THE LITERACY PROGRAMS OF SAVE THE CHILDREN
Results from the 2008-09 School Year

Andrea S. Palmiter
Erikson R. Arcaira
Richard N. White
Elizabeth R. Reisner

Policy Studies Associates, Inc.

September 2009

Prepared for Save the Children, U.S. Programs

Contents

THE LITERACY PROGRAMS OF SAVE THE CHILDREN
Results from the 2008-09 School Year

Overview and Summary

Since the 2003-04 school year, Save the Children, U.S. Programs, has supported programming designed to improve reading skills among struggling students in rural areas. The literacy program includes both afterschool and in-school interventions with small groups of children. These are implemented by well-trained paraprofessionals and are designed to provide a range of activities aimed at increasing literacy achievement.

Central to Save the Children's literacy program is the guided independent reading program (GIRP). It provides regular opportunities to read "just right" books independently, under the guidance of a trained adult, providing children with increased motivation for reading and vocabulary growth. GIRP uses Accelerated Reader, a reading management software program by Renaissance Learning, to monitor guided reading practice. Change in reading proficiency is assessed using the STAR Reading assessment.

In the afterschool setting, daily literacy activities occur within an hour-long literacy block. Primarily for children in grades 2-6, these activities include 30 minutes of guided independent reading practice, 20 minutes of fluency-building support, and a 10-minute read-aloud. The 20-minute fluency building component stresses repeated reading in a supportive environment. Text is practiced aloud so that students pay attention to the phrasing, punctuation, and tone of the text. Unfamiliar vocabulary words are heard and repeated daily in context, thus increasing a child's likelihood of success in learning them. All of these activities are intended to work in concert to increase the achievement of struggling readers. GIRP is used for both monitoring and evaluating program outcomes.

For children in kindergarten and first grade who are having difficulty learning how to read, literacy activities that target beginning reading skills are available in the afterschool Emergent Reader Literacy Block, as offered by Save the Children. This includes a 30-minute extended read-aloud with developmentally appropriate follow-up activities, a 15-minute reading together period when easy text is practiced daily over a week's time, and 15 minutes of hands-on learning to support increased growth in phonemic awareness, letter recognition, sound-symbol correspondence, and/or beginning sight words. When assessments show a student has successfully mastered the material targeted in the Emergent Reader Literacy Block, the student moves on to the guided independent reader program. As a result, some kindergarten and many first-grade students participate in guided independent reading and take STAR Reading assessments.

In addition, some schools choose to provide children participating in GIRP with an in-school program that provides small-group tutorials that target phonics, sight word, vocabulary, and comprehension growth in addition to the afterschool Literacy Block.

During the 2008-09 school year, Save the Children selected 118 sites in rural communities to implement its reading initiative. Twenty of the 1`8 sites continued to offer literacy programming begun in 2003-04 or 2004-05, fifty-one began operating during 2005-2006 or 2006-07, 19 during 2007-08, and 26 site2 operated for the first time in 2008-09. Services at the 118 local sites included the delivery of integrated in-school and afterschool literacy activities for children in kindergarten through sixth grade.

Save the Children's assistance includes the following:

- Funding for literacy learning materials, including Accelerated Reader (AR), Emergent Reader (ER) resources, small-group reading tutorial resources, library resources, computer equipment, and standardized testing (the STAR Reading assessment and the STAR Early Literacy assessment) to permit periodic assessment of literacy proficiency

- Training of local staff and volunteers, including 23 hours of training and 37 hours of technical assistance per site, on average; 12 sites received, on average, an additional four hours of technical assistance, from Renaissance Learning

- 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 growth

- Feedback and training based on program data collected

Key findings from this analysis include the following:

- Across the 118 sites, approximately 12,001 children in grades K-6 participated in one or both reading interventions: the guided independent reading program and Emergent Reader.

- Most of the participating children demonstrated reading skills significantly below grade level at the start of the school year. On average, 84 percent of participants scored below grade level on the first STAR Reading assessment administered during 2008-09.

- On average, participating children were engaged in at least one literacy program on 67 days over the 2008-09 school year. While participating in the program, children attended an average of 73 percent of the days the program was offered.

- Measures of engagement with the program's literacy activities indicate that participants read 64 books during 2008-09 on average, and the median number of books read was 48.

- Overall, participants passed 92 percent of the AR quizzes they completed about the books they had read.

- Children involved in GIRP improved their literacy skills substantially over the course of the program period, moving closer to grade-level performance. The average literacy gain was 5.8 normal curve equivalents (NCEs).[1]

- Sixty percent of participants increased their literacy performance by at least 2 NCEs, which is considered to be a meaningful learning gain by Renaissance Learning, the publisher of the STAR Reading assessment.

- Sixty-three percent of children who attended programming for 55 days or more and took two assessments 90 days apart gained 2 NCEs.

- Sixty-six percent of children reading below grade level on the initial STAR Reading assessment who attended programming for 55 days or more and took two assessments 90 days apart gained 2 NCEs.

- The proportion of participants reading at a level appropriate for their grade or above increased during the 2008-09 school year. On the initial STAR Reading assessment, 16 percent of participants were reading at grade level or higher (50 NCEs or more). By the time the final STAR Reading assessment was administered, 29 percent read at grade level or higher.

- The level of children's participation in guided independent reading activities was clearly associated with the level of gains in reading proficiency. The correlation between number of days attended and gain in scores was +0.12, and was a statistically significant relationship. There was also a significant correlation between the number of books read and gains in scores on the STAR Reading assessment (+0.07) and between the proportion of Accelerated Reader quizzes passed and gains in scores (+0.17). In addition, 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 Appendix A.

- Among the participants in the Emergent Reader program (serving primarily students in kindergarten and first grade) who completed a STAR Early Literacy (SEL) assessment in fall 2008 and spring 2009, the proportion scoring in the transitional or probable reading classification levels increased by 48 percentage points during the school year, from 11 percent in the fall to 59 percent in the spring. Students in these reading classification levels demonstrate sufficient reading proficiency to be considered prepared for beginning reading instruction.

- Emergent Reader participants achieved substantial growth on all seven literacy domains tested by the STAR Early Literacy assessment. Of the seven literacy domains, participants demonstrated the largest growth in structural analysis, phonemic awareness, and comprehension.

1. What Opportunities Were Provided for Targeted Children to Participate in Literacy Instruction and for Staff to Develop Instructional Skills?

The Save the Children literacy programs provide support to partner schools located in areas of high rural poverty to serve struggling readers in grades K-6. The 118 literacy programs are located in 12 states. Thirty-two are located in Kentucky, 14 in Tennessee, 13 each in Mississippi and New Mexico, 11 in South Carolina, 10 in California, 6 each in Arkansas and Arizona, 5 each in Louisiana and Nevada, 2 in Colorado, and 1 in Alabama.

In most instances, a literacy program was affiliated with a host school located in a rural school district. The enrollment of the host schools ranged from under 100 children to more than 800. The majority of the schools served kindergarten through eighth grade, but several served only two or three grade levels.

Services Provided

Each of the 118 sites offered literacy support programming either afterschool or during the school day. All but four sites offered an hour-long literacy block in the afterschool program. The literacy block consisted of 30 minutes of guided independent reading practice using the Accelerated Reader program, 20 minutes of fluency-building activities, and a 10-minute read-aloud. Ninety-three sites offered guided independent reading practice using Accelerated Reader materials during the school day. Sixty sites offered Emergent Reader support, targeting the early literacy needs of children in kindergarten and first grade, either afterschool or in school, or both. In addition, some schools provided an optional tutoring component to participating students.

Description of Services

Guided Independent Reading Practice – 30 minutes daily using Accelerated Reader books and software that monitors progress

Emergent Reader Support – Individual or small-group reinforcement for struggling readers in grades K-1 to accelerate beginning reading skills such as phonemic awareness, letter recognition, sound-symbol correspondence, and sight word growth

Frequency of Services Provided

Among the different literacy services, the most frequently offered in the afterschool program was guided independent reading practice, coupled with fluency-building activities and read-alouds, provided on an average of 109 days among the 114 sites that provided this service. Guided independent reading practice was offered in school on an average of 93 days in 95 sites. As a supplement to GIRP, some schools chose to offer supplemental tutoring services during and after school hours.

Emergent Reader support, an optional program component, was offered afterschool by 40 sites on an average of 92 days per site, and the program was offered during school by 52 sites on an average of 86 days per site.

Program Staff Training

Save the Children and Renaissance Learning provided training, technical assistance, and ongoing support in literacy services for the staff of the literacy programs. On average, staff and volunteers involved in literacy instruction at each program received 23 hours of training and 37 hours of technical assistance from Save the Children, for a total of 60 hours. Staff in 12 of the sites also received an average of four hours of technical support and coaching from Renaissance Learning specialists.

2. What Were the Enrollment Patterns in the Literacy Programs?

The total number of individual participants enrolled in the 118 programs over the course of the 2008-09 school year was 12,001.[2] The number of participants at individual sites ranged from 34 to 250 children. The average number of participants per program overall was 102.

More than half of the participants were enrolled in first through third grade, as shown in Exhibit 1.

[2] A participant is defined as a child who attended a literacy activity on at least one day between August 1, 2008, and June 30, 2009.

Exhibit 1
Grade Level of Participants

School Grade	Percent of Participants *(N=12,001)*
K	8
1	11
2	18
3	21
4	18
5	15
6	9

Exhibit reads: Eight percent of literacy program participants during 2008-09 were in first grade.

The monthly enrollment in the literacy programs ranged from a low of 638 in June 2009 to a high of 9,323 in March 2009, and averaged 7,062 children per month, as shown in Exhibit 2.

Exhibit 2
Number of Participants, by Month

Month	Total Enrollment
August-08	2,702
September-08	7,083
October-08	7,886
November-08	8,041
December-08	7,831
January-09	8,811
February-09	9,118
March-09	9,323
April-09	9,029
May-09	7,224
June-09	638
Monthly Average	7,062

Exhibit reads: In August 2008, a total of 2,702 children participated across all of the 118 literacy programs.

3. Did Children Participate in Available Literacy Learning Opportunities?

The following discussion examines attendance at the literacy programs and the level of participants' engagement in literacy learning activities.

Measures of Participation in Literacy Programming

The analyses conducted for this evaluation are limited to children who attended the program at least one day during the 2008-09 school year. A child's participation in any combination of program activities on the same calendar day during the school day or afterschool session is counted as a single day of attendance. The types of program activities, as already described, are guided independent reading practice and work with the Emergent Reader materials, offered both during the school day and afterschool.

Average Daily Attendance

Over the 2008-09 school year, an average of 56 children attended each program on a typical day (Exhibit 3). Across the 86 literacy programs, an average of 4,939 total children participated each day.

Exhibit 3
Average Number of Children in Attendance Each Day per Program, by Month

	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June	2008-09
Average Number of Children Attending Each Day, Per Program	49	54	60	61	57	60	63	65	62	51	32	56

Exhibit reads: During August 2008, 49 children, on average, were in attendance each day at each of the literacy programs offering services that month.

Program Days Attended

Among individual literacy program participants, the number of days that children attended during the 2008-09 school year ranged from 1 to 182 days and averaged 67 days, with a median of 65 days, as shown in Exhibit 4.

Exhibit 4
Number of Days Attended per Participant

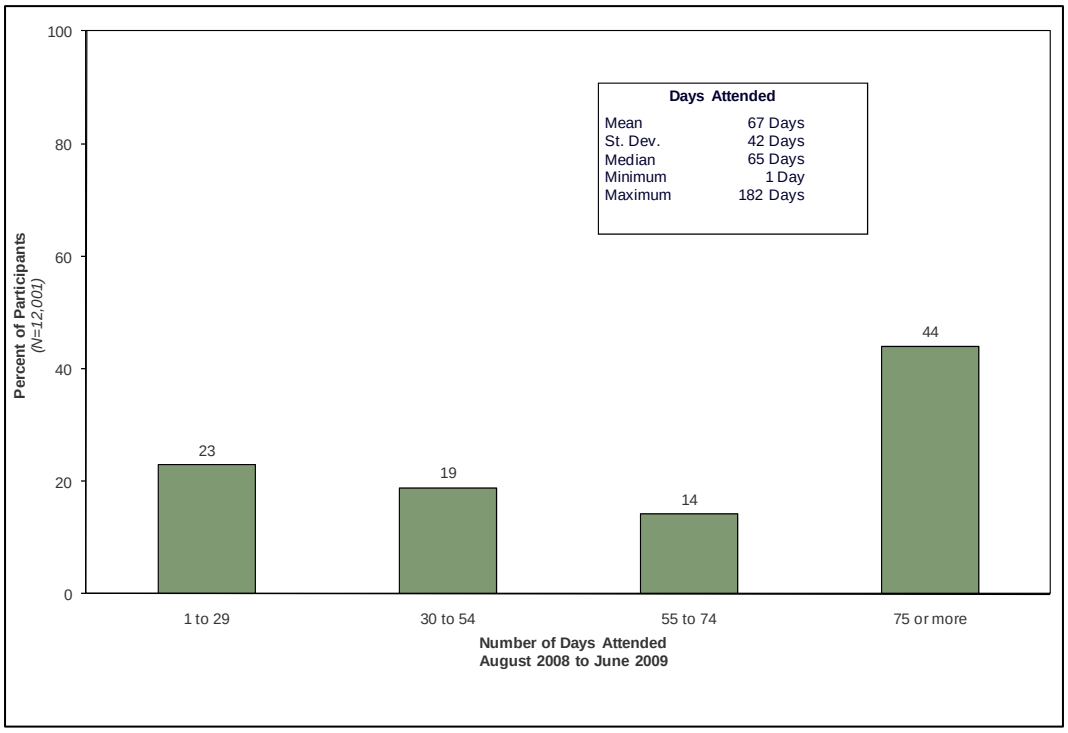


Exhibit reads: Twenty-three percent of children attended a literacy program between 1 and 29 days during the 2008-09 school year.

Program Attendance Rate

Each child's program attendance rate is computed as the number of days the child attended the program divided by the number of days it was possible for that child to attend. Attendance rates ranged from 1 percent to 100 percent and averaged 73 percent, with a median of 80 percent, as seen in Exhibit 5. Twenty-nine percent of participants attended 90 percent or more of the days possible for them to attend.

Exhibit 5
Individual Participant Program Attendance Rate

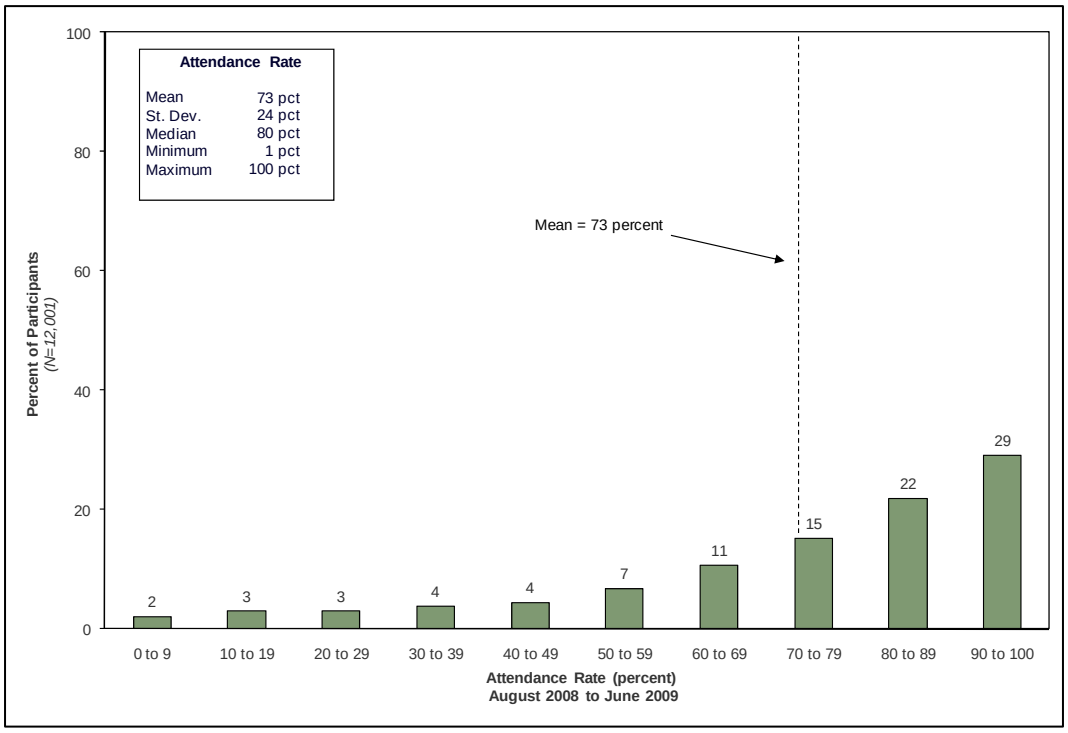


Exhibit reads: Two percent of participants attended between 0 and 9 percent of the days possible to attend.

Measures of Participation in Guided Independent Reading Activities

The guided independent reading 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 an AR 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.

A 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 was for the average of the number of books read by children at each site to be 25 books or more during the school year.

Among the 10,602 participating children who read at least one book, the number of books read ranged from a low of one book to a high of 688 books, with an average of 64 books and median of 48 books (Exhibit 6). Overall, 74 percent of participants met or exceeded the standard by reading 25 or more books during the school year. Across the 117 sites for which the number of books read is available, the average number of books read was 25 or more in 102 sites (87 percent of sites).

Exhibit 6
Number of Books Read

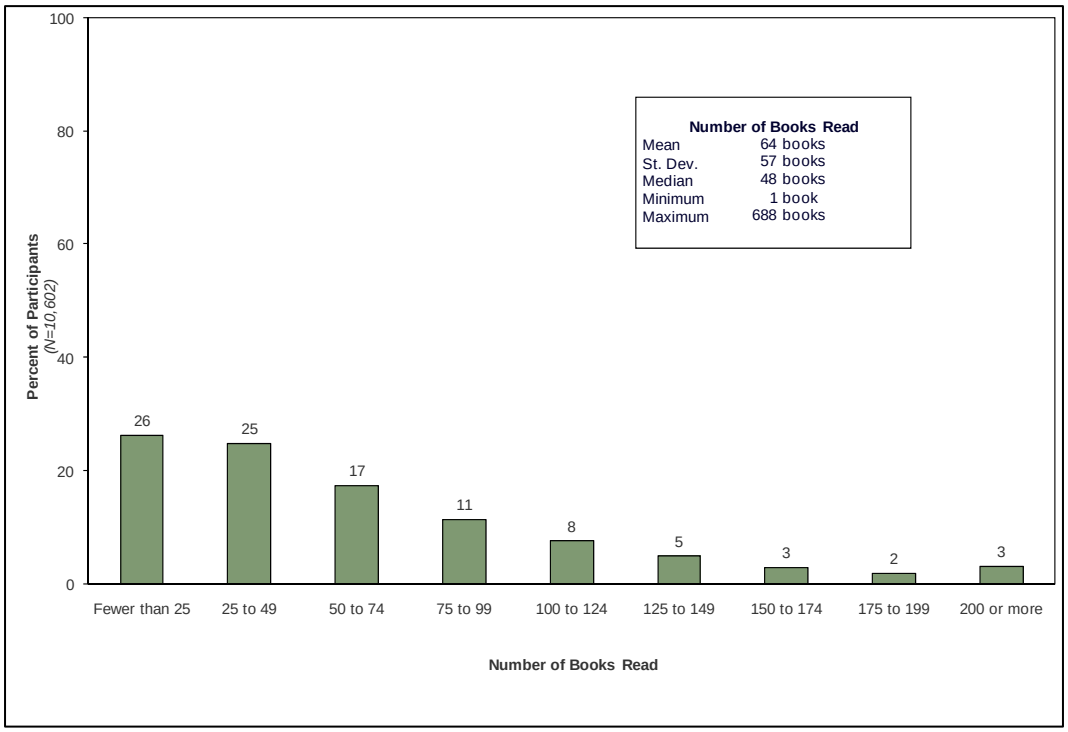


Exhibit reads: Twenty-six percent of participants read fewer than 25 books in the 2008-09 school year.

On average, participants read one book every 2.4 days they attended the program. The 8,261 students who attended the program for at least 30 days read an average of 70 books during the school year, and 81 percent of them read 25 or more books or more. Among the 6,271 students who attended at least 55 days, the average number of books read was 76 and 87 percent of students read at least 25 books. Among students attending for 55 days or more at the 117 sites for which the number of books read was available, the average number of books read was 25 or more at 111 sites (95 percent of sites).

Participants' Performance on AR Quizzes

On average, participants attempted 64 AR quizzes during the 2008-09 school year and achieved a passing score on 59 of them (92 percent). The criterion for passing a quiz is answering 60 percent of the questions correctly.

One of the objectives of the literacy programs 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 apply the content they have read. More than three-quarters (81 percent) of participants passed 85 percent or more of the quizzes they attempted during the 2008-09 school year.

Across all of the quizzes that each participant attempted during the 2008-09 school year, participants correctly answered, on average, 85 percent of the total number of questions that comprised the Accelerated Reader quizzes.

4. Did Participants' Literacy Skills Improve through the Guided Independent Reading Program?

The primary measure of the reading proficiency of the children participating in the literacy programs was their performance on the STAR Reading assessment, which is a standardized, norm-referenced assessment appropriate for children in grades 1 to 12 and is published by Renaissance Learning, the developers of the AR program. The assessment 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. Participants in the literacy programs complete a STAR Reading assessment two or three times per year.

The computer software that administers the STAR Reading assessment scores each child's reading proficiency relative to national norms. The results of the assessment 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 assessment 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 any analysis of change in reading proficiency as measured by standardized assessments, is the need to allow a large enough interval of time between the initial and final administrations of the assessment. Too short an interval does not allow enough time for any real growth in reading proficiency to show up on measures of reading proficiency. If the interval between assessments is too short, any change in score on the reading assessment is more likely to reflect measurement error than change in proficiency. The publishers of the STAR Reading assessment state 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 2008-09 school year. Analyses were limited to children in grades 1 through 6 who completed two STAR Reading assessments at least 90 days apart.

Baseline Reading Proficiency

The distribution of scores from the first STAR Reading assessment that each child completed is the best indicator available of the child's baseline reading proficiency. Analysis of the distribution of the baseline scores indicates that, overall, the programs succeeded in targeting children who were poor readers relative to their grade in school, as shown in Exhibit 7. Among the 11,055 children in first through sixth grade participating in the guided independent reading program, 10,153 completed one or more STAR Reading assessments, and 9,090 (82 percent of the 11,055) completed two or more assessments at least 90 days apart. The STAR Reading assessment is designed to be administered to children in the spring of first grade and higher.

The average of participants' initial scores on the STAR Reading assessment among the students who took at least one assessment was 33 NCEs the equivalent of the 21st percentile using national norms.[3] A score of 50 NCEs is generally interpreted as performance at grade level. Using this definition, 84 percent of participants performed below grade level on their initial STAR Reading assessment.

Exhibit 7
Initial STAR Reading Assessment Scores

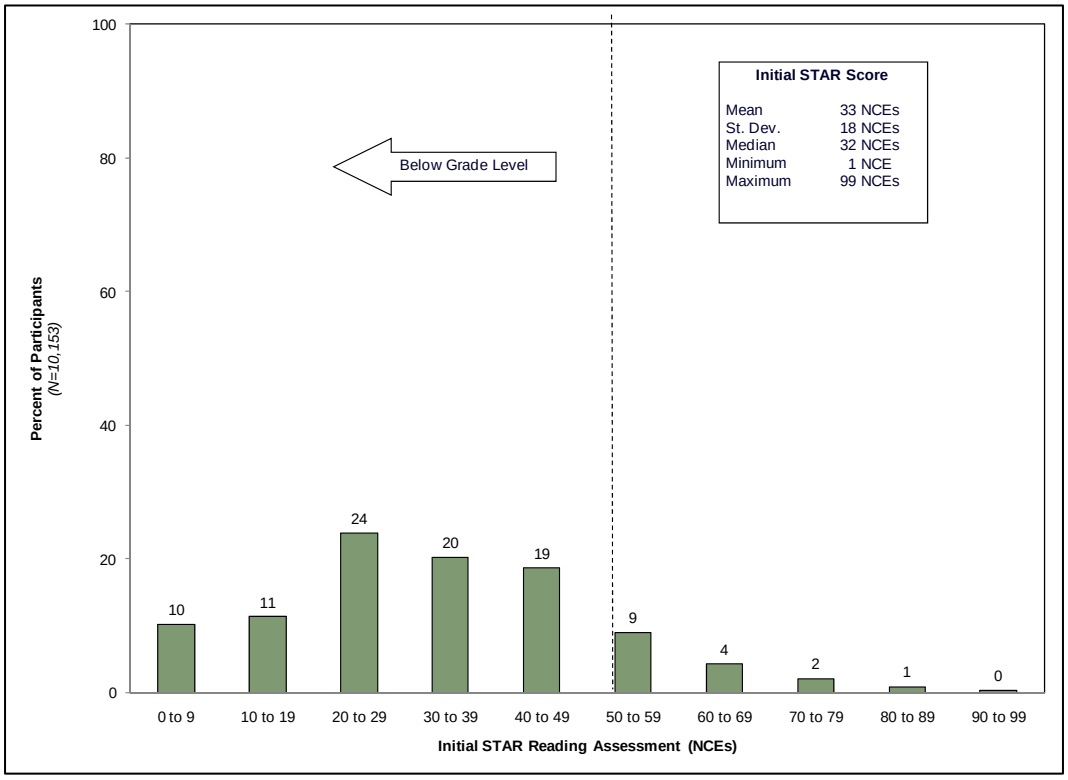


Exhibit reads: Ten percent of participants scored between 1 and 9 NCEs on their initial STAR Reading assessment.

[3] Included in this analysis are all children in grades 1 to 6 who attended one of the 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.

Children in the upper grades were somewhat more likely to score below grade level than were children in the primary grades, as shown in Exhibit 8. Eighty-seven percent of sixth-grade participants scored below grade level (below 50 NCEs) on their initial assessment and their median Grade Equivalent (GE)[4] achieved on the initial assessment was 4.3, the equivalent of November of fourth grade. Among first-grade participants, 74 percent scored below grade level, and the median GE was 0.9, or the July before first grade.

Exhibit 8
Performance on Initial STAR Reading Assessment, by Grade Level

Grade Level	Average STAR Reading Assessment Score (NCEs)	Percent Below Grade Level	Median Grade Equivalent	*N*
1	34	74	0.9	*716*
2	35	85	1.5	*2105*
3	31	85	2.2	*2444*
4	33	84	2.7	*2140*
5	35	85	3.6	*1703*
6	33	87	4.3	*1045*

Exhibit reads: Among literacy program participants enrolled in the first grade, the average score on their initial STAR Reading assessment was 34 NCEs; 74 percent scored below 50 NCEs, indicating they were performing below grade level; and the median GE was 0.9.

Change in Reading Proficiency

In order to examine the change in reading proficiency achieved by participants in the reading programs, it is necessary to restrict the analysis sample to children who completed at least two STAR Reading assessments at least 90 days apart. Among these 9,090 participants, the average score on the initial assessment was 33 NCEs (equivalent to the 21st percentile), and ranged from a low of 1 NCE to a high of 99 NCEs (the lowest and highest NCE scores possible). Fifteen percent of participants performed at 50 NCEs or higher, as seen in Exhibit 9. At the other end of the distribution, 21 percent performed below 20 NCEs.

[4] A Grade Equivalent score (GE) indicates the grade placement of students for whom a particular score on the STAR Reading assessment is typical. The grade placement is described in terms of the grade level (expressed as an integer) and the month of the school year (expressed as a decimal, ranging from 0 for September to 9 for June). A GE of 5.1, for example, indicates that the child's level of performance is consistent with the median level of performance achieved among children in October of their fifth-grade year.

Exhibit 9
Initial Score on STAR Reading Assessment in NCEs, for Children Who Completed Two Assessments at Least 90 Days Apart

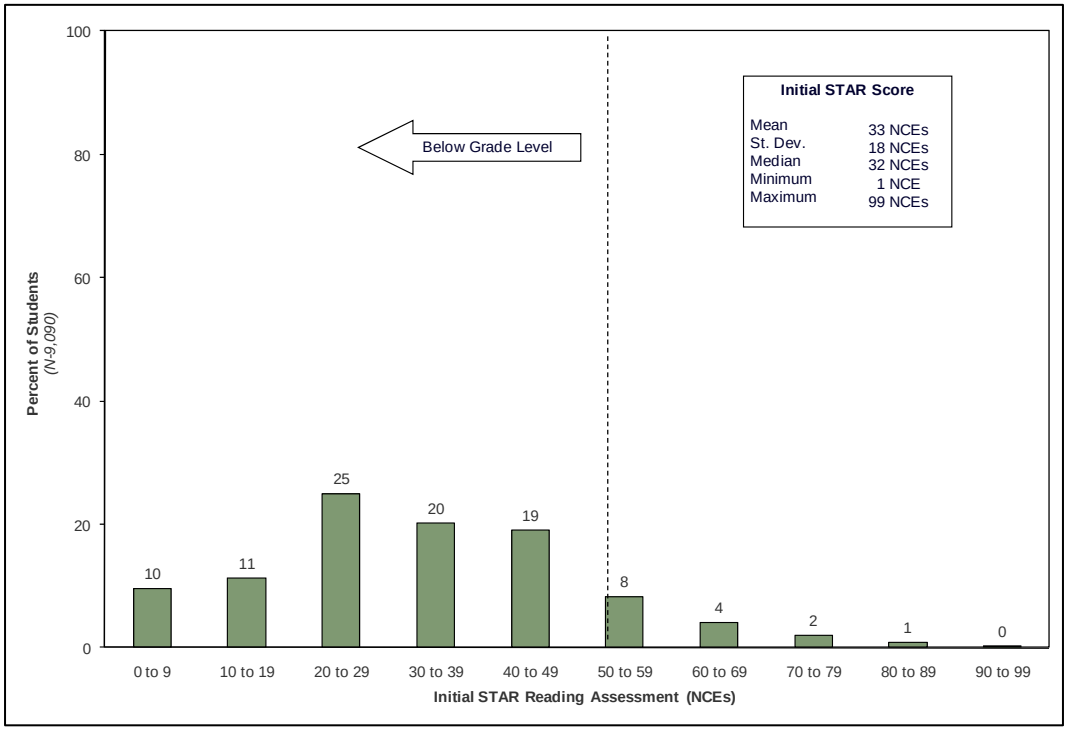


Exhibit reads: Nine percent of children with two STAR Reading assessments at least 90 days apart scored between 1 and 9 NCEs on their initial STAR Reading assessment.

Overall, the average score increased from 33 NCEs on the first STAR Reading assessment of 2008-09 to 39 NCEs on the final STAR Reading assessment in this period, the equivalent of the 25th and 31st percentiles respectively, as shown in Exhibit 10.

Exhibit 10
Initial and Final Scores on STAR Reading Assessment in NCEs,
for Children Who Completed Two Assessments at Least 90 Days Apart

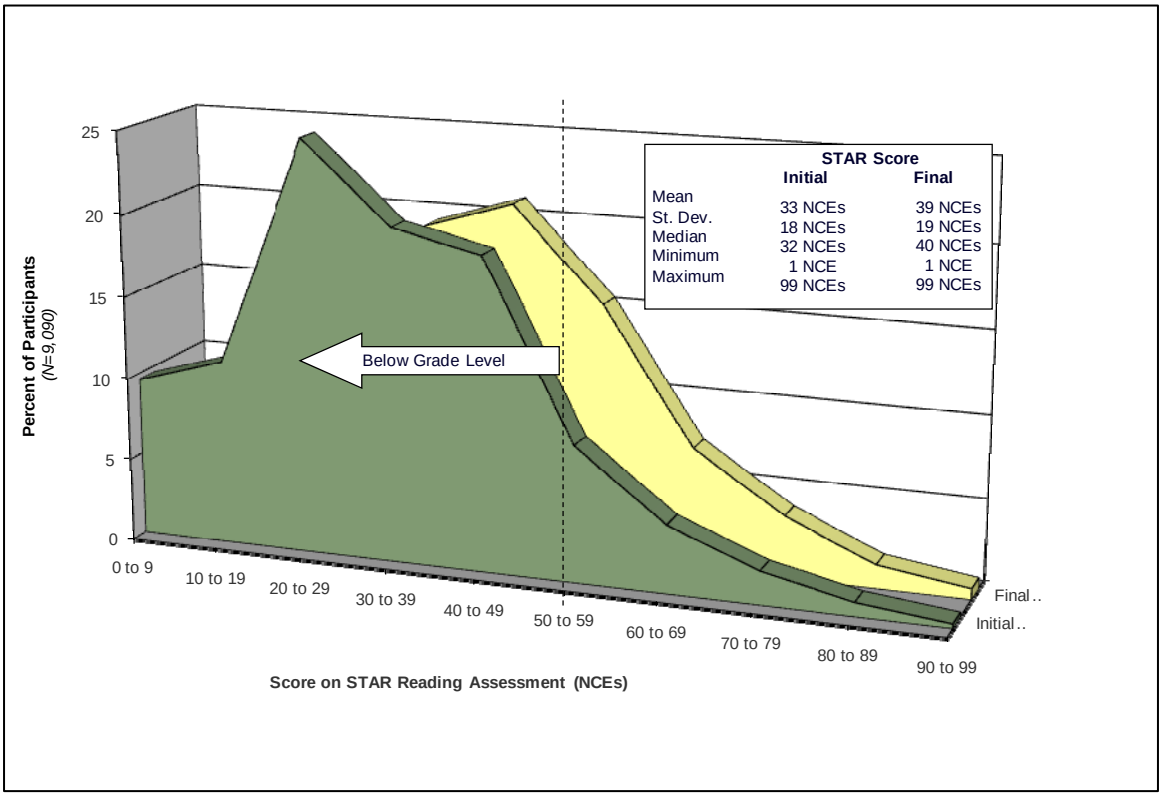


Exhibit reads: Ten percent of participants with at least two STAR Reading assessments at least 90 days apart scored between 1 and 9 NCEs on their first STAR Reading assessment, decreasing to 7 percent on their final STAR Reading assessment.

Among the participants in the literacy program during the 2008-09 school year, the change in STAR Reading assessment scores ranged from a decrease of 69 NCEs to an increase of 80 NCEs. The average change in scores was a gain of 5.8 NCEs, as shown in Exhibit 11. The change is statistically significant ($p<.05$). The effect size of the change is +0.32, indicating that the change is large enough to be considered substantive.

Exhibit 11
Change in Score on STAR Reading Assessment, in NCEs

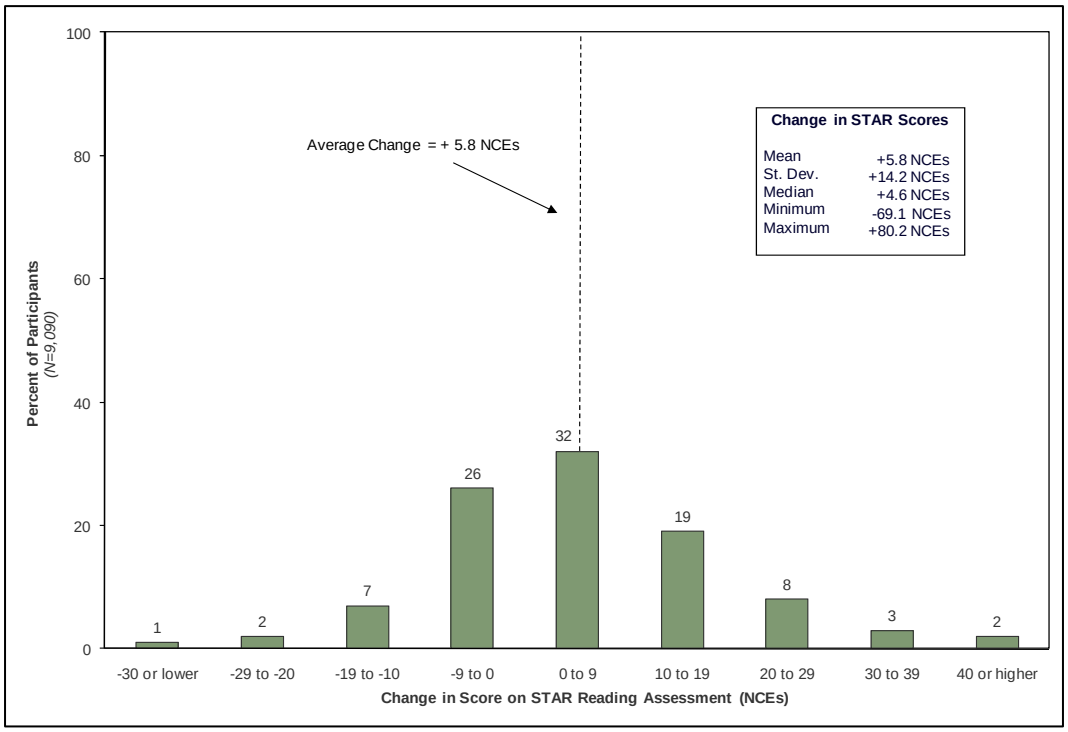


Exhibit reads: One percent of participants experienced a decrease of more than 30 NCEs on the STAR Reading assessment between their first and last administration.

The proportion of participants reading at a level appropriate for their grade or above increased during the 2008-09 school year. On the initial STAR Reading assessment, 16 percent of participants who completed two STAR Reading assessments 90 days apart were reading at grade level or higher (50 NCEs or more). By the time the final STAR Reading assessment was administered, 29 percent read at grade level, as shown in Exhibit 12. The difference is statistically significant ($p<.05$).

Exhibit 12
Proportion of Participants Reading at Grade Level or Above at Initial and Final Administrations of the STAR Reading Assessment

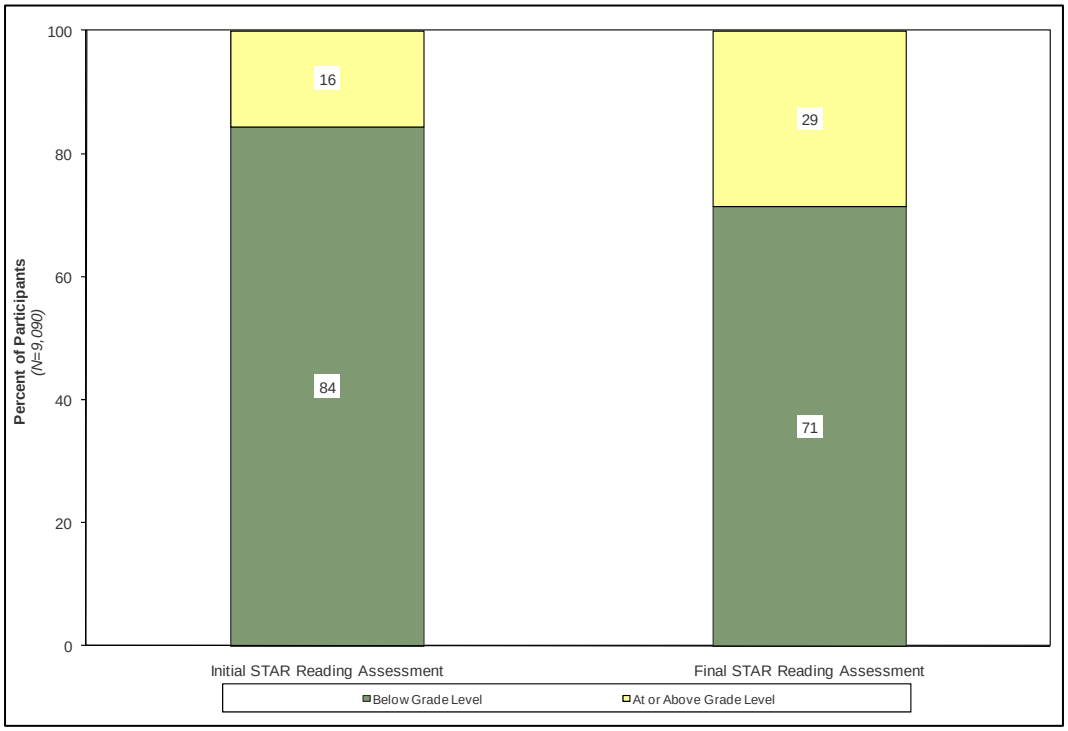


Exhibit reads: Sixteen percent of participants read at grade or above on their initial STAR Reading assessment.

A month of schooling represents attendance every day during the month and approximately five hours of instruction per day, a total of 100 hours of instruction per month. A year of schooling would involve approximately 900 hours of instruction. Participation in the literacy program, in contrast, involves between 30 minutes and one hour of instruction per day, and the average number of days attended was 67 days, representing a total of between 34 and 67 hours of instruction, or between one-third to thirds of a school month. An estimate of the effect size expected from this amount of additional instructional time would be between +0.03 and +0.07. The effect size actually observed among participants was +0.32, or between four and 11 times the expected effect. This gain could have been driven by the literacy program, children's regular school-day learning experiences, or both. As discussed later in this report, it is very likely that the literacy program was a major contributor to this gain.

Comparisons with other studies of educational programs suggest that the gains associated with the Save the Children literacy initiative are substantial.

- A study of the impact of the Teach For America program on math achievement found an effect size of +0.15 on math and +0.03 on reading over a year (Decker, Mayer, & Glazerman, 2004).

- A study of the impact of the reduction in class size by eight students per class found an effect size of +0.23 (Finn & Achilles, 1999).

- In a review of four studies of afterschool programs, Kane (2004) concluded that the expected impact of an extra hour of instruction delivered in an afterschool setting can be estimated to be an effect size of +0.05.

- An evaluation of the 21st Century Community Learning Centers Program in Louisiana found that the impact of this afterschool program was an effect size of +0.13 (Jenner & Jenner, 2007).

A model for expressing the magnitude of change in scores on a nationally-normed assessment, such as the STAR Reading assessment, emerges from this literature. The findings consistently estimated that the effect size of the change in scores over a full school year was one standard deviation, or an effect size of 1.0. If we assume that the change in student performance occurs evenly over a school year, then each increment in effect size of 0.1 is the equivalent of the gain expected from one month of schooling. Among the participants in the literacy program during the 2008-09 school year, the effect size of the change of scores on the STAR Reading assessment was +0.32, which, according to this model, could be interpreted as the equivalent of an additional three months of school for the average participant.

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 assessment publisher. Among all of the children participating in the literacy programs who completed two STAR Reading assessments at least 90 days apart, 60 percent achieved a gain of 2.0 NCEs or greater, an indication of a meaningful gain in reading proficiency.

Similarly, more than 60 percent of the participants in grades 1 through 4 achieved gains of two NCEs or greater, as shown in Exhibit 13, and more than 50 percent of participants in grades 5 and 6 did so.

Exhibit 13
Percent of Participants Achieving Gains of Two NCEs or More, by Grade

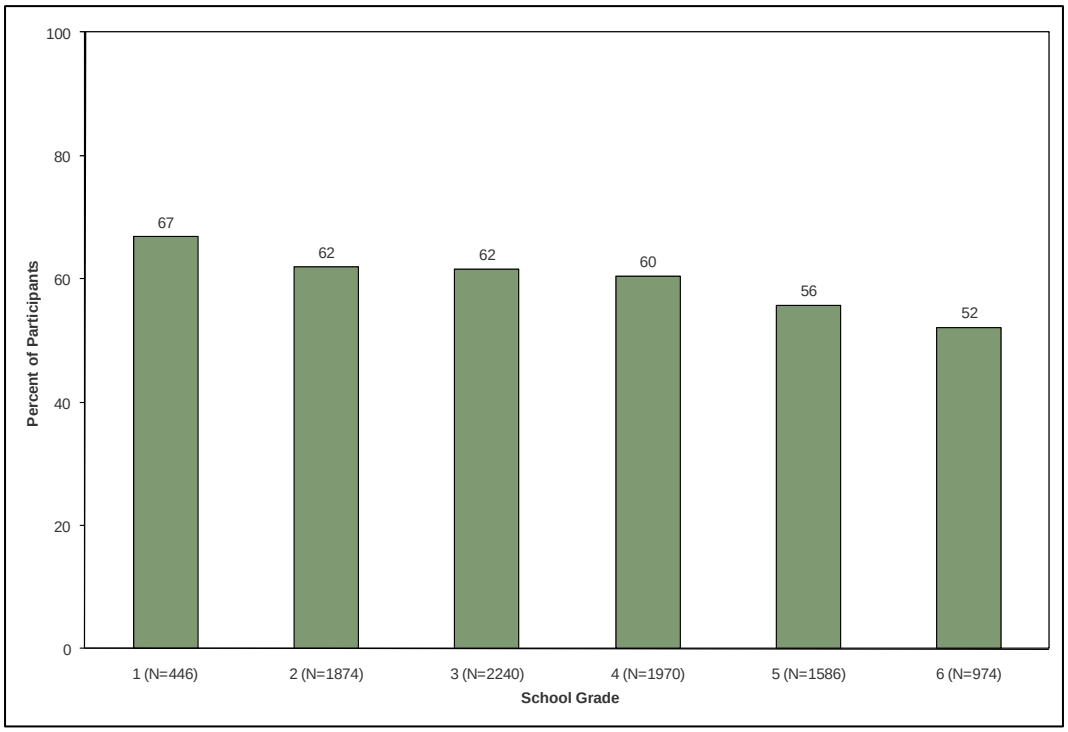


Exhibit reads: Sixty-seven percent of participants enrolled in first grade achieved gains of 2 NCEs or greater on the STAR Reading assessment.

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

To determine whether some groups 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 guided independent reading using AR. This information permitted the identification of key factors associated with participants' changes in STAR Reading assessment scores.

Participant Characteristics

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 + 0.71 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 2008-09 was -0.29 and was 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 14. The difference between the initial and final STAR Reading assessments was statistically significant among the initial performance groups below 50 NCEs.

The proportion of participants who gained 2.0 NCEs or more between their initial and final STAR Reading assessment was higher among participants who scored below 50 NCES (below grade level) on their initial assessment, as shown in Exhibit 14. In addition, 57 percent of participants who scored below 50 NCEs on the initial assessment gained 2.0 NCEs or more, compared to 39 percent for those who scored above 50 NCEs on the initial assessment.

Exhibit 14
Change in Score on STAR Reading Assessment, by Performance on Initial Assessment

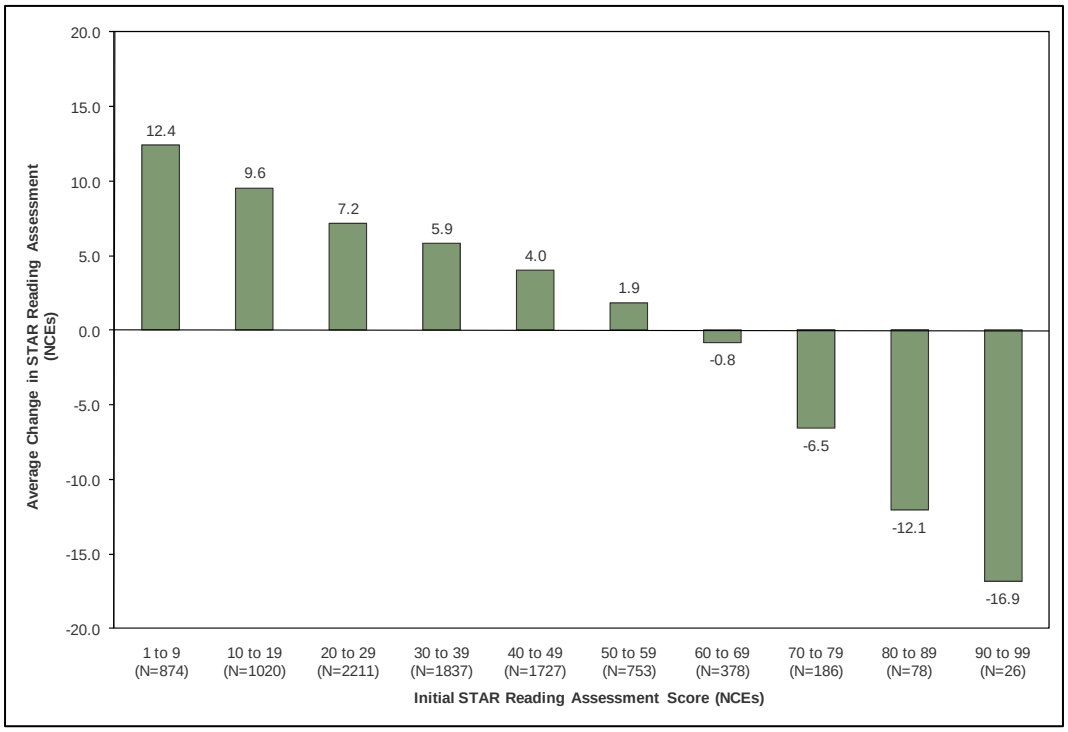


Exhibit reads: Children with initial scores between 1 and 9 NCEs averaged a gain of 12.4 NCEs between their initial and final STAR Reading assessments.

The proportion of participants who gained two NCEs or more between their initial and final STAR Reading assessments also varied by their level of performance on the initial assessment. Nearly two thirds of participants (63 percent) who scored below 50 NCEs on the initial assessment gained two NCEs or more, while only 41 percent of those who scored above 50 NCEs on the initial assessment did so, as shown in Exhibit 15.

Exhibit 15
Proportion of Participants Gaining Two NCEs or More, by Performance on Initial Assessment

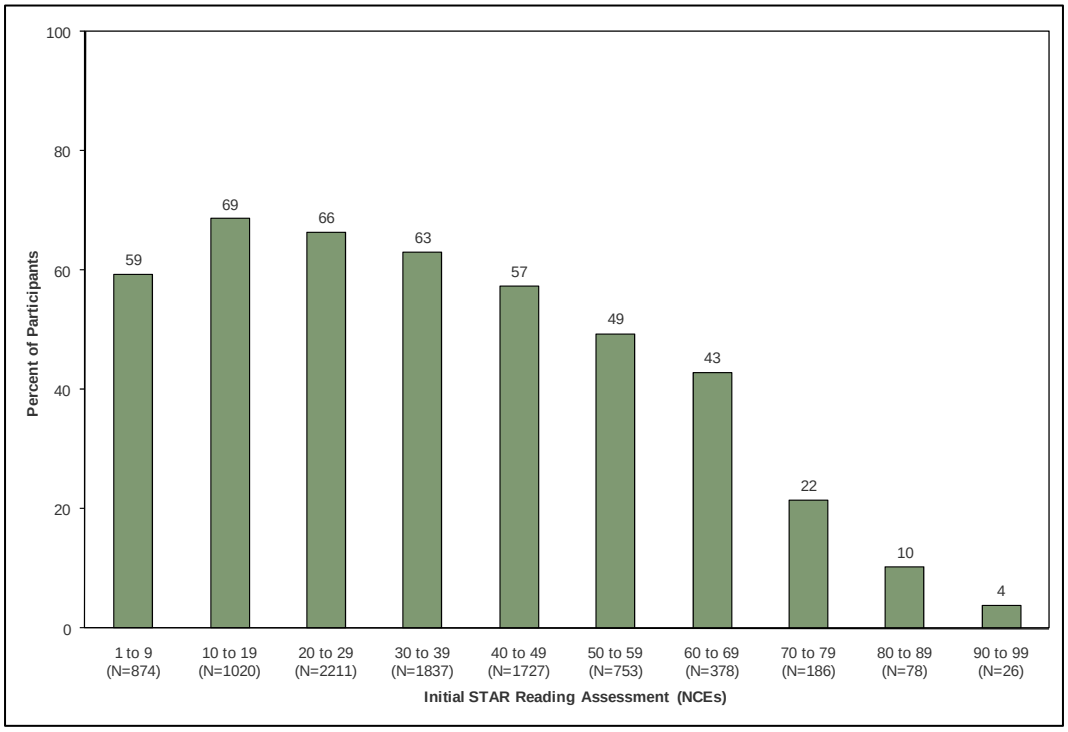


Exhibit reads: Fifty-nine percent of participants who scored below 10 NCEs on their initial STAR Reading assessment gained 2.0 NCEs or more on their final STAR Reading assessment.

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 participants to reach grade level (Exhibit 16).

Exhibit 16
Proportion of Participants Reading at Grade Level on Their Final STAR Reading Assessment, by Performance on Initial Assessment

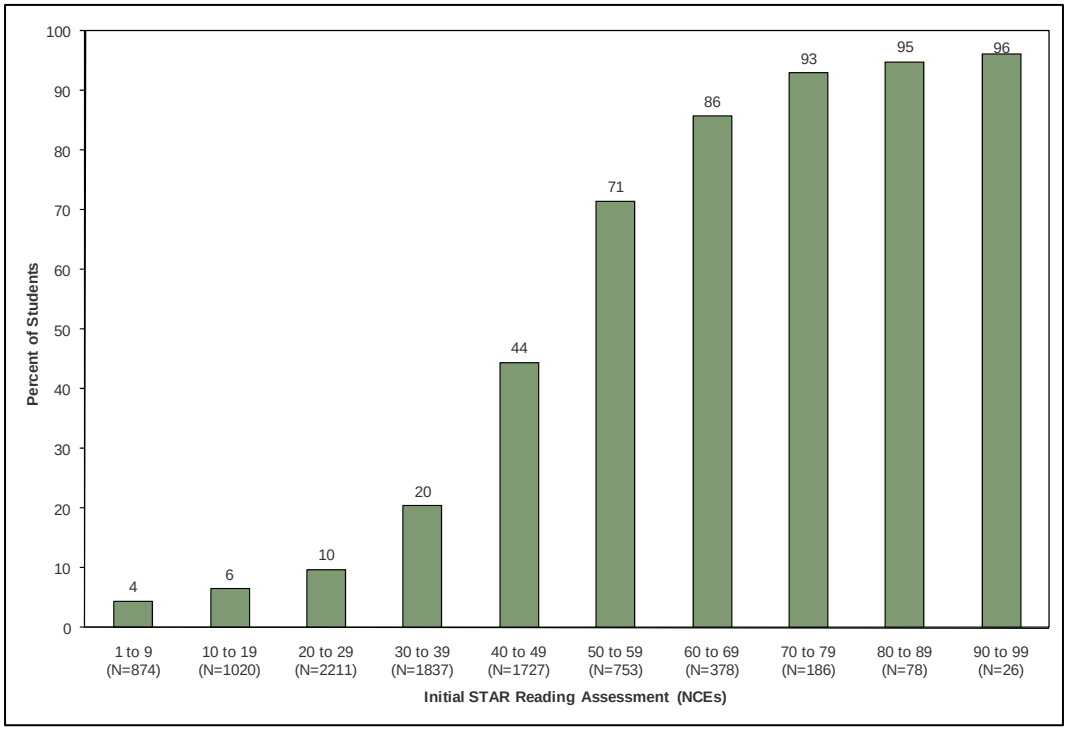


Exhibit reads: Four percent of the participants scoring below 10 NCEs on their initial STAR Reading assessment read at grade level (50 NCEs or higher) on their final assessment.

Patterns of Program Attendance

Overall, the correlation between the number of days attended and gains in reading proficiency was +0.12 and a statistically significant relationship, indicating that there is a small tendency for children who attended more days to achieve higher gains in proficiency. The correlation between the percent of days attended (the attendance rate) and the change in STAR Reading assessment scores was 0.07 and was a statistically significant relationship ($p<.05$).

The positive relationship between program attendance during the 2008-09 school year and gain in reading proficiency is shown in Exhibit 17. In particular, among children who attended 55 days or more, the average change in STAR Reading assessment scores was 7.1 NCEs, while children attending fewer than 55 days averaged a gain of 3.6 NCEs. The difference was statistically significant ($p<.05$).

Exhibit 17
Change in STAR Reading Assessment Score (NCEs), by Days Attended

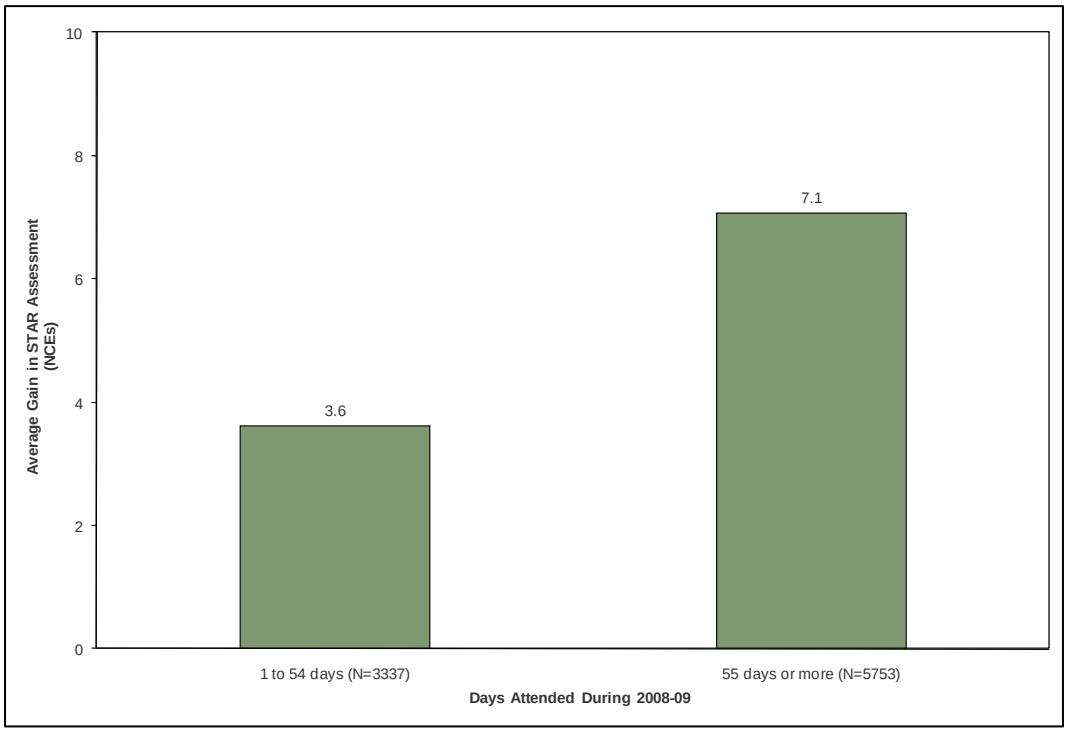


Exhibit reads: Children who attended a literacy program for fewer than 55 days during 2008-09 averaged a gain of 3.6 NCEs on the STAR Reading assessment.

Generally, a larger proportion of children who attended the program on more days gained 2 NCEs or more on the STAR Reading assessment. Sixty-three percent of children who attended 55 days or more achieved a gain of at least 2 NCEs, compared with 53 percent of children who attended fewer than 55 days. The difference was statistically significant ($p<.05$).

Participants' Performance on AR Quizzes

A key activity for guided independent reading practice is reading books. Among participants in the literacy programs, 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 Reading assessment score) between the number of books read during the 2008-09 school year and the change in reading proficiency on the STAR Reading assessment was +0.07, and was a statistically significant relationship ($p<.05$).

Evaluators found a substantial and positive relationship between the percent of AR quizzes each participant passed during the 2008-09 school year and their gains on the STAR Reading assessment. Controlling for each child's initial level of reading proficiency, the partial correlation between the percent of AR quizzes passed and gains on the STAR Reading assessment was +0.17, and was a statistically significant relationship.

As noted, a goal of the literacy program is that each participant will pass 85 percent of the quizzes they attempted. Participants who passed at least 85 percent of the quizzes they attempted gained an average of 6.3 NCEs on the STAR Reading assessment compared with 3.8 NCEs for participants who passed less than 85 percent of their quizzes. The difference was statistically significant ($p<.05$).

In addition to the relationship between gains in STAR Reading assessment scores and the percent of AR quizzes passed, there was also a substantial and positive relationship between the average percent of the total number of items answered correctly on all of the AR quizzes and participants' gains on the STAR Reading assessment. The partial correlation, controlling for each child's initial level of reading proficiency, was +0.20 and statistically significant.

As seen in Exhibit 18, children who correctly answered 95 percent or more of the AR quiz items they attempted during 2008-09 averaged gains of 8.5 NCEs on the STAR Reading assessment, while those answering less than 80 percent of items correctly averaged gains of 3.8 NCEs. The difference was statistically significant.

Exhibit 18
Change in Score on the STAR Reading Assessment, by Percent of AR Quiz Items Answered Correctly

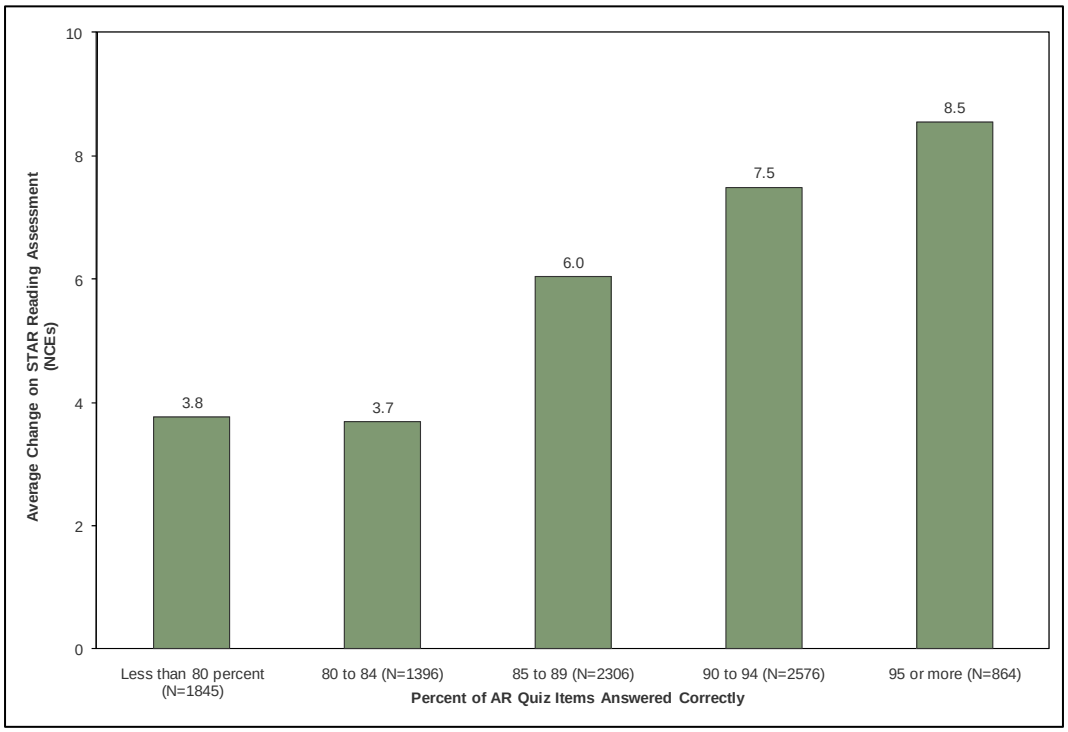


Exhibit reads: Among children who answered less than 80 percent of the AR quiz items they attempted correctly, the average gain on the STAR Reading assessment was 3.8 NCEs.

6. What Were the Academic Outcomes of Students Who Read Below Grade Level on Their Initial STAR Reading Assessment?

The following discussion examines the attendance and academic outcomes of participants who read below grade-level on the initial STAR Reading assessment. Because most participants overall were reading below grade level (see Exhibit 7), the results of this subgroup tended to reflect the trends seen earlier in this report.

Measures of Participation in Literacy Programming

Among the 8,545 students who scored below grade level on their initial STAR Reading, the number of days attended during the 2008-09 school year ranged from 1 to 182 days and averaged 68 days. Twenty-one percent attended the program for fewer than 30 days, 18 percent between 30 and 54 days, 15 percent between 55 and 74 days, and 46 percent 75 days or more.

Attendance rates ranged from 1 percent to 100 percent and averaged 74 percent, with a median of 81 percent. Nearly one-third of participants (30 percent) attended 90 percent or more of the days possible for them to attend.

Measures of Participation in AR Activities

Among participating children reading below grade level on their initial STAR Reading assessment, the number of books read ranged from a low of 1 book to a high of 688 books, with an average of 65 books and median of 49 books.

Participants' Performance on AR Quizzes

On average, participants reading below grade level attempted 65 AR quizzes during the 2008-09 school year and achieved a passing score on 59 of them (91 percent). One of the objectives of the literacy programs was for participants to pass at least 85 percent of the quizzes they attempted. More than three-quarters (79 percent) of participants reading below grade level passed 85 percent or more of the quizzes they attempted.

Baseline Reading Proficiency

As expected, participants reading below grade level scored lower on average than other participants on their initial STAR Reading assessment. Ninety percent of the 8,545 participants completed at least two STAR Reading assessments at least 90 days apart. Among the students completing two assessments, the average score on their initial assessment was 28 NCEs (equivalent to the 15th percentile rank) and ranged from a low of 1 NCEs to a high of 49 NCEs.

More than half of the participants (54 percent) reading below grade level scored below 30 NCEs on the initial STAR Reading assessment.

Change in Reading Proficiency Scores

The proportion of participants reading at a level appropriate for their grade or above increased during the 2008-09 school year. Eighteen percent of the participants who were reading below grade level on the initial STAR Reading assessment were reading at grade level or higher (50 NCEs or more) on the final assessment. Among participants reading below grade level on their initial STAR Reading assessment, the average score increased from 28 NCEs on the initial assessment to 34 NCEs on the final assessment. The difference is statistically significant ($p<.05$).

Sixty-three percent of all participants who scored below grade level on their initial STAR Reading assessment increased their score by 2 NCEs or more, and the average gain was 7.1 NCEs. The change in scores was statistically significant ($p<.05$) and the effect size of the change was +0.53, indicating that the average gains were substantial. Among those who attended the program for at least 55 days, 66 percent gained at least 2 NCEs, and the average gain was 8.3 NCEs.

7. What Were the Academic Outcomes of Students Who Participated in the Emergent Reader Literacy Program?

For several years, some kindergarten and first-grade children at the programs participated in Emergent Reader activities, either afterschool or during the in-school program. This program grew out of the need to target younger students by providing daily opportunities to receive systematic support in skill areas they are struggling to master, such as phonemic awareness and letter recognition, that will lead to successful beginning reading achievement. Because these components are not always available in small group settings for struggling students during regular language arts instruction, the Emergent Reader program allows students to participate in targeted activities to build a strong foundation towards future reading success.

In order to determine the effectiveness of the Emergent Reader program, the participating sites began administering a standardized assessment, the STAR Early Literacy assessment (SEL), to the children participating during the 2008-09 school year. Published by Renaissance Learning, the SEL is a computer administered instrument designed to assess the early literacy skills of beginning readers. The results from the assessment provide program staff with information to guide instruction. The assessment is designed to be taken three times per year– fall, winter, and spring– providing staff with a means for tracking each student's progress over time.

In contrast to the STAR Reading assessment, which is a norm-referenced assessment, the SEL is a criterion-referenced assessment. Norm-referenced assessments allow the comparison of the performance of each student with the performance of other, similar students, who constitute

the norming sample. Comparison of the scores of tested students can be expressed in terms of national percentile ranks or normal curve equivalents. Criterion-referenced assessments like the SEL, are intended to measure how well a person has learned a specific body of knowledge, and skills scores are expressed in terms of whether a student has achieved levels of performance that demonstrate the mastery of specific skills or concepts. For the SEL, scores are reported as scale scores, reflecting the level of a student's performance after taking into account the difficulty of the specific assessment items each student answers but not the performance of other students.

Enrollment and Attendance

The SEL assessment was administered in 46 sites during 2008-09. The SEL was administered primarily to students in kindergarten and first grade, and to a few students in higher grades who were not proficient readers. Overall, 1,538 children, 55 percent of those participating in the Emergent Reader program, completed at least one SEL assessment over the course of the 2008-09 school year.

Participants in the Emergent Reader Program

Grade level. Nearly half of the participants who completed an SEL assessment (45 percent) were enrolled in kindergarten during 2008-09, and 52 percent were enrolled in first grade. Three percent of participants were in second grade.

Literacy level. Performance on the SEL is reported in three forms. One is a scaled score ranging from 300 to 900 scaled score points, with the higher score indicating greater reading proficiency. [5] A second is a literacy skills classification level, which places each assessment score into one of four levels. The classification levels are labeled "early emergent" (the lowest level of reading proficiency), "late emergent," "transitional," and "probable reader." Performance is also reported in terms of risk levels. The risk level into which a child is classified is based on the child's grade level and the time of year the SEL is administered, in addition to the child's scaled score. These levels indicate the degree of risk that the child will not gain reading proficiency. The risk levels are labeled as "low risk" (indicating performance at or above the level expected for a student in that grade at that time of the school year), "some risk," and "at risk."

Overall, assessment scores indicate that the children attending the Rural Literacy Initiative sites participating in the Emergent Reader program were appropriately selected, based on their level of literacy skills, as shown in Exhibit 19. Seventeen percent scored at the early emergent reading level, and 63 percent at the late emergent level, the level of skills for which the Emergent Reader program is intended. Fourteen percent scored at the transitional reader level and 6 percent at the probable reader level. Students reading at these two higher levels can be considered for movement to beginning reading instruction.

[5] The SEL scaled scores reflect the difficulty of the questions presented to the student and the number of items answered correctly. A scaled score of 300 is roughly equivalent to the performance expected of a beginning kindergarten student, and a score of 900 to the performance of a student finishing fifth grade.

Exhibit 19
Literacy Skills Classification Level on Initial SEL Assessment

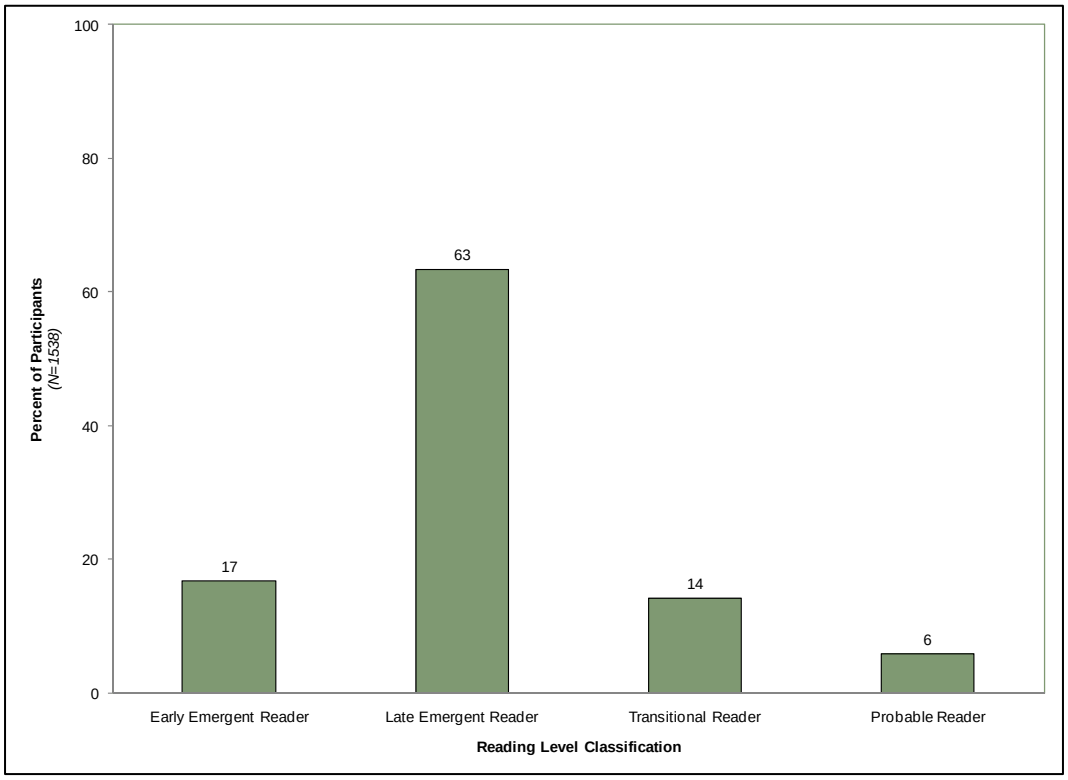


Exhibit reads: Seventeen percent of participants who completed an SEL assessment were classified as early emergent readers on their initial assessment.

Domain scores. To assist in the identification of each child's specific strengths and weaknesses, scores are also reported for seven literacy domains: general readiness, graphophonemic knowledge, phonemic awareness, phonics, comprehension, structural analysis, and vocabulary. Domain scores are reported on a scale from 1 point to 99 points, with higher scores indicating greater proficiency on that domain of reading skills.

On their initial SEL assessment, participants had higher scores in the more basic domains—general readiness and graphophonemic knowledge—than in other domains (Exhibit 20). These two domains encompass early literacy skills, such as differentiating shapes, naming letters, recognizing letter sounds, and applying alphabetical order. Participants demonstrated less proficiency on higher order skills, such as vocabulary (including matching words and pictures and identifying synonyms and antonyms) and structural analysis (including finding words, building words, and identifying compound words).

Exhibit 20
Initial Scores on the Seven Literacy Domains

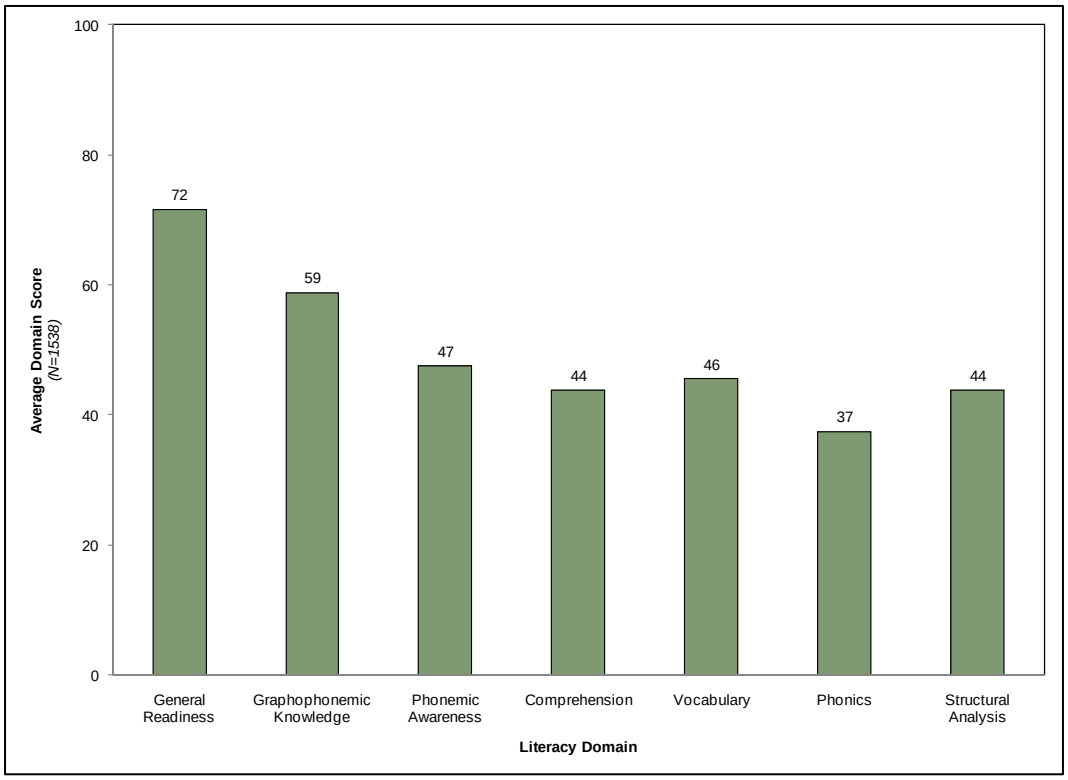


Exhibit reads: Participants had an average General Readiness score of 72 scaled score points on their first SEL assessment.

Risk level. On their initial SEL assessment, approximately three-quarters of Emergent Reader participants were classified as being "at risk" or "some risk" of delayed development or academic failure, indicating that continued participation in the Emergent Reader program was appropriate (Exhibit 21). Twenty-eight percent of participants were classified as being "at risk," the highest level of risk and 46 percent as being at "some risk." This classification took into account each child's SEL scale score, grade level, and the time in the school year when the initial SEL was administered.

Exhibit 21
Risk Status of Participants on Their Initial SEL Assessment

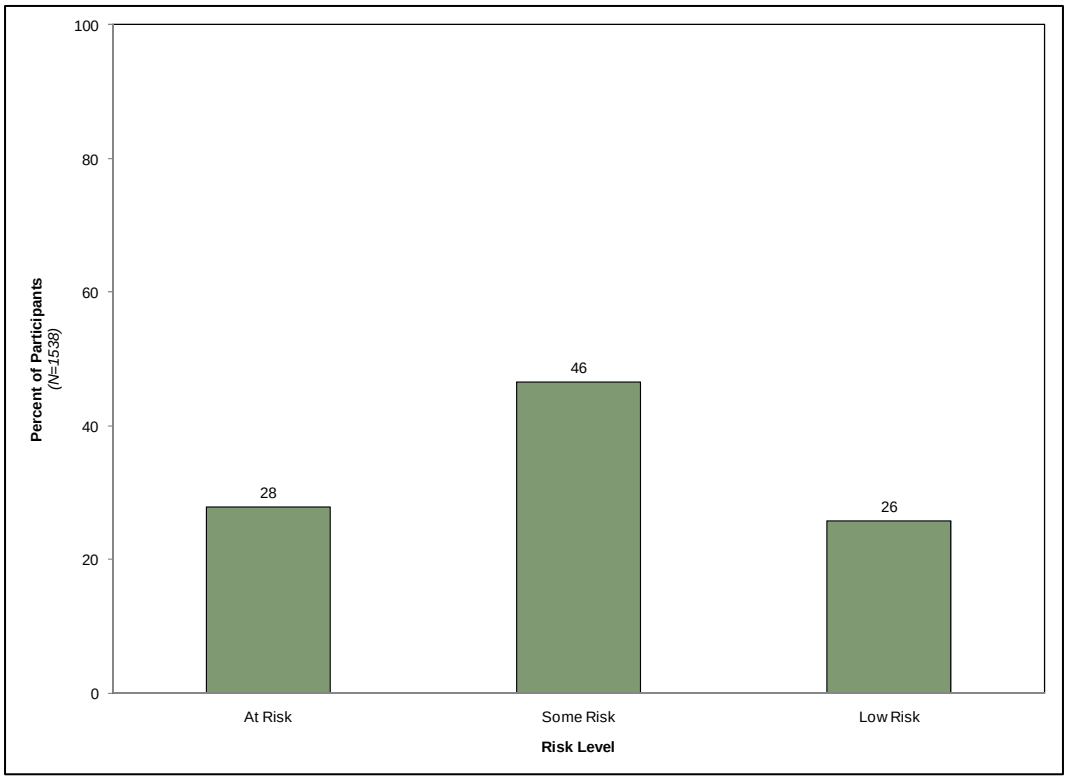


Exhibit reads: Twenty-eight percent of participants were classified as at risk of delayed development or academic failure, based on their SEL scale score, grade level, and the time of year they completed their initial SEL assessment.

Changes in Reading Performance

Eighty-four percent of the children who completed one SEL assessment during 2008-09 also completed at least one additional assessment, allowing an analysis of the change in reading proficiency for these participants, as shown in Exhibit 22. For this study, the estimate of the change in reading proficiency was limited to children in kindergarten and first grade who completed at least two SEL assessments. The change was computed by subtracting each child's score on the first SEL assessment completed during the 2008-09 school year from the child's score on the last assessment.

Exhibit 22
Time Pattern of SEL Completion

Assessment Administration	Percent of Participants *(N=1484)*
Fall, winter, and spring	52
Fall and spring	6
Fall and winter	8
Winter and spring	18
Fall only	5
Winter only	4
Spring only	7

Exhibit reads: Fifty-two percent of participants completed three SEL assessments during 2008-09, in the fall, winter, and spring.

SEL scale scores. Among the children who completed a SEL assessment during both fall 2008 and spring 2009, the average change in scaled score was +117 scaled score points (Exhibit 23).[6] The difference in scores between the fall and the spring was statistically significant ($p<.05$).

[6] An increase in scale score reflects an increase in reading proficiency. Because the SEL is a criterion-referenced assessment, there is no way to express this increase in terms of percentiles or normal curve equivalents or performance at grade level. An estimate provided by the assessment publisher is that a gain of 123 scale score points is approximately the gain expected from 25 weeks of full time participation in the Emergent Reader program.

Exhibit 23
Distribution of Initial and Final SEL Scale Scores

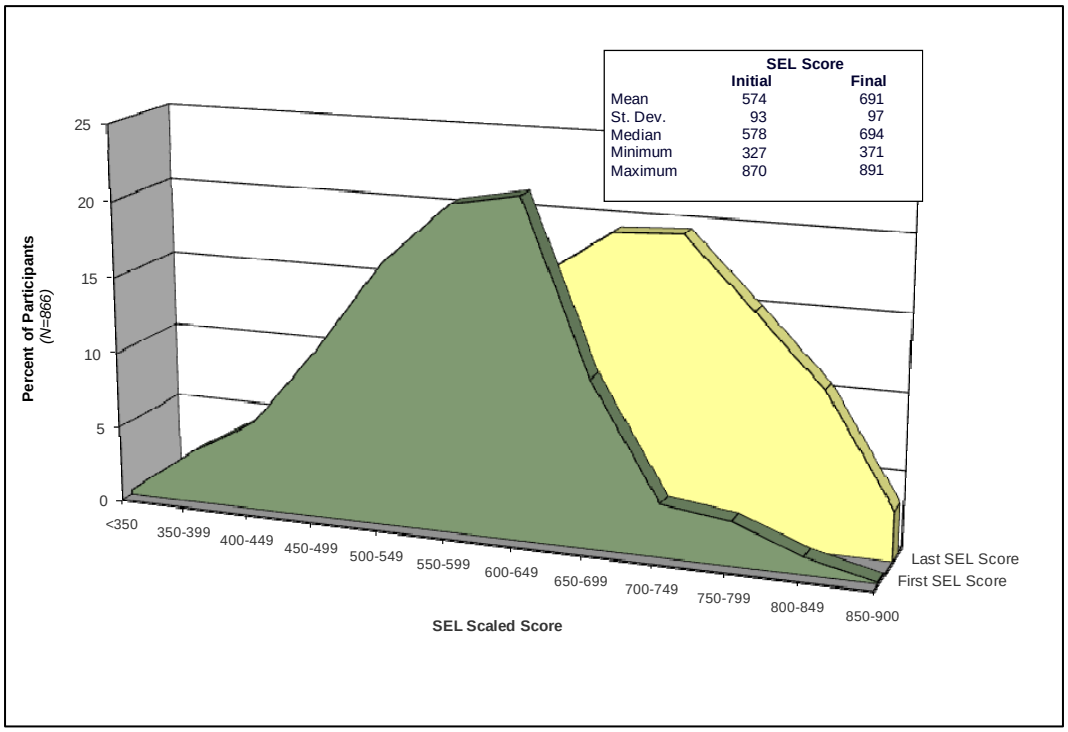


Exhibit reads: Three percent of participants scored between 350 and 399 scale score points on their first SEL assessment, decreasing to 0 percent on their final assessment.

The children who completed SEL assessments only in the fall and winter or only in winter and spring also achieved gains in scale scores, although smaller than the gains observed between fall and spring, as expected. Those completing the assessment only in the fall and winter averaged a gain of 59 scale score points. The group who completed it only in the winter and spring averaged a gain of 50 scale score points.

Reading classification levels. The proportion of children who completed the SEL in both the fall and spring who were classified as early emergent readers decreased from 18 percent in the fall to 3 percent in the spring, while the proportion classified as probable readers increased from 3 percent to 21 percent (Exhibit 24). Eleven percent of the children who completed an SEL in the fall were transitional or probable readers on their initial assessment, and thus able to participate in a beginning reading program. On the spring assessment, this proportion increased to 59 percent.

Exhibit 24
Reading Classification Level Based on the Fall and Spring SEL

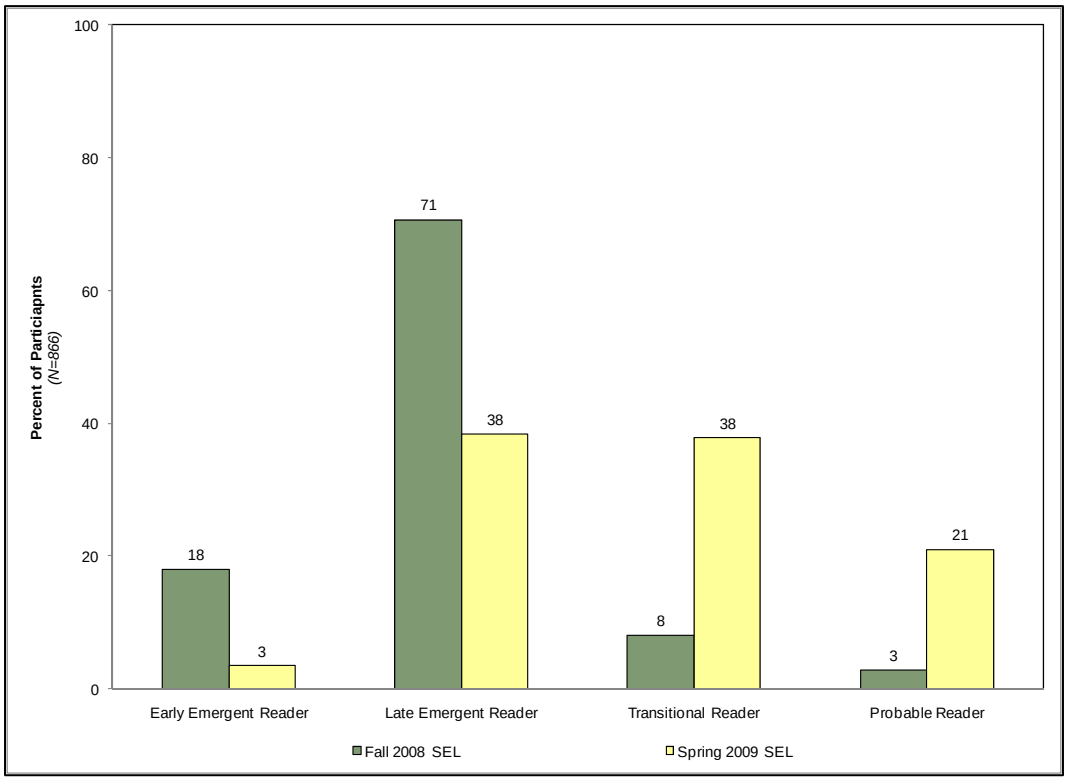


Exhibit reads: Fifteen percent of children were classified as early emergent readers on their fall 2008 SEL assessment.

Children who completed the SEL in fall and winter only demonstrated a pattern of gains in reading classification level as well. The proportion classified as transitional or probable readers increased from 18 percent on the initial SEL to 38 percent on the final SEL. Similarly, among children completing the SEL in winter and spring only, the proportion classified as transitional readers or above increased from 29 percent to 44 percent.

Literacy domain scores. Among the seven literacy domains measured by the SEL, the largest gains between fall 2008 and spring 2009 were observed on structural analysis, phonemic awareness, and comprehension, as shown in Exhibit 25.

Exhibit 25
Change in Literacy Domain Scores, Fall and Spring SEL
(N=866)

Literacy Domain	Average Change in Domain Score	Average Fall Score	Average Spring Score
Structural Analysis Change	21.1	34.1	55.1
Phonemic Awareness Change	20.4	44.3	64.7
Comprehension Change	20.2	42.3	62.5
Phonics Change	19.8	40.7	60.6
Vocabulary Change	18.9	40.8	59.8
Graphophonemic Knowledge Change	17.6	56.1	73.7
General Readiness Change	14.2	69.5	83.6

Exhibit reads: Among students who completed an SEL in fall 2008 and spring 2009, the average change in score on the Structural Analysis literacy domain was 21.1 scale score points.

Risk benchmarks. Both the SEL scale scores and reading classification levels are absolute measures of reading proficiency, that is, they are not adjusted for a child's school grade level or the time of the school year. So, while changes in the scores indicate the degree of literacy growth, they do not tell us whether children participating in the Emergent Reader program are making gains that are greater than would be expected from all students given the additional months in school. The Risk Level Benchmarks suggested by the assessment publishers do, however, take into account the expected growth in literacy skills, and thus allow the estimation of whether the gains observed among participants are greater than expected.

The percentage of program participants considered at risk increased slightly between the fall 2008 and spring 2009 administrations of the SEL (Exhibit 26). Twenty-four percent of participants were considered at risk in the fall and 25 percent in the spring. However, the percentage of students identified as having some risk of academic failure decreased significantly from 47 percent in the fall to 44 percent in the spring. The percentage of students identified as being on track or having low risk of academic failure significantly increased from 30 percent in the fall to 31 percent in the spring. This suggests that despite the low levels of reading proficiency on their initial SEL assessment, participants in the Emergent Reader program make gains consistent with a national sample of students in the same grade levels, and that the pace of gain in proficiency was slightly greater than that of the national sample of students. Taken together, these findings regarding change in risk benchmarks indicate that the Emergent Reader program may be more effective with students who are not initially rated at the highest risk level.

Exhibit 26
Percentage of Participants in Each Risk Benchmark on the Fall and Spring SEL

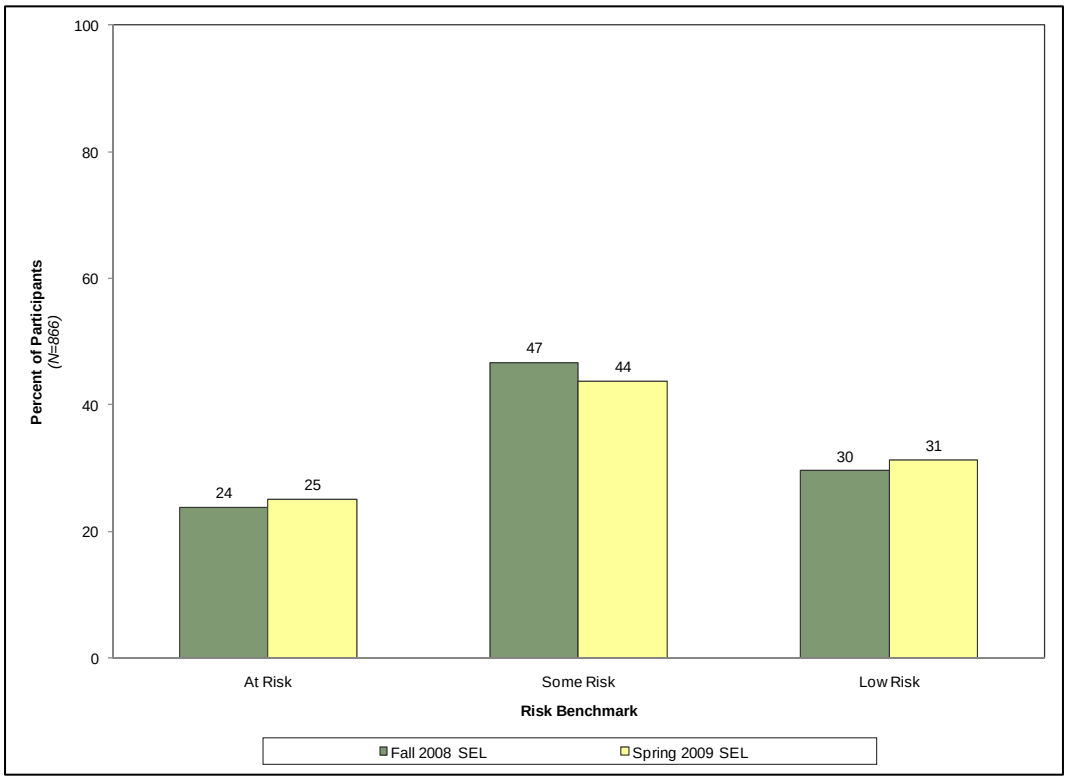


Exhibit reads: Twenty-four percent of participants were considered at risk on their fall 2008 SEL assessment, increasing to 21 percent on their spring 2009 assessment.

Thirty percent of the children who completed at least one SEL assessment made a transition to the guided independent reader program and completed at least one STAR Reading assessment during 2008-09. Forty-nine percent of the children who were in first grade who started the year in the Emergent Reader program made this transition during the school year.

References

Cooc, N., White, R.N., & Reisner, E.R. (2008). *The Literacy Programs of Save the Children: Results from the 2007-08 school year*. Washington, DC: Policy Studies Associates, Inc.

Decker, P., Mayer, D., & Glazerman, S. (2004, June). *The effects of Teach for America on students: Findings from a national evaluation.* Princeton, NJ: Mathematica Policy Research, Inc.

Finn, J., & Achilles, C. (1999). Tennessee's class size study: Findings, implications, misconceptions. *Educational Evaluation and Policy Analysis, 21(2)*, 97-109.

Jenner, E., & Jenner, L. (2007). Results from a first-year evaluation of academic impacts of an after-school program for at-risk students. *Journal of Education for Students Placed at Risk, 12(2)*, 213-237.

Kane, T. (2004, January). *The impact of after-school programs: Interpreting the results of four recent evaluations* (Working paper). New York: William T. Grant Foundation.

Renaissance Learning. (2007). *Understanding STAR Assessments.* Wisconsin Rapids, WI: Renaissance Learning.

Appendix
Using Normal Curve Equivalent Scores

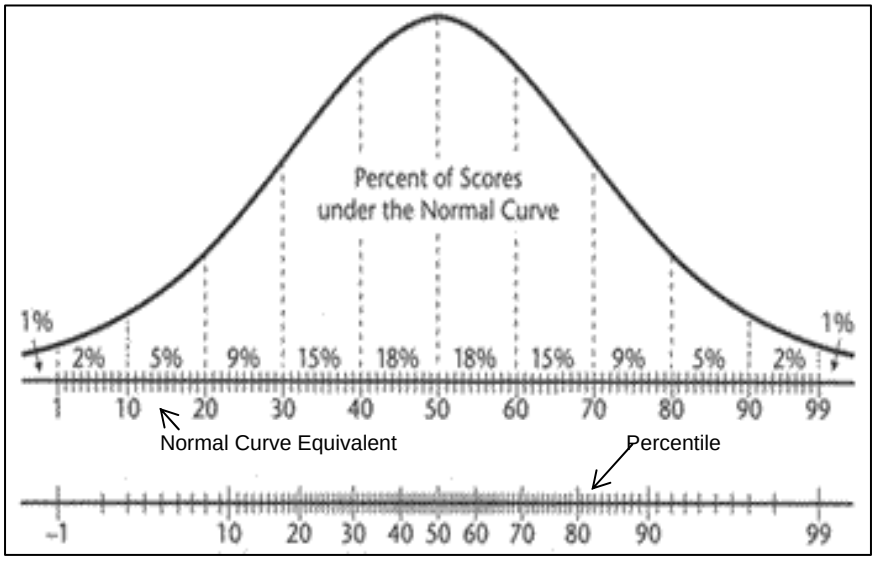
Using Normal Curve Equivalent Scores

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.

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



Children 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.