

Fourth Graders' Growth in Reading Comprehension and Fluency

A Pretest-Posttest Randomized Control Study Comparing Reading Mastery and Scott Foresman Basal Reading Program



NATIONAL INSTITUTE FOR DIRECT INSTRUCTION

2010

Jean Stockard, NIFDI Director of Research

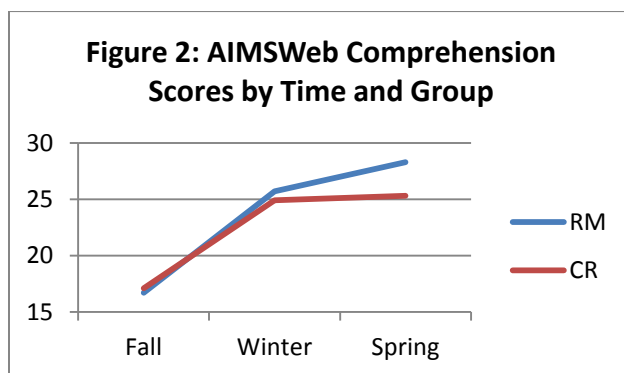
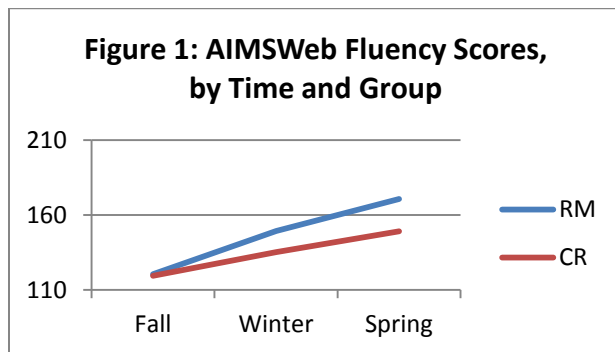
NIFDI Technical Report 2010-3

Fourth Graders' Growth in Reading Comprehension and Fluency

A Pretest-Posttest Randomized Control Study Comparing Reading Mastery and Scott Foresman Basal Reading Program

Executive Summary

The purpose of this study was to examine growth in reading comprehension and fluency of fourth grade general education students receiving instruction in *Reading Mastery, Signature Edition* or Scott Foresman's *Celebrate Reading* program. Students in a Midwestern elementary school were randomly assigned to one of the programs and received instruction for ninety minutes a day. Each group had two teachers (four teachers in all), who were randomly assigned to teach the different groups. The analysis is based on data from 52 students (26 in each group). The curriculum-based AIMSweb measurement system was used to gather fluency and comprehension data. While there were no significant differences between the groups at baseline, the RM group had significantly stronger growth in both fluency and comprehension skills over the school year. Effect sizes associated with the differences at spring were .75 for fluency and .41 for comprehension, both well beyond the level usually used to denote educational significance. The changes over time for the two groups are illustrated in Figures 1 and 2 below.



Fourth Graders' Growth in Reading Comprehension and Fluency

A Pretest-Posttest Randomized Control Study Comparing Reading Mastery and Scott Foresman Basal Reading Program

The purpose of this study was to examine growth in reading skills of fourth grade general education students in the Direct Instruction program, *Reading Mastery*, and the Scott Foresman Program, *Celebrate Reading*, using a randomized control trial. Sections below describe the methodology and the results of the study.

Methodology

The study was conducted in a Midwestern elementary school throughout one academic year with high-achieving (Tier 1) fourth grade students. At the end of third grade, before the study began, eighty percent of the students in the sample scored above the 50th percentile on AIMSweb tests of reading fluency and comprehension. Half of the students scored above the 75th percentile. (See below for more detail on the measures.) Lower achieving (tiers 2 and 3) fourth grade students in the school were not included in the study but instead received intensive remedial instruction. All of the study participants were classified as non-Hispanic white.

In the spring of the students' third grade year all of the high-achieving students were randomly assigned to two conditions, using an alternative assignment with random start technique. In the fall of their fourth grade year, one group of students received instruction in *Reading Mastery Signature Edition*, and the other group received instruction in *Celebrate Reading*, which the school had been using for the previous 8 years. Both groups received instruction for ninety minutes a day. Each group had two teachers (four teachers in all), and teachers were randomly assigned to teach the different programs.

The analysis reported here was limited to 52 students (26 in each group) who had data available for all testing periods. Substantively identical results appeared when those with missing data at one time point (n=8) were included. Standard processes for obtaining approval from the school district and school were followed.

The curriculum-based AIMSweb measurement system was used to gather data on reading skills. Two measures were used: a measure of reading fluency in the form of number of words read correctly within one minute and a measure of reading comprehension in the form of the number of correct answers on a maze test. Data were gathered at the time periods recommended by the AIMSweb group: in the fall, shortly after the start of the school year and the beginning of instruction; again in winter, approximately halfway through the school year; and in the spring near the end of the school year.

Results

Table 1 reports the data for the fluency measure for each group at each testing period as well as associated inferential tests and effect sizes. As would be expected given the random assignment procedure, there were no significant differences in scores at baseline. At the winter testing the *RM* group had significantly higher fluency scores than the *CR* group and this difference widened in the spring testing. The significant interaction effect in the repeated measures analysis of variance results indicates that the increase over time was significantly greater for the *RM* students than for the *CR* students. The effect size of .75 associated with the difference between the groups in the spring is considered large (Cohen, 1988). On average, the students in the *RM* group had a weekly increase of 1.52 words per minute in fluency scores (s.d. = .61), equal to the “ambitious” rate identified by Deno and associates (n.d.). In contrast, those in the *CR* group had an increase of .90 words per minute (s.d. = .42), equivalent to Deno and associates’ “reasonable” pace. The difference between these average rates was statistically significant ($t = 4.26$, $df = 50$, $p < .001$).

Table 2 gives data for the maze test of reading comprehension for the two groups at each testing period. There were no significant differences between the groups at the fall or winter testing periods, but the differences at the spring testing approached traditional levels of significance ($p = .07$), and the associated effect size of .41 surpassed the level of .25 usually seen as educationally important (Fashola & Slavin, 1997). The interaction effects associated with the analyses of variance were significant for the analysis that included only the fall and spring test scores.

Conclusion

The data from this study of high achieving, intermediate level students support evidence from previous studies regarding the effectiveness of Reading Mastery in enhancing both reading fluency and reading comprehension.

Table 1

AIMS Web Fluency Scores by Group and Testing Period, T-Tests, Effect Sizes and Repeated Measures Anovas

<i>Group Means by Testing Period, T-Tests and Effect Sizes</i>							
<u>Testing Period</u>	<u>CR</u>		<u>RM</u>		<u>t</u>	<u>p</u>	<u>d</u>
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>			
Fall	119.3	28.6	120.4	22.3	0.15	0.44	0.04
Winter	135.2	28.6	149.3	28.4	1.80	0.04	0.50
Spring	149.0	27.3	170.6	29.9	2.72	0.005	0.75
<i>Repeated Measures Analyses of Variance</i>							
<u>Comparisons</u>	<u>F-Gp.</u>	<u>prob.</u>	<u>F-Time</u>	<u>prob.</u>	<u>F-Int.</u>	<u>prob.</u>	
Fall - Winter	1.09	0.30	143.53	<.001	12.39	0.001	
Fall-Spring	2.50	0.12	275.66	<.001	18.15	<.001	
Winter-Spring	5.55	0.02	59.62	<.001	2.61	0.11	
Fall - Winter - Spring	2.88	.10	147.03	<.001	10.43	<.001	

Note: Degrees of freedom for the analyses of variance with three time periods were 2,49 for time and the interaction with time and 1,50 for group. Degrees of freedom for analyses of variance with two time periods were 1,50. T-test probabilities are one-tailed to match the substantive hypothesis. There were 26 cases in each group.

Table 2

AIMS Web Maze Scores by Group and Testing Period, T-Tests, Effect, Sizes, and Repeated Measures Anovas

<i>Group Means by Testing Period, T-Tests and Effect Sizes</i>							
<u>Testing Period</u>	<u>CR</u>		<u>RM</u>		<u>t</u>	<u>p</u>	<u>d</u>
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>			
Fall	17.1	3.7	16.7	3.5	-0.35	0.37	-0.10
Winter	24.9	5.1	25.7	6.2	0.51	0.31	0.14
Spring	25.3	7.1	28.3	7.4	1.49	0.07	0.41

<i>Repeated Measures Analyses of Variance</i>						
<u>Comparison</u>	<u>F-Gp.</u>	<u>prob.</u>	<u>F-Time</u>	<u>prob.</u>	<u>F-Int.</u>	<u>prob.</u>
Fall - Winter	0.05	0.83	114.33	<.001	0.54	0.47
Fall-Spring	0.94	0.34	148.88	<.001	4.25	0.04
Winter-Spring	1.54	0.22	2.37	0.13	1.30	0.26
Fall - Winter - Spring	0.87	0.36	101.36	<.001	2.10	0.13

Note: Degrees of freedom for the analyses of variance with three time periods were 2,49 for time and the interaction with time and 1,50 for group. Degrees of freedom for analyses of variance with two time periods were 1, 50. T-test probabilities are one-tailed to match the substantive hypothesis.

References

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd edition)*. Hillsdale, NJ: Erlbaum.
- Deno, S., Lembke, E., & Anderson, A. R. (n.d.) *Progress Monitoring Study Group Content Module. Expanding-Educational-Horizons*.
- Fashola, O. S. & Slavin, R.E. (1997). Promising programs for elementary and middle schools: Evidence of effectiveness and replicability. *Journal of Education for Students Placed at Risk*, 2, 251-307.