Read Better Be Better

Literacy Program

Analysis of Program Data – 3rd Grade Students

Final Version

August, 2018

Prepared for Read Better Be Better

Submitted by:
Melissa Kovacs, Ph.D., GStat
mkovacs@firsteval.com
602.492.1406

firsteval
Finding Meaning in Data

List of Figures, Tables, and Appendices

Executive Summary

This report presents results from analyzing Read Better Be Better's program data from the Spring, 2015 through Spring, 2018 semesters.

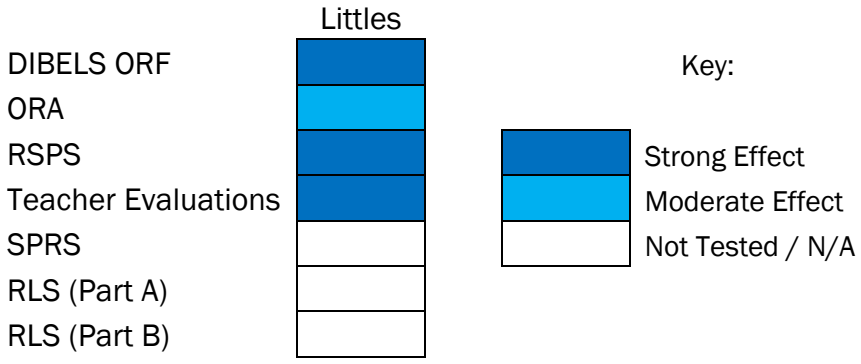
The Read Better Be Better ("RBBB") program uses trained 6th, 7th, and 8th grade student volunteers ("Bigs") to implement a structured reading program to 3rd grade students ("Littles"). The program's mission is to help children improve their literacy skills and become better learners. The program targets Tier II students according to the Arizona State Literacy Plan.

This report represents an analysis of all available outcome data. This report is not a comprehensive evaluation report, and does not address RBBB processes or program implementation. Evidence of program improvement was obtained using the following educational assessments:

- DIBELS ORF
- Oral Reading Analysis
- Reader Self-Perception Scale
- Teacher Evaluations

The RBBB program has strong positive effects on its 3rd grade participants' literacy skills. The table below summarizes the findings of this report:

Summary of Data Results, by Assessment Instrument



Introduction

This report presents results from analyzing Read Better Be Better's program data from the Spring, 2015 semester through the Spring, 2018 semester.

In January of 2015, Read Better Be Better ("RBBB") piloted an after-school literacy program that pairs 8th grade students to help 3rd grade students become better readers. The program has grown steadily since the pilot, and is now operating in 18 schools. The RBBB program uses trained 6th, 7th, and 8th grade student volunteers to implement a structured reading program to 3rd grade students. The older students ("Bigs") work one-on-one with paired younger students ("Littles") to model comprehension strategies, help with reading skills, and provide activities that improve focus and concentration.

RBBB's mission and vision is shown in Figure 1 below, and the RBBB Logic Model is included in Appendix A.

Figure 1 – Read Better Be Better Mission and Vision

Mission: Read Better Be Better helps children improve literacy skills and become better learners.

We create change by:

1. Improving concentration
2. Encouraging an active enjoyment of reading
3. Helping develop a deeper understanding of what is being read

Vision: A Society in which children master the foundational skills necessary to become independent learners.

The RBBB program is intended to target Tier II students, as defined by the Arizona State Literacy Plan. In the Arizona State Literacy Plan, Tier I students are in need of "universal instruction," consisting of a core reading program and benchmark testing. Tier II students need an additional small group intervention beyond Tier I instruction. Tier III students need intensive instruction and remediation services.[1]

Commonly, literacy programs address Tier III students, while RBBB specifically targets Tier II students. Tier II students are often the ones "falling through the cracks," in that they do need additional literacy help, but are not the students in their schools who struggle the most with

[1] The Arizona State Literacy Plan can be found here: http://www.azpromisingpractices.com/LiteracyforRTI.pdf and the definitions of Tier I, II, and III are found on pages 6-7.

reading. Therefore, when resources are scarce, Tier II students' needs often remain minimally addressed at best.

Two schools from two school districts participated in the RBBB pilot program in the spring, 2015 semester, with five schools participating in the program during the fall, 2015 semester. The RBBB program has grown steadily since these early semesters, and Table 1 below details participating schools and the program's growth:

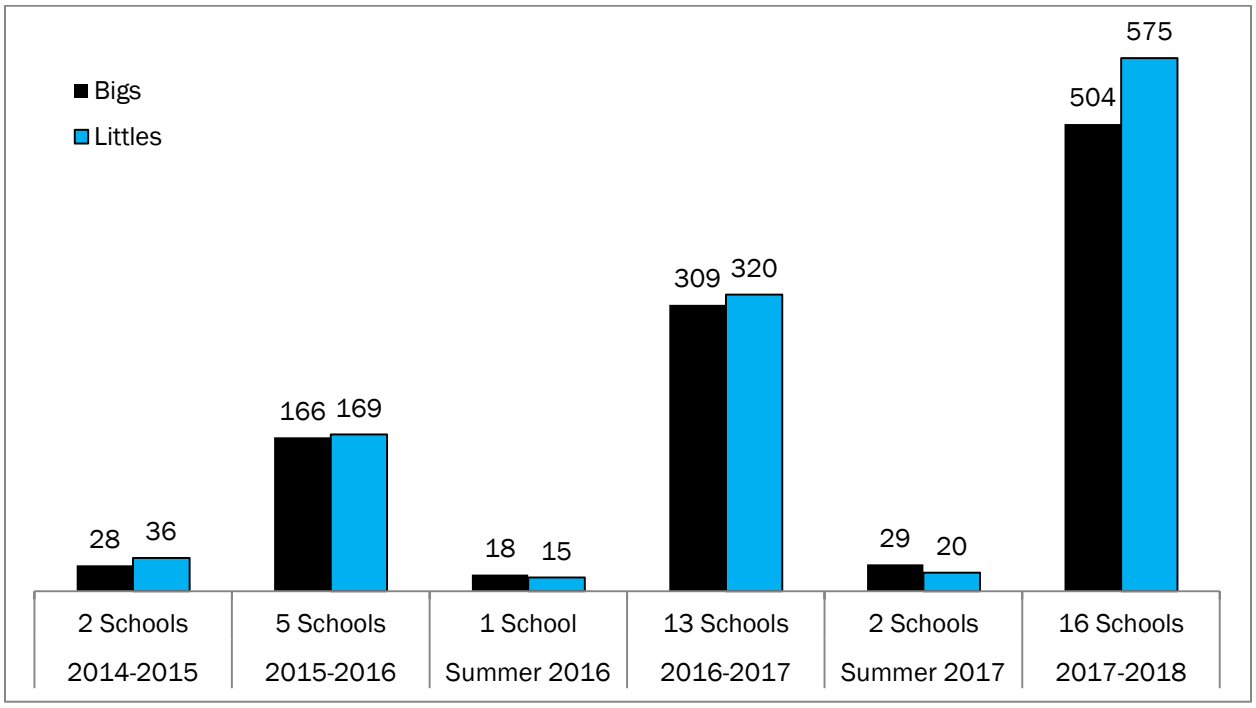
Table 1 - Participating Schools and Program Dates

School	District	Spring 2015 (Pilot)	2015-2016	Summer 2016	2016-2017	Summer 2017	2017-2018
Lattie Coor	Avondale	X	X		X		X
Michael Anderson	Avondale		X		X		X
Edison	Phoenix		X				
Garfield	Phoenix		X				
Whittier	Phoenix	X	X				X
Tri-City/Thornwood BGC	Avondale			X		X	
Arizona Desert	Tolleson				X		X
Centerra Mirage	Avondale				X		X
Copper Trails	Avondale				X		X
Desert Star	Avondale				X		X
Desert Thunder	Avondale				X		X
Encanto / Clarendon	Osborn				X		X
Longview	Osborn				X		X
PH Gonzales	Tolleson				X		X
Sheely Farms	Tolleson				X		
Solano	Osborn				X		X
Wildflower	Avondale				X		X
Parsons BGC	Balsz					X	
Arroyo	Washington						X
Chaparral	Washington						X
Mountain View	Washington Elementary						X
Shaw Butte	Washington Elementary						X
Sunnyslope	Washington Elementary						X
Sunset	Washington Elementary						X

Desert Oasis	Tolleson						X

As well, Figure 2 below shows the growth in the number of program participants and participating schools since the program pilot began during the 2014-2015 school year.

Figure 2: Number of RBBB Program Participants and Schools by Academic Year



RBBB in Context

RBBB's cross-age peer tutoring model comes from a strong foundation of proven effectiveness. Numerous studies find that cross-age peer tutoring is beneficial for both the younger and older grade participating students.[2] Some studies also found that participation in these programs boosts views of oneself as a reader and leads to less negative thinking about reading.[3] A study in

[2] See, for example, Loretta Abassi, Cleveland State University, "Effects of Cross-Age Tutoring on Reading Attitudes of Elementary School Students;" John Hattie, 2006, "Cross-Age Tutoring and the Reading Together program," in *Studies in Educational Evaluation*; Van Keer et al., 2005, "Effects of Explicit Reading Strategies Instruction and Peer Tutoring on Second and Fifth Graders' Reading Comprehension and Self-Efficacy Perceptions," in the *Journal of Experimental Education*; Wright and Cleary, 2006, "Kids in the Tutor Seat: Building Schools' Capacity to Help Struggling Readers Through a Cross-Age Peer-Tutoring Program," in *Psychology in the Schools*; and Slavin and Madden, 1989, "What Works for Students at Risk: A Research Synthesis," in *Educational Leadership*.

[3] See Abassi and Van Keer

Syracuse found that tutees' participation in a cross-age peer reading program engendered bigger gains than their tutors experienced.[4]

One researcher found that the effects of participating in a cross-age peer tutoring reading program are stronger in later phases of the program.[5] This is important evidence for RBBB to monitor its long-term outcomes. RBBB's outcomes are displayed in the logic model in Appendix A.

Scope of This Report

This report represents an analysis of outcome data for all existing RBBB program data, covering 2015 through present. Table 1 above describes the academic years and program sites included in existing RBBB data. This report is not a comprehensive evaluation report. FirstEval did not evaluate RBBB processes or program implementation. FirstEval did, however, analyze all existing RBBB program participant and comparison group data provided by RBBB. In this sense, this report addresses most of the short-term outcomes in the RBBB logic model, and all of the medium-term outcomes in the logic model. This report does not address any of the long-term outcomes in the logic model. The RBBB logic model, highlighted to show this report's scope, is attached as Appendix A.

Instruments

Data from numerous educational assessment instruments were provided by RBBB for analysis by FirstEval. This section describes those instruments and how this report is structured around the instruments.

This report examines data from Littles and Bigs separately, and examines data among the different assessment instruments separately. For the DIBELS ORF assessment, RBBB has assessment data for 3rd grade students and 6th/7th/8th grade students who are not part of the RBBB program. For the Oral Reading Analysis ("ORA") assessment, RBBB has assessment data for 3rd grade students who are not part of the RBBB program. This data of non-participants provided an ideal group with which to construct a matched comparison group to assess the impact of the RBBB program on DIBELS ORF and ORA progress.

Because RBBB program participation was not randomly assigned, these non-participants do not represent a true control group, but rather a matched comparison group.

The following table details the assessment instruments available for each examined group.

[4] See Wright and Cleary
[5] See Hattie

Table 2 – Instruments and Data Availability by Group

Student Group	Dibels ORF	ORA	Reader Self-Perception Scale	Teacher Evaluation of Students
Littles	X	X	X	X
3rd Grade Comparison Group	X	X		

DIBELS

RBBB receives data from participating schools on students' DIBELS scores. DIBELS is an acronym for Dynamic Indicators of Basic Early Literacy Skills. The DIBELS family of instruments is widely used, and was developed at the University of Oregon. For purposes of this report, the DIBELS ORF data provided useful information to gauge RBBB participants' progress. The Dibels ORF (oral reading fluency) is individually administered to test for reading fluency. Students read passages while an administrator records miscues, then the student retells the passage, hitting certain highlights in the passage to prove comprehension.

More information about the Dibels ORF instrument can be found here: https://dibels.uoregon.edu/assessment/dibels/measures/orf.php

Oral Reading Analysis

RBBB receives data from Osborn School on students' Oral Reading Analysis ("ORA") scores. The ORA scores stem from two assessments - Fountas & Pinnell Benchmark Assessment System and Rigby's Reading Assessment. These are individual assessments of students, and data was provided with goal benchmarks at the student level.

More information about the Fountas & Pinnell Benchmark Assessment System can be found here: http://www.fountasandpinnell.com/bas/

Reader Self-Perception Scale

The Reader Self-Perception Scale ("RSPS") is RBBB's primary reading efficacy tool, and is a tool to measure how children feel about themselves as readers.[6] The instrument consists of 33 items, and is divided into 5 subscales for analytic purposes. The subscales measure general perception of one's own reading, progress (how one's perception of present reading performance compares with past performance), observational comparison (how a child perceives her or his reading performance to compare with the performance of classmates), social feedback (direct or indirect input about reading from teachers, classmates, and people in the child's family), and

[6] See Henk & Melnick, 1995. "The Reader Self-Perception Scale (RSPS): A New Tool for Measuring How Children Feel About Themselves as Readers," in *The Reading Teacher*, Vol. 48 No. 6.

physiological state (internal feelings that the child experiences during reading). These subscales have been shown to have high internal consistency and reliabilities.[7] The instrument is included as Appendix B.

The RSPS fits neatly into measuring progress towards RBBB's mission and vision, by measuring an active enjoyment of reading.

Teacher Evaluation of Students

Teachers are asked to answer three brief questions of RBBB program students to assess whether the student is able to focus on the reading material; whether the student appears to enjoy reading; and whether the student understands reading material in class. This data is examined for changes between the beginning of the semester and the end of the semester, and these three questions are included as Appendix C.

Methodology

RBBB provided clean data to FirstEval to test for progress among program participants and non-participants, as measured by the instruments discussed in the previous section. RBBB receives the data itself from the participating schools and by RBBB staff conducting some of the assessments themselves (RSPS). Data is provided at the individual student level.

FirstEval compiled and analyzed data to test for differences between baseline scores and post-program scores among participants. We also tested for improvement among the comparison groups, and compared improvement rates between participant groups and comparison groups. When sample sizes allowed, and when appropriate, paired-samples t-tests and repeated measures general linear models were employed.[8]

For the DIBELS ORF data, RBBB provided FirstEval with benchmark data. These benchmarks are published data points that show the cut points for risk among students. These benchmarks are visualized along with the achievement results in the DIBELS ORF reporting section of this report. As well, individual students at Osborn School have goal benchmark ORA data. Comparing program effects to benchmarks helps account for outside factors and isolate the RBBB program effect.

[7] See Henk & Melnick mentioned previously.
[8] Paired samples t-tests were used to compare student growth from pre-test to post-test time. Repeated measures general linear models were used to test the differences in growth between program participants and non-participants in the matched comparison group. See Appendix D for more on the methodology.

Results

DIBELS ORF

As stated earlier, RBBB has rich data consisting of pre-test and post-test scores for both program participants and non-participants, for Littles and Bigs, of their DIBELS ORF scores. RBBB also has benchmark cut-point data for DIBELS ORF, which was incorporated into this analysis. The benchmark data is specific for beginning-of-year levels, middle-of-year levels, and end-of-year levels. In order to align with the benchmark cut-scores, results are reported separately for all fall semesters and all spring semesters. The same middle-of-year benchmark score was used twice to compare with the fall semester's post-test and the spring semester's pre-test. The following table describes this alignment:

Table 3: Alignment of RBBB Data with DIBELS ORF Benchmark Cut-Points

RBBB Data Time	DIBELS ORF Cut-Point Time
Fall Semester, Pre-Test	Beginning of Year
Fall Semester, Post-Test	Middle of Year
Spring Semester, Pre-Test	Middle of Year
Spring Semester, Post-Test	End of Year

FirstEval applied a matched-comparison group approach to analyzing the effect of RBBB program participation on DIBELS ORF scores. A comparison group of non-program participants was crafted from available data to match the pre-test starting point mean DIBELS score of RBBB program participants, within one standard deviation of the participants' mean.

Three repeated-measures general linear models ("GLM") were performed on each of the following subgroups' DIBELS ORF scores: Bigs, Littles in any Fall semester, and Littles in any Spring semester. Littles were separated into the Fall / Spring semester groups so as to compare their progress with DIBELS ORF benchmarks. A GLM analysis can test whether the score increase for RBBB program participants is significantly greater than the score increase among comparison group students.

Other statistical methods were explored (hierarchical linear modeling and propensity score matching) and found inappropriate for this data. A full description of the statistical testing methods follows in Appendix D.

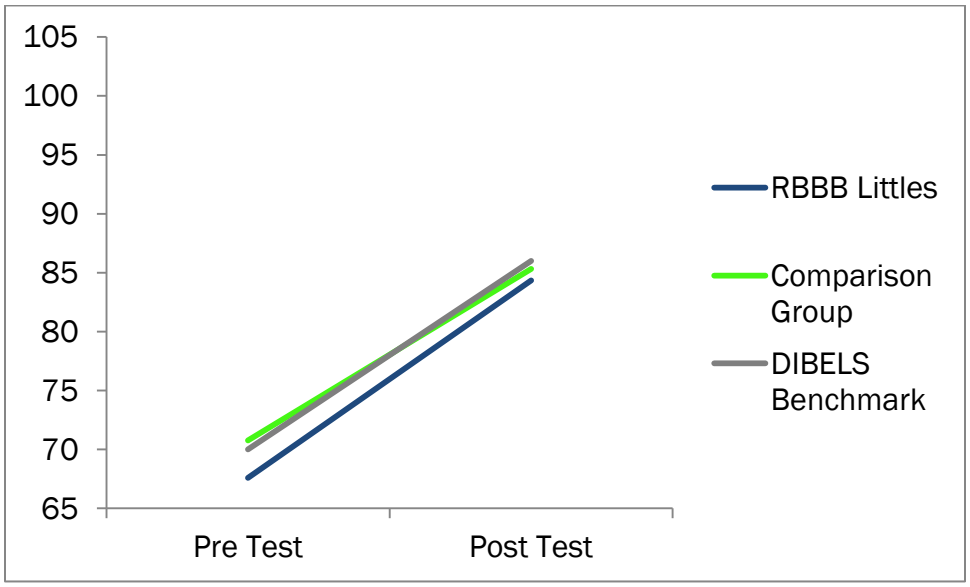
Finally, when sample size allowed, the data was tested for effects that differed by school year.

Littles, Fall semesters only:

For Littles in the RBBB program, during all Fall semesters, their DIBELS ORF scores increased 16.77 points, on average. This is greater than the matched comparison group's score increase,

and greater than the benchmark increase. The following visualization and table show these results. These results are not statistically significant.

Figure 3: DIBELS ORF Average Scores, Pre and Post Test, RBBB "Littles," Fall Semesters only



	Pre Test	Post Test	Score Increase	Group Size
RBBB Littles	67.58	84.35	16.77	n=292
Comparison Group	70.76	85.31	14.55	n=614
DIBELS Benchmark	70	86	16.00	

Littles' fall semester data was tested separately during the 2016-2017 school year and the 2017-2018 school year. Neither of these school years showed significant effects compared to the comparison groups for that year. The 2015-2016 school year data wasn't tested separately due to the small sub-sample of participants (n=25).

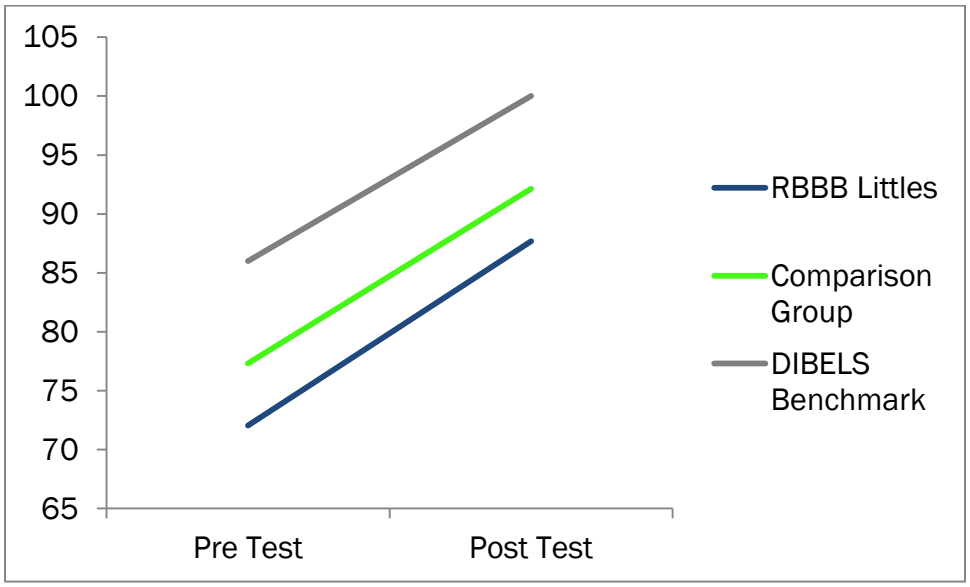
However, in every year during fall semesters, RBBB program participants significantly increase their DIBELS ORF scores.

Littles, Spring semesters only:

For Littles in the RBBB program, during Spring semesters, their DIBELS ORF scores increased 15.63 points, on average. This is greater than the comparison group's score increase, and greater

than the benchmark increase. The following visualization and table show these results. These results are statistically significant.

Figure 4: DIBELS ORF Average Scores, Pre and Post Test, RBBB "Littles," Spring Semesters only



	Pre Test	Post Test	Score Increase	Group Size
RBBB Littles	72.05	87.68	15.63*	n=289
Comparison Group	77.3	92.13	14.83	n=625
DIBELS Benchmark	86	100	14.0	

** Littles' score increase is statistically significantly greater than the comparison group's score increase at the α=.05 level, according to GLM test.*

Littles' spring semester data was tested separately during the 2016-2017 school year and the 2017-2018 school year. The 2017-2018 school year data shows significant effects compared to the comparison group for that year. The effect is strong in the 2016-2017 school year data, but not statistically significant. The 2015-2016 school year data wasn't tested separately due to the small sub-sample of participants (n=32).

However, in every year during spring semesters, RBBB program participants significantly increase their DIBELS ORF scores.

Oral Reading Analysis ("ORA")

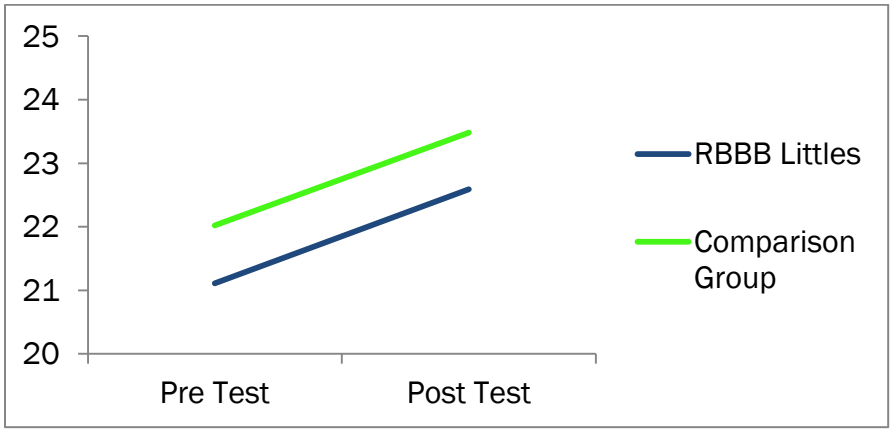
RBBB provided FirstEval with ORA data from the Osborn School. ORA data is for Littles only, of both program participants and non-participants. ORA data also includes a goal score for some students in the data. Similar to the DIBELS ORF score analysis, FirstEval applied a matched-

comparison group approach to analyze the effect of RBBB program participation on ORA score improvement. A comparison group of non-program participants was crafted from available data to match the pre-test starting point mean ORA score of RBBB program participants, within one standard deviation of the participants' mean.

A repeated-measures GLM was performed to compare Littles' progress against the comparison group's progress. This methodology follows the methodology used for the DIBELS ORF assessment data, and is described in Appendix D. The GLM results are reported, followed by a further look at the students' goal scores.

For Littles in the RBBB program, their ORA score increased by 1.48 points from pre-test to post-test time. This is greater than the matched comparison group's score increase. The following visualization and table show these results. In spite of the participants' score increase being only slightly greater than the non-participants' score increase, the results are statistically significant.

Figure 6: ORA Average Scores, Pre and Post Test, RBBB "Littles"



	Pre Test	Post Test	Score Increase	Group Size
RBBB Littles	21.11	22.59	1.48*	n=52
Comparison Group	22.02	23.48	1.46	n=184

** Littles' score increase is statistically significantly greater than the comparison group's score increase at the α=.05 level, according to GLM test.*

As described earlier, some students had individualized goal scores within the data. This allows for a comparison between the proportion of RBBB participant students who met or exceeded their ORA goal scores with the proportion of non-participant students who met or exceeded their ORA goal scores. While RBBB participant students were slightly more likely to meet or exceed their goal scores by post-test time than non-participants, this proportion was not statistically significant.[9] The table below shows these results.

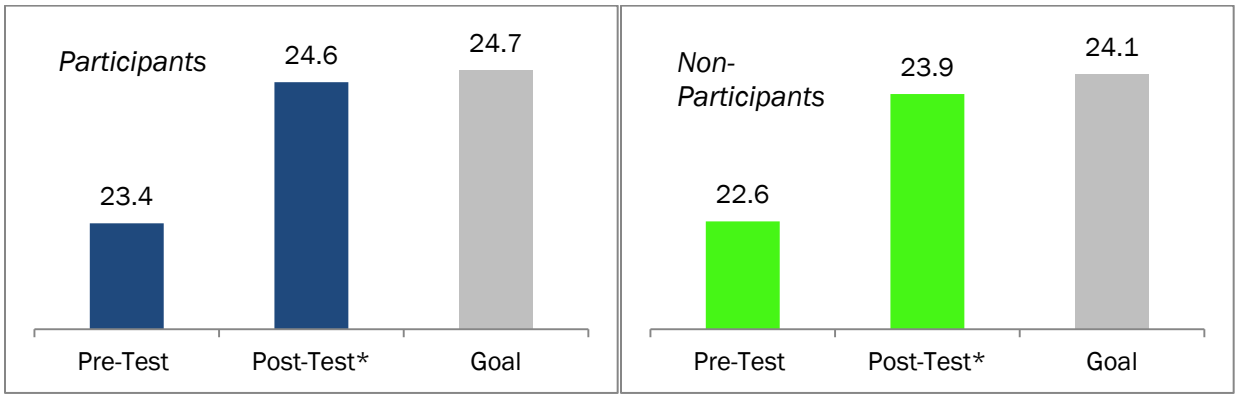
[9] According to a z-test of proportions.

Table 4: Proportion of Students At or Above ORA Goal Score at Post-Test Time

	Proportion	Group Size
RBBB Littles	85%	n=20
Comparison Group	84.2%	n=76

On average, for both RBBB participants and non-participants, all students significantly increased their ORA scores from pre-test to post-test time.[10] Yet, when averaged, neither the RBBB participants nor the non-participants met their goal scores at post-test time, taken as a group. The figure below shows the average scores for these groups compared to the averaged goal scores.

Figure 7: Average ORA Test Scores for RBBB Participants and Non-Participants



** Denotes a statistically significant improvement from pre- to post-time measures, at the α =.05 level, according to paired samples t-tests.*

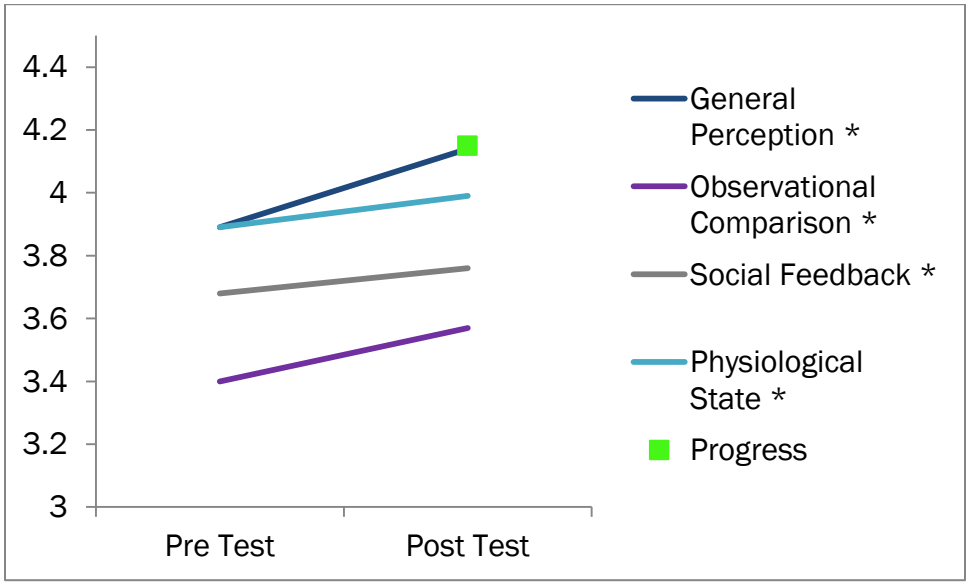
In sum, the ORA data provides evidence of strengthened reading scores compared with the matched comparison group.

Reader Self-Perception Scale ("RSPS")

Figure 8 below shows results of 3rd grade participants' changes in scores on the RSPS. The RSPS subscales denoted here are scored between zero and five, with a higher score indicating greater achievement. The Progress subscale was only scored at post-test time, since its purpose is to quantify how one's perception of present reading performance compares with past performance. For all other subscales (Social Feedback, Physiological State, General Perception, and Observational Comparison), RBBB participating students demonstrated statistically significant improvement in their scores between pre-test time and post-test time.

[10] According to paired-samples t-tests.

Figure 8 – RSPS Subscale Component Score Averages for 3rd Grade Participants (Scored from zero to five)



	Pre Test	Post Test	Score Increase
General Perception	3.89	4.14	0.25 *
Progress	0	4.15	N/A
Observational Comparison	3.4	3.57	0.17 *
Social Feedback	3.68	3.76	0.08 *
Physiological State	3.89	3.99	0.10 *

** Denotes a statistically significant improvement from pre- to post-time measures, at the α =.05 level, according to paired samples t-tests.*

This analysis also explored the potential impact that attendance in RBBB sessions could have on progress, when data was available. For the RSPS, on average, attendance was positively correlated with greater increases on the RSPS subscale scores. In other words, the more RBBB program sessions a student attended, the higher the student increased their own rated General Perception, Observational Comparison, Social Feedback, and Physiological State related to reading. These correlations are not statistically significant.[11]

Some changes in subscale improvement among RBBB program participants by year are noted in the RSPS instrument. Table 5 below shows the differences in subscale score increases by year for the RSPS instrument:

[11] Bivariate Pearson correlation found to be statistically significant at the α=.05 level.

Table 5: RSPS Subscale Score Increases by Year

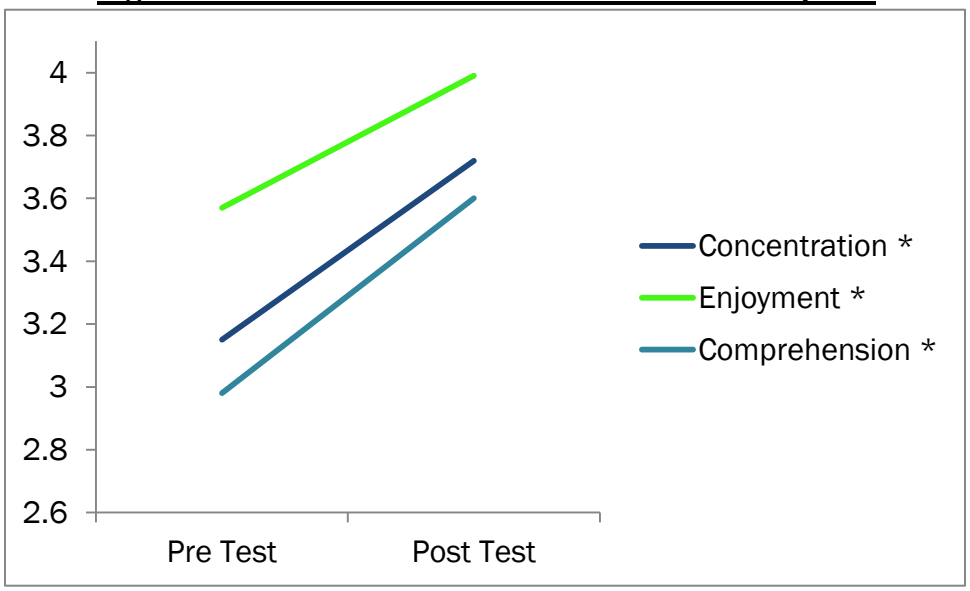
	2015-16 Score Increase	2016-17 Score Increase	2017-18 Score Increase
General Perception	0.49 *	0.25 *	0.17 *
Observational Comparison	0.25 *	0.24 *	0.08
Social Feedback	0.13	0.07	0.07
Physiological State	0.16 *	0.06	0.10

** Denotes a statistically significant improvement from pre- to post-time measures, at the α =.05 level, according to paired samples t-tests.*

Teacher Evaluation of Students

Teachers evaluated their 3rd grade students who participated in the RBBB program on three factors – Focus, Enjoyment of Reading, and Reading Comprehension. Figure 9 below shows the average score for participants from the beginning of the semester ("pre") to the end of the semester ("post"). On average, participants improved on all three subscales to a statistically significant extent. The subscales were scored from 1 to 5, with a higher score denoting greater achievement.

Figure 9 – Teacher Evaluations of 3rd Grade Participants



	Pre Test	Post Test	Score Increase
Concentration	3.15	3.72	0.57*
Enjoyment	3.57	3.99	0.42*
Comprehension	2.98	3.6	0.62*

** Denotes a statistically significant improvement from pre- to post-time measures, at the α =.05 level, according to paired-samples t-test.*

On average, attendance was positively correlated with greater increases on the teacher evaluation scores. The more sessions a student attended, the higher their increase in teacher evaluation score from pre-test to post-test time. Attendance was significantly positively correlated[12] with increases in concentration and enjoyment of reading.

Notable, however, was the analysis of teacher evaluations by year. For every school year (including the 2014-15 pilot year), RBBB Littles statistically significantly increased their teacher evaluation scores on every subscale. Table 6 below shows these results:

Table 6: Teacher Evaluation of Students' Subscale Score Increases by Year

	2014-15 Score Increase	2015-16 Score Increase	2016-17 Score increase	2017-18 Score Increase
Concentration	0.94 *	0.87 *	0.76 *	0.36 *
Enjoyment	0.88 *	0.88 *	0.34 *	0.30 *
Comprehension	0.91 *	1.13 *	0.72 *	0.40 *

** Denotes a statistically significant improvement from pre- to post-time measures, at the α =.05 level, according to paired-samples t-test.*

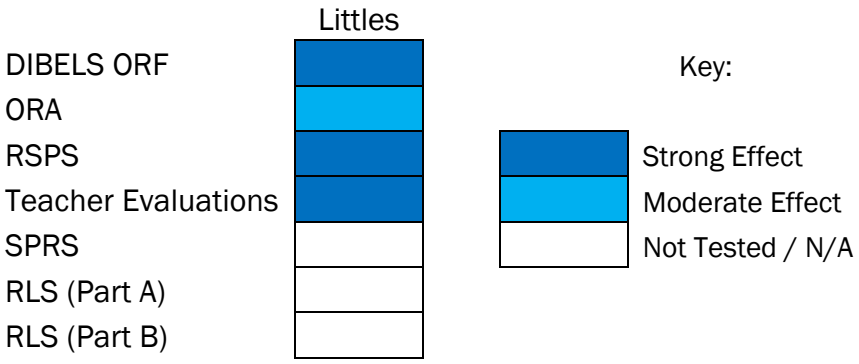
[12] Bivariate Pearson correlation found to be statistically significant at the α=.05 level.

Summary

The RBBB program has notable effects on its participants' literacy skills. Most importantly, for 3rd grade participants, or "Littles," program participation has strong effects on their social feedback skills, physiological state, and their rating score of general perception. Also importantly, 3rd grade participants show significant improvement in reading according to the DIBELS ORF and ORA scores. And, teacher evaluations' of the Littles show significant improvement for RBBB program participants.

The following table summarizes this report's results:

Table 7: Summary of Data Results, by Assessment Instrument



When available, attendance information was incorporated into this report's analyses. Attendance data often showed a correlation with stronger assessment results. This trend of overall positive correlations is an indication of program effectiveness.

RBBB has complex program data, in that they have tracked their students' progress over multiple years and locations with multiple assessment instruments. The data is made richer with the availability of comparison students' data. This report reflects the examination of the RBBB data and the comparison data. The numerous assessment instruments that the RBBB program employs is commendable, in that RBBB truly aims to measure its program participant progress through many angles. Overall, the RBBB program is significantly improving reading for its 3rd grade participants, as evidenced by numerous assessment instruments.

Appendix A – RBBB Logic Model with Report Scope Highlighted in Red

Target Population:
- 3rd grade
- Tier II
- Fluency >60wpm
- Not receiving additional intervention

Read Better Be Better

Read Better Be Better Logic Model

The red-shaded area represents the data examined for this report.

Inputs	Activities	Outputs	Short	Medium	Long
Staff expertise and experience Research Collaborative Partnerships Time Funds	Proprietary reading comprehension curriculum. o 2x/week o Minimum one semester 8th grade "reading leaders" trained to implement curriculum Be A Better Reader Be A Better Thinker Be A Better Mover	# Students received program # Student Hours % Students completing program # 8th grade trained"reading leaders"	Attendance – RBBB and school (RBBB log and school attendance data with grade average as control) Retention (RBBB log) Increased enjoyment of reading (3rd grade teacher pre-post evaluation) Enjoyment of RBBB program (post self-evaluation) Increased self-efficacy(pre-post self-evaluation: The Reader Self-Perception Scale)	Improved reading comprehension (Galileo testing with grade average as control) Improved concentration (3rd grade teacher pre-post evaluation) Increased sense of personal responsibility (Social & Personal Responsibility Scale and pre-post self-evaluation) Improved general academic performance (longitudinal study of school Math and ELA scores)	Better learners (Approaches to Learning) Increased high school graduation rates (longitudinal study of school graduation data)

Assumptions
Achieving a reading age of 10 allows independent learning and positively impacts earning potential

External Factors
Current poor AZ literacy stats < MOWR law < AZ Literacy Plan. Current high school drop-out rates. Rise of service-learning model.

Appendix B: Reader Self-Perception Scale ("RSPS")

Instruction: Below are statements about reading. Please read each statement carefully. Then fill in the bubbles that show how much you agree or disagree with the statement.
Example:
If you are really positive that pepperoni pizza is best, fill in the bubble under "**Strongly Agree**".
If you think that is good but maybe not great, fill in the bubble under "**Agree**".
If you can't decide whether or not it is best, fill in the bubble under "**Undecided**".
If you think that pepperoni pizza is not all that good, fill in the bubble under "**Disagree**".
If you are really positive that pepperoni pizza is not very good, fill in the bubble under "**Strongly Disagree**".
Now, please fill in the bubbles that show how much you agree or disagree with each of the following statement.

1. I think I am a good reader…… ☐☐ ☐ ☐ ☐ ☐
2. I can tell that my teacher likes to listen to me read…… ☐☐ ☐ ☐ ☐ ☐
3. I read faster than other kids…… ☐☐ ☐ ☐ ☐ ☐
4. My teacher thinks that my reading is fine…… ☐☐ ☐ ☐ ☐ ☐
5. I like to read aloud…… ☐☐ ☐ ☐ ☐ ☐
6. When I read, I can figure out words better than other kids…. ☐☐ ☐ ☐ ☐ ☐
7. My classmates like to listen to me read…… ☐☐ ☐ ☐ ☐ ☐
8. I feel good inside when I read…… ☐☐ ☐ ☐ ☐ ☐
9. My classmates think that I read pretty well…… ☐☐ ☐ ☐ ☐ ☐
10. When I read, I don't have to try as hard as I used to…… ☐☐ ☐ ☐ ☐ ☐
11. I seem to know more words than other kids when I read. ☐☐ ☐ ☐ ☐ ☐
12. People in my family think I am a good reader…… ☐☐ ☐ ☐ ☐ ☐
13. I am getting better at reading…… ☐☐ ☐ ☐ ☐ ☐
14. I understand what I read as well as other kids do…… ☐☐ ☐ ☐ ☐ ☐
15. When I read, I need less help than I used to…… ☐☐ ☐ ☐ ☐ ☐
16. Reading makes me feel happy inside…… ☐☐ ☐ ☐ ☐ ☐
17. My teacher thinks I am a good reader…… ☐☐ ☐ ☐ ☐ ☐
18. Reading is easier for me than it used to be…… ☐☐ ☐ ☐ ☐ ☐
19. I read faster than I could before…… ☐☐ ☐ ☐ ☐ ☐
20. I read better than other kids in my class…… ☐☐ ☐ ☐ ☐ ☐
21. I feel calm when I read…… ☐☐ ☐ ☐ ☐ ☐
22. I read more than other kids…… ☐☐ ☐ ☐ ☐ ☐
23. I understand what I read better than I could before…… ☐☐ ☐ ☐ ☐ ☐
24. I can figure out words better than I could before…… ☐☐ ☐ ☐ ☐ ☐
25. I feel comfortable when I read…… ☐☐ ☐ ☐ ☐ ☐
26. I think reading is relaxing…… ☐☐ ☐ ☐ ☐ ☐
27. I read better now than I could before…… ☐☐ ☐ ☐ ☐ ☐
28. When I read, I recognize more words than I used to…… ☐☐ ☐ ☐ ☐ ☐
29. Reading makes me feel good…… ☐☐ ☐ ☐ ☐ ☐
30. Other kids think I'm a good reader…… ☐☐ ☐ ☐ ☐ ☐
31. People in my family think I read pretty well…… ☐☐ ☐ ☐ ☐ ☐
32. I enjoy reading…… ☐☐ ☐ ☐ ☐ ☐
33. People in my family like to listen to me read…… ☐☐ ☐ ☐ ☐ ☐

Appendix C: Teacher Evaluation of Students

Please make selections:

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
The student is able to focus on reading material from the beginning to the end	○	○	○	○	○
The student appears to enjoy reading self-selected texts	○	○	○	○	○
The student understands reading material in class	○	○	○	○	○

Additional Notes

Appendix D – Statistical Methodology for DIBELS ORF and ORA Assessments

The rich DIBELS ORF and ORA data provided an opportunity to isolate the effects of the RBBB program because it contained a large comparison group. This allowed for the construction of a matched comparison group. As well, the data contained repeated measures (the DIBELS ORF and ORA scores measured at pre-test and post-test time), matched at the student level.

Hierarchical linear modeling ("HLM") is a popular design model in educational evaluation. HLM considers students to be "nested" within schools, as a way of controlling for potential school-level effects while isolating program effects on students. While our overall dataset was large (n > 1,000), the multiple school locations resulted in small sub-samples. These sub-sample sizes were not large enough to analyze the data at a nested level, and therefore it was deemed ineligible for HLM.

The data was also extensively explored for its eligibility for propensity score matching methods ("PSM"). Precedence exists for conducting PSM instead of HLM in educational evaluation settings.[13] PSM mimics randomization in a quasi-experimental setting by matching treated results (here, the RBBB program participants' DIBELS ORF and ORA score changes) with untreated results (the non-participants' DIBELS ORF and ORA score changes). In order to conduct PSM on a dataset, certain assumptions must be met (conditional independence, common support, and stable unit treatment values). As well, the beauty of using PSM is the ability to control for varying factors and isolate the treatment effect. In our data, the only two eligible control factors among both the treated and untreated groups are gender and race/ethnicity. RBBB participants don't significantly differ on gender make-up compared to non-participants at any grade level. As well, RBBB participants don't significantly differ on racial / ethnic make-up from non-participants at the 3rd grade level. Among Bigs, however, Hispanics were more likely to be found in the non-participant group. Given this, and that the data did not pass all of the PSM assumptions, it wasn't deemed eligible for PSM methods.

For these reasons, the best approach was to employ general linear modeling ("GLM"). GLM provides a way to test whether the change in a repeated measure (our DIBELS ORF and ORA scores pre and post-test time) differs between a treated group (RBBB program participants) and an untreated group (non-participants).

RBBB provided FirstEval with a rich comparison group dataset. With this, FirstEval was able to construct a comparison group that closely matched the pre-test DIBELS ORF and ORA scores of the participants. By applying this matched comparison group approach, the RBBB participants and the non-participants had similar DIBELS and ORA starting points. The original comparison group datasets were culled down to be within +/- 1.0 standard deviations of the participants'

[13] See Lane et al., 2012, An Illustrative Example of Propensity Score Matching with Education Research, https://www.researchgate.net/publication/273061804_An_Illustrative_Example_of_Propensity_Score_Matching_with_Education_Research

DIBELS ORF and ORA pre-test mean scores. This GLM matched comparison group approach is reported in this document.