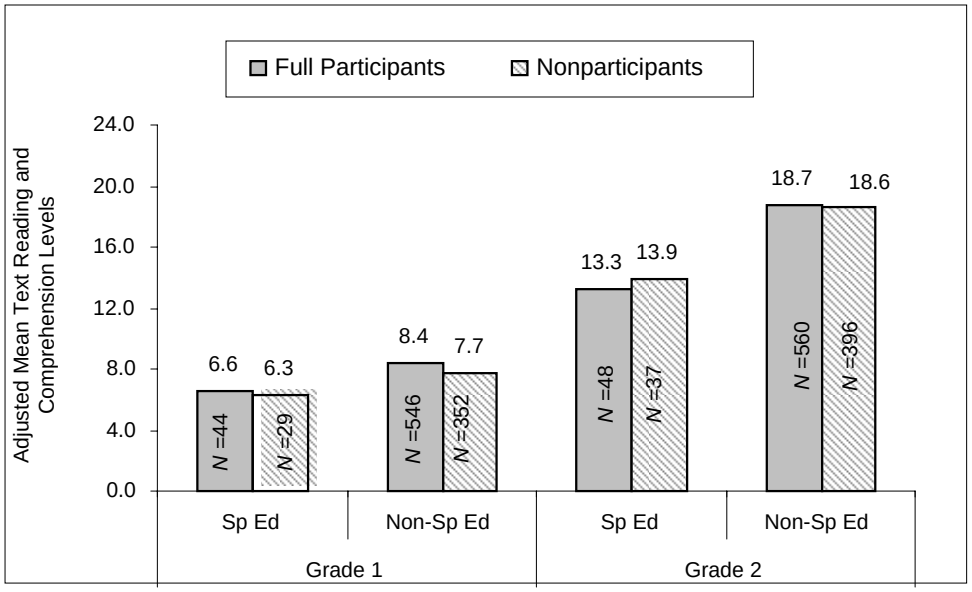

Figure 3. Adjusted means for fall 2007 text reading and comprehension levels, by special education and ELO SAIL participation.

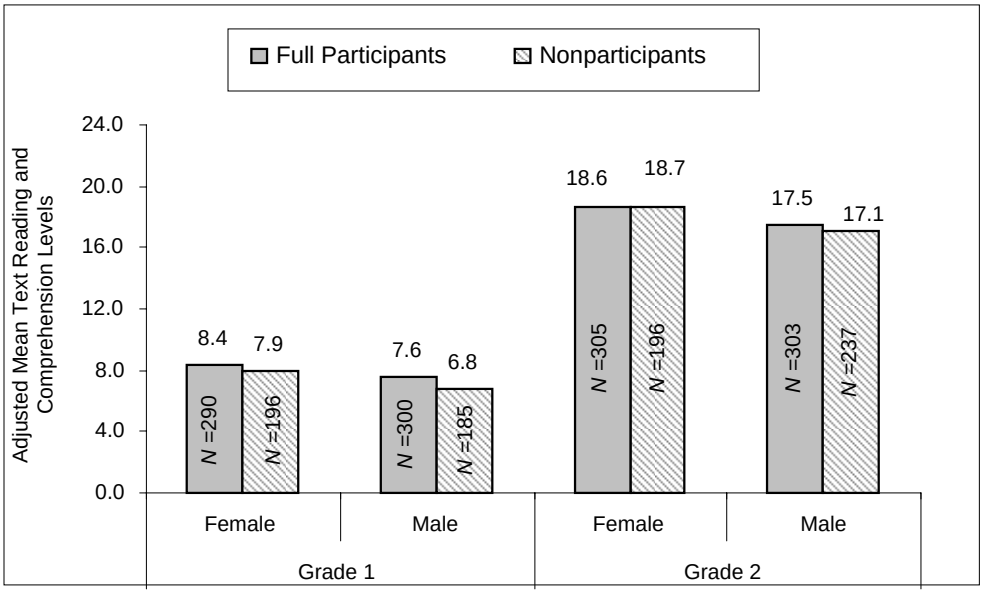

Figure 4. Adjusted means for fall 2007 text reading and comprehension levels, by gender and ELO SAIL participation.

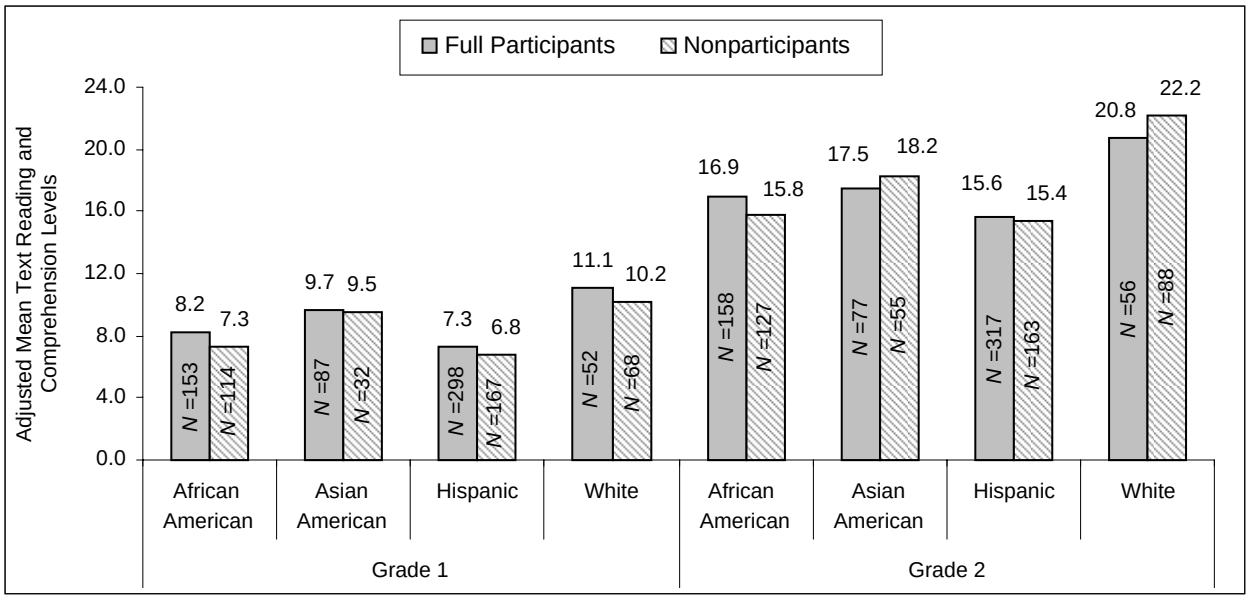

Figure 5. Adjusted means for fall 2007 text reading and comprehension levels, by race/ethnicity and ELO SAIL participation.

Grade 2 fall 2007 text reading and comprehension levels by student subgroups

The results in Figures 1 to 5 also indicate that none of the student subgroups appeared to significantly benefit from participating in ELO SAIL, with respect to fall text reading and comprehension levels. For African American students, benefit in text reading and comprehension levels approached significance with a between-group difference of 1.04 (F=3.58; p=.06) and an effect size of 0.16. Among White students, the effect size of 0.17 approached being practically significant.

Grade 4 fall 2007 mathematics scores by student subgroups

Among students receiving FARMS services, full participants scored 1.52 points (F=21.81; p<.0001) higher on average than nonparticipants in the fall mathematics test (Figure 6). This statistically significant difference was also practically meaningful, with an effect size of 0.31. However, the between-group difference was not obtained for students not receiving FARMS services. These results suggest an ELO SAIL benefit in number relations and computation for students receiving FARMS services but not for students not receiving these services.

For the ESOL groups, the between-group differences of 1.42 (F=7.96; p<005) on the fall mathematics scores among services recipients and 1.11 (F=11.23; p<.001) among non-recipients were both significant in favor of full participants (Figure 7). The mathematics benefit was more substantive for services recipients, as suggested by the effect sizes of 0.29 and 0.22, respectively. Similarly, full participants scored significantly higher in both subgroups of special education, with 1.93 (F=5.10; p<.05) for recipients and 1.17 (F=16.16; p<.0001) for non-recipients (Figure 8). But a more substantive between-group difference for recipients than for non-recipients was confirmed by their effect sizes of 0.45 and 0.23, respectively.

Looking into gender groups (Figure 9), the between-group difference of 1.59 (F=19.15; p<.0001) for female students was significant, with an effect size of 0.32, whereas the difference of 0.80 (F=3.55; p=.061) for male students approached significance, with an effect size of 0.16.

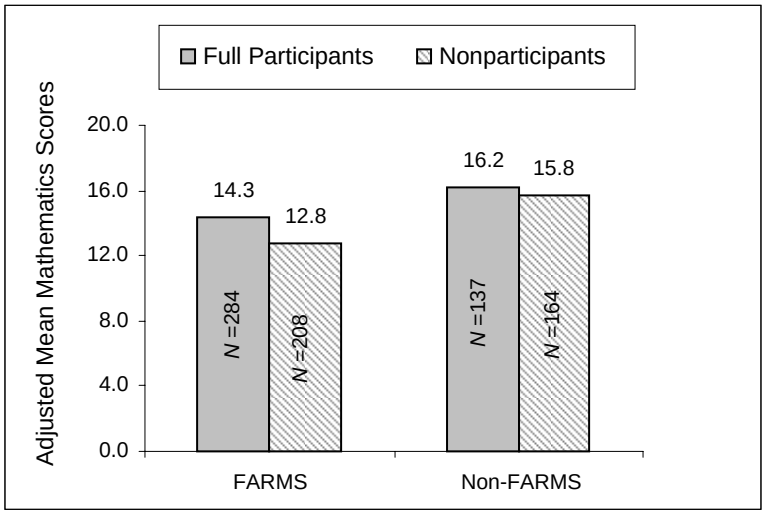

Figure 6. Adjusted means for Grade 4 fall 2007 mathematics scores, by FARMS and ELO SAIL participation.

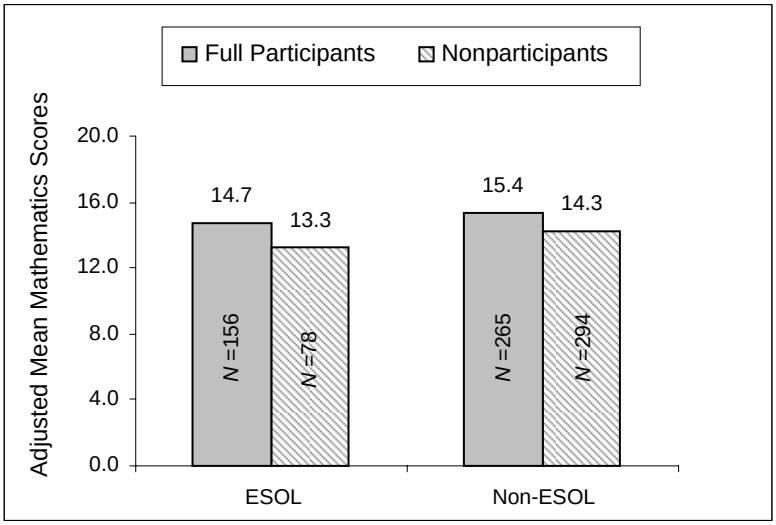

Figure 7. Adjusted means for Grade 4 fall 2007 mathematics scores, by ESOL and ELO SAIL participation.

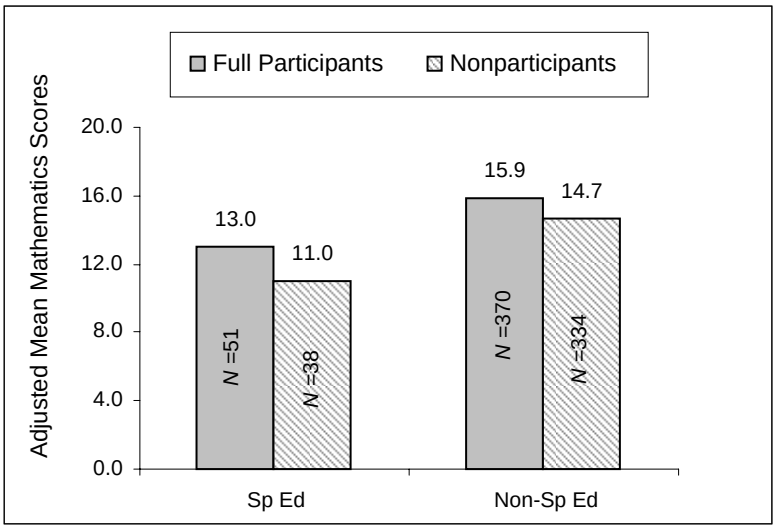

Figure 8. Adjusted means for Grade 4 fall 2007 mathematics scores, by special education and ELO SAIL participation.

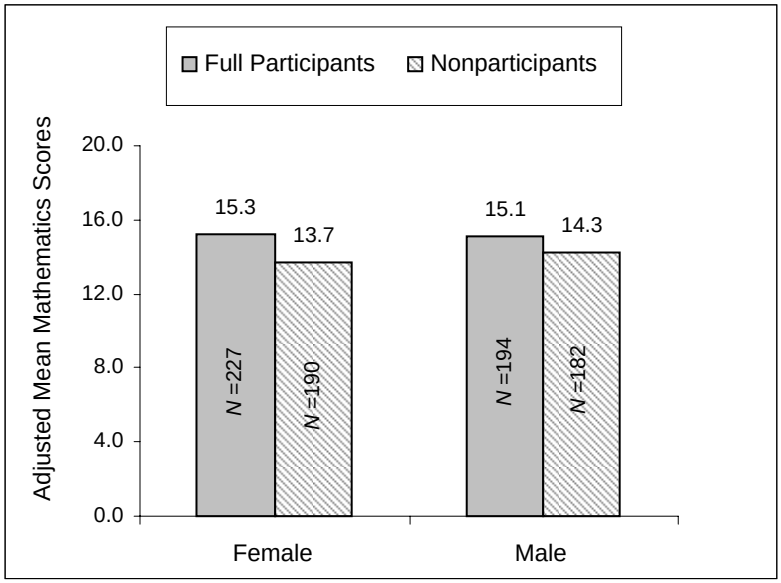

Figure 9. Adjusted means for Grade 4 fall 2007 mathematics scores, by gender and ELO SAIL participation.

Moreover, ELO SAIL affected racial/ethnic groups differently (Figure 10). Full participation was highly beneficial for Asian American and Hispanic students but not for African American and White students, with respect to number relations and computation. The between-group difference was 1.99 (F=5.27; p<.05) for Asian American students and 1.50 (F=16.99; p<.0001) for Hispanic students, with relatively large effect sizes of 0.43 and 0.32, respectively.

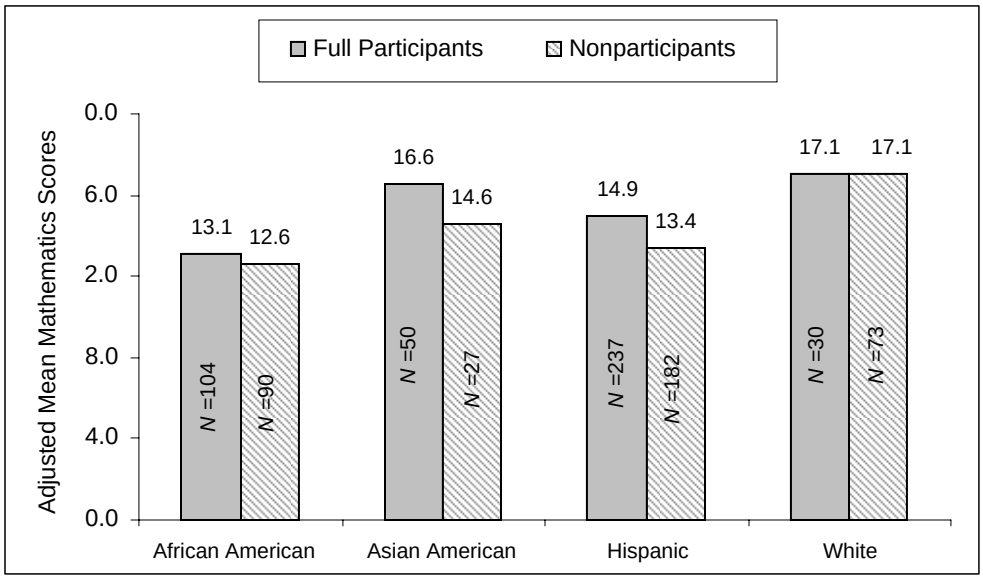

Figure 10. Adjusted means for Grade 4 fall 2007 mathematics scores, by race/ethnicity and ELO SAIL participation.

Conclusion

This evaluation summarizes the implementation and academic outcomes of the 2007 ELO SAIL program. The primary goal of the program was to alleviate summer academic loss and promote continued learning in students. The results obtained from program stakeholder surveys, classroom observations in selected grades, and student outcome measures for selected grades address whether the program was operated as planned and accomplished its academic goal. The information provided by this evaluation may also help understand whether the program is on the track of reducing the academic achievement gaps among students based on socio-demographic groups.

Findings

Administrator/principal surveys

The 32 administrators who responded viewed the 2007 ELO SAIL implementation positively, in general. In particular, 85% or more of them received support and information for administration, program implementation, and procuring material and supplies. They were able to recruit qualified teachers and paraeducators and receive adequate support from multiple staff members. Most summer administrators had prior-program communication with the principal, but there was less agreement about the effectiveness of assigning responsibilities to summer administrators (60%). Also, 80% or more agreed that the procedures of student arrival and dismissal went well; and the afternoon session enhanced the program effectiveness. Though more than 80% agreed that the administration of mathematics assessments worked well, more than 25% believed the administration of the running records during the program did not go smoothly. Further, more than 90% agreed the classroom observations were minimally disruptive to instruction; and between 80% and 100% also agreed that the program met the needs of students in each grade.

Classroom teacher surveys

Teachers also held a positive view of the ELO SAIL implementation, in general. Almost all of the 271 teachers perceived the program as successful in providing review and preview sessions to facilitate continued learning in reading. There was less agreement about the effectiveness of the mathematics curriculum among the teachers (about 75%). With respect to differentiated instruction, teachers responded more favorably to the reading curriculum (81%) than to the mathematics curriculum (60%). Nearly 75% agreed that sufficient time was provided for planning and setup, and 65% agreed that supplies were adequate. Among non-Reading First teachers, almost 80% agreed that the reading material was adequate. More than 90% agreed that the library media center resources were available, and the media centers were more often used for checking out books by teachers and students than for lesson planning.

Fidelity of curriculum implementation

Findings from classroom observations demonstrated that the mastery objectives of the ELO SAIL curricula were appropriately communicated to students through implementation of the instructional components, such as alignment with the mastery objectives of the lesson, connection to student prior knowledge, and modeling of effective reading strategies and mathematics problem solving.

In almost all the 16 Grade 1 and the 16 Grade 2 reading classes, components of the instructional block, including whole group, small groups, independent practice, and lesson warm-up and reflection, were observed. Differentiated instruction included small-group instruction at different paces and levels, differentiated assignments, and center activities that allowed students to select activities of various types and levels. The small-group instruction in the reading classes was well organized and involved most students. However, there was inconsistent implementation of the instructional block components in the mathematics classes. For instance, small-group instruction was observed in less than 50% of Grade 3 and slightly more than 60% of Grade 4 mathematics classes. Center activities and differentiated class assignments also were observed more frequently in Grades 1 and 2 reading than in Grades 3 and 4 mathematics classes.

Furthermore, teachers in both reading and mathematics classes supported student engagement through teacher modeling, student practices, discussions, various learning activities, and games. Opportunities for multiple problem solutions were observed in about three quarters of Grade 3 and one half of Grade 4 mathematics focus lessons.

Academic outcomes

The hypothesis about the impact of ELO SAIL on text reading and comprehension levels was supported by the results for Grade 1 but not for Grade 2. Specifically, Grade 1 full participants significantly outperformed their nonparticipating counterparts on fall text reading and comprehension levels. However, no significant association emerged between program participation and Grade 2 text reading and comprehension levels. The hypothesis about the impact of ELO SAIL on mathematics was supported by Grade 4 fall test scores. Specifically, full participants significantly outperformed their nonparticipating counterparts on number relations and computation. The between-group difference was large enough to be of practical significance to educators for Grade 4 mathematics but not for Grade 1 reading. The higher level of practical significance for Grade 4 mathematics than for Grade 1 reading might be explained by the following reasons: 1) compared with the standardized reading assessment, the mathematics assessment developed for ELO SAIL was more sensitive to the score change related to the program; and 2) students from low-income families may be especially vulnerable to summer reading achievement loss (Cooper, 2001), therefore a short summer program would alleviate their summer loss or increase their reading scores to a limited extent.

Also as anticipated, the positive impact of ELO SAIL was more likely to occur in certain socio-demographic subgroups. For Grade 1 text reading and comprehension levels, full ELO SAIL participation had a significant benefit for students receiving FARMS and ESOL services, whereas it did not make a notable change for non-recipients of these services. As for Grade 4 mathematics, full participation was significantly beneficial for students receiving FARMS services but not for students who were not eligible to receive FARMS services. Grade 4 students benefited from full attendance in mathematics but with a greater benefit for ESOL students than for non-ESOL students.

In addition, ELO SAIL participation was associated with a significant reading benefit for students not receiving special education services in Grade 1; it had a greater mathematics benefit for recipients than for non-recipients of these services in Grade 4. Male students in Grade 1 benefited more in reading than females, and female students in Grade 4 had a greater gain in mathematics performance than males. Grade 1 African American and Hispanic students

improved their text reading and comprehension levels significantly; the improvement for same-grade White students approached practical significance. As for Grade 4 mathematics, the program benefit was evident only for Asian American and Hispanic students.

In summary, findings of the outcome evaluation demonstrates that positive academic impacts associated with the 2007 ELO SAIL program remained evident in fall, two months after the summer program ended. The higher Grade 1 fall text reading and comprehension levels and the higher Grade 4 fall mathematics scores for full participants suggest that 2007 ELO SAIL was effective in prevention of summer academic loss and even resulted in more academic gain in these grades. Additionally, findings from the disaggregated data suggest that demographically disadvantaged student groups, especially those impacted by poverty and limited English language proficiency, increased more in Grade 1 fall text reading and comprehension levels and Grade 4 fall mathematics scores as a result of full participation. These results are consistent with the ones found in previous ELO SAIL evaluations (Sunmonu et al., 2002, 2004).

Recommendations for Future Program Implementation

Class size

Investigate why a smaller class was among the most common recommendations for program improvement proposed by the teachers.

Curriculum

Provide sufficient review lessons in reading for all grades, in order to reinforce academic skills already acquired and help students acquire on grade-level concepts and skills. Update the mathematics curriculum to ensure sufficient lessons and resources are provided to challenge students at different levels and accelerate student learning. Investigate and identify factors responsible for the lack of significant program impact on Grade 2 text reading and comprehension levels.

Instruction

Provide teachers with guidance and assistance to support differentiated instruction in Grades 3 and 4 mathematics classes. Provide sufficient instructional materials and time for classroom setup and lesson planning. Encourage teachers to increase student-to-student and student-to-teacher discourse, student sharing of multiple problem-solving strategies, and mathematics journals. Encourage the use of summary activities by students, beyond exit cards. Further, provide support for teachers during the administration of mathematics assessments; and provide training to help teachers administer running records efficiently, if the evaluation is repeated in future years.

Student recruitment

Continue and expand the afternoon program, if possible, because it would enhance the ELO SAIL effectiveness, and also increase program attendance (Wade and Cooper-Martin, 2007; Cooper-Martin and Wade, 2006).

Strengths and Limitations of the Evaluation

Strengths

Data collection in this evaluation involved multiple procedures, such as stakeholder surveys, classroom observations, and the administration of assessments. The results from analyzing the data provide a comprehensive understanding of the 2007 ELO SAIL program operation and its academic impact. Based on stakeholders' feedback in the previous year, the teacher surveys were updated in order to gather more detailed information about the summer curriculum for future improvement. Data obtained from the structured observations depict a picture of instructional activities in the classrooms, which offer a closer look at the core part of program implementation. Moreover, the outcome analysis in this evaluation incorporated a nonrandomized comparison group pre- and posttest quasi-experimental design, which meets both the scientific rigor and ethical standard. Specifically, the full-ELO SAIL participating group was compared with the voluntary nonparticipating group, so the ELO SAIL services were not withheld from eligible students. Between-group differences in intervening variables were adjusted for through the sophisticated analytical procedure selected for this outcome analysis. Propensity scores may further mitigate the limitation of nonrandomized design and improve the comparability of the two groups by taking care of additional variances associated with the between-group differences.

Limitations

Because of the nature of a quasi-experimental study without random assignment, the nonequivalence between the full-participating group and the nonparticipating group may potentially affect the outcome analysis. Although the selected statistical procedure is effective in controlling for some variables potentially associated with the preexisting differences, the degree to which the two groups are comparable is still limited by some unobservable confounding factors occurring before and during the program. A classic example related to this design weakness is the difficulty of knowing the reasons behind nonparticipation and nonparticipants' summer activities. Nonparticipants could do things like attend other academic summer programs, go to the library or visiting museums frequently, get tutoring at home, or travel to another part of the world during the summer. Such activities may have academic benefits among nonparticipants, which might lead to biased conclusions.

Attrition of the student sample due to the selection procedure is another limitation in this outcome study. Only students of selected grades who had a valid score in both spring and fall tests were included in the analysis. The attrition rate from the original sample was about 10% for Grade 1 and 2 text reading and comprehension levels and more than 30% for Grade 4 mathematics. Even though descriptive statistics show that students without complete assessment scores were similar in the distribution of socio-demographic characteristics by program participation level, their performance was not measured and analyzed for a more definitive conclusion.

References

Borman, G. D., Overman, L. T., Fairchild, R., Boulay, M. B., & Kaplan, J. (2004). Can a multiyear summer program prevent the accumulation of summer learning losses? In G. D. Borman & M. Boulay (Eds.), *Summer learning* (pp. 233–253). Mahwah, NJ: Lawrence Erlbaum Associates.

Borman, G. D., & Dowling, N. M. (2006). Longitudinal achievement effects of multiyear summer school: Evidence from the Teach Baltimore Randomized Field Trial. *Educational Evaluation and Policy Analysis, 28*(1), 25–48.

Chaplin, D., & Capizzano, J. (2006). *Impacts of a summer learning program: A random assignment study of Building Educated Leaders for Life (BELL).* Mathematics Policy Research/The Urban Institute.

Cohen, J. (1988). *Statistical power analysis for the behavioral science (2nd ed.).* Hillsdale, NJ: Lawrence Earlbaum Associates.

Cooper-Martin, E., & Wade, J. (2006). *Montgomery County Public Schools 21st Century Community Learning Centers Program: Outcomes for summer 2003 and 2004*. Rockville, MD: Montgomery County Public Schools.

Cooper, H. (2001). Summer school: Research-based recommendations for policymakers. *SERVE Policy Brief*. Greensboro, NC: School of Education, University of North Carolina.

Cooper, H., Charlton, K., Valentine, J. C., & Muhlenbruck, L. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of Society for Research in Child Development, 1*(201), 1–109.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research, 66*(3), 227–268.

Datnow, A., Borman, G. D., Stringfield, S., Overman, L. T., & Castellano, M. (2003). Comprehensive school reforming culturally and linguistically diverse contexts: Implication and outcomes form a four-year study. *Educational Evaluation and Policy Analysis (2)*, (25), 143–170.

Dehejia, R. H., & Wahba, S. (2002). Propensity score matching methods for non-experimental causal studies. *Review of Economics and Statistics, 84* (1), 151–161.

Entwisle, D. R., & Alexander, K. L. (1992). Summer setback: Race, poverty, school composition, and mathematics achievement in the first two years of school. *American Sociological Review, 57*, 72–84.

Green, C. A. (1998*).* The extended school year program consolidated report: Achievement test scores and survey findings. Detroit, MI: Detroit Public Schools, Office of Research, Evaluation, and Testing.

Holdzkom, D. (2002). *Summer academy in the Wake County Public Schools: An evaluation of the 2001 program.* Raleigh, NC: Wake County Public School System, Department of Evaluation and Research.

Joseph, F. H. (1999). *Measuring achievement growth in an 18-day summer school session.* Presented at the Annual Meeting of the American Education Research Association. Montreal, Quebec, Canada.

Joseph, F. H. (2001). *The effectiveness of summer school in getting students to function on grade level in gateway grades.* Presented at the Annual Meeting of the American Education Research Association. Seattle, WA.

Landis, J. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, (33), 159–174.

MCPS. (2007) Extended Learning Opportunities Summer Adventures in Learning (ELO SAIL) program in Montgomery County Public Schools (MCPS). *http://www.montgomeryschoolsmd.org/departments/familycommunity/ELO.shtm*

Paris, S. G., Pearson, P. D., Cervetti, G., and Carpenter, R. (2004). Assessing the effectiveness of summer reading programs. In G. D. Borman & M. Boulay (Eds.), *Summer Learning* (pp. 121–161). Mahwah, NJ: Lawrence Erlbaum Associates.

Portz, John. (2004). Summer school 2000 and 2001*:* The Boston Public Schools Transition Services Program. In G. D. Borman & M. Boulay (Eds.), *Summer Learning* (pp. 103–119). Mahwah, NJ: Lawrence Erlbaum Associates.

Roberts, G., & Nowakowski, J. (2004). Addressing the summer learning loss: An evaluation of the Voyager summer reading intervention program. In G. D. Borman & M. Boulay (Eds.), *Summer Learning* (pp. 165–181). Mahwah, NJ: Lawrence Erlbaum Associates.

Roderick, M., Jacob, B. A., & Bryk, A. S. (2004). Summer in the city: Achievement gains in Chicago's Summer Bridge Program. In G. D. Borman & M. Boulay (Eds.), *Summer Learning* (pp. 73–102). Mahwah, NJ: Lawrence Erlbaum Associates.

Sunmonu, K., Larson, J., Van Horn, Y., Cooper-Martin, E., & Nielsen, J. (2002). *Evaluation of the Extended Learning Opportunities Summer Program.* Rockville, MD: Montgomery County Public Schools, Department of Shared Accountability.

Sunmonu, K., Curry-Corcoran, D., Mordica, J. (2004). *2003 Summer Evaluation of the Extended Learning Opportunities (ELO) Program.* Rockville, MD: Montgomery County Public Schools, Department of Shared Accountability.

Wade, J., & Cooper-Martin, E. (2007). *Montgomery County Public Schools 21st Century Community Learning Centers Program: Implementation and outcomes for summer 2005*. Rockville, MD: Montgomery County Public Schools.

Appendixes

Appendix A

List of Title I Schools Included in 2007 ELO SAIL Program Evaluation

Participating elementary schools (*N*=22)[a]	Reading First schools[b]	Primary schools[c]	Upper-grade schools[d]
Broad Acres			
Burnt Mills			
East Silver Spring		√	
Gaithersburg			
Georgian Forest			
Harmony Hills			
Highland	√		
Kemp Mill			
Montgomery Knolls		√	
New Hampshire Estates		√	
Oak View			√
Rolling Terrace			
Roscoe Nix		√	
Rosemont	√		
Sargent Shriver			
South Lake			
Summit Hall	√		
Twinbrook			
Viers Mill			
Washington Grove			
Weller Road			
Wheaton Woods	√		

[i]ncluding schools receiving Title I funds. [b]Including schools using Reading First notebook for Grades K–3. [c]Including schools with Grades K–2 only. [d]Including schools with Grades 3–5 only.

Appendix B

Inter-rater Reliability for 2007 ELO SAIL Classroom Observation

Grade 1 and 2 reading class observation (Valid items: 79)			Grade 3 and 4 mathematics class observation (Valid items: 32)		
Observer pair	Kappa coef.	*P*-value	Observer pair	Kappa coef.	*P*-value
1. DSA - Title I	0.894	0.000	1. DSA - DSA	0.871	0.000
2. DSA - Title I	0.772	0.000	2. DSA - Title I	0.805	0.000
3. DSA - Title I	0.732	0.000	3. DSA - Title I	0.798	0.000
4. DSA - Title I	0.684	0.000	4. DSA - DSA	0.750	0.000
5. DSA - Title I	0.651	0.000	5. DSA - Title I	0.467	0.002
6. DSA - Title I	0.643	0.000	6. DSA - Title I	0.438	0.012
7. DSA - DSA	0.628	0.000	7. DSA - Title I	0.266	0.027
8. DSA - Title I	0.582	0.000	8. DSA - Title I	0.166	0.089
9. DSA - DSA	0.462	0.000			
Mean	0.672		Mean	0.570	
Std. Deviation	0.121		Std. Deviation	0.270	
Median	0.651		Median	0.609	

Note. Kappa coefficients can range from 1.0, representing perfect agreement, to negative infinity, representing near perfect disagreement, with a kappa of zero indicating that the level of agreement between raters is equivalent to chance. Landis and Koch (1977) suggested descriptive labels for kappa coefficients greater than 0. Kappa coefficients between .41-.60 represent moderate inter-rater agreement, coefficients between .61-.80 represent substantial inter-rater agreement, and coefficients between .81-1.00 represent almost perfect agreement.

Appendix C

Item Analyses for Fall 2007 ELO SAIL Mathematics Test

	Review subtest			Preview subtest	
Item	Item difficulty[a]	Discrimination index[b]	Item	Item difficulty[a]	Discrimination index[b]
Q 1	.93	.189	Q 15	.62	.508
Q 2	.92	.271	Q 16	.58	.510
Q 3	.68	.498	Q 17	.59	.495
Q 4	.67	.507	Q 18	.56	.518
Q 5	.92	.303	Q 19	.59	.482
Q 6	.83	.318	Q 20	.25	.450
Q 7	.88	.427	Q 21	.16	.480
Q 8	.70	.387	Q 22	.14	.467
Q 9	.30	.440	Q 23	.46	.399
Q 10	.28	.477	Q 24	.41	.429
Q 11	.79	.396	Q 25	.08	.274
Q 12	.59	.436	Q 26	.05	.329
Q 13	.47	.534	Q 27	.12	.391
Q 14	.45	.534	Q 28	.09	.398

Note. The Cronbach's Alpha (reliability) of the test is 0.861.

[a]Item difficulty refers to percentage of students who selected the correct answer. The higher the value, the easier an item is. [b]Discrimination Index refers to the relationship between students' total performance and their performance on each item. The higher the value is toward 1.00, the stronger the relationship is. The index above 0.25 is within acceptable range.

Appendix D: Administrator and Teacher Respondents' Background

Table D1
Number and Percentage of Administrators' Positions and Schools in Spring and Summer 2007 (N=32)

Administrative position		*n*	%
Position in spring	Principal	2	6.7
	Permanent Assistant Principal	7	23.3
	Assistant Principal I	16	53.3
	Assistant Principal II	4	13.3
	Other	1	3.3
Position in summer	Principal	4	12.5
	ELO SAIL administrator	26	81.3
	Other	2	6.3
Same school in spring and summer	Yes	13	40.6
	No	19	59.4

Table D2
Timing of Administrators' Involvement with Planning for 2007 ELO SAIL Administration/Management (*N*=32)

Planning time (*n*=32)	*n* of responses	% of responses[a]
Prior to start of 2007 program	26	81.3
During 2007 program	17	53.1
Not involved	2	6.3

[a]Calculated based on the total number of respondents who provided at least one response (*n*). The percentage of responses may exceed 100% because respondents marked more than one response.

Table D3
Grade Levels Teachers Taught in Spring, Summer, and Fall 2007

		Non Reading First teachers (*N*=224)		Reading First teachers (*N*=47)	
	Grade/Position	*n* of responses	% of responses	*n* of responses	% of responses[a]
Spring 2007	Prekindergarten	6	2.7	1	2.2
	Kindergarten	38	17.1	12	26.1
	First grade	38	17.1	8	17.4
	Second grade	36	16.2	7	15.2
	Third grade	29	13.1	6	13.0
	Fourth grade	36	16.2	3	6.5
	Fifth grade	29	13.1	4	8.7
	Specialist/ESOL/MCC/GT	41	18.5	6	13.0
	Other grade/position	19	8.6	3	6.5
		(*n*=222)		(*n*=46)	
Summer 2007	Kindergarten	65	29.0	15	31.9
	First grade	64	28.6	13	27.7
	Second grade	56	25.0	15	31.9
	Third grade	40	17.9	12	25.5
	Fourth grade	50	22.3		
	Fifth grade	39	17.4		
		(*n*=224)		(*n*=47)	
Fall 2007	Prekindergarten	8	3.6		
	Kindergarten	45	20.1	12	26.1
	First grade	42	18.8	8	17.4
	Second grade	45	20.1	7	15.2
	Third grade	25	11.2	5	10.9
	Fourth grade	37	16.5	3	6.5
	Fifth grade	34	15.2	4	8.7
	Specialist/ESOL/MCC/GT	41	18.3	5	10.9
	Other grade/position	12	5.4	5	10.9
		(*n*=224)		(*n*=47)	

[a]Calculated based on the total number of respondents who provided at least one response (*n*). The percentage of responses may exceed 100% because respondents marked more than one response.

Table D4
Teachers' Years of Teaching Experience

		Mean	*SD*	Median	Min	Max
Non Reading First teachers (*N*=224[a])	Years of teaching experience	9.6	8.7	7	1	42
	Years of working at MCPS	7.3	6.8	6	1	35
Reading First teachers (*N*=47[a])	Years of teaching experience	7.4	6.6	5	1	32
	Years of working at MCPS	6.1	6.1	4	1	32

[a]Missing data were not included.

Table D5
Teachers' Grade Levels and Certified Areas

	Non Reading First teachers (*N*=224)		Reading First teachers (*N*=47)	
Grade levels and certified areas	*n* of responses	% of responses*	*n* of responses	% of responses*
Early childhood certification	74	33.6	20	42.6
Elementary certification	155	70.5	33	70.2
Secondary certification	11	5.0		
Special education certification	24	10.9	5	10.6
ESOL certification	18	8.2	6	12.8
Other certification	32	14.5	5	10.6
	(*n*=220)		(*n*=47)	

[a]Calculated based on the total number of respondents who provided at least one response (*n*). The percentage of responses may exceed 100% because respondents marked more than one response.

Table D6
Professional Characteristics of Teachers

		Non Reading First teachers (*N*=224)		Reading First teachers (*N*=47)	
		n	%	*n*	%
Highest degree held	Doctorate	2	0.9		
	Master	118	52.7	25	53.2
	Bachelor	97	43.3	22	46.8
	Other	6	2.7		
	No response	1	0.4		
Years of teaching experience	0–3 years	52	23.2	16	34.0
	4–-6 years	57	25.4	12	25.5
	7–12 years	54	24.1	10	21.3
	13 years or more	58	25.9	8	17.0
	No response	3	1.3	1	2.1
Years of working at MCPS	0–3 years	73	32.6	23	48.9
	4–6 years	55	24.6	9	19.1
	7–12 years	60	26.8	7	14.9
	13 years or more	33	14.7	7	14.9
	No response	3	1.3	1	2.1

Appendix E: Demographic Characteristics of Students Included in 2007 ELO SAIL Outcome Analyses by Grade[a]

Student characteristics	Grade 1				Grade 2				Grade 4			
	Full participants		Nonparticipants		Full participants		Nonparticipants		Full participants		Nonparticipants	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
All students	590	100	381	100	608	100	433	100	421	100	372	100
Services provided												
Current FARMS	403	68.3	208	54.6	423	69.6	226	52.2	284	67.5	208	55.9
Current ESOL	360	61.0	157	41.2	325	53.5	148	34.2	156	37.1	78	21.0
Current Sp Ed	44	7.5	29	7.6	48	7.9	37	8.5	51	12.1	38	10.2
Gender												
Female	290	49.2	196	51.4	305	50.2	196	45.3	227	53.9	190	51.1
Male	300	50.8	185	48.6	303	49.8	237	54.7	194	46.1	182	48.9
Race/ethnicity[b]												
African American	153	25.9	114	29.9	158	26.0	127	29.3	104	24.7	90	24.2
Asian American	87	14.7	32	8.4	77	12.7	55	12.7	50	11.9	27	7.3
Hispanic	298	50.5	167	43.8	317	52.1	163	37.6	237	56.3	182	48.9
White	52	8.8	68	17.8	56	9.2	88	20.3	30	7.1	73	19.6

[a]Grades 1 and 2 were selected for reading while Grade 4 for mathematics outcome analyses. [b]Indian American students were not included in the outcome analysis due to an insufficient number of this group.

Appendix F: Descriptive Statistics of Scores in Reading and Mathematics Assessments

Table F1
Descriptive Statistics for Grade 1 and 2 Spring and Fall 2007 Text Reading and Comprehension Levels

		Full participants				Nonparticipants			
		Min.	Max.	Mean	*SD*	Min.	Max.	Mean	*SD*
Grade 1	Spring 2007	3	27	7.06	3.95	3	38	7.87	5.39
	Fall 2007	3	27	7.78	4.54	3	38	7.98	5.74
	Spring-fall gain (unadjusted)			0.72	(*N*=590)			0.11	(*N*=381)
Grade 2	Spring 2007	3	38	16.60	5.21	3	38	17.98	5.44
	Fall 2007	3	38	17.18	6.59	3	38	18.83	7.33
	Spring-fall gain (unadjusted)			0.58	(*N*=608)			0.85	(*N*=433)

Table F2
Descriptive Statistics for Grade 4 Spring and Fall 2007 Mathematics Scores

Number relations and computations	Full participants (*N*=421)				Nonparticipants (*N*=372)			
	Min.	Max.	Mean	*SD*	Min.	Max.	Mean	*SD*
Fall 2007 test	1	25	13.19	4.91	1	28	14.77	5.37
Spring 2007 test	1	26	14.72	4.81	1	28	14.73	5.34
Spring-fall gain (unadjusted)			1.53				-0.04	

Table F3
Pre- to Posttest Gain in Reading Levels with On-site Grade 1 Running Records (*N*=565[a])

Running records	Range	Mean	*SD*	Pre- to posttest gain	*t*-value	Effect size
Beginning of program	0-38	7.77	6.70	1.27	7.42 (p<.000)	0.44
End of program	0-38	9.04	7.60			

[a]Including non-Reading First Grade 1 full participants who were administrated running records at both beginning and end of the program

Appendix G: Adjusted Mean Differences and Effect Sizes for Reading and Mathematics Assessments, by Student Subgroups

Table G1
Adjusted Mean Differences and Effect Sizes for Grade 1
Fall 2007 Text Reading and Comprehension Levels, by Student Subgroups

Student subgroups	Adjusted mean[a]		Treatment effect	
	Full participants	Nonparticipants	Adjusted mean difference	Effect size
FARMS	6.75 (*N*=403)	5.89 (*N*=208)	0.86 (*F*=26.24; *p*<.00)	0.22
Non-FARMS	10.16 (*N*=187)	9.89 (*N*=173)	0.26 (*F*=0.56; *p*=.455)	0.04
ESOL	7.51 (*N*=360)	6.79 (*N*=157)	0.71 (*F*=13.83; *p*<.00)	0.19
Non-ESOL	9.74 (*N*=230)	9.20 (*N*=224)	0.54 (*F*=3.72; *p*=.06)	0.10
Special education	6.59 (*N*=44)	6.34 (*N*=29)	0.25 (*F*=0.22; *p*=.636)	0.06
Non-special education	8.41 (*N*=546)	7.71 (*N*=352)	0.70 (*F*=15.24; *p*<.00)	0.14
Female	8.43 (*N*=290)	7.94 (*N*=196)	0.49 (*F*=3.74; *p*<.05)	0.09
Male	7.58 (*N*=300)	6.75 (*N*=185)	0.83 (*F*=13.84; *p*<.00)	0.17
African American	8.18 (*N*=153)	7.29 (*N*=114)	0.89 (*F*=8.98; *p*<.005)	0.18
Asian American	9.72 (*N*=87)	9.53 (*N*=32)	0.19 (*F*=0.09; *p*=.769)	0.03
Hispanic	7.26 (*N*=298)	6.75 (*N*=167)	0.51 (*F*=5.91; *p*<.05)	0.13
White	11.13 (*N*=52)	10.19 (*N*=68)	0.94 (*F*=1.98; *p*=.161)	0.16

[a]Student initial ability, propensity score, FARMS, ESOL, race/ethnicity, special education, and gender were controlled.

Table G2
Adjusted Mean Differences and Effect Sizes for Grade 2
Fall 2007 Text Reading and Comprehension Levels, by Student Subgroups

Student subgroups	Adjusted mean[a]		Treatment effect	
	Full participants	Nonparticipants	Adjusted mean difference	Effect size
FARMS	16.17 (*N*=423)	15.61 (*N*=226)	0.56 (*F*=2.87; *p*=.09)	0.09
Non-FARMS	19.69 (*N*=185)	20.06 (*N*=207)	0.36 (*F*=0.56; *p*=.454)	0.05
ESOL	14.31 (*N*=325)	14.13 (*N*=148)	0.17 (*F*=0.26; *p*=.605)	0.03
Non-ESOL	19.33 (*N*=283)	19.19 (*N*=285)	0.14 (*F*=0.12; *p*=.727)	0.02
Special education	13.34 (*N*=48)	13.87 (*N*=37)	0.53 (*F*=0.33; *p*=.567)	0.08
Non-special education	18.68 (*N*=560)	18.56 (*N*=396)	0.12 (*F*=0.16; *p*=.69)	0.02
Female	18.62 (*N*=305)	18.69 (*N*=196)	0.07 (*F*=0.26; *p*=.871)	0.01
Male	17.50 (*N*=303)	17.10 (*N*=237)	0.40 (*F*=1.17; *p*=.279)	0.06
African American	16.88 (*N*=158)	15.84 (*N*=127)	1.04 (*F*=3.58; *p*=.06)	0.16
Asian American	17.46 (*N*=77)	18.15 (*N*=55)	0.69 (*F*=1.09; *p*=.298)	0.11
Hispanic	15.59 (*N*=317)	15.35 (*N*=163)	0.24 (*F*=0.39; *p*=.54)	0.04
White	20.83 (*N*=56)	22.23 (*N*=88)	1.40 (*F*=2.01; *p*=.158)	0.17

[a]Student initial ability, propensity score, FARMS, ESOL, race, special education, and gender were controlled.

Table G3
Adjusted Mean Differences and Effect Sizes for Grade 4
Fall 2007 Mathematics Scores, by Student Subgroups

Student subgroups	Adjusted mean[a]		Treatment effect	
	Full participants	Nonparticipants	Adjusted mean difference	Effect size
Current FARMS	14.32 (*N*=284)	12.79 (*N*=208)	1.52 (*F*=21.81; *p*<.00)	0.31
Not Current FARMS	16.24 (*N*=137)	15.76 (*N*=164)	0.48 (*F*=0.93; *p*=.337)	0.10
Current ESOL	14.68 (*N*=156)	13.27 (*N*=78)	1.42 (*F*=7.96; *p*<.005)	0.29
Not Current ESOL	15.36 (*N*=265)	14.25 (*N*=294)	1.11 (*F*=11.23; *p*<.001)	0.22
Current Sp Ed	12.97 (*N*=51)	11.04 (*N*=38)	1.93 (*F*=5.10; *p*<.05)	0.45
Not Current Sp Ed	15.86 (*N*=370)	14.69 (*N*=334)	1.17 (*F*=16.16; *p*<.00)	0.23
Female	15.27 (*N*=227)	13.68 (*N*=190)	1.59 (*F*=19.15; *p*<.00)	0.32
Male	15.07 (*N*=194)	14.27 (*N*=182)	0.80 (*F*=3.55; *p*=.061)	0.16
African American	13.10 (*N*=104)	12.62 (*N*=90)	0.48 (*F*=0.66; *p*=.419)	0.10
Asian American	16.56 (*N*=50)	14.56 (*N*=27)	1.99 (*F*=5.27; *p*<.05)	0.43
Hispanic	14.94 (*N*=237)	13.44 (*N*=182)	1.50 (*F*=16.99; *p*<.00)	0.32
White	17.08 (*N*=30)	17.09 (*N*=73)	0.01 (*F*=0; *p* =.993)	0.00

[a]Student initial ability, propensity score, FARMS, ESOL, race/ethnicity, special education, and gender were controlled.