

JUNE 2002

Within School Treatment and Control Study

Voyager Universal
Literacy System[®]
Orange County, Florida

2001–2002

Steven A. Hecht, Ph.D.
Florida Atlantic University

Joseph K. Torgesen, Ph.D.
Florida State University

Research Study In Orlando, Florida

Evaluation of an Approximately 11-Week Period of Use of the Voyager Kindergarten Curriculum

This study was conducted for a period of approximately three months to evaluate the effectiveness of the Voyager Program with economically disadvantaged kindergartners enrolled in Orange County Public Schools (Orlando, Florida). To that end, students were given pre and posttests of reading and reading-related outcomes. The effectiveness of the Voyager program, which was implemented for approximately an 11-week period, was compared with control classrooms.

METHOD

Participants

116 students initially participated in this evaluation. Sample attrition from pre to posttest was 8 students (5 control and 3 Voyagers). These 8 students' data were removed from all analyses. This left 58 students who were given the Voyager program as part of the regular educational practices at their school. The control group was composed of 50 students who were not given the Voyager curriculum. Available demographic characteristics of the students are shown in Table 1. The values in Table 1 represent the frequency of students along a range of demographic characteristics.

Table 1

Demographic Characteristics of Children (Values Represent Frequency of Children with Given Characteristic)

	Control	Voyagers
1. Gender		
Girls	25	27
Boys	25	31
2. Race		
White	7	13
African American	30	30
Hispanic	8	5
Other	4	2
3. Limited English Proficiency		
Yes	12	2
No	34	17
4. Eligible for free or reduced lunch		
Yes	27	15
No	5	0
Did not apply	14	5
5. Receives special education service		
Yes	3	1

Note: Values for gender based on 100 percent of cases. Values for race based on approximately 92 percent of sample. Approximately 60 percent of the complete sample were available for tallying the remaining frequency values.

Assignment of Teachers to Voyager and Control Groups

All students attended one of two schools that served economically disadvantaged students. The two elementary schools contained both Voyager and control classrooms. School A contained one Voyager and one control classroom, and School B contained two classrooms per group.

School principals made final decisions regarding which classrooms would receive the Voyager program. The principals were interviewed to determine the basis on which decisions were made regarding group assignment of teachers. The School A principal simply indicated that teachers volunteered. The School B principal assigned two teachers to the control group that she perceived were implementing the existing instructional approaches used in that school “well.” One teacher was not selected to participate in the study because she was “brand new.” Finally, the School B principal said that the remaining kindergarten teachers volunteered and were “excited” to be included in the Voyager group.

Based on school records, control group teachers had more years of employment in Orange County Public Schools than Voyager teachers at both School A (12 versus 6 years) and School B (on average, 7.5 versus 2.5 years). All participating teachers were certified and had earned a bachelor’s degree. One control teacher also had earned a master’s degree in early childhood education. All control teachers reported using the Houghton Mifflin reading program (see e.g., <http://www.eduplace.com/marketing/wins/ca/efficacy.pdf>).

Tests Administered for Pre and Posttest by Non-Teachers

1. **Word Identification.** This subtest, from the Woodcock Reading Mastery Test (Woodcock, 1987), requires students to name 106 individually presented words.
2. **Word Analysis.** This subtest, also from the Woodcock Reading Mastery Test (Woodcock, 1987), requires students to name 45 individually presented nonwords.
3. **Spelling.** This was the Spelling subtest from the Wide Range Achievement Test (WRAT, Wilkinson, 1995). This test requires students to write 13 letters and spell 15 words, including their first name. The total number correct on both the letter-writing portion and the spelling-of-words portion was recorded separately.
4. **Letter Name Knowledge.** Students named the 26 letters of the alphabet.
5. **Letter Sound Knowledge.** Students indicated the sounds that individual letters make in words. The highest possible score on this subtest was 38 correct.
6. **Print Concepts.** This was the Stones - Concepts About Print Test (Clay, 1979). The task has 24 items that measure students’ knowledge about print. The 24 items are embedded in a children’s storybook about a stone rolling down a hill that the student and examiner jointly read.
7. **Phonological Awareness – Elision.** This task was the Elision subtest from the Comprehensive Test of Phonological Awareness (CTOPP; Wagner, Torgesen, & Rashotte, 2000). Students were asked to say a word, then say what the word would be if a specified phoneme in the word were deleted. For example, students were instructed to “Say the word ‘cup.’ Now tell me what word would be left if I said ‘cup’ without saying /k/.” All phonemes to be deleted were consonants, the position of which varied at random. The remaining phonemes formed a word. There was a total of 20 items.
8. **Phonological Awareness – Segmenting.** This task was the Segmenting subtest from the CTOPP. Students listened to words and were instructed to “tell me each sound you hear in the word in the order that you hear it.” There was a total of 20 items.

9. **Phonological Awareness – Blending.** This is the most direct measure of a student’s skill in combining phonological elements to form a word. Phonemes are presented separately by the test giver, and the student is required to blend them together and identify the word they make when pronounced together. There was a total of 20 items.
10. **Expressive Vocabulary.** This was the vocabulary subtest of the Stanford-Binet: Fourth Edition (Thorndike, Hagen, & Sattler, 1986). Students uttered the definition of words. The vocabulary subtest is used by reading researchers as an estimate of general verbal IQ because it is the most highly correlated subtest with overall verbal IQ (Sattler, 1988). The first 36 items were used in the present evaluation.
11. **Rapid Automatized Naming (RAN).** For the practice trial, students were shown one row of five letters and asked to name them; corrective feedback was given. The remaining trials were given if students were able to name the letters on the practice card. For test trials, six rows of five single letters per row were arrayed on a card, and the students were instructed to name the letters as fast as they could beginning at the top and continuing to the bottom. Two test trials were given using two cards with differently arranged letters. Times were recorded on a stopwatch, and scores were based on average time for the two series (rounded to the nearest centisecond). Accuracy also was recorded.

Procedure

Administration of the pretest measures to both control and treatment (i.e., Voyager) students began on February 4, 2002, and ended February 27. Posttesting began April 29, 2002, and ended May 15. All testing activities were primarily supervised by Dr. Timothy Blair, professor, Department of Teaching and Learning Principles, College of Education, University of Central Florida. Testing was done by: a) two retirees hired by Voyager, b) two graduate students sent from Alabama by Voyager, c) one assessment specialist, and d) several graduate and undergraduate students enrolled in Dr. Tim Blair’s department. Pretesting of all but one classroom took approximately three weeks. A fourth week was needed to test one control classroom. This control classroom was tested at the request of the principal of School B Elementary. She felt that a more valid comparison between Voyager and control groups would be enhanced by replacing one control classroom with a more experienced teacher. The use of the Voyager program first began approximately the same day that the pretesting was initiated and continued until the end of the school year. Tasks were administered in random order. There were no consistent associations (i.e., correlations) between the dates that students were tested and performance on the pre or posttest measures.

RESULTS

Voyager and control children were compared with respect to posttest levels of word identification, word analysis, spelling letters, spelling words, letter name and sound knowledge, print concepts, phonemic elision, segmenting and blending, vocabulary, and rapid automatized naming. Analysis of covariance was employed to test group differences in performance, with pretested ability specified as a covariate (i.e., held constant).

Group Differences in Posttest Performance. Table 2 shows the mean levels of performance on all pre and posttest measures, separated by group. In general, there were no statistically significant differences between groups with respect to pretest performance on the reading and reading-related measures. The fact that the classes were not significantly different in all but one literacy measure in February indicates that there were no striking differences in teacher effectiveness so far in the year. In other words, as of the pretest time period, there was no consistent evidence of differential effectiveness between teachers in the control and Voyager groups in terms of their ability to stimulate growth in important early literacy skills.

Table 2

Average Raw Pre and Posttest Performance by Group (Voyager n = 58; Control n = 50)

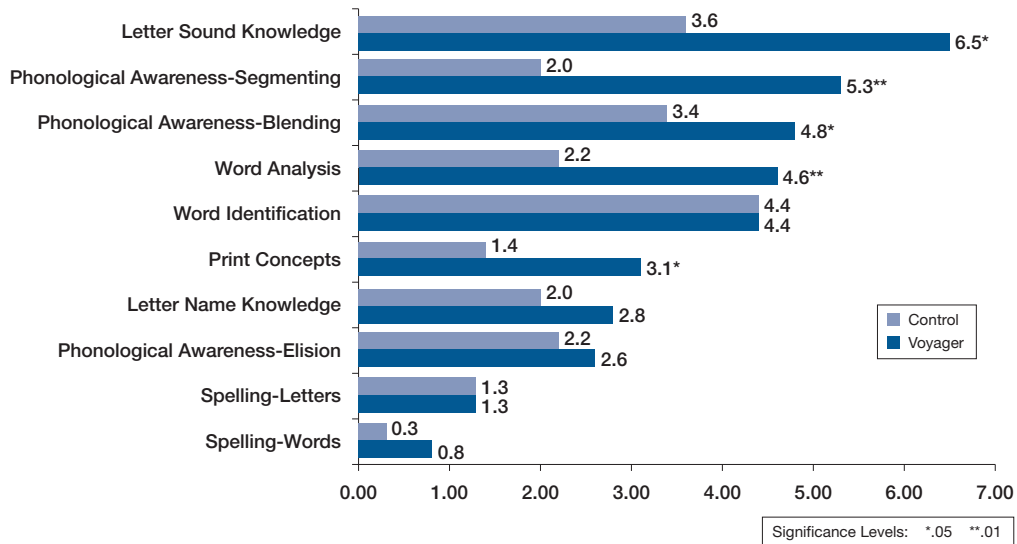
	Control			Voyagers		
	Pretest	Posttest	Gain	Pretest	Posttest	Gain
1. Word Identification	6.4	10.8	4.4	4.4	8.8	4.4
2. Word Analysis	1.6	3.8	2.2	1.0	5.6**	4.6**
3. Spelling — Letters	11.9	13.2	1.3	11.9	13.2	1.3
4. Spelling — Words	1.3	1.6	0.3	1.0	1.8	0.8
5. Letter Name Knowledge	23.1	25.1	2.0	23.1	25.9	2.8
6. Letter Sound Knowledge	19.1	22.7	3.6	19.3	25.8**	6.5*
7. Print Concepts	12.4	13.8	1.4	11.6	14.7*	3.1*
8. Phonological Awareness — Elision	2.2	4.4	2.2	3.5*	6.1	2.6
9. Phonological Awareness — Segmenting	1.5	3.5	2.0	2.5	7.8***	5.3**
10. Phonological Awareness — Blending	4.0	7.4	3.4	5.0	9.8*	4.8*
11. S-B Vocabulary	16.3	19.5	3.2	16.7	19.6	2.9
12. Rapid Automatized Naming — Time^a	50.4	48.3	-2.1	56.9	52.5	-4.4
13. Rapid Automatized Naming — Accuracy	34.2	33.1	-1.1	32.8	34.8	2.0*

Note: Significance value denotes that one group scored significantly higher than the other group. A significant effect for gain indicates that that group showed greater improvements in performance over the pre to posttest time period than the other group.

* p < .05 ** p < .01 *** p < .001

^a Lower times are indicative of better performance (i.e., faster access to name codes of letters in long-term memory).

Orange County, Florida
A Comparison Graph of Control Students to Voyager Students
Raw Score Gains on Ten Measures



Note: Significance value denotes that one group scored significantly higher than the other group. A significant effect for gain indicates that that group showed greater improvements in performance over the pre to posttest time period than the other group.

*p < .05 **p < .01 ***p < .001

Control n=50 Voyager n=58

There were 10 more students known to have limited English proficiency in the control group than the Voyager classrooms. (See Table 1.) There was a very slight, non-significant advantage for Voyagers with respect to both pre and posttest expressive vocabulary (i.e., S-B vocabulary). Importantly, posttest performance was significantly better for students in the Voyager group than controls on several measures that require explicit awareness and access to the sound structure of oral language. Voyager students showed superior performance on the word analysis, letter sound knowledge, print concepts, phonemic segmenting, and phonemic blending tasks. The control students did not score significantly higher than the Voyager students did on any measure.

To get a better sense of the meaningfulness of the outcomes, we compared average performance of the current participants with estimated mean performance of U.S. children. Table 3 depicts posttest performance expressed as age-based standard scores. Effect sizes (Cohen’s standardized mean differences, *d*) also are presented, which should be interpreted as the difference in group means expressed in standard deviation units. Effect sizes of .80 or higher suggest large effects, .50 to .79 suggest medium-sized effects, and .20 to .49 reflects relatively small effects (Cohen, 1988; Murphy & Myers, 1998). Effect sizes for phonemic segmenting and blending were high to medium in size, suggesting a meaningful advantage for Voyagers over controls, particularly for phonemic awareness.

Table 3

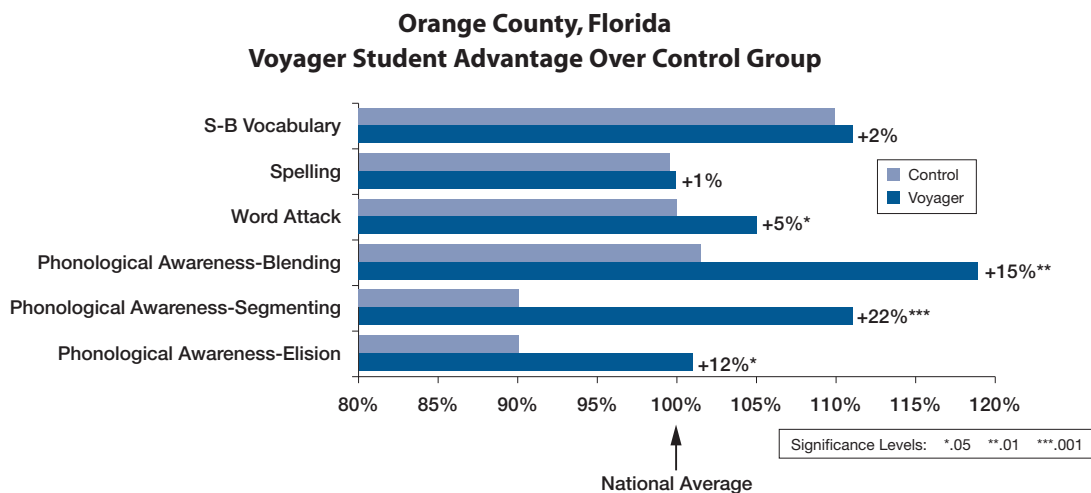
Percentage of Posttest Points Relative to Expected Performance Based on the Standardization Sample, Separated by Group (Voyager n = 58; Control n = 50).

Note: Percentage points reflect the average number of items earned by the current sample relative to the mean number of items scored by the standardization sample. Voyager Advantage stands for how many more percentage points were obtained by Voyagers than the control group. *d* stands for the difference in group means expressed in standard deviation units.

	Control	Voyager	Voyager Advantage	<i>d</i>
1. Phonological Awareness — Elision	90%	102%	12%*^a	.40
2. Phonological Awareness — Segmenting	90%	112%	22%***	.90
3. Phonological Awareness — Blending	103%	118%	15%**	.50
4. Word Attack	100%	105%	5%	.38
5. Word Identification	104%	102%	-2%	.00
6. Spelling	99%	100%	1%	.07
7. S-B Vocabulary	110%	112%	2%	.13

* $p < .05$, ** $p < .01$, *** $p < .001$

^a This Voyager Advantage not significant when pretest age-based standardized scores for elision were held constant.



Note: Percentage points reflect the average number of items earned by the current sample relative to the mean number of items scored by the standardization sample. Voyager advantage stands for how many more percentage points were obtained by Voyager students than the control group.

* $p < .05$ ** $p < .01$ *** $p < .001$

Control n=50 Voyager n=58

SUMMARY

Low-income kindergarten students in an 11-week Voyager study in Orange County, Florida, made substantial progress on several kinds of early literacy skills that are crucial for reading development. Specifically, Voyager students significantly outperformed control students on measures of print concepts, letter sound knowledge, and nationally standardized measures of phonemic awareness. Voyager students also demonstrated superior ability to apply their newly acquired literacy skills toward sounding-out unfamiliar nonwords when compared to control students. The findings from this study are particularly impressive considering that students only were exposed to approximately three months of Voyager instruction. Regarding future studies, the current findings suggest that increasingly larger differences in reading-related skills between Voyager and control students will emerge when students are given more months of exposure to Voyager instruction.

REFERENCES

- Clay, M. (1979). *Stones — Concepts About Print Test*. Auckland, New Zealand: Heinemann Publishing.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Murphy, K. R., & Myers, B. (1998). *Statistical Power Analysis: A Simple and General Model for Traditional and Modern Hypothesis Tests*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Sattler, J. M. (1988). *Assessment of Children* (3rd ed.). San Diego, CA: J. M. Sattler.
- Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale* (4th ed.). Chicago, IL: Riverside Publishing.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2000). *Comprehensive Test of Phonological Processing*. Austin, TX: Pro-Ed.
- Wilkinson, G. S. (1995). *The Wide Range Achievement Test 3*. San Antonio, TX: Psychological Corp.
- Woodcock, R. W. (1987). *Woodcock Reading Mastery Tests—Revised*. Circle Pines, MN: American Guidance Service.