



**MODEL LITERACY PROGRAMS
SAVE THE CHILDREN**

Evaluation Findings from the 2005-06 School Year

Richard N. White
Elizabeth R. Reisner

Policy Studies Associates, Inc.

February 2007

Prepared for Save the Children, U.S. Programs

Contents

	Page
1. Did the Model Literacy Program Increase Opportunities for Children to Participate in Literacy Instruction and for Staff to Develop Instructional Capacity?.....	3
2. Who Participated in the Model Literacy Programs?.....	4
3. Did Children Participate in Available Literacy Learning Opportunities?	7
4. Did Participants' Literacy Skills Improve?.....	12
5. What Factors Were Associated with Changes in Reading Proficiency?	19
Appendix: Using Normal Curve Equivalent Scores	

MODEL LITERACY PROGRAMS SAVE THE CHILDREN

Evaluation Findings from the 2005-06 School Year

During the 2005-06 school year, Save the Children supported 47 local programs in its model literacy initiative. Twenty of the sites operated for the first time during the 2005-06 school year, while 15 began operation in the spring of the 2004-05 school year, and the remaining 12 began operation during the 2003-04 school year. Services at local sites included the delivery of integrated in-school, afterschool, and summer-school literacy activities. Save the Children's support to the participating programs included the following:

- Provision of literacy learning materials, including Accelerated Reader (AR), Fluent Reader (a program designed to help struggling readers increase oral reading skills), creation or expansion of libraries, provision of computer equipment, and a standardized test to permit periodic assessment of literacy proficiency
- Training of local staff and volunteers
- An integrated, web-based system for tracking information about participating children and youth, the program services they receive, and their outcomes, including change in literacy skills

This report describes the implementation of the literacy model in 47 program sites, the characteristics of participating children, and the learning results they achieved during 2005-06.

Key findings from this analysis include the following:

- The 47 programs were successful in establishing programs that provided literacy services to children both after school and during the school day. Each program recruited and employed both a site supervisor and a literacy coordinator. Other paid and volunteer staff were employed to provide an average ratio of one adult for every 10 participating children. Programs served a total of 6,650 children during the 2005-06 school year.
- Save the Children and Renaissance Learning, the publisher of Accelerated Reader, provided substantial amounts of training and technical assistance to local program staff. On average, programs received 43 hours of training by Save the Children, 10 hours of coaching by Renaissance Learning, and 22 hours of technical assistance in literacy programming.
- The model literacy programs served a population of children whose reading skills were significantly below grade level. On average, 79 percent of participants

scored below grade level on the first reading assessment administered during 2005-06.

- In the 27 programs in operation for two years, 41 percent of participants had also participated in 2004-05. In the 12 programs in operation for three years, 33 percent of participants were in the program in both 2003-04 and 2005-06.
- On average, participating children attended a model literacy program for 51 days during the 2005-06 school year. While participating in the program, children attended an average of 54 percent of the days possible for them to attend. Participants in the program for two or three years attended at higher levels in 2005-06 than did first-time participants.
- Measures of engagement with the programs' literacy activities indicate that participants read 63 books during 2005-06, on average, and the median number of books read was 49. Overall, participants passed 89 percent of the AR quizzes they completed about the books they had read.
- Participants improved their literacy skills significantly over the course of the program period, moving closer toward grade-level performance. The average literacy gain was 3.9 normal curve equivalents (NCEs).¹ Fifty-four percent of participants increased their literacy performance by at least 2 NCEs, which is considered to be a meaningful learning gain by the publisher of the test used by the programs.
- Following a year's participation in a program, the percent of participants scoring at grade level grew by 5.5 percentage points, a noteworthy increase given that only 22 percent of those completing two standardized tests of reading proficiency were reading at grade level when measured at baseline.
- Estimates of the cumulative gains across school years suggest that children who continue to participate in the model literacy programs made additional gains in reading proficiency for each year of participation.
- The level of children's participation in model program activities was clearly associated with the level of gains in reading proficiency. If unrelated factors, such as literacy instruction in their school day classrooms, were causing the learning gains, we would not expect to see this relationship. Because we do, this finding suggests that the programs' literacy learning efforts produced positive results for participating children.

¹ Throughout this study, scores are reported in normal curve equivalents (NCEs). An NCE score is a standardized score (based on a normal distribution) in which scores represent equal intervals that range from 1 to 99. Unlike percentiles or grade equivalents, which are not equal interval scales, NCEs can be subtracted and averaged, allowing the computation of summary statistics of group performance. A more detailed explanation of NCE scores appears in the appendix.

1. Did the Model Literacy Program Increase Opportunities for Children to Participate in Literacy Instruction and for Staff to Develop Instructional Capacity?

The 47 model literacy programs were located in 11 states. Forty percent of programs were located in Kentucky, 11 percent each in Arizona and California, 9 percent each in Mississippi and South Carolina, and the remainder in Arkansas, Colorado, Georgia, Nevada, New Mexico, and Tennessee.

Forty-two of the model literacy programs were affiliated with a host school located in a rural school district. The enrollment of the host schools ranged from fewer than 100 children to almost 800. The majority of the schools served kindergarten through eighth grade, but several served only two or three grade levels. In nearly three-quarters of the host schools (73 percent), more than 90 percent of children were eligible for the Free or Reduced-Price Lunch Program, an indication that their families had incomes near or below the poverty line.

Days of Service Provided

Thirty-four of the 47 model literacy programs (72 percent) operating during the 2005-06 school year operated during each of the nine months between September 2005 and May 2006. Ten programs operated for eight months, one for seven months, one for five months, and one for four months. Overall, the 47 programs averaged 8.6 months of operation.

The 47 programs provided an average of 108 days of literacy services as part of their afterschool program. Most (36 of 47 or 77 percent) also provided individual activities related to the AR program to participating children during the school day. Thirty-two of the programs (68 percent) provided in-school tutoring in literacy. On average, the 47 programs provided 100 days of in-school services.

For this study, when a program provided more than one type of service on the same day, providing both in-school and afterschool services, it was counted as a single day of service. On average, each of the 47 programs offered literacy services on 125 days between September 2005 and May 2006.

Program Staffing

The model literacy programs used a combination of paid staff and volunteers for their programs. The total number of adults involved in each program ranged from 4 to 52, with an average of 18. A majority of projects (25 of the 47) had more paid staff than volunteers.

Overall, programs had sufficient staff to maintain a ratio of 10 participants per adult. The participant to staff ratio ranged from a low of 1:1 to a high of 36:1.

The two key staff positions at each program were the site supervisor and the literacy coordinator. The site supervisor was responsible for recruiting other staff, planning the activities of the program, overseeing the day-to-day operation of the program, and serving as the liaison to school staff.

The literacy coordinator's position was full time and involved directing the on-site literacy activities, collecting and monitoring data to guide program improvement, managing the technology initiatives, overseeing the tutoring programs, and working directly with children to increase their reading skill.

Save the Children and Renaissance Learning provided training and technical assistance in literacy activities for the staff of the model literacy programs. On average, the staff and volunteers involved in literacy at each program received 43 hours of training from Save the Children. Staff also received an average of 10 hours of coaching by Renaissance Learning specialists in the use of AR. This was supplemented by an average of 22 hours of technical assistance focused on literacy programming. From all sources, an average of 75 hours of support was provided to the staff in each model literacy program.

2. Who Participated in the Model Literacy Programs?

The following discussion reviews overall enrollment patterns, settings for program services, children's baseline reading proficiency, and relevant patterns of children's participation over multiple years.

Aggregate Enrollment Patterns

The total number of individual participants enrolled in the 47 programs over the course of the 2005-06 school year was 6,650.² The average number of participants per program was 141.

Children enrolled in the model literacy programs were about equally likely to be sponsored Save the Children participants (49 percent) or children without sponsors (51 percent). More than half of the participants were enrolled in kindergarten through fourth grade, as shown in Exhibit 1.

² A participant is defined as a child age six or older, who attended literacy activity on at least one day between September 1, 2005, and April 30, 2006.

Exhibit 1
Grade Level of Participants, 2005-06

School Grade	Percent of Participants (N=6,650)
Pre-kindergarten	0
Kindergarten	6
1	10
2	14
3	19
4	19
5	17
6	9
7	3
8	3

Exhibit reads: Less than 1 percent of model literacy program participants during 2005-06 were enrolled in pre-kindergarten.

The monthly enrollment in the model literacy programs ranged from a low of 4,094 in September 2005 to a high of 4,858 in October 2005, and averaged 4,540 children per month, as shown in Exhibit 2.

Exhibit 2
Number of Participants, by Month, 2005-06

Month	Total Enrollment
September 2005	4,094
October 2005	4,858
November 2005	4,702
December 2005	4,623
January 2006	4,754
February 2006	4,593
March 2006	4,577
April 2006	4,122
Average	4,540

Exhibit reads: In September 2005, a total of 4,094 children participated in the 47 model literacy programs.

Contact Setting

The model literacy programs provided services to participating children after school and during the school day. Twenty-seven percent of children received services both after school and during the school day, 36 percent only after school, and 37 percent only during the school day.

Baseline Reading Proficiency

An important question in this evaluation was whether the model literacy programs targeted children who were poor readers relative to their grade in school. The measure of reading proficiency available for this study was the STAR Reading Assessment, a standardized test that children participating in the AR program complete two to three times per year.

The distribution of scores from the first STAR Reading Assessment that each child completed is the best indicator available of the baseline reading proficiency of program participants. Analysis of this distribution indicates that, overall, the programs succeeded in targeting children who were poor readers relative to their grade in school, as shown in Exhibit 3. The average of participants' initial scores on the STAR Reading Assessment (N=5,281) was 34 NCEs, the equivalent of the 22nd percentile using national norms.³

Exhibit 3
Initial STAR Reading Assessment Scores

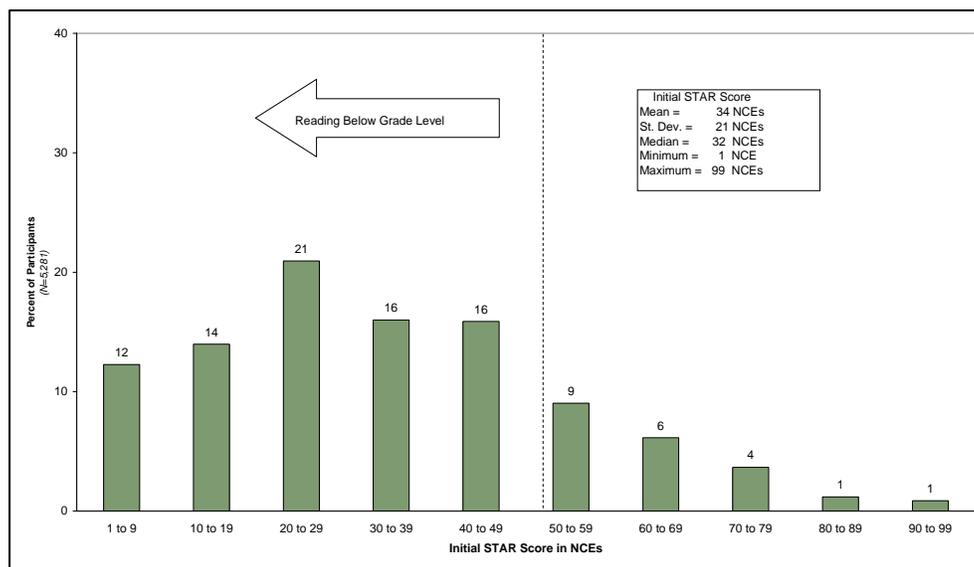


Exhibit reads: Twelve percent of participants scored between 1 and 9 NCEs on their initial STAR Reading Assessment.

A score of 50 NCEs is generally interpreted as performance at grade level. Using this definition of a poor reader relative to the child's grade in school, 79 percent of participants performed below grade level on their initial STAR Reading Assessment.

Initial performance on the STAR Reading Assessment varied among the 46 programs for which STAR Reading Assessment data were available. Participants at seven of the 46 programs averaged below 30 NCEs on their initial assessment, while participants at 11 programs averaged

³ Included in this analysis are all children, age six or older, who attended one of the model literacy programs at least one day and who completed at least one STAR Reading Assessment. The STAR Reading Assessment is not administered to children in kindergarten and is administered only in the spring for children in first grade.

above 39 NCEs. Similarly, at 11 of the programs, more than 90 percent of participants scored below grade level on their initial assessment, while at the other end of the distribution, fewer than 70 percent scored below grade level at nine other programs.

Sponsored participants in the Save the Children program had initial scores that were higher than those of other children. Sponsored participants averaged a score of 36 NCEs, compared with an average of 32 NCEs among other children, a statistically significant difference. Seventy-six percent of sponsored participants had initial scores below grade level, compared with 81 percent of other participants, a statistically significant difference.

Participation in Multiple Years

Of the 47 model literacy programs operating during the 2005-06 school year, 20 were in their first year of operation, 15 in their second year, and 12 had been operating for three years. Among 12 model literacy programs that operated each year between 2003-04 and 2005-06, 1,773 children participated during 2003-04. Among the children who participated in 2003-04, 1,438 (84 percent) continued to participate during 2004-05, and 703 (41 percent) continued to participate during 2005-06. Among the 4,062 children who participated at one of the 27 model literacy programs operating during 2004-05, 1,767 (44 percent) continued to participate in 2005-06.

3. Did Children Participate in Available Literacy Learning Opportunities?

The following discussion examines attendance at the literacy programs and the level of children's engagement in the AR literacy learning activities, including their performance on the AR quizzes.

Participation in Literacy Programming

The analyses conducted for this evaluation are limited to children who were age six or older and who attended the program at least one day between September 2005 and May 2006. As noted, participation in any combination of the three types of program activities on the same calendar day is counted as a single day of attendance: in-school work with a program tutor, in-school individual work with the AR materials, or afterschool program activities.

Average daily attendance. Over the 2005-06 school year, 78 or more children attended each of the programs per day, on average, and the average total daily attendance was 3,646 children across all programs, as shown in Exhibit 4.

Exhibit 4
Average Number of Children Attending per Day per Program,
by Month, 2005-06

Program	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	2005-06
Average number of children attending each day, per program	74	79	76	77	79	82	81	73	78

Exhibit reads: During September 2005, an average of 74 children attended each of the model literacy programs each day.

Program days attended. Among individual literacy program participants, the number of days that children attended between September 2005 and May 2006 ranged from 1 to 145 days and averaged 51 days, with a median of 43 days, as shown in Exhibit 5. Thirty-one percent of participants attended 80 days or more.

Exhibit 5
Number of Days Attended per Participant, 2005-06

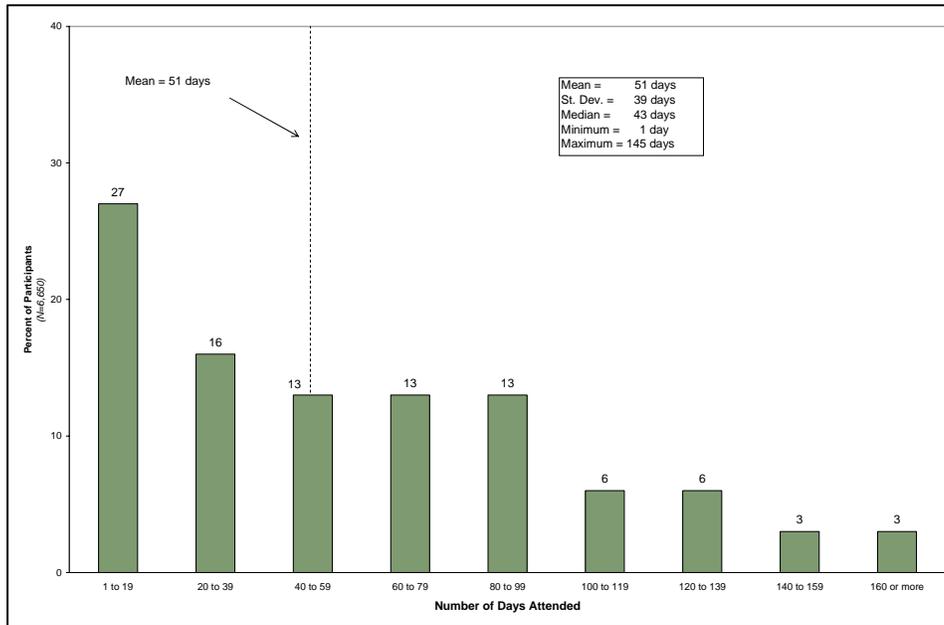


Exhibit reads: Twenty-seven percent of children attended a literacy program between 1 and 19 days during the 2005-06 school year.

Program attendance rate. The average attendance rate, which is computed as the number of days attended divided by the number of days it was possible for a child to attend, ranged from 2 percent to 100 percent and averaged 54 percent, with a median of 55 percent, as seen in Exhibit 6. Fifty-seven percent of participants attended 50 percent or more of the days possible for them to attend.

Exhibit 6
Individual Participant Program Attendance Rate, 2005-06

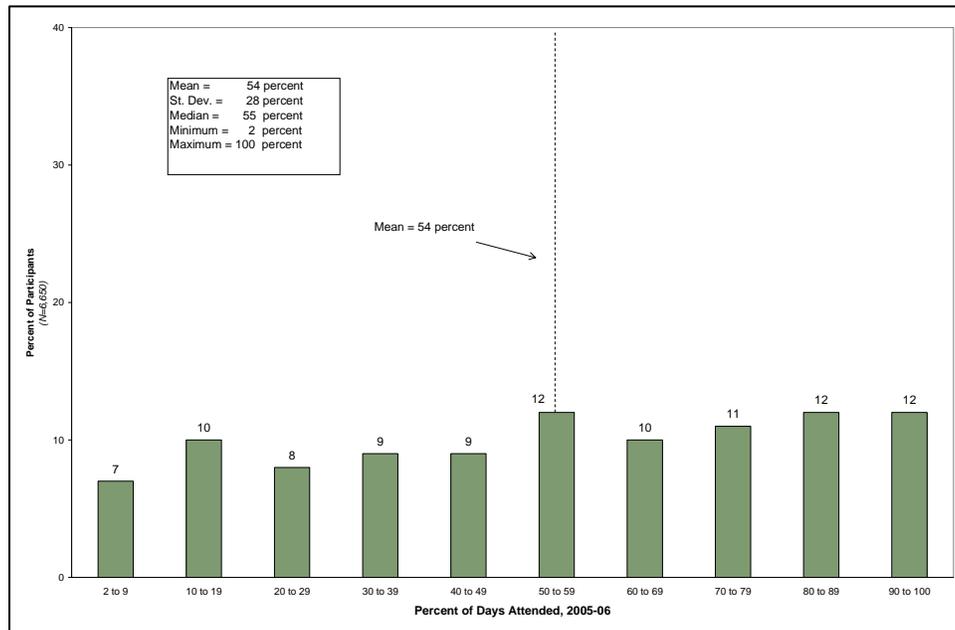


Exhibit reads: Seven percent of participants attended between 2 percent and 9 percent of the days possible.

The average number of program days attended and the program attendance rate varied considerably across the 47 model literacy programs. The average number of days attended ranged from 16 days to a high of 101 days. The average number of days attended was below 30 days at five programs and above 75 days at seven other programs. Similarly, the average percent of days attended at six programs was below 35 percent and above 75 percent at six other programs.

Prior participation. Overall, the children who had participated in the model literacy program for more years attended the program for more days during 2005-06 and attended a higher percentage of the days possible, as shown in Exhibit 7.

Exhibit 7
Days Attended and Program Attendance Rate During 2005-06,
by Years of Participation

Program Attendance 2005-06	Years of Participation		
	One year beginning in 2005-06 (N=4,860)	Two years beginning in 2004-05 (N=1,767)	Three years beginning in 2003-04 (N=703)
Days attended	46	66	71
Percent of days possible attended	51	60	68

Exhibit reads: Children who began participation in 2005-06 attended an average of 46 days during 2005-06 and attended 51 percent of the days that services were offered.

Participation in AR Activities

The AR program is designed to encourage children to read and understand books written at an appropriate level of difficulty for their reading skill level. After completing a book, the child takes a quiz on the content of the book. The results of the quiz are used to track changes in reading proficiency and to help identify additional books that are appropriate for the child's skill level.

Number of books read. One measure of children's involvement in literacy learning activities is the number of books that they read while participating in the literacy program. The goal among the literacy programs was for each child to read at least 25 books during the school year. Among participating children, the number of books read in 2005-06 ranged from a low of one book to a high of 300, with an average of 63 books and median of 49 books, as shown in Exhibit 8. Overall, 73 percent of participants met or exceeded the standard by reading 25 or more books during the school year. More than 50 percent of participants read 25 books or more in 39 of the 47 model literacy programs (83 percent).

Exhibit 8 Number of Books Read, 2005-06

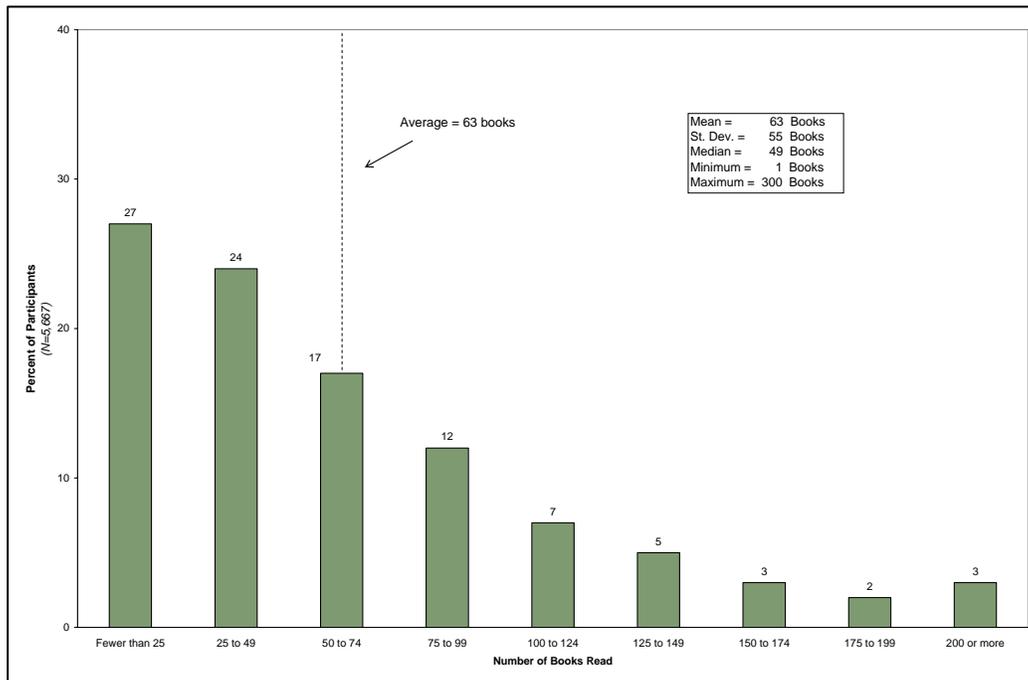


Exhibit reads: Twenty-seven percent of participants read fewer than 25 books during the 2005-06 school year.

Participants' Performance on AR Quizzes

On average, participants attempted 64 AR quizzes during the 2005-06 school year and achieved a passing score on 57 of them (89 percent). The criterion for passing a quiz is answering 60 percent of the questions correctly.

One of the objectives of the model literacy program was for participants to pass at least 85 percent of the quizzes they attempted. This threshold indicates that children are reading materials that are appropriate for their levels of reading proficiency and that they are sufficiently engaged in their reading to retain and use the content they have read. Nearly three-quarters (71 percent) of participants passed 85 percent or more of the quizzes they attempted.

Across all of the quizzes each participant in the 47 programs attempted during the 2005-06 school year, participants answered an average of 82 percent of the questions correctly. Participants answered 85 percent or more of the questions correctly on 56 percent of the quizzes attempted.

Participations in prior years. Children who first participated in 2003-04 and continued through 2005-06 participated in more AR activities during 2005-06 than the children who participated for fewer years (Exhibit 9). Three-year participants read more books, passed a higher percentage of AR quizzes, and answered a higher proportion of quiz items correctly.

Exhibit 9
2005-06 AR Activities, by Years of Participation

AR Activities	Years of Participation		
	One, Beginning in 2005-06 (N=4,039)	Two, Beginning in 2004-05 (N=955)	Three, Beginning in 2003-04 (N=644)
Books read	65	55	68
Percent of quizzes passed	86	88	92
Percent of quiz items correct	81	82	86

Exhibit reads: The children who first participated during 2005-06 read an average of 65 books during 2005-06, passed 86 percent of the AR quizzes they attempted, and answered 81 percent of the quiz items correctly.

4. Did Participants' Literacy Skills Improve?

The primary measure of the reading proficiency of the children participating in the model literacy programs was their performance on the STAR Reading Assessment. The STAR is a standardized, norm-referenced test appropriate for children in grades 1 through 12 and is published by Renaissance Learning, the developers of the AR program. The test is administered on a computer and is adaptive, adjusting the difficulty of each question based on how well the child performed on the preceding questions.

The computer software that administers the STAR Reading Assessment scores each child's reading proficiency relative to national norms. The results of the test are available in a variety of metrics, including scale scores, grade equivalents, percentiles, and normal curve equivalents (NCEs). An important characteristic of the method used to compute STAR Reading Assessment results is that the child's grade level and the month of school within that grade are factored into the scores. As a result, the performance of a child in the fifth month of fourth grade is computed against the estimate of the performance of a national sample of children in the fifth month of fourth grade. This method of scaling test scores means that any increase in NCE scores on the STAR Reading Assessment represents an increase in reading proficiency *in addition to* the increase that would be expected from maturation and simply attending school for the period of time between assessments.

One issue inherent in the analysis of change in reading proficiency as measured by standardized tests is the need to allow a large enough interval of time between the initial and final administrations of the test. Too short an interval does not allow enough time for any real growth in reading proficiency to show up on test items. If the interval between assessments is too short, any change in score on the reading test is more likely to reflect measurement error than change in proficiency. The publishers of the STAR Reading Assessment report that the minimum interval likely to produce valid results is 90 days.

For this study, the estimate of a child's change in reading proficiency was computed by subtracting his/her first score on the STAR Reading Assessment from the child's score on the last assessment completed during the 2005-06 school year. Analyses were limited to children who completed two STAR Reading Assessments at least 90 days apart during a single school year.

Baseline Reading Proficiency

As already explained, a component of the estimation of change in reading proficiency among the children participating in the literacy program is the measured level at which the children were reading when they began participation. Baseline estimates calculated here for purposes of computing change over time are slightly different from the baseline descriptions presented earlier in this report (e.g., Exhibit 3), because the change estimates are limited to the subset of program participants who completed at least two STAR Reading Assessments at least 90 days apart. Among the 6,650 children participating in the program, 5,281 completed one or more STAR Reading Assessments, and 4,316 (65 percent of the 6,650) completed two or more assessments at least 90 days apart. A large portion of the participants who did not complete two STAR Reading Assessments were in pre-kindergarten or kindergarten, so they did not complete any STAR Reading Assessments, or they were in first grade and likely to have completed only one assessment at the end of the school year.

A review of the initial STAR Reading Assessment scores of the pertinent subgroup clearly confirms that, while the model literacy programs served children at the full range of initial reading proficiency during the 2005-06 school year, the majority read far below the level expected for children at their grade in school. Among the 4,316 participants who completed at least two STAR Reading Assessments at least 90 days apart, the average score on the initial test was 34 NCEs, and ranged from a low of 1 NCE to a high of 99 NCEs (the lowest and highest NCE scores possible). A score of 50 NCEs, equivalent to the fiftieth percentile on national norms, is generally accepted as performing at grade level. Twenty-two percent of participants performed at 50 NCEs or higher, as shown in Exhibit 10. At the other end of the distribution, 25 percent performed below 20 NCEs.

Exhibit 10
Initial Score on STAR Reading Assessment in NCEs, 2005-06,
for Children Who Completed Two Assessments at Least 90 Days Apart

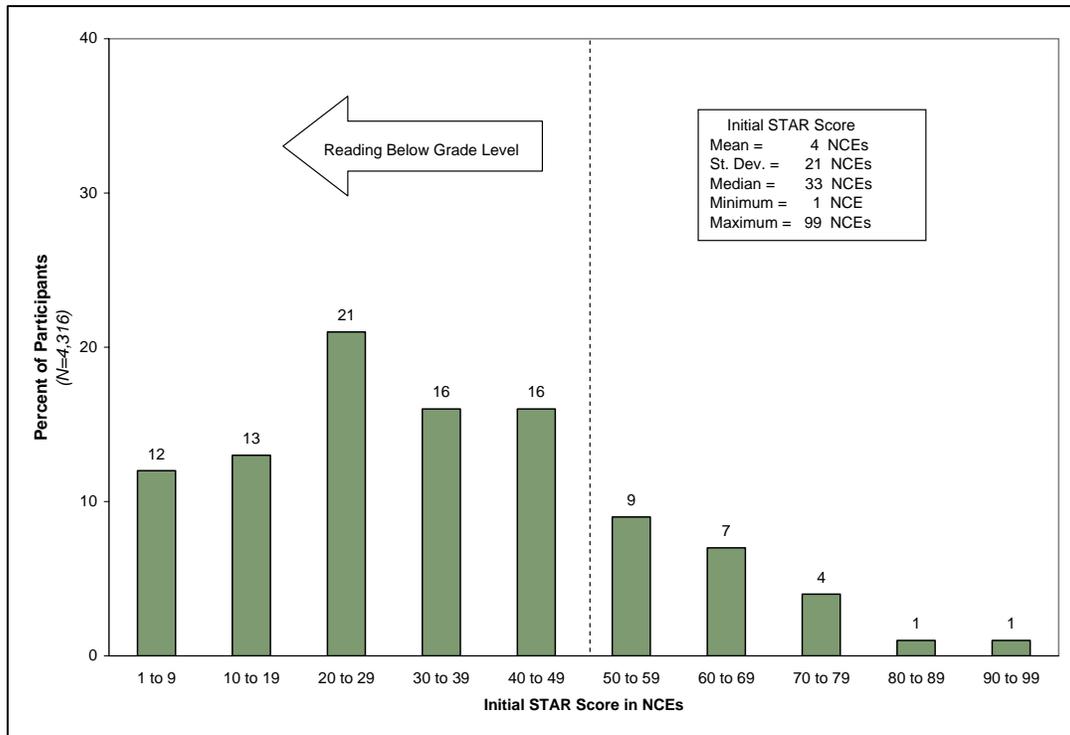


Exhibit reads: Twelve percent of children with two STAR Reading Assessments at least 90 days apart scored between 1 and 9 NCEs on their initial STAR Reading Assessment.

Change in Reading Proficiency Scores

A comparison of the distribution of the first and last STAR Reading Assessments achieved by participants is shown in Exhibit 11. Overall, the average score increased from 34 NCEs on the first assessment to 38 NCEs on the final assessment.

Exhibit 11
Initial and Final Scores on STAR Reading Assessment in NCEs, 2005-06,
for Children Who Completed Two Assessments at Least 90 Days Apart

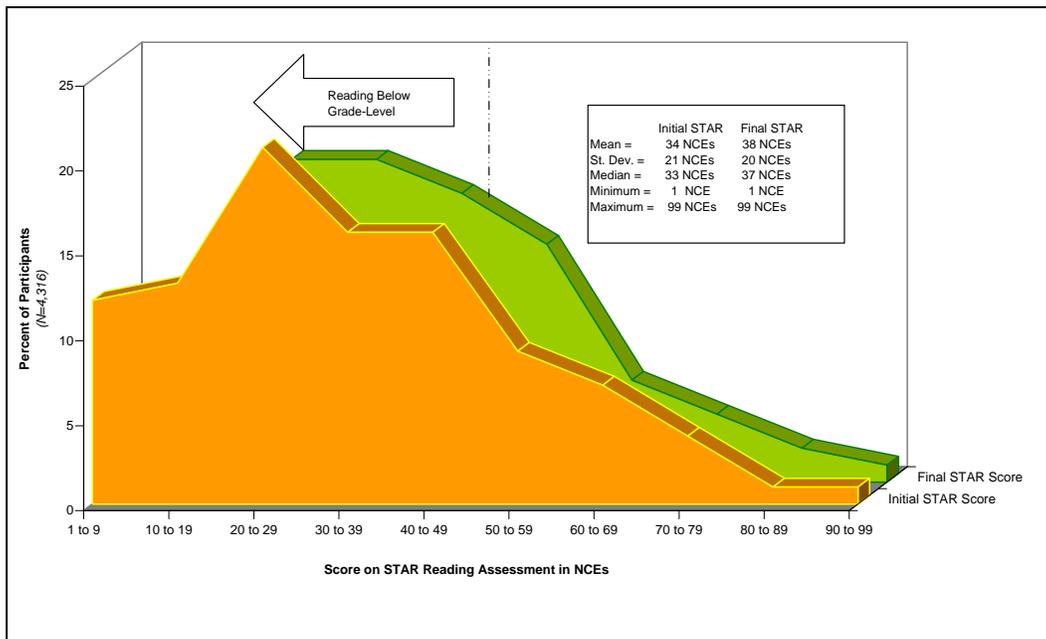


Exhibit reads: Twelve percent of participants with at least two STAR Reading Assessments at least 90 days apart scored between 1 and 9 NCEs on their first STAR Reading Assessment, decreasing to 7 percent on their final STAR.

The proportion of participants reading at a level appropriate for their grade level or above increased during the 2005-06 school year. On the initial STAR Reading Assessment, 21.8 percent of participants were reading at grade level or higher (50 NCEs or more). By the time the final STAR Reading Assessment was administered, 27.3 percent read at grade level.

Among participants in the literacy program during the 2005-06 school year, the change in STAR Reading Assessment scores ranged from a decrease of 34 NCEs to an increase of 43 NCEs. The average change in scores was a gain of 3.9 NCEs, as shown in Exhibit 12. The change was statistically significant.

Exhibit 12
Change in Score on STAR Reading Assessment, in NCEs, 2005-06,
for Children Who Completed Two Assessments at Least 90 Days Apart

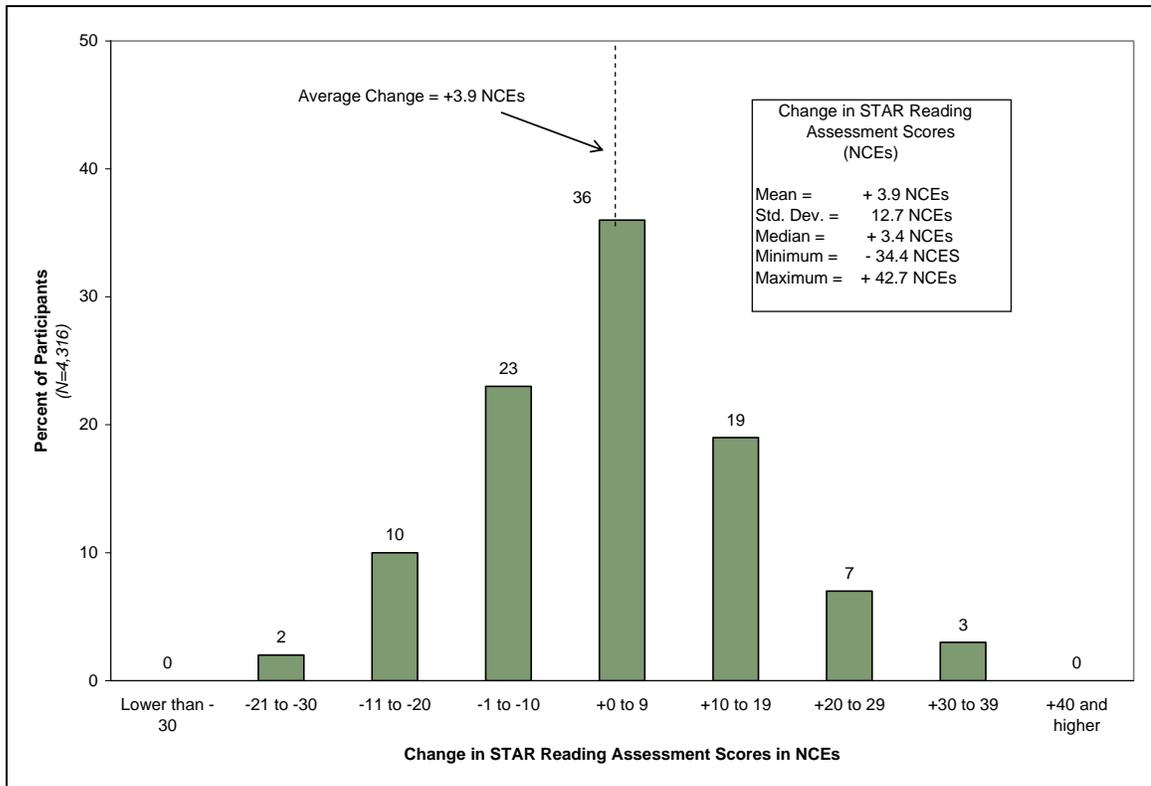


Exhibit reads: Less than 1 percent of participants experienced a decrease of 30 NCEs or more.

An issue in the analysis of change in scores on the STAR Reading Assessment is determining how large a change must occur before it is considered to represent a meaningful gain in reading proficiency. For the STAR Reading Assessment, the minimum meaningful gain is 2.0 NCEs, according to the test publisher. Among all of the children participating in the model literacy programs who completed two STAR Reading Assessments at least 90 days apart, 54 percent achieved a gain of 2.0 NCEs or greater, an indication of a meaningful gain in reading proficiency.

The average change in reading proficiency achieved by participants differed among the 46 programs. The average gain on the STAR was 10 NCEs or more at five programs, while performance decreased (by any amount) on average in six programs.

At four sites, more than 80 percent of the participants gained 2.0 NCEs or more, while at five sites, fewer than 40 percent gained at least 2.0 NCEs.

Children enrolled in grades 1-4 achieved larger gains on the STAR Reading Assessment than did children in grades 5-8, as shown in Exhibit 13.

Exhibit 13
Changes in Performance on the STAR Reading Assessment,
in NCEs, by Grade, 2005-06

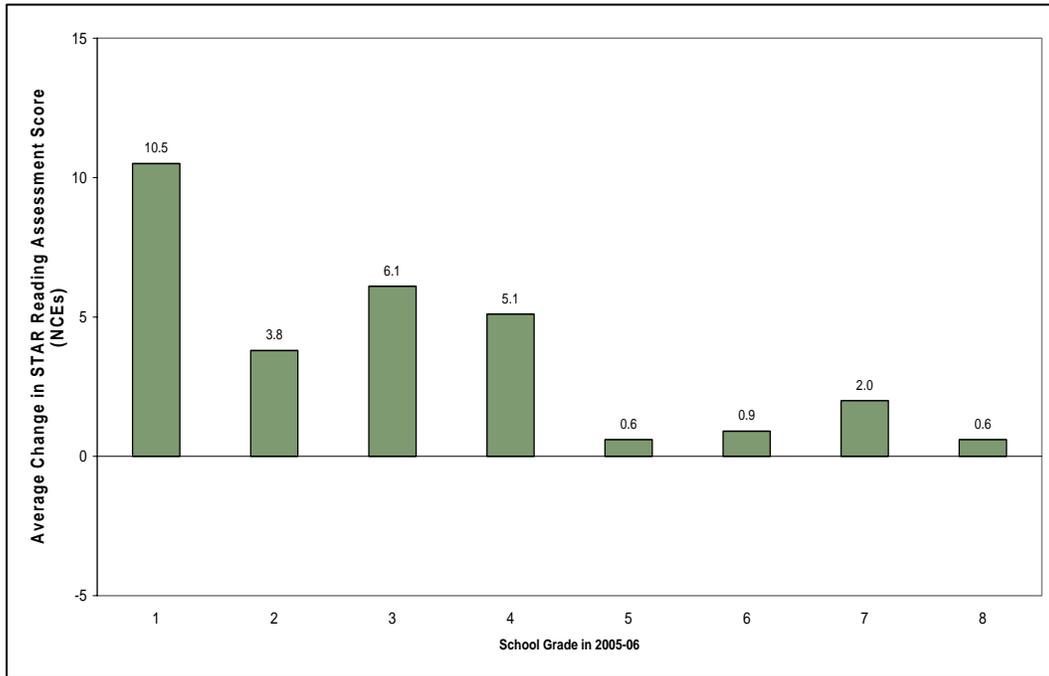


Exhibit reads: Children enrolled in first grade averaged a gain of 10.5 NCEs on their STAR Reading Assessments.

Similarly, as shown in Exhibit 14, more than 50 percent of the participants in grades 1 to 4 achieved gains of 2 NCEs or greater, while fewer than 50 percent of participants in other grades gained that amount.

Exhibit 14
Percent of Participants Achieving Gains of Two NCEs or More,
by Grade, 2005-06

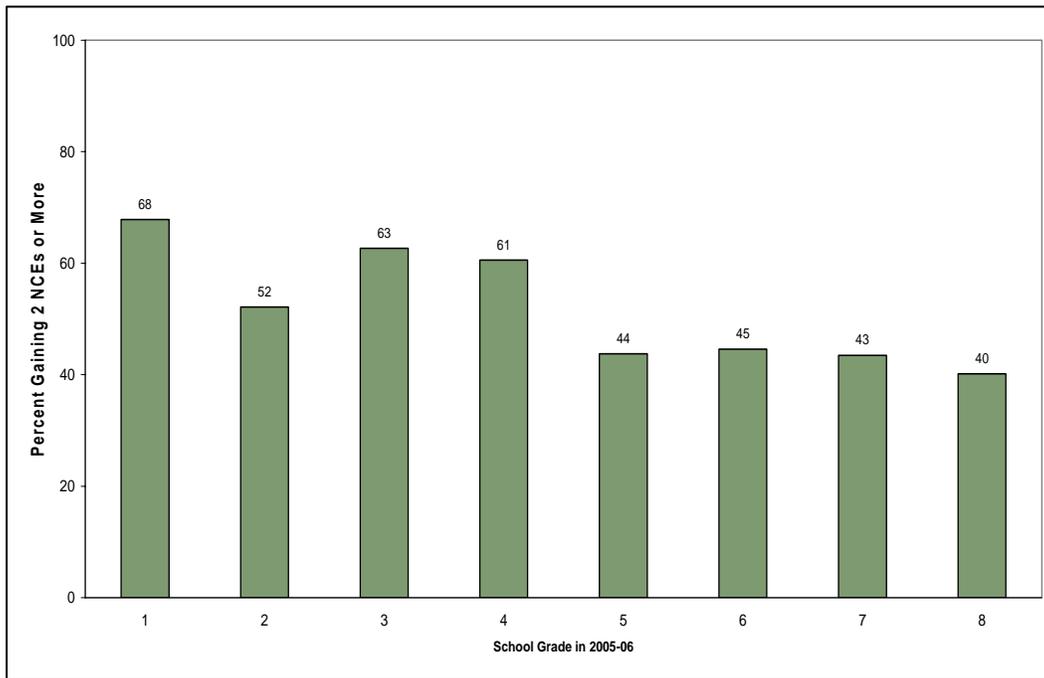


Exhibit reads: Sixty-eight percent of participants enrolled in first grade achieved gains of 2 NCEs or greater on their STAR Reading Assessments.

The publishers of the STAR Reading Assessment compute the results of each child's performance in a variety of forms, including NCEs, scale scores, percentiles, and Instructional Reading Level. The Instructional Reading Level (IRL) is an estimate of the most appropriate grade level of reading material that a child should use for optimal learning. The IRL is the level at which a child can recognize words and comprehend instructions, with some assistance from a teacher. Scores range from 1.0, indicating the first month of first grade (September), through 12.9, indicating the ninth month of twelfth grade (June), plus additional scores indicating pre-school and post-high school performance. A child with an IRL of 4.0 will do best with materials at the fourth-grade level, regardless of the grade in which the child is currently enrolled.

During 2005-06, the children participating in model literacy programs averaged a gain of 1.0 IRLs, the equivalent of one grade level (Exhibit 15). Forty-four percent of children gained more than one grade level in instructional reading level during one year of program participation, and 19 percent more than two grade levels.

Exhibit 15
Change in Score on STAR Reading Assessment, in IRLs, 2005-06,
for Children Who Completed Two Assessments at Least 90 Days Apart

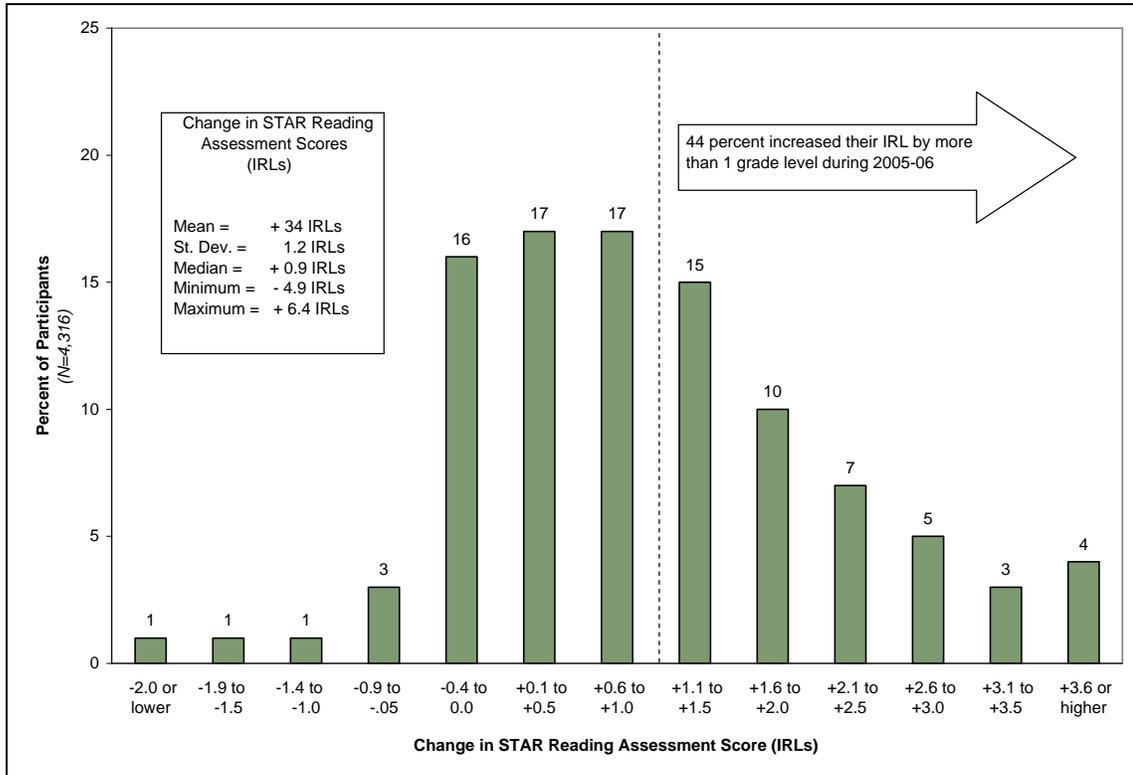


Exhibit reads: Less than 1 percent of participants experienced a decrease of 2 IRLs or more.

5. What Factors Were Associated with Changes in Reading Proficiency?

To determine whether some subgroups of children benefited more than others from participation in the program, evaluators explored the relationship between change in reading proficiency and participant characteristics, patterns of program attendance, and details of children's participation in AR activities. This information permitted the evaluators to assemble information capable of predicting likely patterns in participants' final STAR scores.

Participant Characteristics

Initial reading proficiency. The factor that was most strongly related to a participant's final score on the STAR Reading Assessment was that child's initial score. The overall pattern is that children with high scores on their initial assessment had high scores on their final assessment, and children with low initial scores had low final scores. The correlation coefficient

is + .80 and is statistically significant. A correlation coefficient of this magnitude is generally described as indicating a strong relationship.

The correlation between a participant’s initial STAR Reading Assessment score and the change in score during 2005-06 was -0.36, a statistically significant relationship. A negative correlation indicates that the children who scored lowest on their initial STAR Reading Assessment achieved larger gains than children who had higher initial scores, as shown in Exhibit 16. The differences among the different initial performance groups were statistically significant.

Exhibit 16
Change in Score on STAR Reading Assessment,
by Performance on Initial Assessment, 2005-06

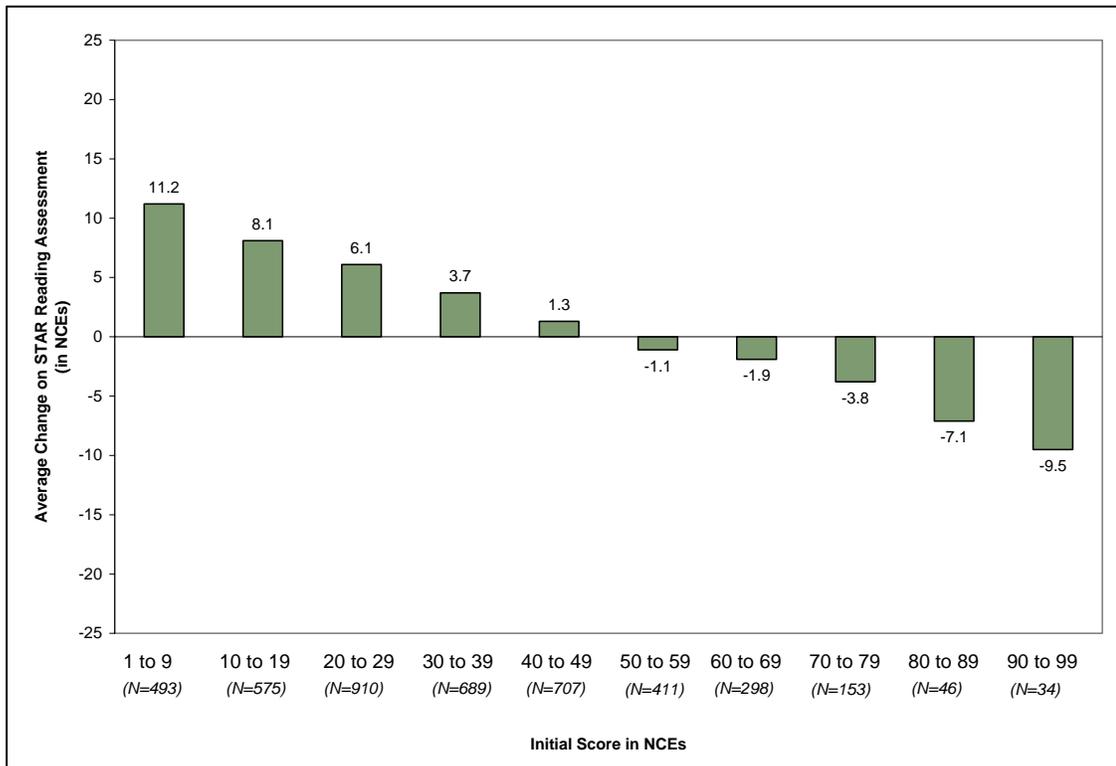


Exhibit reads: Children with initial scores between 1 and 9 NCEs averaged a gain of 11.2 NCEs between their initial and final STAR Reading Assessments.

The proportion of participants who gained 2 NCEs or more between their initial and final STAR Reading Assessments also varied by their level of performance on the initial assessment. A larger proportion of participants who scored below 30 NCEs on the initial assessment gained 2 NCEs or more, compared with participants in the other initial performance groups, as shown in Exhibit 17.

Exhibit 17
Proportion of Participants Gaining Two NCEs or More
Between Initial and Final STAR Reading Assessments,
by Performance on Initial Assessment, 2005-06

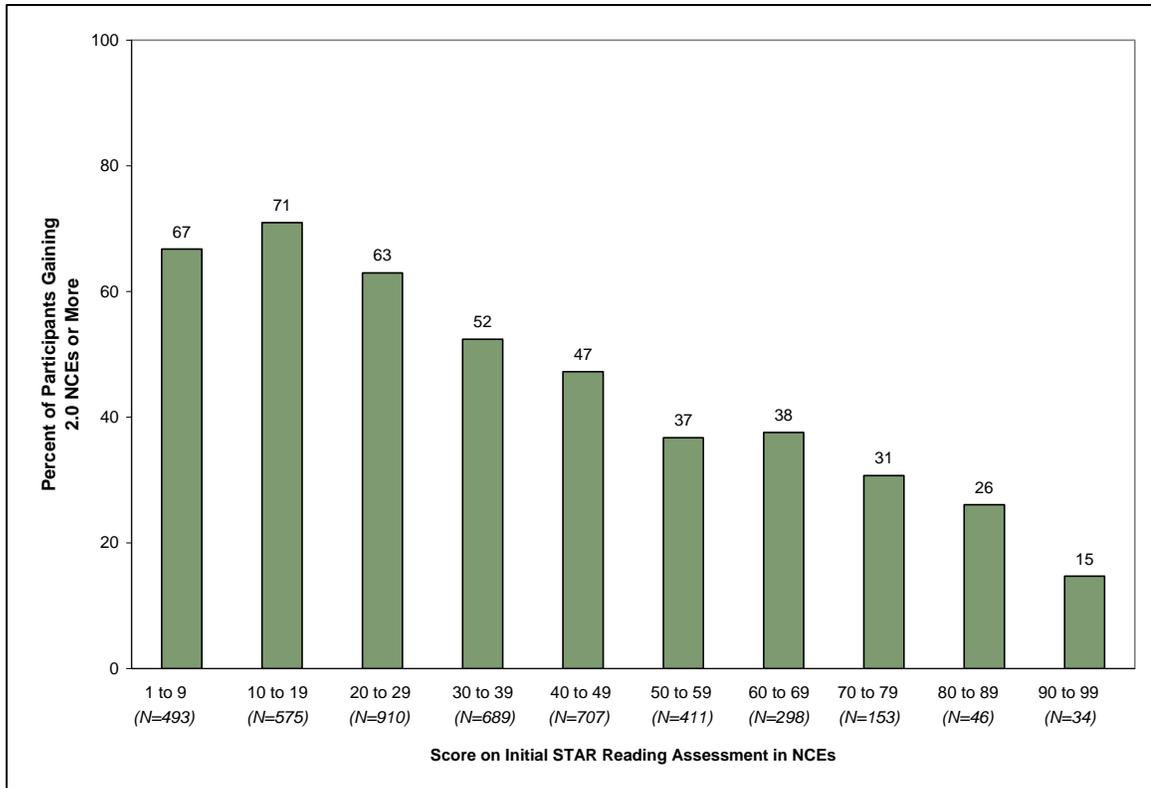


Exhibit reads: Sixty-seven percent of participants who scored below 10 NCEs on their initial STAR Reading Assessment gained 2.0 NCEs or more on their final STAR Reading Assessment.

Not surprisingly, the proportion of participants performing at or above grade level on their final STAR Reading Assessment was smallest among children who scored below 10 NCEs on their initial assessment, reflecting the magnitude of the increase in reading proficiency required for these children to reach grade level (Exhibit 18).

Exhibit 18
Proportion of Participants Reading at Grade Level
on Their Final STAR Reading Assessment,
by Performance on Initial Assessment, 2005-06

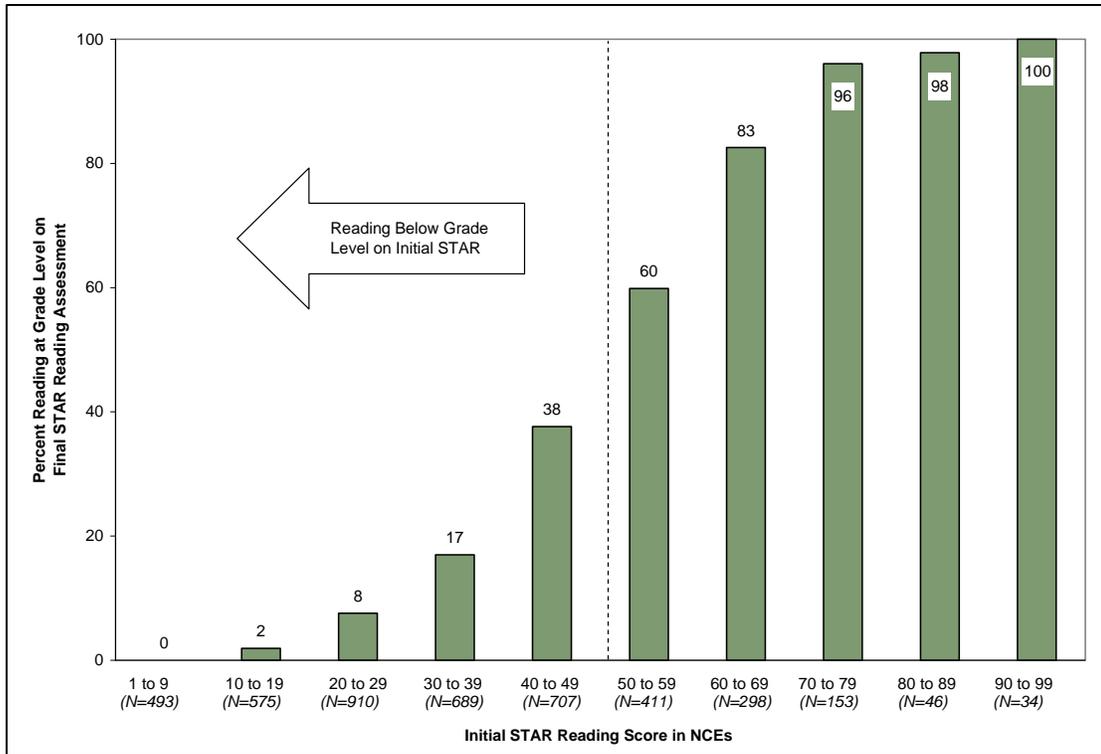


Exhibit reads: Zero percent of the participants scoring below 10 NCEs on their initial STAR Reading Assessment read at grade level (50 NCEs or higher) on their final STAR Reading Assessment.

Sponsorship status. Sponsored children participating in the model literacy programs averaged gains of 3.2 NCEs between the initial and final STAR Reading Assessments, compared with an average gain of 4.5 NCEs among other participants, a statistically significant difference. Fifty-one percent of sponsored children achieved a gain of 2.0 NCEs or larger, compared with 57 percent of other children, also a statistically significant difference.

Patterns of Program Attendance

Days of program attendance. Overall, the correlation between the number of days attended and gains in reading proficiency was +0.06, a small but statistically significant relationship, showing that there is a slight tendency for children who attended more days to achieve higher gains in proficiency. The correlation between the *percent* of days attended and gains in reading proficiency was +0.06, also a small and statistically significant relationship. Taking into account each participant's initial score on the STAR Reading Assessment, the relationship between program attendance and participants' final scores remained small but

statistically significant. The partial correlation between the days attended and the final STAR Reading Assessment score was +0.07, and between the percent of days attended and the final STAR the correlation was +0.07, both statistically significant correlations.

Among the children at the 46 model literacy programs who attended 75 percent or more of the days it was possible for them to attend, the average gain on the STAR Reading Assessment was 4.6 NCEs compared with 3.6 NCEs among other participants, a statistically significant difference.

There was a relationship between attending the program for 30 days or more during the 2005-06 school year and achievement gain. Among children who attended 30 days or more, the average gain in reading proficiency on the STAR Reading Assessment was 4.4 NCEs, while among other children the average gain was 3.0 NCEs, a statistically significant difference. Fifty-six percent of the children who attended the program 30 days or more achieved a gain of 2 NCEs or more, compared with 51 percent of other children, a statistically significant difference.

Years of program participation. Children who participated in the model literacy program for three years averaged higher initial 2005-06 STAR Reading Assessment scores than did the students who had participated for a shorter time, but they achieved smaller gains during 2005-06 (Exhibit 19). On average, children in their third year of participation averaged 39 NCEs on the first STAR Reading Assessment they completed during 2005-06, while children in their second year averaged 36 NCEs, and children in their first year averaged 33 NCEs. The differences among the groups were statistically significant.

Exhibit 19
Score on the Initial STAR Reading Assessment, 2005-06,
by Years of Participation

Score on STAR Reading Assessment, 2005-06	Years of Participation		
	One, Beginning in 2005-06 (N=3,042)	Two, Beginning in 2004-05 (N=729)	Three, Beginning in 2003-04 (N=529)
Initial STAR Reading Assessment score, NCEs	33	36	39
Final STAR Reading Assessment, NCEs	37	41	41
Average change in STAR scores, NCEs	4.1	4.4	2.3
Percent of children gaining 2 NCEs or more	55	55	47

Exhibit reads: The children who first participated during 2005-06 scored an average 33 NCEs on their first STAR Reading Assessment during 2005-06.

Eighty percent of the children who began participating in a model literacy program in 2005-06 performed below grade level (below 50 NCEs) on their initial STAR Reading assessment during 2005-06, compared with 75 percent of the children in their second year of

participation, and 72 percent among those participating for a third year (Exhibit 20). The differences were statistically significant.

Among the children in their first year of participation, 5.5 percent fewer scored below grade level on their final STAR Reading Assessment of 2005-06 than they did on their initial assessment. Among children in their second year of participation, 7.3 percent fewer scored below grade level on their final assessment, and 3.0 percent fewer of the children in their third year of participation did so.

Exhibit 20
Proportion of Children Performing Below Grade Level, 2005-06,
by Years of Participation

Performance Below Grade Level, 2005-06	Years of Participation		
	One, Beginning in 2005-06 (N=3,042)	Two, Beginning in 2004-05 (N=729)	Three, Beginning in 2003-04 (N=529)
Initial STAR Reading Assessment	80	75	72
Final STAR Reading Assessment	75	68	69
Change in proportion scoring below grade level	-5.5	-7.3	-3.0

Exhibit reads: Eighty percent of the children who first participated during 2005-06 scored below grade level (below 50 NCEs) on their initial STAR Reading Assessment in 2005-06.

Trends in Literacy Proficiency Among Three-Year Participants

The procedure used to convert participants’ performance on the STAR Reading Assessment into NCEs uses a different conversion factor for each grade level, making the NCE scores inappropriate for comparing performance across school years. However, test results in the form of scale scores can be compared across school years. The scale-score measure is a continuous measure of reading proficiency that is not adjusted for a child’s grade in school or for the month of the school year in which the STAR Reading Assessment is taken.

The children who participated for three years at one of the 12 model literacy programs that operated in 2003-04, 2004-05, and 2005-06, achieved a steady increase in reading proficiency throughout the administrations of the STAR Reading Assessment, as shown in Exhibit 21. Between their first score in 2003 and their final score in 2006, these children averaged a gain of 223 scale-score points.

Exhibit 21
Progression in Reading Proficiency Over Time among
Children Participating in All Three Years

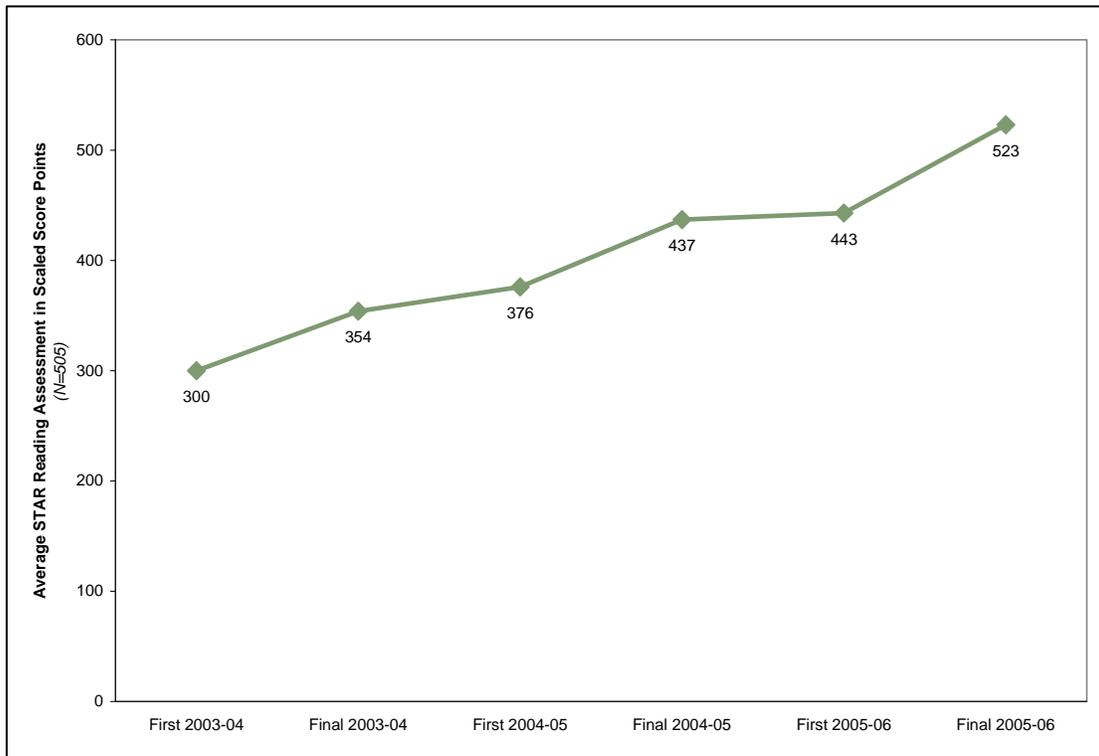


Exhibit reads: Among the children who began participation in 2003-04 and continued through 2005-06, the average score on the first STAR Reading Assessment was 300 scale-score points.

Regardless of the year in which children began participation in the model literacy program, those who participated for more than one year achieved approximately the same gain in reading proficiency during each year of participation (Exhibit 22).

Exhibit 22
Cumulative Gains in Scores on STAR Reading Assessment,
by Years of Participation

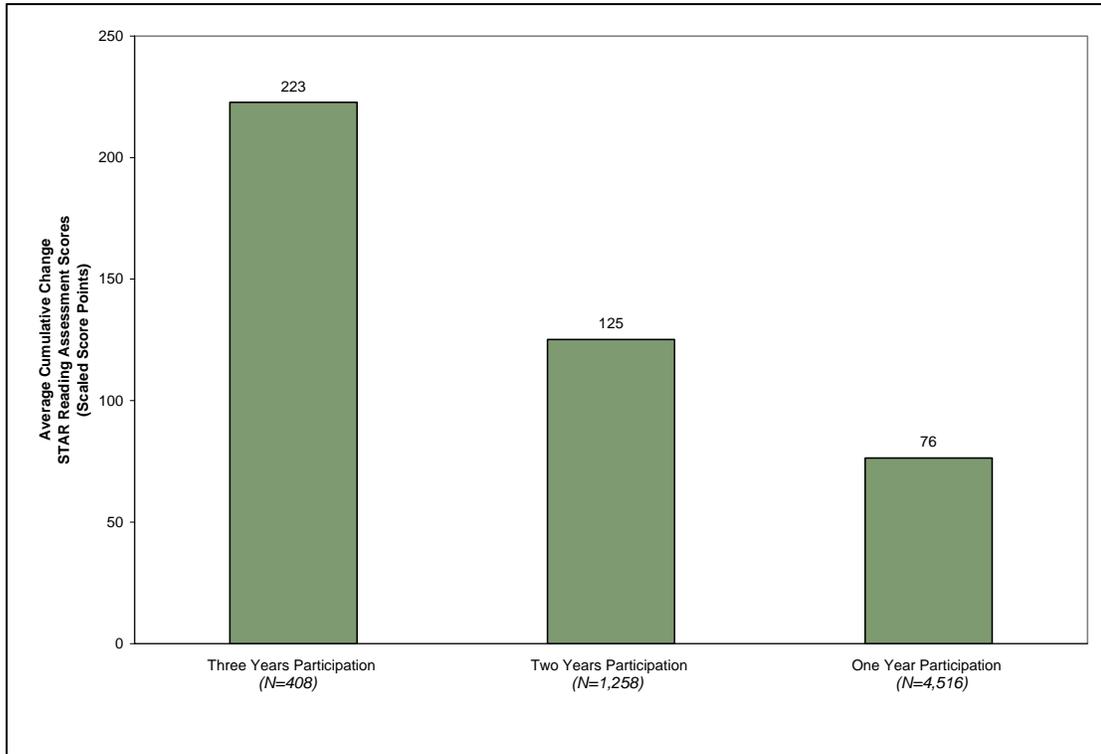


Exhibit reads: Among the children who participated in the model literacy program for three years, the sum of the gains in reading proficiency achieved each year averaged 223 scale-score points, nearly three times the average gain achieved by children participating in the program for one year.

A disadvantage of assessing performance in terms of scale-score points is that there is no established method for estimating whether a child is performing at grade level, or whether the gains achieved between administrations of the STAR Reading Assessment are greater than would be expected from an additional period of schooling and maturation alone.

One approach to interpreting gains achieved by model literacy program participants is to estimate the change in the proportion of participants who performed at grade level or above for each year of participation. Using this method, analyses found an increase in the proportion of children performing at grade level for each year of participation. As shown in Exhibit 23, among the children who first participated during 2005-06, 5.5 percent more children performed at grade level or above on their final STAR Reading Assessment than on their first assessment. Among the children who first participated in 2003-04 and continued to participate for three years, 2.4 percent more performed at grade level or above during their first year of participation, 5.5 percent during their second year, and 3.0 percent at the end of their third year.

Exhibit 23
Within-Year Change in the Proportion of Model Literacy Program Participants Performing Above Grade Level, by Duration of Participation

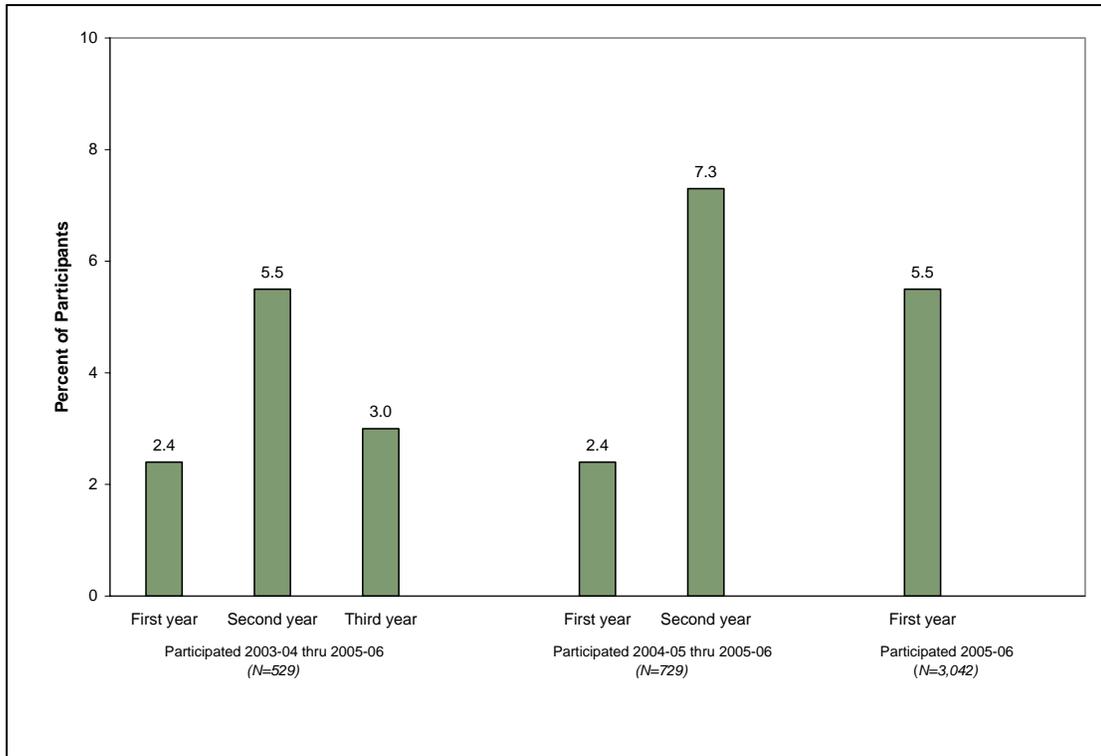


Exhibit reads: Among the children participating in the Model Literacy Program for three years, 2.4 percent more performed at or above grade level on the final STAR Reading Assessment than on their initial STAR Reading Assessment during their first year of participation.

Another approach for interpreting the importance of the change in participants' STAR Reading Assessment scores is to examine the change in the level of instruction appropriate for each child, the IRL. Among the children who first participated during 2005-06, the gain in performance on the STAR Reading Assessment indicated that average change in the level of reading materials appropriate for participants increased by 1.02 IRLs, approximately one grade level (Exhibit 24). Among the children who first participated in 2003-04 and continued to participate for three years, the average increase after their first year of participation was 0.70 IRLs, 0.88 IRLs during their second year, and 0.95 during their third year.

Exhibit 24
**Within-Year Change in Average Instructional Reading Level,
 by Duration of Participation**

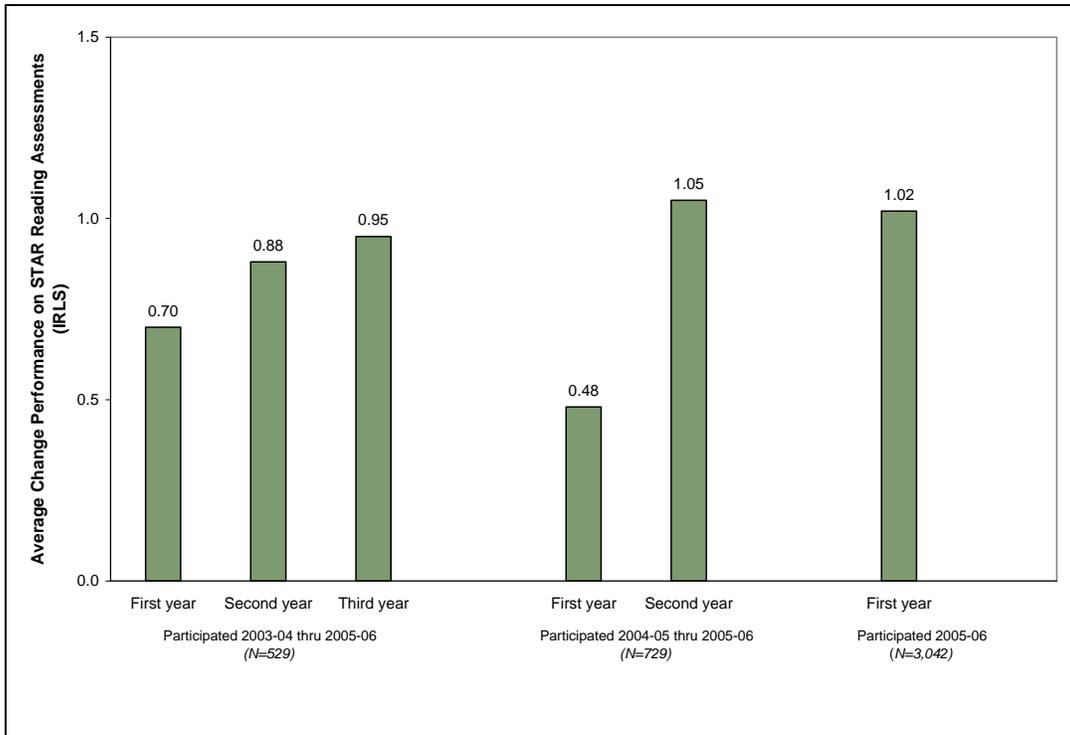


Exhibit reads: Among the children participating in the Model Literacy Program for three years, the average gain in IRL after one year was 0.70 IRLs, slightly less than one school year.

Participation in AR Activities

Among participants in the 46 model literacy programs for which STAR Reading Assessment results are available, there was a small but substantial relationship between the number of books each child read and the gain in reading proficiency achieved. The partial correlation (controlling for each participant’s initial STAR score) between the number of books read during the 2005-06 school year and the change in reading proficiency on the STAR Reading Assessment was +0.10, a statistically significant relationship.

Among participants who read 25 books or more during the school year, the average gain on the STAR Reading Assessment was 4.4 NCEs, while among children who read fewer books, the average gain was 2.2 NCEs. The difference was statistically significant.

Evaluators also found a substantial and positive relationship between the percent of AR quizzes each participant passed and their gains on the STAR Reading Assessment. The partial correlation, controlling for each child’s initial level of reading proficiency, was +0.18, a statistically significant and substantial relationship.

Further, there was a substantial and positive relationship between the average percent of items answered correctly on AR quizzes and participants' gains on the STAR Reading Assessment. The partial correlation, controlling for each child's initial level of reading proficiency, was +0.21, a statistically significant relationship.

Seventy percent of participants passed 85 percent or more of the AR quizzes they attempted. Those who passed 85 percent or more of the AR quizzes they attempted averaged a gain of 4.6 NCEs on their STAR Reading Assessments, while those who passed a lower proportion of their quizzes averaged 2.6 NCEs, a statistically significant difference.

Prediction of Final STAR Reading Assessment Scores

A multivariate linear regression procedure allowed evaluators to estimate the relationship among multiple dimensions of participation in the model literacy programs, participants' characteristics, and performance on the STAR Reading Assessment. The statistical procedure explores the relationship between multiple independent, or predictor, variables and a dependent, or criterion, variable. Regression describes the relationship between the variables in the model by calculating a single, best-fitting straight line that can be expressed as an equation.

The elements of the equation, a constant that indicates where the line crosses the Y axis, and the betas, which describe the slope of the line, can be combined in a linear way to show each child's predicted performance on their final STAR Reading Assessment. The beta coefficient for each independent variable in the regression equation provides an estimate of the contribution of each of the independent variables in the model, holding constant the impact of each of the other elements of the model.

Among all participants in the 46 programs, the model that best predicted a child's score on the child's final STAR Reading Assessment involved the child's initial STAR score, the percent of quizzes passed, and the number of books read. The equation for the model is shown in the box below

Final STAR Scale-Score Points =
+ 0.17 (the constant)
+ Initial STAR NCEs * .75
+ Percent of AR Quizzes Passed * .13
+ Number of Books Read * .02

A measure of how well this regression model predicted participants' final STAR score is the model's R^2 . For this model, the R^2 is .66 and is statistically significant. This indicates that the model explains 66 percent of the variance in a child's final scale score on the STAR Reading

Assessment. An R^2 of this magnitude indicates that the model is a moderately accurate tool for estimating a participant's final STAR score.⁴

The model demonstrates that there is a statistically significant relationship between level of participation and level of increase in a child's reading proficiency. The model is useful in estimating the potential benefit, in terms of increases in test scores, of specific components of program participation.

The model indicates that:

- The best predictor of a child's final STAR scale score is his or her first scale score. For every additional STAR Reading Assessment, the final scale score can be expected to increase by 0.75 NCEs.
- For every increase of 1 percentage point in the percent of AR quizzes passed, the final STAR score can be expected to increase by 0.13 NCEs.
- For every additional book read, the final STAR score can be expected to increase by 0.02 NCEs.

⁴ The coefficients of the multiple regression coefficients for participants are:

Coefficient	Beta	Standard Error of Beta	Standardized Beta
Constant	0.17	1.00	
Score on First STAR Reading Assessment (in NCEs)	0.75	0.09	0.77 *
Percent of AR Quizzes Passed	0.13	0.01	0.10 *
Number of Books Read	0.02	0.00	0.04 *

* The Standardized Beta is statistically significant at $p < 0.05$.

Appendix
Using Normal Curve Equivalent Scores

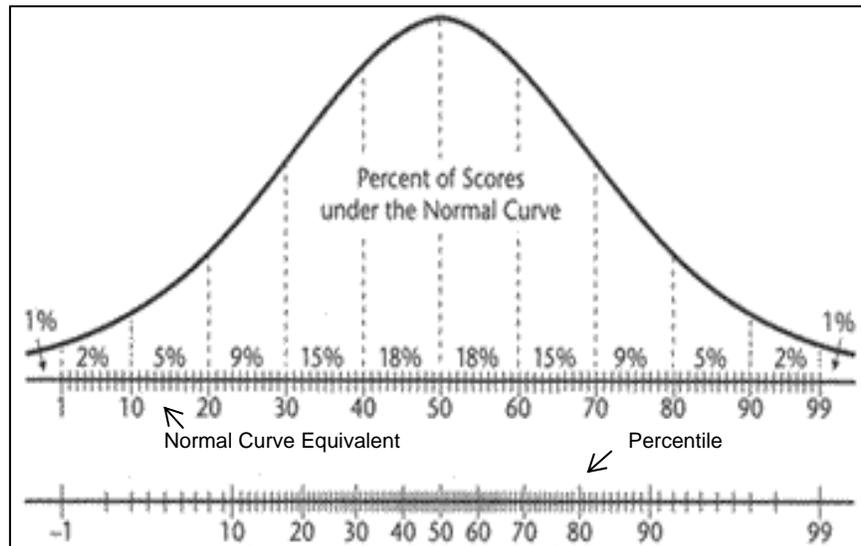
Using Normal Curve Equivalent Scores

A normal curve equivalent (NCE) score is a standardized score (based on a normal distribution) that makes it possible to compare scores across grades and to compute statistics that describe the performance of a group of children, such as average scores, that are not possible using other ways of expressing test scores. NCEs consist of 99 equal units, an advantage over percentiles and grade equivalents in analyzing changes in test scores. Because NCE scores are equal units, they can be averaged for a group of children. In addition, an average score for a group of children can be computed as can the average of the change in scores achieved by groups of children.

NCEs are computed by dividing the normal curve distribution of student scores on an assessment into 99 units, with each NCE unit spanning the same number of test points. NCE scores have a mean of 50 and a standard deviation of 21.06. In a normal distribution, an NCE of 1.0 is equivalent to a percentile rank of 1.0, and an NCE of 99.0 is equivalent to a percentile rank of 99.0.

Percentile ranks, as can be seen in Exhibit A1, tend to cluster around the middle of the distribution of test scores, so that a change in performance of one percentile represents a smaller change in the raw score on an assessment at the center of the distribution than at the extremes of the distribution. Because NCEs comprise 99 units of equal size, a change of one NCE represents the same change in the raw score at any point in the distribution.

Exhibit A1
Distribution of Student Scores, Normal Curve Equivalent Scores,
and Percentiles



Students performing at grade level will have an NCE of 50, regardless of their grade in school. If a student makes exactly one year of progress after one year of instruction, his/her NCE score will be exactly the same both years, and the change in NCE scores will be zero. A student with a one-year change in NCE scores that is greater than zero has achieved larger gains than the general population, and a student with a negative change in NCE scores has made less progress.