

United Way
of the Wine Country
Schools of Hope
Literacy Initiative
Evaluation Report

Fall 2015



Schools of Hope Evaluation

Prepared by Carlos Ayala and Stephanie Simon
Sonoma State University
Fall 2015

Contents

Executive Summary	5
Importance of Early Language and Literacy Development	5
Context and Origin of Schools of Hope	5
Schools of Hope in Sonoma County	5
Introduction	6
Schools of Hope Project Description	6
Schools of Hope Goals	6
Evaluation Plan	6
Methods Used	7
Participants	7
SOH Tutee Selection	8
Let’s Go Learn Study	8
Tutor Training Observation Study	9
Tutor Observation Study	9
Results and Discussion	9
Analysis of DIBELS Next Student Data	9
DIBELS Next assessment scores	9
Using Let’s Go Learn Scores for Tutee Selection	11
Correlational Analysis	11
Selection Criteria Analysis.....	13
Error Amelioration	16
Tutor Training Observation Study	18
Background	18
Training Session Structure.....	18
Observation Highlights	18
Training Observation Conclusions.....	19
Recommendations Based on Training Observation Study	19
Tutor Observation Study	19
Observation Findings	20
Teaching Reading.....	22
Conclusions.....	22
Recommendations based on Tutor Observation Study Only	22
Overall Evaluation Report Conclusion	23
Next Step Recommendations	23
References	24
APPENDIX 1: Executive Summary of Educational Reforms in California	25
New Standards and Assessments for California Students	25
Computer-Based Assessments for Math and ELA	25

What do these changes mean for the Schools of Hope program?	26
APPENDIX 2: California Department of Education Information Regarding the Smarter Balanced Assessments	27
APPENDIX 3: Tutor Observation Form	30
APPENDIX 4: Tutoring Training Checklist.....	32

Executive Summary

Importance of Early Language and Literacy Development

Early grade reading mastery is one of the best predictors of children’s success in school. Early language and literacy development plays a key role in supporting learning experiences that are linked with academic achievement (United Way’s Education Research Overview, 2011). Schools of Hope (SOH) intends to support reading in the primary grades so that 90% of Sonoma County’s third graders are reading at grade level. Third grade reading has been found to be a strong predictor for adult success. To date, over **2,000 tutors** have served 2,000 students in kindergarten, first, second, and third grades. The data collected as part of this project provides positive results of the efficacy of the program. Several indicators in the past three years point toward positive results, whereas, previously, we had inconclusive evidence about the direct link of the program to reading achievement. This suggests that the adjustments undertaken in the last three years may have been effective. The DIBELS Oral Reading Fluency and DIBELS Composite scores reflect increased reading achievement for tutees.

Context and Origin of Schools of Hope

The Schools of Hope program originated in Dane County WI with the stated goal to close the achievement gap. Despite 20 years and millions of dollars dedicated to the effort, an achievement gap persisted in the community. With community input, the United Way of Dane County began the Schools of Hope program in 1995. “Using community input, as well as hard data and research, the group determined to place volunteer tutors in schools to work in concert with teachers. Volunteer coordinators were placed in the schools with the most students in need of reading assistance” (unitedwaydanecounty.org, 2014). Within ten years, the third grade reading achievement gap was reduced from 12% to 2%. The Dane County program has been a model for other counties across the nation. United Way of the Wine Country used the Wisconsin model as a starting point for its own, county-specific model.

Schools of Hope in Sonoma County

In 2009, United Way of the Wine Country published *On Our Watch: Opportunity for All*, a call to action examining local education data and describing promising practices that could enable all children to achieve their full potential. Following the release of the report, United Way pulled together focus groups and working groups of local experts from six elementary school districts, the Sonoma County Office of Education, and other key non-profit partners to further develop a literacy initiative meeting our region’s unique needs. Based upon best practices from the Wisconsin model, Schools of Hope in Sonoma County launched in October 2010.

The Schools of Hope (SOH) program is an early reading intervention model for children who are struggling with reading. Using DIBELS Next composite scores, schools identify “strategic” readers who have the most potential to benefit from weekly one-to-one literacy tutoring.

Community volunteer tutors are trained by site coordinators at each school site with a lesson plan focused on basic reading skills (sounds, letters, high-frequency words, fluency, comprehension, and vocabulary). From October to May, tutors meet on a consistent basis with tutees in sessions that last 30 minutes, usually twice per week.

Introduction

Schools of Hope Project Description

A strong reading ability in third grade is a good predictor of high school graduation and college enrollment. For example, Lesnick, Goerge, Smithgall and Gwynne (2010) found that students who are “at or above” grade level for reading in third grade graduate and enroll in college at higher rates than students who are “at or below” grade level. We know that students who fall behind in school are more likely than students who are at grade level to be unemployed or to earn incomes below the poverty line as adults (Kutner et al., 2007) and more likely to be incarcerated (Harlow, 2003). Analysis of school teaching practices in third and fourth grade suggest that students stop learning to read in third grade and read to learn in fourth grade (Annie Casey Foundation, 2010). In Sonoma County, over 50% of third grade students are reading below grade level as measured by the California Standards Test (CST). Nearly 80% of third grade students who are also English Language Learners are reading below grade level.

Schools of Hope Goals

When the program began in 2010, 54% of Sonoma County third graders were not reading at or above grade level as measured by the 2009 California Standards Tests (CST). Last year 2014-15 third grade students 32% did not meet standard and 28% nearly met standard on the new Smarter Balance exam (SBAC). Since SBAC is a different exam than the CST, it is unwise to make comparisons to the previous CST tests especially since the SBAC break up the students in different categories than the CST. We should expect that as K-12 curriculum changes to align with SBAC, all scores will go up. A major goal of the Schools of Hope has been that by 2020 90% of third grade students will demonstrate the knowledge and skills in English language arts/literacy needed for likely success in future coursework. Given SBAC’s abilities levels, it would be important for the SOH policy organization to rearticulate the literacy goals in terms of the new achievement levels.

Evaluation Plan

The School of Education at Sonoma State University (SSU) will serve as External Evaluator for United Way of the Wine Country’s Schools of Hope (SOH) program with a focus on the literacy component. The overall evaluation plan will focus on the progress toward the program’s goals and propose reasonable adjustments to ensure the effectiveness and integrity of the program’s outcomes. Therefore, the overall evaluation plan focuses on the program goal of impacting

primary students' reading abilities through a tutoring intervention with K-3 students in Sonoma County schools.

This evaluation report is based on the following questions:

- 1) Who are the SOH students?
- 2) Do students who participate in the study learn to read better than students who do not participate in the study?
- 3) What do we know about the tutors and the quality of their tutoring?

Methods Used

Participants

In order to determine the efficacy of the SOH program, student performance of students who participated in the program (SOH students) will be compared to students at schools that do not have SOH tutors (non-SOH students). In order to select a comparison group with similar characteristics to the treatment group, we applied the same criteria used for Schools of Hope student selection to students outside of SOH programs. We refer to this method as students who are in the model. We also remove from the model any students with missing data. Some SOH students fall outside of the model. Table 1 contains information about the comparison groups across grade levels. It is important to note that the percent of English Learners in the SOH group is significantly larger than the percent of students in the comparison groups across all grade levels. It is also important to note that the percent of students participating in the Free and Reduced Lunch Program in the SOH group is significantly larger than the percent of students in the comparison groups across all grade levels. Moreover, it is important to note that the 2014-15 reading scores for students who are in model but not in the SOH is about the same as the students who were selected for the SOH program.

Table 1: Description of 2014-15 SOH Tutees and Comparison Group (Students in the model)

Grade Level and Group 14-15		Reading Score ¹	% FRLP ²	% English Learner	% Special Education	Gender (F/M)
First Grade	SOH	94	65³	63⁴	4	54/46
	Comparison Group	94	48	34	7	46/54
Second Grade	SOH	117	58⁶	63⁵	0	56/44
	Comparison Group	115	49	31	5	49/51
¹ Average DIBELS Next Composite Score: DIBELS Next comparison score is used to determine program eligibility. For 1 st grade the value of 73 and 112 are used, while for 2 nd grade the value of 82 and 140. Calculated values for 1 st grade composite scores were used rather than the reported values. ² Free and Reduced Lunch Program Eligible ³ p = 0.00 ⁴ p = 0.00 ⁵ p = 0.00 ⁶ p = 0.00						

SOH Tutee Selection

Who gets SOH support? An evaluation committee made up of SOH educators and evaluators decided that students’ DIBELS Next composite scores would be used to determine whether or not a student gets tutoring services. Based on pre-literacy skills, a student’s DIBELS Next score can be ranked at three levels. The highest ranking is *At or Above Benchmark* and those students need “Core Support.” The next level down is considered *Below Benchmark* and those students need “Strategic Support.” The lowest level is considered *Well Below Benchmark* and those students need “Intensive Support.” The SOH program targets the students who need “Strategic Support,” as well as the top 25% of those who need “Intensive Support.” The committee hypothesized that those needing “Strategic Support” and the top 25% of “Intensive Support” would benefit the most from tutoring, and students who score lower need additional support by professionals already in place at the school. Therefore, the SOH selection level for first grade is a Beginning First Grade Composite Score between 73 and 112 inclusive and a Beginning Second Grade Composite Score between 82 and 140.

Let’s Go Learn Study

As California transitions to new assessment tools to measure state standards, the Schools of Hope program will also need to adapt to K-3 changes in assessments. Some schools have already begun using Let’s Go Learn (LGL) in place of DIBELS Next as their primary reading assessment tool. This year’s report includes a study of the correlation between the two assessments (see full analysis in the next section, Results and Discussion).

Tutor Training Observation Study

Observations were made of trainings conducted by site coordinators for the volunteer tutors. The trainings were held at the school sites during September and October 2014. Using a pre-established protocol (see Appendix 3), SSU faculty and United Way staff conducted observations of the training sessions to assess the quality and consistency of instruction where the Schools of Hope program is in place (see Appendix 4).

Tutor Observation Study

In order to determine the efficacy and implementation of the program, observations were made in fall 2014 and spring 2015 of volunteer tutors conducting sessions with their tutees. Trained observers from SSU and United Way sat passively near tutor/tutee pairs while conducting their observations using a pre-established protocol (see Appendix 3).

Results and Discussion

Analysis of DIBELS Next Student Data

During the 2011-2012 school year, data was reported for 170 SOH students who had complete data sets. In 2012-2013, data was reported on 323 SOH students across four grade levels with the number of students for each of the three scoring sessions. For 2013-14, a total of 188 students had complete data sets. This year we had 284 SOH students with useable data across first and second grade. **While in previous years we reported out DIBELS Oral Reading Fluency only, this year the results continue to be positive and can be found in others areas of the DIBELS assessments.**

DIBELS Next assessment scores

This year the evidence is not so clear that SOH is resulting in higher gains in reading compared to non-SOH students as measured by the DIBELS Next scores. Based on the results of the DIBELS Next reading scores, students in second grade benefited from the SOH tutoring program (Table 3); however, this year we see a negative effect on gain scores for first graders (Table 2). However, we have found evidence of a SOH benefit over the last four years with significant results since the 2012-2013 school year. Last year the robustness of the data and conclusion were very strong and we found the positive results exactly where we expected them to appear. This year, however, data are less robust, more inconsistent and required additional cleaning steps to make the data useful for analysis. The process that we use in this analysis is to combine three data sets together sent by the districts. We combine student demographic information, DIBELS Next Data and SOH implementation data (attendance etc.). We use the student's CALPADS identification number to link the data sets. The lack of robustness becomes apparent when we combined these files together. For example some districts did not report using the CALPADS number and reported only district IDs for students. Other districts identified a single student

with two sets of DIBELS Next scores and other districts identified a student at more than one school site. Finally, in the first grade comparison group data there were remarkable gains in non-SOH student scores leading us to question the testing methodology. All of this inconsistency in data throws into question this year's data sets.

Therefore given this year's results and considering the results from the last few years, we are inclined to suggest that this year's data muddled the results rather than seeing a new trend in the data and we believe that SOH still closes the achievement gap for students.

Table 2: DIBELS Composite Scores for First Grade

DIBELS Scores	SOH	N	Mean	Std. Deviation
Gain Score Beg-End Students in model only	Non-SOH	366	54.1	66.1
	SOH	192	50.2	62.7
Gain Scores Beg-End All Students	Non-SOH	1450	65.2*	60.5
	SOH	266	50.0*	60.2

*P = .00 for both homogeneity of variance assumed or not.

Table 3: DIBELS Composite Scores for Second Grade

DIBELS Scores	SOH	N	Mean	Std. Deviation
Gain Score End - Mid Students in model only	Non-SOH	186	41.6	49.1
	SOH	94	42.8	50.0
Gain Score Beg-Mid Students in model only	Non-SOH	199	41.0	55.5
	SOH	94	47.3	46.5
Gain Scores Beg-End All Students	Non-SOH	1372	71.7*	52.4
	SOH	92	87.0*	60.0

*P < .05 for both homogeneity of variance assumed or not.

Another way to considering the data is to look at the number of students who are proficient in Schools of Hope at the three testing points. Across both 1st and 2nd grade we see huge growth of the percent of SOH students who become proficient over the course of academic year Table 4. Schools of Hope help closes the gap between those who are proficient and those who are not.

Table 4: Comparison of SOH students and nonSOH students who are proficient across testing periods.

2014-15 First Grade Composite	% At or above benchmark Fall	% At or above benchmark Mid	% At or above benchmark Spring
SOH	13.5	31.5	38.6
nonSOH	37.4	43.3	40.5
2014-15 Second Grade Composite	% At or above benchmark Fall	% At or above benchmark Mid	% At or above benchmark Spring
SOH	21.3	26.6	31.9
nonSOH	51.2	46.8	44.2

Using Let's Go Learn Scores for Tutee Selection

The purpose of this section of the report is to explore how LGL might be used for student selection and project evaluation instead of DIBELS Next. Several districts are already moving from DIBELS Next to LGL for reading assessments. If this continues and more districts make this change, we ask them to use LGL to select and track students.

There are two methods that I will use to explore this relationship. A caveat as we move forward is that the developers state that LGL is not a tool that was designed for diagnostic purposes beyond giving feedback to teachers on what is best to teach students next based on a computer assessment. Since we are using LGL as a diagnostic tool to select students, the purposes are somewhat aligned; we are using the tool for the purpose of what do to next with the student. Our first exploration will be to look at the correlations between LGL and DIBELS followed by an exploration across grade levels to see if there are differences in the relationship across grades. The second exploration will be to use the SOH full selection model of identifying tutees and see which of two methods of using LGL might be the best selection method. Finally, we must be very cautious about this method and explore it fully as we progress. Standard setting is a very difficult procedure requiring extensive research and evaluation. We are use a relative rough estimation here and consider ways to reduce false selections and false exclusions in our procedures.

Correlational Analysis

Lets Go Learn and DIBELS Next scores for 123 first and second graders reveals the relationship between the scores (See Table 5). The LGL administration dates range from 1/15/15 to 2/11/15 while the DIBELS Next administration dates are late Sept to early October for the DIBELS Beg and late January to early February for the DIBELS Mid. Since SOH uses the DIBELS Next composite score for tutee selection, we should use the parallel composite score in LGL (Weighed Score). The correlations are remarkably strong and significant for the all the measures of LGL

and for either Beg or Mid DIBELS scores. Based on this data alone, further exploration is warranted.

Table 5: Correlation between DIBELS Next and LGL scores.

N=123	DIBELS Beg	DIBELS Mid	HF	WR	PH	PA	Weighted Score
DIBELS Beg	1	.619**	.503**	.659**	.620**	.359**	.706**
DIBELS Mid	.619**	1	.518**	.520**	.616**	.340**	.660**
LGL HF	.503**	.518**	1	.526**	.620**	.421**	.744**
LGL WR	.659**	.520**	.526**	1	.776**	.399**	.880**
LGL PH	.620**	.616**	.620**	.776**	1	.449**	.899**
LGL PA	.359**	.340**	.421**	.399**	.449**	1	.523**
LGL Weighted Score	.706**	.660**	.744**	.880**	.899**	.523**	1

Since we have both first and second graders in the data set, we investigate them separately. Upon review of the data for first graders, the correlation between LGL and DIBELS scores is very strong between the LGL scores and DIBELS Mid scores (See Table 6). We would expect to see a stronger relationship at the middle of the year since both exams were given at about the same time (see above). Based on these correlations, the tests are measuring the same domain or content. We often substitute one test for another when the scores are in the .6 range. The same is true for the second grade scores, with a significant and strong correlation at the middle of year Table 7.

Table 6: LGL and DIBELS Score Correlation Grade 1

Grade 1/ N =70	DIBELS Beg	DIBELS Mid	HF	WR	PH	PA	Weight Score
Beginning	1	.634**	.264*	.370**	.297*	.140	.492**
Middle	.634**	1	.600**	.647**	.701**	.269*	.831**
LGL HF	.264*	.600**	1	.434**	.502**	.432**	.702**

LGL WR	.370**	.647**	.434**	1	.794**	.294**	.842**
LGL PH	.297*	.701**	.502**	.794**	1	.326**	.882**
LGL PA	.140	.269*	.432**	.294**	.326**	1	.451**
LGL Weighted Score	.492**	.831**	.702**	.842**	.882**	.451**	1

Table 7: LGL and DIBELS Scores Grade 2

Grade 2/ N =50	DIBELS Beg	DIBELS Mid	HF	WR	PH	PA	Weight Score
Beginning	1	.878**	.575**	.673**	.733**	.442**	.751**
Middle	.878**	1	.617**	.660**	.747**	.503**	.781**
LGL HF	.575**	.617**	1	.412**	.573**	.253	.651**
LGL WR	.673**	.660**	.412**	1	.694**	.392**	.867**
LGL PH	.733**	.747**	.573**	.694**	1	.487**	.872**
LGL PA	.442**	.503**	.253	.392**	.487**	1	.521**
LGL Weighted Score	.751**	.781**	.651**	.867**	.872**	.521**	1

Selection Criteria Analysis

In this section we explore what would happen if we applied the SOH selection criteria to the weighted scores. How can we use the weighted scores to match the DIBELS selection criteria? Since the correlations are strong we would expect very tight football shaped scatter plots (

Figure 1 and Figure 2).

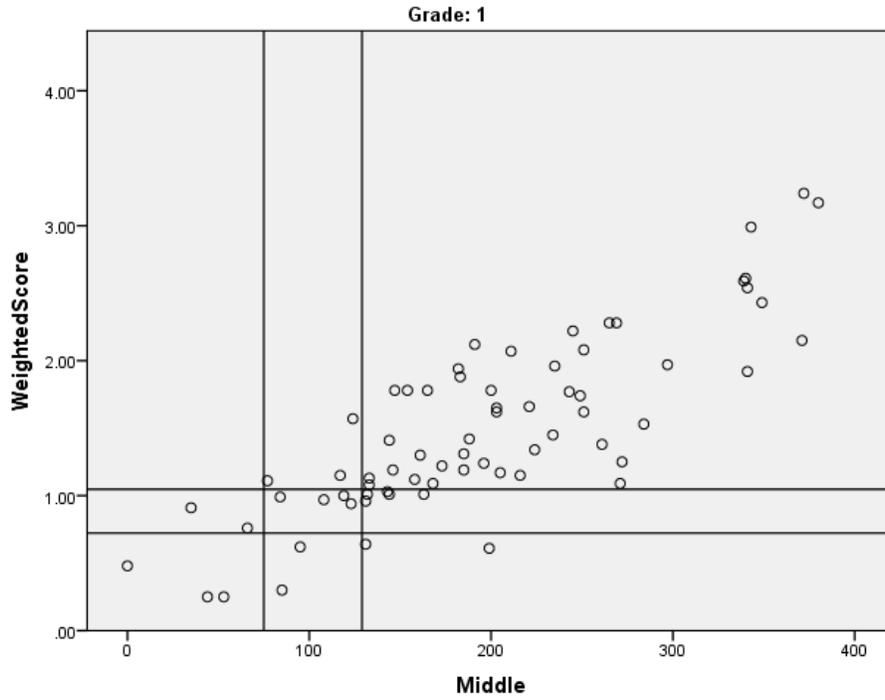


Figure 1: Scatter Plot First Grade DIBELS and LGL scores.

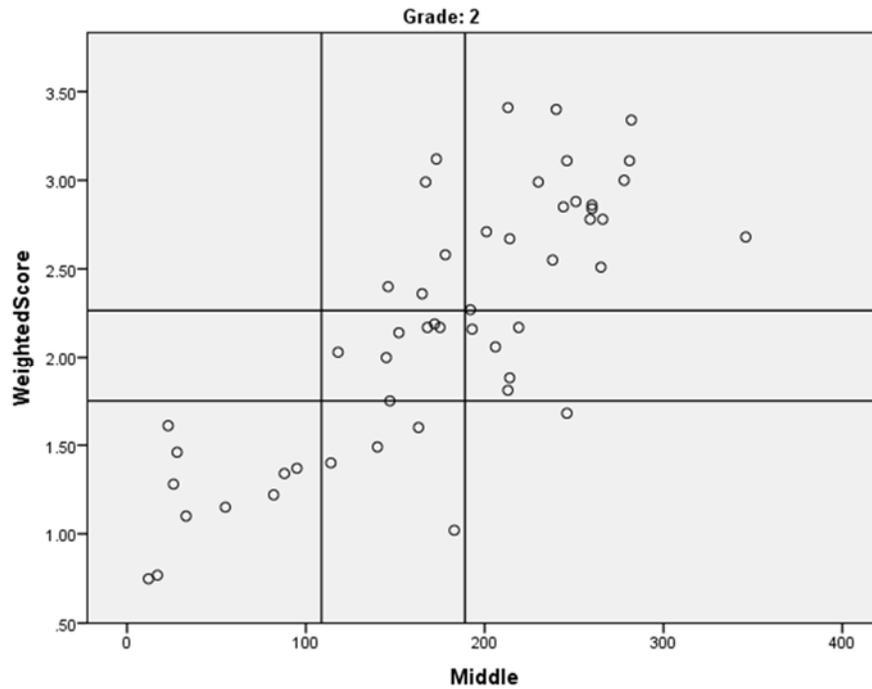


Figure 2: Scatter Plot Second Grade LGL and DIBELS Data

Given that the correlations are strong and the scatter plots are football shaped, we can hypothesize that the LGL can be used to estimate a DIBELS score and most importantly that we can provide a standard for student selection by using a LGL score. A regression analysis reveals coefficients that might be used for the selection of the students (Table 8)

Table 8: Coefficient parameters for estimating DIBELS scores from LGL scores.

Coefficients^a

Grade	Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	
		B	Std. Error	Beta			
1	1	(Constant)	.273	.108		2.515	.014
		Middle	.006	.001	.831	12.296	.000
2	1	(Constant)	.942	.152		6.188	.000
		Middle	.007	.001	.791	9.051	.000

a. Dependent Variable: Weighted Score

Based on a linear regression analysis above, we can estimate the value of the Weighted LGL score from a DIBELS Next Score. The estimations for these purposes are problematic because the error in both of the tests is amplified by this estimation; however, since this is a relatively low stakes situation, teacher input can be used to ameliorate problems, and students will be retested three times per year, we can identify students using LGL scores. Moreover, it will be important to reevaluate our hypothesis about the estimate with more data as we proceed forward, especially the data where we look at students' growth across the year with both LGL and DIBELS scores. Given the coefficients above, we can create the equations below for tutee selection based on DIBELS scores for students in the strategic group and 25% of the students in the intensive group. These are the criteria that we have used successfully before.

$$WS_{1st} = .006 \text{ Mid} + .273$$

$$WS_{2nd} = .007 \text{ Mid} + .942$$

Table 9: DIBELS and Let's Go Learn Comparisons

Beginning of Year Selection Criteria		DIBELS Scores	Estimated LGL scores
1st Grade	Lower Strategic Plus		
	Top 25% of Intensive	72.75	0.710
	Upper Strategic	112	0.945
2nd Grade	Lower Strategic Plus		
	Top 25% of Intensive	81.75	1.514
	Upper Strategic	140	1.922
Middle of Year Selection Criteria		DIBELS Scores	Estimated LGL scores
1st Grade	Lower Strategic Plus		
	Top 25% of Intensive	75.00	0.723
	Upper Strategic	129.00	1.047
2nd Grade	Lower Strategic Plus		
	Top 25% of Intensive	108.75	1.703
	Upper Strategic	189	2.265

Error Amelioration

It is important to remember that the values provided here are estimates based on a relatively small amount of data; we used middle of the year scores to estimate beginning of the year scores; we find the first grade DIBELS scores to be higher than expected; and the standard error from both assessments is embedded in our estimation. We should consider other sources of data when we select students for tutoring. Moreover, DIBELS has been researched for many years and the selection of the strategic group has been thoroughly supported in practice. There are three suggestions to help ameliorate the error in student selection. First, listen to teachers. Teachers can help identify students who might be too low or too high for the tutoring session. If a teacher feels strongly that a student is incorrectly placed, then follow his or her advice. Second, we must

retest the students at the middle of the year and reevaluate our decision there and track the growth of these students carefully. Again we are hoping to see tutee reading scores go up. Finally, we must make sure that the LGL scores are complete. If a student has missing data in his or her scores, then that student must be retested before the decision can be made about his or her selection.

Tutor Training Observation Study

Background

SSU and United Way staff observed six training sessions in September and October 2014. The trainings were conducted by each school's site coordinator(s), who had prior training and preparation by United Way staff. Each site's volunteer tutors had a choice of attending one of two offered sessions, given on different days. The trainings were held in classrooms, faculty lounges, or offices at the school sites.

Training Session Structure

Site coordinators followed a prescribed format based upon their training by United Way staff. However, each site coordinator varied the format with various pedagogical techniques, such as modeling a book walk, hands-on practice, power point presentations, and videos. The trainings lasted between one and two hours. All trainings were held at the school sites, but training locations varied from classrooms to offices to faculty lounges. Of the six trainings observed, three of the instructors were new, while the other three had conducted prior SOH trainings in previous years.

Observation Highlights

-All site coordinators presented aspects of the theoretical literacy concepts behind the Schools of Hope curriculum. Of the six site coordinators observed, four closely followed the training curriculum, using either a power point presentation of the material or actually handing out the training manual to the tutors while leading the tutors, page by page, through the manual. Two coordinators (both experienced) focused more on their own programs and merely skimmed the SOH training curriculum. Some site coordinators called upon experienced tutors to share observations or tips with new tutors.

-The level of interaction among tutors and the use of hands-on practice of concepts varied widely among the presentations. One site coordinator had the tutors practice each of the six aspects of the SOH curriculum. Another showed a video of a tutoring session. A third used the site's tutoring notebook to lead tutors step-by-step through a typical tutoring session. It appears that when training sessions used the techniques described, the level of engagement and questions by tutors went up and most likely improve tutor knowledge and skills.

-One tutor presented background information during her introduction about the Schools of Hope program and its efficacy. This information was well received by the tutors and relating this information to other tutors set an important tone for the tutors' future work with the kids.

Training Observation Conclusions

The structured training curriculum is being followed by site coordinators. All six schools incorporated the SOH model for organizing training sessions. Every school addressed, to one degree or another, the reading skills outlined in the training (phonemic awareness, phonics, high-frequency words, fluency, comprehension, and vocabulary/background knowledge). In addition, all six schools hit nearly every bullet point/detail under the “typical tutoring day,” logistics,” and “working with a 3rd grader” sections of the training manual.

Recommendations Based on Training Observation Study

- 1) New tutors seemed to benefit from the experience and advice of seasoned tutors. SOH trainers should incorporate best practices of experienced tutors. United Way could support this approach by tapping into long-term volunteers for their expertise, creating materials to be shared throughout the network of partner schools (such as YouTube videos, interviews, case studies, etc.).
- 2) The level of engagement by tutors went down during the presentation of theoretical concepts about literacy and went up during question/answer sessions and any activity that was interactive or hands-on. Varied pedagogical techniques should be encouraged for the trainings, such as role-playing, hands-on practice with the materials, and site coordinators modeling good teaching.
- 3) During training of the site coordinator instructors, there should be more focus on the structure of a typical tutoring session, including how to align tutoring activities with the “three-part tutoring model.”
- 4) SOH should consider adding a brief reminder about the tutors’ value to SOH in the training introduction. This information would set a positive tone for the tutors, show the importance of their work and their time, and help them feel valued—all goals that United Way has for its volunteers.

Tutor Observation Study

In order to determine the efficacy of the program and the training tutors received, we carried out observations of tutors conducting sessions with their tutees. During fall 2014 and spring 2015, we conducted 27 observations of schools in various cities throughout the SOH service area in Sonoma County. SSU faculty and United Way staff used an observational checklist and reporting tool.

An observation protocol was developed in 2012 based upon open-coded observations and the list of practices and procedures given to tutors during their initial training. The observation goal was to look for patterns in the structure and practices of Schools of Hope tutoring. The data from the 2012 observations combined with the program director’s list of activities resulted in the refinement of the observational checklist (see Appendix 2).

Twenty-seven observations were conducted at eight different schools throughout Sonoma County. With multiple observers and back-to-back sessions, more than one observation was done at every school, and one school had a repeat visit. School selection for the observations was made with an eye toward diversity of demographics, location, school size, and new versus established Schools of Hope programs. Observations were conducted between November 2014 and April 2015. From the observations conducted in fall 2014 and spring 2015, several patterns arose from the data relating to the site organization and management, session structure, and reading instruction.

Observation Findings

Site Organization and Management

Tutoring sessions occurred in a variety of locations at school sites, including SOH-dedicated classrooms, central workrooms in the middle of classroom buildings in “pod-style” schools, a faculty lounge, an office, and a room off of the school library.

All schools had a central organizational location for SOH with bulletin boards devoted to tutor communication, folders for each student, and reading materials for tutoring sessions. Some site coordinators created spaces that were visually appealing and well-organized, making the space easy for tutors to navigate and utilize. Upon arrival, all tutors used the space to retrieve their students’ binder and reading materials. In addition, they logged in using the site coordinators’ system and proceeded to set up for the tutoring session. At the conclusion of each tutoring session, tutors again used the record-keeping system to make notes and returned the binders and reading materials.

Session Structure

Tutors arrived early for their sessions and used the established check-in routines at the front office and in the tutoring locations. All tutors appeared to know established routines and processes. Although there is much variation (to be discussed in detail in the time analysis, below), most tutoring sessions followed a typical pattern:

1. Transition time (arrival, greeting, opening books and folders, etc.)
2. Phonics and sight word practice
3. Reading
4. Other activities (writing, discussion, games, record-keeping, etc.)

In analyzing the approximate time spent on each instructional component during tutoring sessions, observers logged the start and end times of each portion of the session. Categories specified on the observation protocol include: phonics, sight words, reading, writing, questions/discussion, and other (See Figure 3).

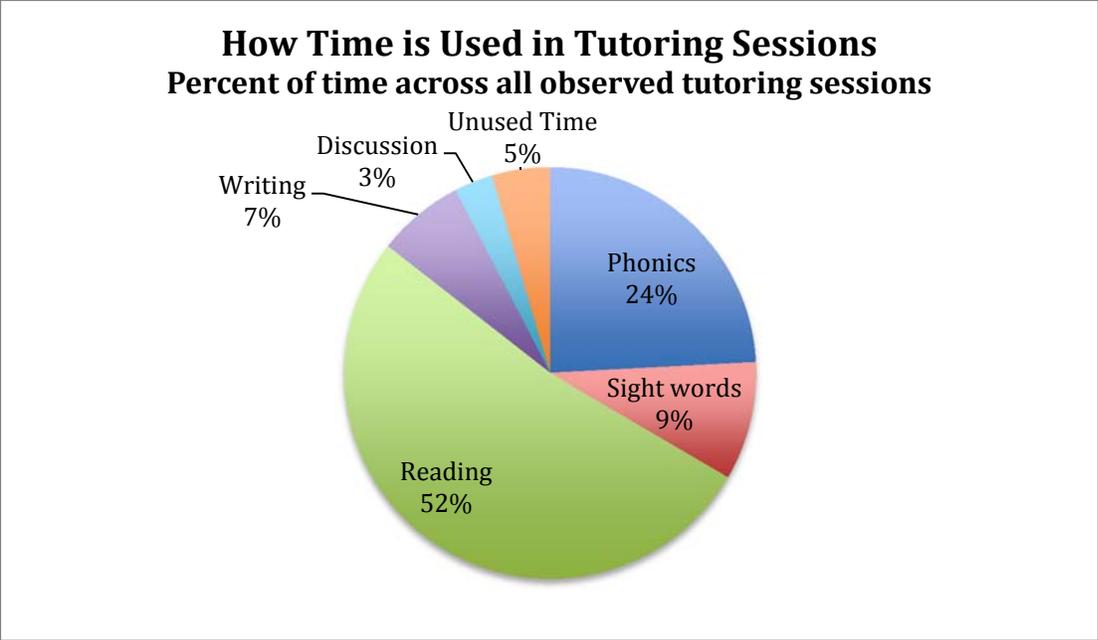


Figure 3: Distribution of how tutoring time is used as a percent of total time across all observations.

The number of minutes and percentage of total time from this year’s observations are as follows:

Table 10: Distribution of minutes observed in all tutor observations.

	Minutes used	Percent of time used
Phonics	159	20%
Sight words	61	8%
Reading	343	44%
Writing	45	6%
Questions/discussion	20	3%
Unused Time	30	15%

In addition, it should be noted that 117 minutes of the total time, defined by SOH allotted time of 30 minutes per session, was non-instructional time spent in transition or for late arrival/early dismissal from the sessions. This represents 15% of the total time. Observers logged the time spent on each activity. From the total time dedicated for all sessions, 15% was non-instructional time (27 observations x 30 minute sessions=810 minutes. Subtracting the total observed minutes—693—from that total leaves 117 minutes of non-instructional time).

Teaching Reading

Students: Almost all students were observed to be engaged in their tutoring sessions. They were described as “excited,” “enthusiastic,” and “interested.” Only two students were described as “bored” and “uncooperative.” Overall, they listened attentively and interacted with their tutors by responding and giving reading details. Few students were observed asking questions. All students used one or more reading strategy, and many students used multiple strategies throughout a session, such as: sounding out, rereading, using context, pausing/hesitating, and skipping.

Tutors: During skills practice, tutor interactions included giving examples and hints for unknown sight words, praising, correcting, allowing for wait time, and making it a game (“Find the three rhyming words”). During reading practice, tutors employed the following techniques: pre-reading conversation (“What do you think it means to have a whale of a problem?”), discussing the author or story content, linking the story to previous stories read during a session), stopping to summarize/discuss plot/check for understanding, using illustrations to predict plot or to connect with challenging words, offering hints, tracking reading with a pencil, defining vocabulary, partner reading, and stopping to correct pronunciation. Also during reading, we observed tutors praising and pointing out grammatical concepts (“I like how you paused when you saw commas, “That’s a compound word.”), using wait time, and emphasizing the positives of reading (“Reading should be fun!” “You should be really proud of yourself!”).

Conclusions

Based upon the variety of techniques employed by tutors during the skills and reading practice, tutors appear to be well-trained in the structure, sequence, and purpose of the tutoring at their school sites. Based upon the sites’ visual appeal, organization, and utilization by tutors of materials and supplies, the systems and processes seem to work efficiently.

Recommendations based on Tutor Observation Study Only

- 1) Whenever possible, encourage tutors to maximize their session time so that tutees receive full use of their 30-minute sessions. Encouraging a full 30-minute session will also increase fidelity to the program across school sites
- 2) Session locations can affect session noise level, ability of tutor to avoid being alone with a student, and distractions. Whenever possible, more attention and creativity should be employed to provide a quiet environment where tutors can work together to comply with insurance requirements and that still has minimal distractions.
- 3) Session time/activities should be more closely aligned with stated goals (three-part session). Some tutors used the entire 30-minute session to read with a student, and other tutors used the majority of the time for phonics and skills practice. Work with tutors to create balanced sessions that adhere more closely to the three-part tutoring model.

Overall Evaluation Report Conclusion

Based on the data analyzed, we find mixed evidence of the efficacy of the Schools of Hope program. However, given both the positive and negative results this year and given the significant results in the previous two years, we believe that the SOH is moving students to grade level proficiency in 3rd grade faster than just regular schooling.

Next Step Recommendations

- 1) Address the variation in tutor trainings and tutoring sessions among the school sites.
- 2) Use best practices from experienced site coordinators and tutors to fine tune and enhance the program. Many of them now have multiple years of experience with SOH and can share their best practices.
- 3) Seek strategies to maximize the 30-minute instructional timeframe for sessions so that students consistently and fully receive all aspects of the SOH curriculum.
- 4) Highlight the three-part tutoring model in the site coordinators' trainings. This seems to be a key point in the intended overall structure of tutoring sessions, but site coordinators and tutors do not seem to be aware of it. Breaking the tutoring sessions into three overall curriculum goals makes sense and is a logical, easy-to-follow benchmark for tutors; it should be effectively communicated to them during trainings and advanced tutor trainings.
- 5) Begin the preliminary analysis of the Smarter Balanced assessment and SOH tutees.
- 6) Reevaluate the Let's Go Learn tutor selection with SOH data.

References

- Allington, R. (2013) What works for struggling readers, *The Reading Teacher* 66(7), 520–530
- Almasi, J. F. Garas, K., & Shanahan, L. (2005). Qualitative research and the report of the National Reading Panel: No methodology left behind?, *Elementary School Journal*.
- Annie E. Casey Foundation (2010). Kids Count Special Report: Early Warning! Why Reading by the End of Third Grade Matters.
- D'Agostino, J. V. and J. A. Murphy (2004). "A meta-analysis of Reading Recovery in United States schools." *Educational Evaluation and Policy Analysis* 26(1): 23-38.
- Guthrie, J. T. (2004). Teaching for literacy engagement, *Journal of Literacy Research*, 36 (1), 1-28.
- Guthrie, J. T. and N. M. Humenick (2004). Motivating students to read: Evidence for classroom practices that increase motivation and achievement. *The Voice of Evidence in Reading Research* P. McCardle and V. Chhabra. Baltimore, Paul Brookes Publishing
- Harlow, C. (2003). Education and correctional populations. Bureau of Justice Statistics Special Report. Washington, DC: U.S. Department of Justice.
- Lesnick, Joy, Robert M. Goerge, Cheryl Smithgall, and Julia Gwynne, (2010) Reading on Grade Level in Third Grade: How Is It Related to High School Performance and College Enrollment? A Longitudinal Analysis of Third-Grade Students in Chicago in 1996-97 and their Educational Outcomes. Report to the Annie E. Casey Foundation.
- Kamil, M. L. (2004). Vocabulary and comprehension instruction. In P. McCardle and V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 213-234).
- Kutner, M., Greenberg, E., Jin, Y., Boyle, B., Hsu, Y., and Dunleavy, E. (2007). Literacy in Everyday Life: Results From the 2003 National Assessment of Adult Literacy (NCES 2007-480). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
- McKenna, M.C. & Kear, D. J. (1990). Measuring attitude toward reading. A new tool for teachers. *The Reading Teacher*, 43(9), 626-639.
- “Moving Beyond the Bubble.” (California State PTA). Web 15 August 2015.
- Pressley, M. (2002). Improving comprehension instruction: A path for the future. In C.C. Block, L. Gambrella, I M. Pressley (Eds.) *Improving Comprehension Instruction* (pp. 385-399). San Francisco, Jossey-Bass.
- Priciotta, Dan. (2011). United Way Education Research Overview.
- “Testing and Accountability.” (CA Department of Education). Web 15 August 2015.
- Tierney, R. J. and T. Shanahan (1991). Research on reading-writing relationships: Interactions, transactions and outcomes *Handbook of Reading Research*. R. Barr, M. Kamil, P. Mosenthal and P. D. Pearson. New York, Longman. 2: 246-280.

APPENDIX 1: Executive Summary of Educational Reforms in California

New Standards and Assessments for California Students

The state of California--like many other states nationwide--is in the midst of a extensive educational reform process that will affect the curriculum, pedagogy, and assessment of our students. Every district, school, teacher, and student is currently preparing for these changes, and the purpose of this executive summary is to put the Schools of Hope program in the context of this reform, as well as to anticipate how the 3rd grade assessment changes will affect the program directly.

In 2010, California adopted new standards for math and English/language arts (ELA) at all grade levels, K-12. The standards are known as the “Common Core State Standards” (CCSS) and are intended to prepare students for “college and career readiness”.

Hand in hand with the new standards comes a process for assessing them. California has developed the “Smarter Balanced Assessment System,” which is a testing program for all grades, K-12, with particular attention to the following grades: 3rd-8th and 11th, at which point students will take a “summative assessment” at the end of those grades. Smarter Balanced utilizes computer-adaptive tests and performance tasks that allow students to show what they know and are able to do in math and ELA.

Computer-Based Assessments for Math and ELA

While some of the former tests will be used to assess other curricular areas, such as science, the new Smarter Balanced Assessment System will help determine what students in grades 3-8 and 11 have learned in math and ELA. It is a long list of specific skills, knowledge, and thinking that students are expected to know, but Schools of Hope focuses on reading in the primary elementary grades, K-3. In that context, the list below outlines the four categories of reading standards for kindergarten through 3rd grade for literature, informational text, and foundational skills. A complete chart of these standards can be found in the Appendix of the Standards.

1. Key Ideas and Details
2. Craft and Structure
3. Integration of Knowledge and Ideas
4. Range of Reading and Level of Text Complexity

The new Smarter Balanced Assessments are computer-based tests that will replace the paper and pencil tests of the past. Some schools have been piloting these tests, and all California public schools will begin using them in the spring of 2016. The tests are “adaptive,” meaning that the questions students answer will vary depending on correct or incorrect responses given. The

computer-adaptive section includes a range of items, such as selected response, constructed response, table, fill-in, graphing, etc.

In addition to the computer section, a performance task will include extended activities that measure a student's ability to integrate knowledge and skills across multiple standards. The test will be administered in the spring, but each school and district will have flexibility to give the test within a 7-week window of time.

What do these changes mean for the Schools of Hope program?

The Schools of Hope program encompasses Kindergarten through 3rd graders; however, most services are in place for 1st and 2nd grade classes at the partner schools. The Smarter Balanced Assessment will take place for the first time in the spring of a student's 3rd grade year. How does the timing of the new state test affect Schools of Hope and its students? While some circumstances are evolving, the effects can be predicted with a high level of certainty, and the Schools of Hope team should have a plan in place to adjust to the curricular, pedagogical, and assessment changes that will occur beginning in spring of 2016.

The ways in which the new CA state standards and the Smarter Balanced Assessments will change K-3 classrooms include the following:

1. A tangible measure of a student's reading ability at the end of 3rd grade, which marks the time a student shifts from "learning to read" to "reading to learn." The goal of the Schools of Hope program is for all students to read at grade level by the end of 3rd grade.
2. New, specific standards for all K-3 classes, which will impact curriculum (text selection, reading genres, types of assignments given) and pedagogy (teaching approaches, types of student learning).

A direct impact of these changes is the assessment tool used not only to gauge students' reading progress but also to identify students for the program. Schools of Hope is considering shifting this tool from DIBELS ("Dynamic Indicators of Basic Early Literacy Skills," a reading assessment tool developed by the University of Oregon in the 1960's and used widely by California elementary schools) to Let's Go Learn, which is aligned with the new CA state standards and Smarter Balanced Assessments. The anticipated Smarter Balanced Assessment changes and resulting questions for SOH will need to be addressed beginning in the 2015-16 school year.

APPENDIX 2: California Department of Education Information Regarding the Smarter Balanced Assessments

CAASPP Communication FAQs About Smarter Balanced

Addresses common questions that are asked about the Smarter Balanced Assessments including how they were developed and what the tests measure.

The following common questions and answers may be useful as California educators communicate about the Smarter Balanced Assessment System with students, parents, the media, local school boards, and the general public.

Part of a Comprehensive Plan for Student Success

Knowing that public schools are preparing students for the challenges of the future, California has developed a comprehensive plan for high-quality teaching and learning in every school. We have a long way to go, but our work is well under way, with higher academic standards, more decision-making in the hands of schools and communities, and more resources dedicated to schools and to students with the greatest needs.

Gradually, we're providing more support for teachers, more resources for students and more access to technology. As a result, exciting changes have begun to take place inside our classrooms. Along with reading to follow a story, students are learning to read to cite evidence and draw logical conclusions. They are learning to use math to solve real-world problems rather than merely pick out the right multiple-choice answer.

The system-wide changes we have begun are focused on helping students succeed in the long run, achieving their dreams of college and a career. They will take considerable time and effort to carry out. That's why the course we've set in California is to carefully phase in change as state and local capacity grows.

What are the Smarter Balanced Assessments?

The Smarter Balanced assessments are new computer based tests that measure student knowledge of California's English language arts/literacy (ELA) and mathematics standards. These new assessments replace the former paper-based, multiple-choice assessments for students in grades three through eight and grade eleven. The first statewide administration of these assessments will take place in spring 2015.

Why are new assessments needed in California?

- California has adopted more rigorous academic standards that emphasize not only subject knowledge, but also the critical thinking, analytical writing, and problem-solving skills students need to be successful in college and career. These standards set a higher bar for

California students to help ensure they are prepared to succeed in the future. Because what students need to know and be able to do has changed, our tests must change as well.

- California’s new assessment system represents the next step in a comprehensive plan to promote high-quality teaching and learning and improve student outcomes. This plan recognizes that assessments can play an important role in promoting and modeling high-quality instruction.

How were these new assessments developed?

- California worked with 21 other states as part of the Smarter Balanced Assessment Consortium to develop these new assessments, as well as the professional and instructional resources provided to help students, teachers and schools throughout the year.
- Educators – from K–12 to higher education – were deeply involved in the design, testing and scoring of these new assessments. California conducted both pilot and statewide field tests of Smarter Balanced assessments over the last two years.
- Last year, California suspended its use of existing assessments in ELA and Math to give teachers time to focus on refining instruction of California’s more rigorous standards, to give students time to begin deeper learning, and to give schools time to prepare to administer these computer-based tests effectively.

How are Smarter Balanced Assessments an improvement over previous statewide tests?

- The Smarter Balanced assessment system uses both computer-based and computer-adaptive tests, providing students with a wider range of questions tailored to more accurately identify the knowledge and skills individual students have mastered.
- The tests include performance tasks that challenge students to demonstrate critical thinking and problem-solving, and to apply their knowledge and skills to real-world problems.
- The tests measure standards that our K–12 and higher education systems all agree address appropriate expectations for the preparation of high school graduates who are ready for success.
- The tests are taken online, and results are available to teachers, schools and school districts much more quickly than results from previous tests.
- The new computer-based tests include supports for English learners and students with special needs, allowing these students the ability to effectively demonstrate their knowledge and skills.
- The new assessment system is designed to measure student growth over time, which was not possible in California’s previous system, and will provide teachers and schools important information to guide learning.

What will the Smarter Balanced tests measure?

- Smarter Balanced tests provide one measure of student knowledge of the subject matter, critical thinking, analytical writing, and problem solving skills they need to prepare for and succeed in today’s world.

- These assessments provide important information as to whether students are on track to pursue college and career by the time they graduate from high school. The tests provide timely and actionable student information so that teachers and schools can adjust and improve teaching to ensure students have the knowledge and skills they need to succeed in school and beyond.

What results can we expect from the Smarter Balanced Tests?

- The new tests are too fundamentally different from the old exams to make any reliable comparisons between old scores and new. This year's results will establish a new baseline for the progress we expect students to make over time.
- Based on trial runs of the new assessments in California and other states, many if not most students will need to make significant progress to reach the standards set for math and literacy that accompany college and career readiness.
- Over time, as students experience multiple years of instruction related to the skills tested by the new assessments, California's results will show improved achievement. (In 2002, California's new Standardized Testing and Reporting (STAR) tests also set a new baseline for achievement and student results improved quickly over time.)
- Parents will receive a report of their child's scores. But no student, parent or teacher should be discouraged by these scores, which will not be used to determine whether a student moves on to the next grade. Rather, the scores will represent an opportunity to focus on the needs of students and support teachers and schools in their work to achieve college and career readiness.

How will this system help improve teaching and learning?

- The Smarter Balanced assessments are an academic check-up, designed to give teachers the feedback they need to improve instruction. The tests measure critical thinking, analytical writing, problem solving, and subject area knowledge, providing teachers with multiple sources of actionable information about student strengths and areas where students need additional support.
- The system provides two types of interim assessments that teachers and schools can use to assess student learning at key points in the instructional year and to measure student preparedness for year-end summative tests. Both of these tests provide information for teachers to adjust and differentiate teaching in response to the results.
- The system provides a Digital Library of professional learning and instructional resources to help teachers assess individual student learning during instruction, provide feedback to students in a timely manner, and adjust teaching and learning as needed.

Source:

California Department of Education. (March 2, 2015). *CAASPP Communication FAQs About Smarter Balanced*. Retrieved from <http://www.cde.ca.gov/ta/tg/sa/sbcommonqa.asp>

APPENDIX 3: Tutor Observation Form

Circle responses when observed. Record minutes and note interesting behaviors when appropriate.							
Tutor: __ Male __ Female	__ Experienced Tutor __ New Tutor __ Do Not Know	Student Gender: Female Male Student Ethnicity: White Hispanic Other Not Sure				Grade: first second	
Behaviors	Frequency/Minutes	Descriptors or Notes					
Set up time/actions		None Little Some Frequent					
Rapport Building		Established Friendly Professional Distant					
Student Appears		Circle all that apply: Excited Enthusiastic Interested Sullen Bored Uncooperative Inattentive					
Tutor teaching/explaining		Pre-activity Throughout Debrief Closure					
Tutor reading		Shared Reading For Providing Answers Storytelling					
Tutor talking		Giving Directions Probing Responses Giving Answers Rapport					
Tutor writing		Recordkeeping Modeling Other					
Tutor positive encouragement		Yes Good Very Good Giving specific info on correct response					
Tutor listening		While child reads While child talks While child figures it out					
Tutor offering hints		Initial During When child is confused When answer is wrong					
Tutor responding to student questions		Inattentively Somewhat attentive Attentively Conversationally					
Student listening		Inattentively Somewhat attentive Attentively Conversationally					
Student reading		Most of the session Half of the session Third of the session Less					
Student talking		Off Task On Task Questioning Responding Summarizing					
Student writing		Filling in worksheet Composing Recordkeeping					
Student asking questions		Never Rarely Frequently When confused					
Student working on own		Never Rarely Frequently					
Student using strategies		Never Rarely Frequently					
During this session, the student engaged in the following tasks:	Number the tasks in order of the time on task from 1 to 6. Use 0 for never.	Phonics Work	Isolated Word Work	Reading	Writing	Discussing	Answering
List other tasks not listed above Use back of form for additional notes if necessary.							

APPENDIX 4: Tutoring Training Checklist

Trainer Name:	Number of Tutors: _____	Tutor Gender: # Female: _____ # Male: _____	Tutor Ethnicity: White _____ Hispanic _____ Not Sure _____															
		Check if Covered	Scaled Descriptors of Amount of Perceived Emphasis															
Introduction and Welcome			<table style="width: 100%; border: none;"> <tr> <td style="width: 20%;">1</td> <td style="width: 20%;">2</td> <td style="width: 20%;">3</td> <td style="width: 20%;">4</td> <td style="width: 20%;">5</td> </tr> <tr> <td style="text-align: center;">none</td> <td></td> <td></td> <td></td> <td style="text-align: center;">key</td> </tr> <tr> <td colspan="5" style="text-align: center;">topic</td> </tr> </table>	1	2	3	4	5	none				key	topic				
1	2	3	4	5														
none				key														
topic																		
Learning Objectives of Training (Explained)			<table style="width: 100%; border: none;"> <tr> <td style="width: 20%;">1</td> <td style="width: 20%;">2</td> <td style="width: 20%;">3</td> <td style="width: 20%;">4</td> <td style="width: 20%;">5</td> </tr> <tr> <td style="text-align: center;">none</td> <td></td> <td></td> <td></td> <td style="text-align: center;">key</td> </tr> <tr> <td colspan="5" style="text-align: center;">topic</td> </tr> </table>	1	2	3	4	5	none				key	topic				
1	2	3	4	5														
none				key														
topic																		
Pre-Reading Skills – Phonemic Awareness			<table style="width: 100%; border: none;"> <tr> <td style="width: 20%;">1</td> <td style="width: 20%;">2</td> <td style="width: 20%;">3</td> <td style="width: 20%;">4</td> <td style="width: 20%;">5</td> </tr> <tr> <td style="text-align: center;">none</td> <td></td> <td></td> <td></td> <td style="text-align: center;">key</td> </tr> <tr> <td colspan="5" style="text-align: center;">topic</td> </tr> </table>	1	2	3	4	5	none				key	topic				
1	2	3	4	5														
none				key														
topic																		
• Definition																		
• Why important																		
• Strategies or activities																		
• Demonstration or hands-on practice																		
Reading Skills- Phonics			<table style="width: 100%; border: none;"> <tr> <td style="width: 20%;">1</td> <td style="width: 20%;">2</td> <td style="width: 20%;">3</td> <td style="width: 20%;">4</td> <td style="width: 20%;">5</td> </tr> <tr> <td style="text-align: center;">none</td> <td></td> <td></td> <td></td> <td style="text-align: center;">key</td> </tr> <tr> <td colspan="5" style="text-align: center;">topic</td> </tr> </table>	1	2	3	4	5	none				key	topic				
1	2	3	4	5														
none				key														
topic																		
• Definition																		
• Why important																		
• Strategies or activities																		
• Demonstration or hands-on practice																		

Reading Skills- High Frequency Words		1	2	3	4	5
		none				key
• Definition						
• Why important						
• Strategies or activities						
• Demonstration or hands-on practice						
Student Reading- Fluency		1	2	3	4	5
		none				key
• Definition						
• Why important						
• Strategies or activities						
• Demonstration or hands-on practice						
Student Reading- Comprehension		1	2	3	4	5
		none				key
• Definition						
• Why important						
• Strategies or activities						
• Demonstration or hands-on practice						
Student Reading- Vocabulary and Background Knowledge		1	2	3	4	5
		none				key
• Definition						
• Why important						
• Strategies or activities						

• Demonstration or hands-on practice						
Typical Tutoring Day		1	2	3	4	5
		none				key
• Where to sign in & get badge						
• Where to get student						
• Where to get materials						
• Follow lesson plan in binder						
• Fill out sticker sheet and communication log						
• Bring student back to class						
• Sign out & report hours						

Logistics		1	2	3	4	5
		none				key
• Date tutoring begins						
• What happens if a student is absent						
• What happens if the tutor is absent						
• Insurance policy (what to do if alone)						
• What to do if fire drill, lockdown, etc.						
• Bathrooms						
• How to communicate with Site Coordinator and/or teacher						
Working with a K-3 rd grader		1	2	3	4	5
		none				key
• How students are chosen						

• Attention span of 5-8 year old						
• Appropriate contact						
• No gifts, letters okay or gifts to whole class						
Time for questions (Please note tutor questions below):		1	2	3	4	5
		none				key topic
Observer comments re training:						
Feedback Forms Collected		Yes	or	No		