

Summary of Experimental Study of the Ready to Teach *Developmental Electronic Field Trip Reader Project*

Helene Jennings
Leilani Lucca
ORC Macro¹

Background

Maryland Public Television's (MPT) Ready to Teach reading project is aimed at students who did not learn to read in the primary grades. These students enter the upper elementary and middle grades without the reading skills they need to master the subject matter that is covered in the higher grades. Increasingly, Maryland middle schools have found the need to focus on reading, and for most, improving reading is a top priority. Thus, teachers in all content areas are being called upon to provide instruction that will build reading skills while simultaneously covering the subject matter of the curriculum.

MPT initiated the Developmental Electronic Field Trip (DEFT) Reader project to build a set of rich Web-based resources in core content areas such as social studies, language arts, and science. These resources contain embedded reading supports, offer explicit instruction in reading strategies, and engage students in learning activities that build reading skills. These reading tools were designed by reading specialists who worked closely with MPT and were based on the findings of the National Reading Panel (National Institute of Child Health and Human Development, 2000). The project focused on developing and distributing innovative programming linked to State standards; delivering professional development to teachers in the use of digital content resources to assist poorly performing middle-school readers; and most importantly, improving the reading skills of middle school students, especially those reading below grade level.

Study Design

As part of their evaluation of this grant, ORC Macro designed and carried out a school-based experimental study to examine the effect that use of three Electronic Field Trips (EFTs)

¹ ORC Macro, 11785 Beltsville Drive, Calverton, MD 20705. <http://orcmacro.com>

(Knowing Poe, Exploring Maryland's Roots, and Pathways to Freedom) has on student reading comprehension skills and content knowledge. In this experimental study, the three online field trips were integrated into the curriculum at two middle schools in two Maryland districts (one urban and one rural) at the time of the academic year when the content would traditionally be covered.²

The reading performance of the districts of the participating schools fell below the Maryland state average on the State's 8th grade reading assessment, and the selected middle schools, in turn, performed below the average for their districts. Both schools have high percentages of non-white students and serve many children living in poverty (more than half receive Free and Reduced Meals). At the beginning of the school year, 422 students were enrolled in the classes designated for the study at both schools – 270 at one school and 152 at the other. However, due to students leaving the school and new students entering the school, a varied number of students were included in the analysis of content and reading tests, depending on completion of both pre- and post-tests. A total of 375 students at both schools (236 at Salisbury and 139 at Booker T. Washington) took part in the study from beginning to end.

This study implemented an experimental (treatment/control group) design. The small size of the study (involving 9 teachers) required the evaluators to carefully consider the randomization method for the evaluation. In order to control for variance in teacher quality, the evaluation team chose to randomize the groups based on classroom, rather than by teacher. This meant that all participating teachers at the two schools would be responsible for teaching both treatment and control classes. The researchers randomly assigned all language arts and social studies classes at the appropriate grade level to either treatment or control status.³

Implementation

Teachers were given a list of specific learning objectives to cover in each of the three units to be used with both the treatment and control classes. These learning objectives were closely aligned to the language arts and social studies curricula for the state and the two participating districts.

² At Salisbury Middle School (SMS) the study involved all 7th grade students and at Booker T. Washington (BTW) Middle School the study involved all 8th grade students.

³There were a total of 21 classes: Salisbury = 8 Treatment classes, 6 Control classes; Booker T. Washington = 4 Treatment classes, 3 Control classes

For each unit, a 20-item pre- and post-test was developed and externally validated. These content tests were scored by the evaluators. Both the treatment and control students took the pre-test on the same date. Once the pre-test had been administered, each teacher could begin using the online field trips with treatment classes and teaching the unit to control classes. They did not begin the unit (in either the treatment or control status) prior to the pre-test date. Teachers were asked to complete implementation logs for the treatment classes and control classes. Analysis of these logs revealed that teachers averaged approximately 450 minutes (the equivalent of 10 45-minute periods) on each unit for both the experimental classes (using the field trips) and control classes (where traditional approaches, enriched with supplementary resources provided by the project, were used) to meet the established learning objectives.

Because teachers began and finished each unit on different days, the post-tests were administered using Scantron forms by the participating teachers as soon as they completed each unit and scored by the researchers. Teachers were given a timetable and a deadline for completion of each unit.

- The unit on *Knowing Poe* was completed at both schools in November 2003.
- The unit on *Exploring Maryland's Roots* was completed at both schools in March 2004.
- The unit on *Pathways to Freedom* was completed at both schools in May 2004.

In addition to the content assessments, the evaluators also analyzed student gains on a standardized reading test. The evaluators met with reading experts prior to the school year to identify specific instruments that could be used to study middle school students' reading comprehension. Both agreed that the Gates-MacGinitie Reading Test (GMRT) was a strong assessment and suitable for use in this project. This norm-referenced test is designed for use with K-12 students and adults to measure reading ability. It is divided into two sections, Vocabulary and Comprehension, and is often used to evaluate the effectiveness of instructional programs. In the spring of each year, all students are required to take the GMRT at Salisbury Middle, but not at Booker T. Washington Middle. (No standardized reading test was used in the 2003-2004 school year in the latter school.) The evaluation team decided to take advantage of

the existing implementation of the GMRT at Salisbury, and initiated the administration of the GMRT at Booker T. Washington.⁴ Major findings from this study are described below.

Classroom Content Test Results

Students who used the EFTs performed better on the unit tests than the students using only traditional methods. The table below presents the means for all three assessments, for both treatment and control groups, and shows the changes in knowledge before and after the study of the unit—in terms of the change score and the percentage change that represents.

Table 1. Overall Results for the Three Content Assessments

	Knowing Poe ⁵		Exploring Maryland's Roots		Pathways to Freedom	
	Treatment (N=252)	Control (N=101)	Treatment (N=271)	Control (N=105)	Treatment (N=270)	Control (N=80)
Pretest mean	3.8	3.8	8.4	7.6	9.8	9.7
Posttest mean	5.2	4.7	12.0	10.1	14.9	13.4
Change score mean⁶	1.4	0.9	3.6	2.5	5.1	3.7
Percent change	37%	24%	43%	33%	52%	38%

Classroom observations (in treatment and control classrooms) as well as a focus group held with treatment group students at the end of the study confirm a higher degree of engagement and interest on the part of the students when the online field trips were used. It is hypothesized that in addition to the reading supports, this is one reason for better content retention by students in the treatment classrooms.

Students in poverty who used the online field trips performed better on the unit tests than the students using only traditional methods. The researchers analyzed the performance on the classroom assessments for students in poverty (i.e., those who receive free and reduced meals—FARMS) and found that for the high-poverty students there were statistically significant

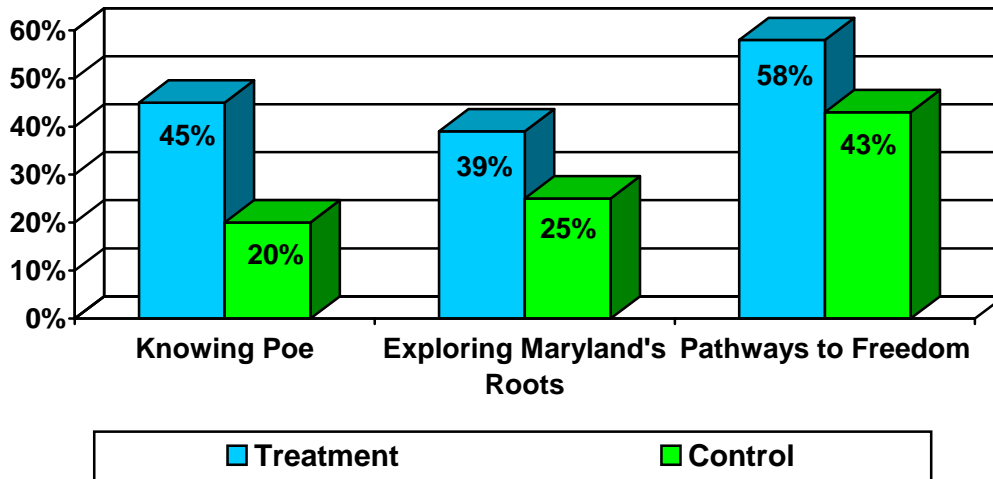
⁴ At SMS the baseline GMRT was conducted in May 2003 and the post-intervention GMRT was given in May 2004. At BTW the baseline GMRT was conducted in October 2003 and the post-intervention GMRT in June 2004.

⁵ Only 11 of the questions on the pre- and post-tests for Poe were matched and the total possible score comparable score was 11 (the other two tests matched 20 items).

⁶ All differences were statistically significant (for Poe $p < .05$, for Roots $p < .01$, and Pathways $p < .001$) with small to moderate effect sizes.

differences in change scores between treatment and control groups with all three field trips.⁷ The graphic below presents the percentage change in the scores between the initial test and the end-of-unit test for the treatment and control groups on the three EFTs.

Figure 1. Percent Change in Change Score for High Poverty Students

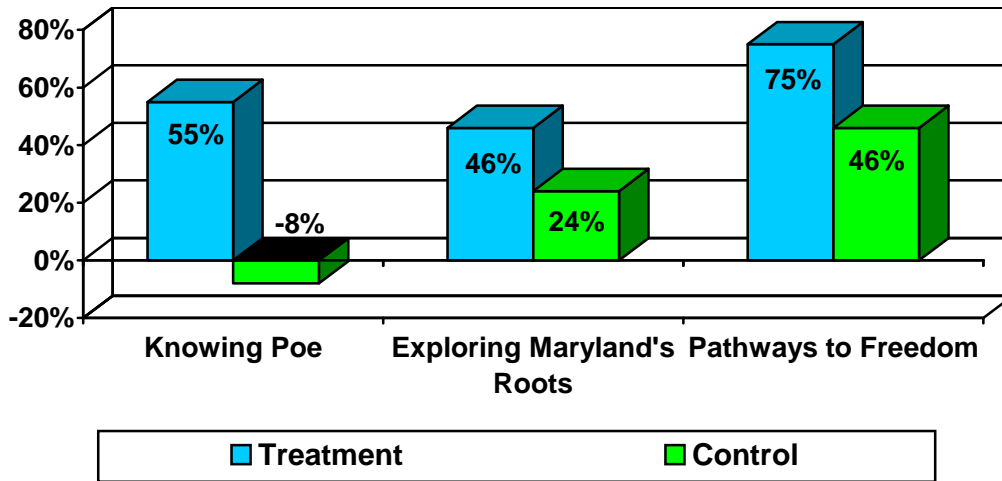


Students who are struggling readers and used the online field trips learned the content of each unit better than students in the traditional classrooms. The researchers analyzed the performance on the classroom assessments for students who scored in the lowest quartile of the baseline reading assessment with the Gates-MacGinitie Reading Test (GMRT). It is these students who have the most difficulty with reading and were the target audience for this enhanced product. The analysis showed that for these students there were statistically significant differences in change scores between treatment and control groups for all three field trips.⁸ Figure 2 displays the percentage change data.

⁷ All differences were statistically significant (for Poe and Roots $p < .01$, for Pathways $p < .05$) with moderate effect sizes.

⁸ All differences were statistically significant (for Poe $p < .001$ and for Pathways and Roots $p < .05$) with moderate to large effect sizes.

Figure 2. Percent Change in Change Score for Poor Readers⁹



Standardized Reading Test Results

Beyond performance on classroom-based knowledge tests similar to the tests that teachers give at the end of a unit of study, the critical question was whether longer-term reading skills were improved from the use of the reading supports provided to the students and teachers through the enhanced online field trips. The following two sections present the important findings on reading comprehension for students overall and for the target group.

Overall Results

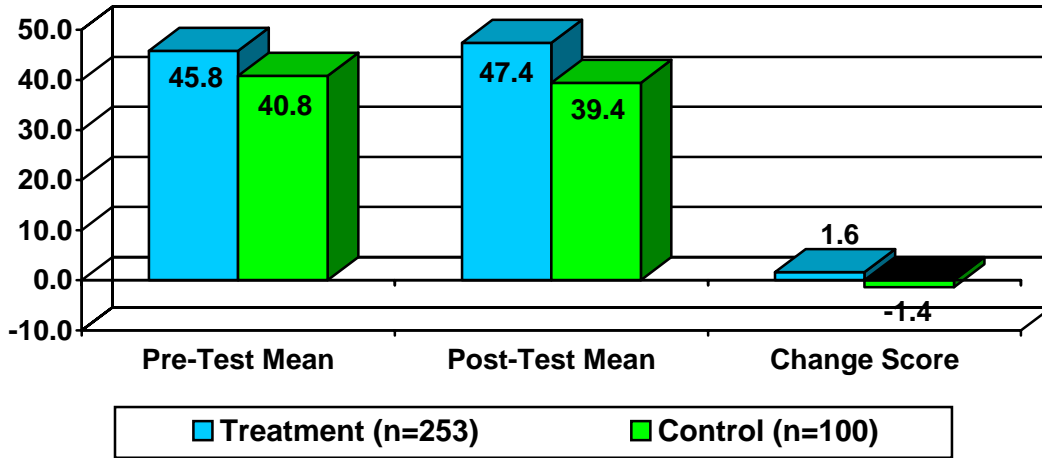
Students using the EFTs in the treatment group demonstrated an improvement in their reading comprehension skills on a standardized reading test compared to students in the control group. Analysis of the baseline and post-intervention administration of the Gates-MacGinitie Reading Test (GMRT) demonstrates a statistically significant improvement in reading comprehension scores (measured in NCEs—Normal Curve Equivalents¹⁰) by the

⁹ As measured by scores on the baseline GMRT—i.e., students in the lowest quartile (n=85).

¹⁰ Normal Curve Equivalent is a statistical (normalized) transformation of percentile ranks in which the range of reading achievement is divided into 99 equal units with a mean of 50. If a student proceeds from one year to the next with his/her normed group, the NCE will not change.

treatment group compared to the control group. Figure 3 presents the data for the pre/post reading tests for both sets of students.¹¹

Figure 3. Reading Comprehension Scores on GMRT Measured by Normal Curve Equivalent (NCE)



Looking at the Grade Equivalent (GE) score¹² is another way to understand progress measured by standardized reading tests. One would expect students to gain, on average, approximately eight months between October and June or approximately 10 months from one month of the school year to the same month in the following year. Table 2 provides data using this measure.

Table 2. Reading Comprehension Scores on GMRT Measured by Grade Equivalent (GE)

		Baseline Mean	Post-Intervention Mean	Gain
SMS				<i>Time = 10 months</i>
	Treatment (n=176)	6.7	7.8	11 months
	Control (n=46)	6.0	6.5	5 months
BTW				<i>Time = 8 months</i>
	Treatment (n=77)	6.0	7.0	10 months
	Control (n=54)	6.2	6.5	3 months

¹¹ Change score is statistically significant at the $p < .10$ level ($p = .053$) and effect size is small ($d = 0.2$).

¹² The GE number is read as the first number being the grade and the number after the decimal as the month. So, 7.8 is the eighth month of the seventh grade.

Therefore, when translated to GE scores, the control students advanced from a baseline GE of 6.0 or 6.2 to a GE of 6.5 on the post-intervention test (an advancement of 3 or 5 months, respectively). In one school, the treatment students advanced from a baseline GE of 6.7 to a post-intervention GE of 7.8 (an advancement of 11 months), and in the other they advanced 10 months—from a GE of 6.0 to 7.0. Thus, the treatment groups in the two different schools achieved more than the expected amount of gain between the two testing periods, while the control groups fell behind in their reading comprehension.

Results for Poor Readers

Struggling readers in the treatment group showed the most significant gains on the standardized reading test from one year to the next. Table 3 presents for the two testing periods the results of the students who had the worst performance in reading comprehension at the baseline testing. It is expressly for these students that the enhanced online field trips were developed. Students in the lowest quartile in the baseline measurement of the GMRT who used the EFTs achieved a notable gain in their comprehension score—twice as many NCEs as their counterparts in the control group.

**Table 3. Gates-MacGinitie Reading Tests: Baseline and Post-Intervention Comprehension Scores, Measured by NCE
Students in the Lowest Quartile – Treatment Group vs. Control Group**

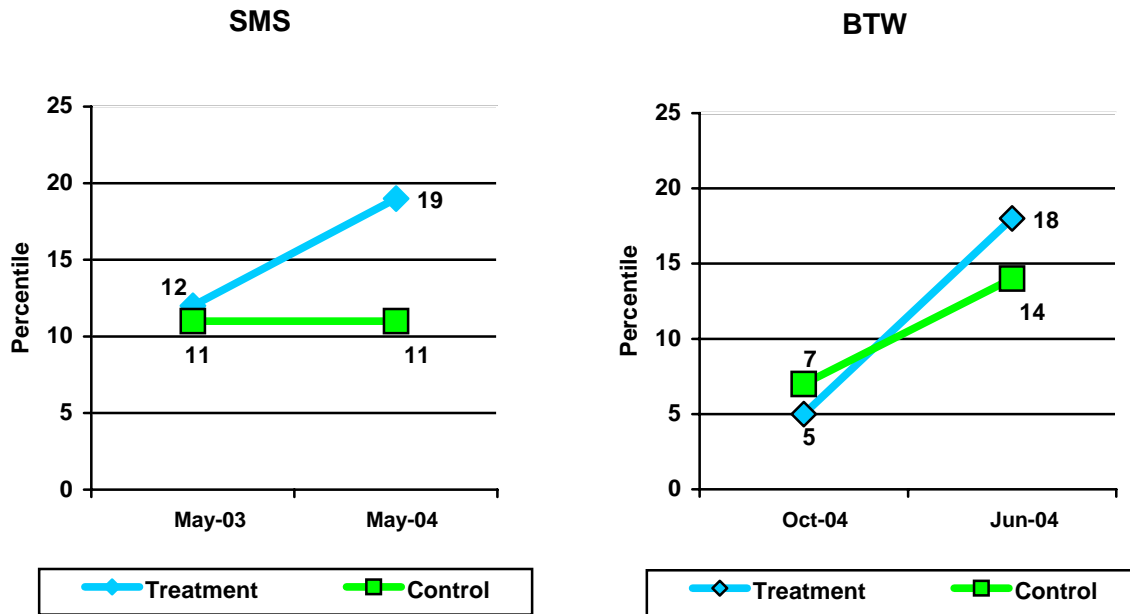
	Lowest Quartile Control (N=31)	Lowest Quartile Treatment (N=54)
Baseline Mean Comprehension NCE	21.3	20.1
Post-Intervention Mean Comprehension NCE	26.4	30.5
Mean Difference in Comprehension (NCE)¹³	5.1	10.4

Another way to assess change among the target group (the low-performing reader) is to look at how those students perform in terms of percentiles.¹⁴ Figure 4 displays these results by school. (They are displayed separately because a different grade is considered in each school.)

¹³ The change for the treatment group was statistically significant at the $p < .10$ level ($p = .056$) and shows a moderate effect size ($d = .4$).

¹⁴ Percentile Rank (PR) indicates where a raw score fits within a range of scores. The PR for a particular raw score provides the percentage of students whose raw scores were lower.

Figure 4. Change in Reading Comprehension Scores by the Lowest Quartile (National Percentile Rank)



At the first school, the poorest readers in treatment group improved on their performance from one year to the next, but those in the control group remained constant. In the other school, both groups improved on their performance, but as the line for the treatment group portrays, that group had a decidedly stronger performance from the start to end of the school year.

Discussion

The U.S. Department of Education defines Evidence-Based Education as “the integration of professional wisdom with the best available empirical evidence in making decisions in how to deliver instruction” (Whitehurst, 2002). Since the passage of No Child Left Behind, there is increased emphasis for evaluators in the field of education to utilize “scientifically-based” methodologies and to conduct random-assignment research. Educators and policymakers are relying on these studies and student achievement outcome data to obtain the empirical evidence of “what works”. With collaboration from two schools in conducting an experimental study over a full school year and measuring student achievement both within the classroom and on a standardized instrument, we have been able to study the potential benefits of a new educational intervention.

In *Teaching Children to Read*, the NRP noted that there is a striking absence of research on the incorporation of Internet applications into reading instruction and that additional research is needed on the value of multimedia applications in reading instruction (National Institute of Child Health and Human Development, 2000). This experimental study makes an important contribution to this research through the evaluation of high-quality content resources specifically designed to improve the reading performance of middle school students—and especially those who are struggling to perform in their academic classes.

References

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Whitehurst, G.J. (2002, October). *Evidence-based education (EBE)*. PowerPoint presented at the Student Achievement and School Accountability Conference, Washington, DC. Retrieved from <http://www.ed.gov/nclb/methods/whatworks/eb/edlite-index.html>

Funding for the development of products by Maryland Public Television and evaluation conducted by ORC Macro was provided by a Ready to Teach grant (#R286b020008) from the U.S. Department of Education.