

Reading Rescue® in Inner City Schools: An Experimental Study Examining Reading Outcomes of a One-on-One Tutoring Intervention

Patricia Muller, Ph.D., Associate Director

Center for Evaluation and Education Policy

Indiana University

509 E Third Street

Bloomington, Indiana 47401-3654

Phone: (812)855-4438

Fax: (812) 856-5890

E-mail: pmuller@indiana.edu

July 10, 2004

Reading Rescue® in Inner City Schools: An Experimental Study Examining Reading Outcomes of a One-on-One Tutoring Intervention

Abstract

The purpose of this study was to examine the effectiveness of the one-on-one tutoring component of Reading Rescue®, specifically addressing the following research question: Do students who receive one-on-one Reading Rescue® tutoring achieve better reading results compared to similar students who do not receive one-on-one Reading Rescue® tutoring? A randomized pretest-posttest control group experimental design was used to study the effects of the Reading Rescue® tutoring program on reading achievement in six elementary schools located in one inner-city school district in New York City with predominantly Hispanic students. Findings of this study support those from research conducted on other structured one-on-one tutoring programs, showing that the reading achievement of elementary students at risk of reading failure can be improved through the use of supplemental, adult-instructed one-to-one reading intervention. Results of the ANCOVA analyses comparing treatment and control groups on posttest reading achievement, after controlling for pretest differences, provides evidence that Reading Rescue® increases students' reading achievement. Students who received one-on-one Reading Rescue® tutoring scored significantly higher ($p < .05$) on the posttest measure of reading achievement than did similar students who were randomly assigned to the control group. The findings from this study are particularly significant given that the majority of research on literacy interventions is based on program implementation in one particular setting with primarily monolingual students, whereas this study is based on data drawn from six separate elementary schools with large percentages of English Language Learners.

Reading Rescue® in Inner City Schools: An Experimental Study Examining Reading Outcomes of a One-on-One Tutoring Intervention

Educators and researchers agree that good reading skills are important to successful learning and that such skills are best learned at an early age (Kennedy & Others, 1986). Early reading success is also important for developing children's self-confidence and motivation to learn (Ambruster, Lehr, & Osborn, 2003). However, according to the U.S. Department of Education's National Center for Education Statistics (2003), the overall reading achievement of the nation's fourth grade students has remained constant since 1992. In 2003, 37% of fourth grade students performed below basic levels of proficiency on the National Assessment of Educational Progress (NAEP) reading test. Responding to this trend, the Elementary and Secondary Education Act (ESEA) of January 2002, or *No Child Left Behind Act*, has renewed expectations for the improvement of reading achievement and evidence that reading programs are effective. Research has identified early identification of, and early intervention for, at-risk readers as one of the best ways to increase reading achievement (Hiebert, 1994; Torgesen, 2002; Wasik & Slavin, 1993). As a consequence, interest in remedial reading and preventative assistance programs for students with reading difficulties continues to be a focus for educational reform in schools across the country.

In particular, one method that has proven effective in providing remediation for struggling readers in the primary grades is one-to-one pull out intervention in the form of tutoring (Allington, 2004; Elbaum, Vaughn, Hughes & Moody, 2000; Institute of Education

Sciences, 2003; Pinnel, Lyons, Deford, Bryk, & Seltzer, 1994; Shanahan, 1998; Wasik & Slavin, 1993). Given the increasing national focus on scientifically-based research and evaluation, and the U.S. Department of Education's identification of experimental design as one of the most rigorous and valid methods of evaluating program impact, the purpose of this study was to examine the effects of a structured one-on-one tutoring program on students' reading achievement. Specifically, this study used a randomized experiment to examine the impact of Reading Rescue® tutoring on the reading achievement of inner city students in six New York City elementary schools serving low socio-economic and primarily Hispanic students.

Background

Reading Rescue®, a comprehensive early intervention model developed in cooperation with the University of Florida and now sponsored by a charitable, non-profit organization (Literacy Trust, Inc.), is designed to prevent reading failure through one-on-one tutoring for the lowest performing students, while also improving the quality of literacy instruction provided for all students in the classrooms of participating teachers. Reading Rescue® consists of the following components: (1) a three-year professional development sequence delivered on site, within a school or district, that trains a school's staff (typically classroom teachers, but also reading and language teachers, paraprofessionals, media specialists, administrators, etc.) to provide research-based one-on-one tutoring for at-risk students in the early grades, (2) technical support for the school restructuring required to implement a multi-tiered assessment and instruction delivery model comprised of large and small group instruction and one-on-one tutoring during the regular school day, (3) large group screening and individual assessments of emergent literacy and reading, (4) Summer

Institutes that prepare each school's Reading Rescue® Coordinator as a Literacy Coach and in-house trainer, (5) formation of a school-based Peer Coaching team under the direction of each school's Reading Rescue® Coordinator, and (6) continuous data collection monitoring program effectiveness. Each component supports the goal of fluent reading for the lowest performing students, as well as enhanced instruction for all students, in the early grades. The model is based on several diverse bodies of research in the fields of teacher professional development and adult learning; school reform and restructuring; and early literacy acquisition, reading disabilities, and reading instruction.

Over the three-years of adoption, the tutors participate in professional development sessions provided by The Literacy Trust, Inc. Tutors' theoretical understanding and pedagogical skills are developed through reading and discussion, analysis of tutoring videotapes, and hands-on practice with feedback. In addition, the professional development is on-going through meetings of each school's peer coaching group. During the second year of adoption, the Coordinator observes tutors to ensure compliance with the instructional model, and leads the tutors in the analysis of the Coordinator's own tutoring recorded on videotapes. Over the three year implementation process, which includes the training of an in-school Literacy Coach during the Summer Institute, capacity is built for self-maintenance of the program within the school.

Students' performance on large group, pencil-and-paper assessments (i.e. Reading Rescue® Classwide Literacy Screening Assessments) are used to identify students who can be provided remediation effectively in large and small groups, and those students who

can be taught most effectively through one-on-one tutoring during the first semester. Students whose scores on these screening assessments place them at the top of the bottom quartile are considered the best candidates for one-on-one tutoring, with levels of emergent literacy skills sufficient to enable them to accelerate with skilled tutoring. Meanwhile, Reading Rescue® recommends the very lowest performing students (bottom of the bottom quartile) receive small group instruction designed to develop the basic knowledge and skills needed to benefit from the one-on-one tutoring in the second semester. One-on-one tutoring is provided through daily thirty minute tutoring sessions, conducted by trained and monitored classroom teachers and paraprofessionals during times when other students are engaged in non-instructional activities (e.g. recess, lunch, etc.). Each Reading Rescue® lesson includes work on phonological awareness in keeping with students' needs, sequential phonics instruction informed by assessment data, fluency building, and the development of vocabulary and comprehension.

Originally developed in 1993, Reading Rescue® has trained teachers and other committed staff in over one hundred schools in seven states to provide research-based tutoring.

Preliminary evidence suggests that Reading Rescue® has been effective in improving students' reading achievement at these schools. In addition to strong anecdotal and self-report evidence from long-standing programs, evaluations and statistical analyses have been conducted by individual schools and school districts (Capella, 2003; Parman, 2001; Price, 2002; Pugh, 2001). All schools who adopt Reading Rescue® are also contractually required to regularly submit test results for tutored students and their peers to the Literacy Trust Inc., who in turn uses these data to continuously monitor program effectiveness.

Although the primary focus of Reading Rescue® has been on program improvement, the Literacy Trust Inc. has also periodically used these data to conduct statistical analyses that document program impact (Britt, 2002; Gibson, 2002; Hoover, 1996; Hoover, 1999; Hoover & Sullivan, 1996).

However, as with most education programs, the type of scientifically-based evaluation discussed in the *No Child Left Behind* legislation has not been conducted to validate the program's effectiveness. This study addresses the need for such evaluation through the use of a rigorous, valid methodology (i.e. random, experimental-control group design) to examine the impact of the one-on-one tutoring component of Reading Rescue® on students' reading achievement.

Participants/Method

The purpose of this study is to examine the effectiveness of the one-on-one tutoring component of Reading Rescue®, specifically addressing the following research question:

Do students who receive one-on-one Reading Rescue® tutoring achieve better reading results compared to similar students who do not receive one-on-one Reading Rescue® tutoring?

A randomized pretest-posttest control group experimental design (Gay & Airasian, 2000) was used to study the effects of the Reading Rescue® tutoring program on reading achievement in six elementary schools located in one inner-city school district in New York City with predominantly Hispanic students. Three of the schools were in the first

year of the three-year implementation process for Reading Rescue®, and three schools were in the second year of implementation of the program.

In terms of demographics, five of the six schools serve students in Kindergarten through fifth grade, with enrollment sizes ranging from approximately 1100 to 1800 students; and one elementary school serves students from Kindergarten through second grade, with an enrollment size of approximately 400 students. All schools are Title I elementary schools, with approximately 82-99% of enrolled students eligible for free lunch (Mean = 92.0%, Median = 93.5%). The majority (approximately 87-98%) of the student body at each school is Hispanic; and approximately 27-41% of students at these schools are English Language Learners with limited English ability, receiving Bilingual or English as a Second Language instruction (Mean = 35.0%, Median = 37.0%). Across these schools approximately 26-46% of students met or exceeded state standards in English Language Arts during the 2001-2002 academic year (Mean = 35.3%, Median 34.0%).

Each of the participating schools (N=6) administered the Reading Rescue® large group, pencil-and-paper screening assessment (i.e. Reading Rescue® Classwide Literacy Screening Assessments) to its first grade students, and subsequently identified those students in the bottom quartile. From the students identified as potential participants for one-on-one Reading Rescue® tutoring, students within each school were randomly selected for participation in the reading intervention. Subsequently, students in the treatment group received one-on-one Reading Rescue® tutoring in addition to their regular classroom instruction, and students in the control group received regular classroom

instruction without one-on-one Reading Rescue® tutoring. Some control group students also received small group instruction in literacy and/or ESL, although this was not a systematically applied intervention for these students. Reading Rescue coordinators from these schools reported that approximately 90 to 100% of the first graders who received one-on-one tutoring were English Language Learners (i.e. students who come from homes in which a language other than English is spoken, in this case primarily Spanish); and given the random assignment of treatment and control group students, similar percentages of control group students who did not receive Reading Rescue® tutoring are also likely to have been English Language Learners.

The study used the Gates-MacGinitie Reading Test (GMRT), Fourth Edition, as the dependent variable to provide a standardized measure of reading achievement. The GMRT is a timed multiple-choice test administered in groups, providing valid and reliable norm-referenced scores that include extended scale scores and Normal Curve Equivalents (NCEs). As opposed to studies that use outcome measures that specifically target the strategies taught by the program, this study used the GMRT as the outcome measure in order to provide a more valid and standardized measure of actual impact on students' reading achievement, regardless of the specific reading strategies or technique used. In other words, if an evaluation uses a measure that is closely aligned with the strategies taught by a particular program, results are biased in favor of the program (e.g. children taught to use context to predict words rather than sounding them out will score better on reading measures based on predictable text rather than measures based on authentic text). Therefore, the use of a norm-referenced, standardized measure of achievement in this

particular study provides a more fair measure of achievement for both treatment and control group students.

As part of this study, the GMRT Fourth Edition, BR Level (Beginning Reading Level) was administered to all first grade students, including regular classroom students (i.e. those students who did not score in the bottom quartile on the large group assessment, and therefore were not identified as potential candidates for the one-on-one tutoring assistance). This test was administered during Fall 2002, prior to any of the students receiving the one-on-one tutoring, and constitutes the pretest data for the study. In Spring 2003, when most students had completed the one-on-one Reading Rescue® tutoring, posttest data were collected using the GMRT, Level 1, Form S. In addition, the GMRT was administered to the treatment and control groups at the end of the 2002-2003 school year to provide follow-up data for the examination of any residual effects of the tutoring. Given the costs of purchasing and scoring the GMRT, and the ability to examine the residual effect without the test score data from regular classroom students, a decision was made to not administer the GMRT to this particular group of students at the end of the school year.

To determine the impact of one-on-one Reading Rescue® tutoring on reading achievement, this study compared students in the treatment group to similar students assigned to the control group. For a treatment-control design with both premeasures and postmeasures, analysis of covariance (ANCOVA) is the most appropriate analysis (Tabachnick & Fidell, 1996). Therefore, ANCOVA was used to compare the posttest

reading achievement of students who received one-on-one Reading Rescue® tutoring to similar students in the control group who did not receive the tutoring intervention, with the pretest treated as the covariate to control for random individual differences¹. Unlike research on Reading Recovery that has been criticized for only including those students who are determined to have completed the program successfully, resulting in inflated reports of impact on achievement (Elbaum, Vaughn, Hughes & Moody, 2000; Shanahan & Barr, 1995), this study includes all students who were initially identified as part of the treatment group and received any Reading Rescue® one-on-one tutoring. This paper also provides descriptive data that compares both treatment and control group students to regular classroom students, in order to provide a context for better understanding students' reading achievement².

In addition to the pretest-posttest analysis, repeated measures analysis of variance (ANOVA) was used to examine differences between treatment and control groups using data from all three test occasions: the pretest, posttest and follow-up scale score achievement data. Given that two of the six participating schools did not administer the

¹ Potential candidates for the tutoring intervention (i.e. those students in the top of the bottom quartile on the large group reading assessment) were randomly assigned to groups. However, this random assignment does not guarantee equality among groups, but rather only assures there are no *systematic* differences between groups to begin with, within probability limits. The experimental design does not diminish the effects of *random individual differences* that can both spread out scores among subjects within a group, and create differences among groups that are not associated with the treatment. ANCOVA diminishes the effects of these individual differences by adjusting for them statistically, thereby providing a more powerful test of differences (Tabachnick & Fidell, 1996).

² Given the focus of the study, and to protect against inflated Type I error, the data for regular classroom students is provided as contextual data rather than incorporated into the primary analysis. Although the study could have used ANCOVA to make comparisons among the three groups (treatment, control, and regular classroom students), the post hoc comparisons necessary to determine where the actual differences occur would inflate the Type I error rate. In addition, differences in reading achievement are to be expected between those students identified as candidates for reading intervention (i.e. scoring in the bottom quartile of the class on a large group reading assessment), and regular classroom students (i.e. primarily those students in the top three quartiles on the large group reading assessment). Therefore, running this larger number of comparisons is unwarranted.

third testing due to logistical and scheduling issues, sample sizes for both the treatment and control groups are much more limited for this particular analysis, and therefore are presented separately.

For all analyses, data was aggregated across schools given that the purpose of the study was to examine program effects across schools, rather than individual differences that might exist between schools in the implementation or effectiveness of Reading Rescue® . Preliminary data screening and analyses supported this decision given that no outliers based on school groupings were detected that might affect the results.

Results

Preliminary analyses were conducted to ensure that there were, in fact, no statistically significant differences between the randomly assigned treatment and control groups. Results indicated that there were no statistically significant differences in reading achievement between students randomly assigned to the treatment and control groups, as measured by the GMRT , with $F(1, 128) = .12, p >.05$. The extended scale score pretest group means for the treatment and control groups were 306.8 (SD = 22.9, N = 58) and 308.2 (SD = 25.1, N=71), respectively.

Preliminary analyses also indicated that the large group screening assessment used by Reading Rescue® (i.e. Reading Rescue® Classwide Literacy Screening Assessments) was effective at differentiating students' reading achievement levels, and identifying low performing students. Students identified through the Reading Rescue® screening assessment as low performing also scored significantly lower on the GMRT than other

students in their respective classrooms, with statistically significant differences between the treatment/control group students and regular classroom students [$F(1, 753) = 40.49, p < .001$]. Whereas the extended scale score pretest group mean for the combined treatment and control groups was 307.4 ($SD = 23.8, N = 129$), the extended scale score pretest group mean for the regular classroom students was 329.8 ($SD = 38.5, N = 625$).

Comparison of Pretest and Posttest Achievement Data. Analysis of covariance

(ANCOVA), with pretest reading achievement (extended scale scores) as the covariate, was performed on students’ posttest GMRT reading achievement (extended scale scores). As summarized in Table 1, results indicate pretest reading achievement scores were statistically significant as a source of variance in posttest reading achievement, with $F(1, 126) = 63.49, p < .001$. After adjustment by covariate, posttest achievement scores varied significantly with treatment/control group status, as also summarized in Table 1, with $F(1, 126) = 4.75, p < .05$. The standardized effect size, or standardized mean difference³, of the effect of treatment/control group status was .325.

Table 1. Analysis of Covariance of Posttest Reading Achievement

Source of Variance	Adjusted SS	Df	MS	F
Corrected Model	30597.4	2	15298.7	33.62*
Pretest Reading Achievement (Covariate)	28895.99	1	28896.99	63.49*
Treatment/Control Group Status	2163.76	1	2163.76	4.75**
Error	57342.63	126	455.10	

* $p < .001$

** $p < .05$

³ Computed as the difference between the mean post-test score of the treatment group minus the mean posttest score for the control group, divided by the standard deviation of the control group.

Table 2 provides descriptive statistics for the posttest reading achievement measures. In addition to the unadjusted means⁴ and standard deviations for the extended scale scores, Normal Curve Equivalent (NCE) mean scores⁵ are provided to enhance the interpretation of the scores. In addition, unadjusted posttest data are included for regular classroom students in order to provide a context for understanding how the treatment and control groups’ performance compares to other students in their respective classrooms who were not identified as potential candidates for Reading Rescue® .

Table 2. Descriptive Posttest Reading Achievement Data

	Treatment Group (Reading Rescue®) N = 58		Control Group (Non-Reading Rescue®) N =71		Regular Classroom Students N = 625	
	Mean	SD	Mean	SD	Mean	SD
Mean Scaled Score	363.2	28.7	355.9	22.4	378.7	38.3
Normal Curve Equivalent (NCE)	35.8	11.2	32.3	10.4	41.6	14.9

Repeated Measures Analyses Using Follow-Up Data. Repeated measures ANOVA was conducted using GMRT reading achievement extended scale scores obtained for the same students at three separate points in time: prior to any students receiving Reading Rescue® tutoring (Fall 2002, pretest data), after the majority of students had completed Reading

⁴ Unadjusted means are presented for the descriptive statistics to allow comparison to the descriptive statistics for regular classroom students not included in the ANCOVA analysis (i.e. there are no comparable adjusted means and standard errors for these students given that they are not included in the analyses). However, the adjusted means for treatment and control groups are very similar to the unadjusted means: 363.6 adjusted mean (2.5 SE) for the treatment group, as compared to 363.2 for the unadjusted mean; and 355.4 adjusted mean (2.8 SE) for the control group, as compared to 355.9 for the unadjusted mean.

⁵ Normal Curve Equivalent Scores (NCE) are standard scores with a known mean and standard deviation, such that NCE scores of 1, 50 and 99 correspond to percentile ranks of 1, 50 and 99. The score 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 and a standard deviation of 21.06.

Rescue® tutoring (Spring 2002, posttest data), and at the end of the 2002-2003 academic year (follow-up data). As noted in Table 3, the analyses indicate that there were statistically significant differences both between test occasions, with $F(1,76) = 567.4$, $p < .001$, and between treatment and control groups, with $F(1,76) = 7.91$, $p < .05$. The between groups analysis indicates that the difference in improved reading achievement for students receiving Reading Rescue® assistance ($N = 32$) was significantly greater than that of students in the control group ($N=46$).

Table 3. Summary Table for Repeated Measures ANOVA

Source of Variance	SS	Df	MS	F	Partial Eta Squared
Between Tests	212654.0	1	212654.0	567.4*	.892
Between Groups	8659.3	1	8659.3	7.91**	.094

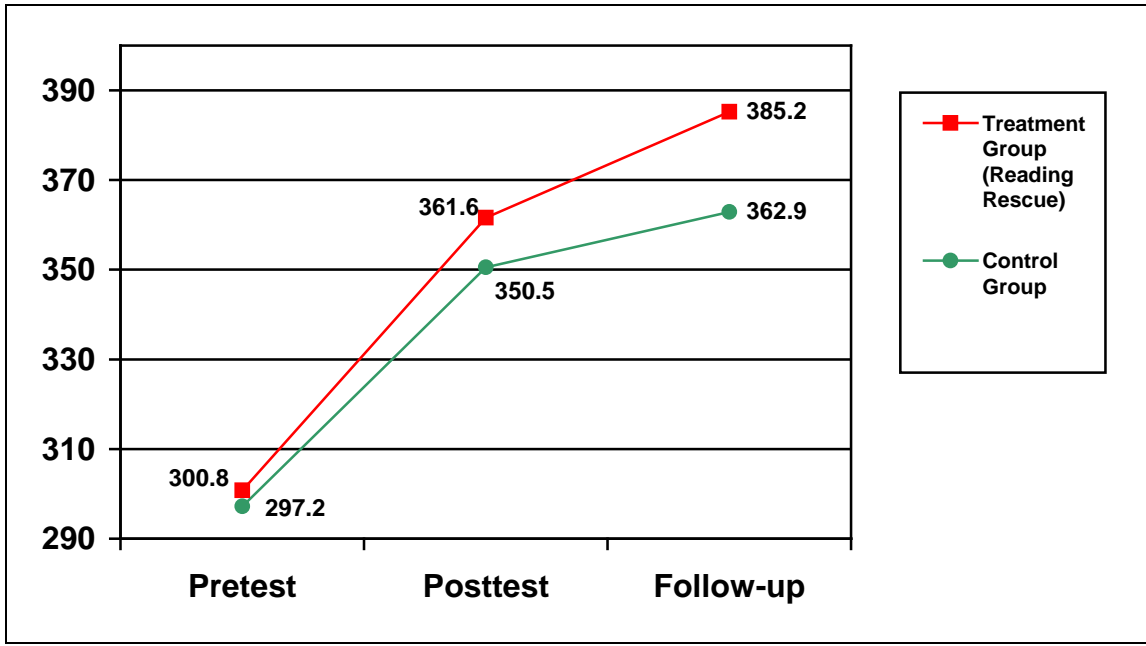
* $p < .001$

** $p < .05$

The reading achievement, over time, for both the treatment and control groups is also graphically displayed in Figure 1 using the GMRT extended scale scores. As visually depicted in this figure, the achievement gap between the treatment and control group students appears to increase over time, with students who receive the one-on-one Reading Rescue® intervention improving in reading achievement at a greater rate than the control group, even after the intervention has concluded. For example, the Reading Rescue® students scored, on average, 11.1 scaled score points higher than the control group students at the time of the posttest. However, at the end of the 2002-2003 academic year, those students who had received Reading Rescue® tutoring scored, on average, 22.3 scaled

score points higher than the control group students who began the year similar in terms of reading achievement.

Figure 1: Longitudinal GMRT Extended Scale Score Comparison



Discussion and Conclusions

As noted previously, the purpose of this study was to examine the effectiveness of the one-on-one tutoring component of Reading Rescue®, specifically addressing the following research question: *Do students who receive one-on-one Reading Rescue® tutoring achieve better reading results compared to similar students who do not receive one-on-one Reading Rescue® tutoring?* The randomized pretest-posttest control group experimental design of this study not only provides the type of evidence encouraged by the *No Child Left Behind Act*, but also allows causal links to be established between the Reading

Rescue® program and its effects on reading achievement. Without such a design, it would be difficult to discern whether any differences found between those students who receive Reading Rescue® tutoring, and similar students who do not receive the tutoring intervention, are directly attributable to the one-on-one tutoring program.

Results of the ANCOVA analyses comparing treatment and control groups on posttest reading achievement, after controlling for pretest differences, provides evidence that Reading Rescue® increases students' reading achievement. Students who received one-on-one Reading Rescue® tutoring scored significantly higher ($p < .05$) on the posttest measure of reading achievement than did similar students who were randomly assigned to the control group. After controlling for pre-program differences, students who received Reading Rescue® tutoring ($N = 71$) scored an average of 363.2 (extended scale score), as compared to 355.9 (extended scale score) for the control group students ($N = 58$). Despite a smaller sample size, the repeated measures ANOVA using three points of data also found statistically significant differences between the treatment and control groups, with students who received Reading Rescue® tutoring having higher levels of reading achievement over time.

The findings from this study are particularly significant given that the majority of research on literacy interventions is based on program implementation in one particular setting with primarily monolingual students (Elbaum, Vaughn, Hughes, & Moody, 2000), whereas this study is based on data drawn from six separate elementary schools with large percentages of English Language Learners. Research based on implementation in one particular

setting can be problematic because the results, and its generalizability, are highly dependent on the actual implementation of the program in one particular context. While it is easier to detect differences in these controlled settings, the practical relevance of the data is more limited. In this particular study, the effects of Reading Rescue® are examined across six distinct settings, each with its own specific contextual factors impacting implementation of the program. Although these variations in implementation and context can make it more difficult to detect differences across settings, findings from such studies are more generalizable, reliable and valid. The results indicate that across the six elementary schools, without any measures of program fidelity included in the study, there was still evidence that Reading Rescue® positively impacts the reading achievement of at-risk students.

These findings are also of particular significance given that English Language Learners (i.e. students who come from homes in which a language other than English is spoken) represent one of the fastest-growing student populations in U.S. schools (Slavin & Cheung, 2004); and the findings from this study suggest that Reading Rescue® is effective in improving the English reading performance of English Language Learners. Both the treatment and control groups across the six schools in this study were predominantly English Language Learners, and the study found that students who received the Reading Rescue® one-on-one tutoring outperformed the control group students who did not receive Reading Rescue® tutoring.

In addition to examining statistical significance, however, it is also important to address the practical significance and meaningfulness of findings. For example, should a difference of 10 points between the means be considered large, medium, or small? For this study, the standardized effect size for the ANCOVA analysis was .325, representing a medium or moderate effect size for this type of intervention. For example, Elbaum, Vaughn, Hughes and Moody's (2000) meta-analysis of 42 samples of students investigated in 29 studies of supplemental, one-to-one reading interventions for elementary students found a mean weighted effect size of .41 when compared to controls. Although the calculated effect size for this study is slightly lower than the mean effect size of literacy interventions included in the meta-analysis, the findings in this study include data across a variety of contexts without checks for treatment fidelity, which likely decreases overall effect size in the aggregated data. In addition, the intensive professional development provided by Reading Rescue® to the classroom teachers and paraprofessionals likely impacts not only tutoring, but also the quality of literacy instruction provided for all students in the classrooms of participating teachers. Therefore, control group students are likely to also be benefiting from the Reading Rescue® program indirectly (as are regular classroom students), mitigating the differences between the treatment and control group students. Also, standardized measures such as the GMRT generally yield smaller effects than the nonstandardized measures used in many of these studies (Elbaum, Vaughn, Hughes, & Moody, 2000); and Reading Rescue® had not yet finished its three year implementation process in these schools (i.e. half the schools were in their first year of implementation, and half the schools were in the second year of implementation).

Although Reading Rescue® students did not reach the same level of reading achievement on the GMRT posttest as did their classroom peers and the national norming sample, the gains made in decreasing the gap might be great enough to allow these students to keep up with classroom instruction and to avoid academic failure. The results of the repeated measures ANOVA regarding residual program effects support this hypothesis. Findings suggest that those students who received Reading Rescue® tutoring continued to increase their reading achievement at a greater rate than did those students who did not receive Reading Rescue® tutoring, even after the tutoring has ended. In other words, even though Reading Rescue® students did not reach the reading achievement level of their classmates by the time of the posttest administration of the GMRT, the increase in reading achievement due to the tutoring program might have provided these students with enough of an academic gain to take better advantage of the regular classroom instruction. Thus, there appear to be residual gains even after the one-on-one tutoring ends, with Reading Rescue® students appearing to have an academic advantage over similar students who do not receive tutoring. These findings are particularly meaningful given that their relevance to impoverished urban school districts, with large percentages of English Language Learners with limited English ability. The positive program effects in this study resulted from a program that was implemented across six low socio-economic urban elementary schools enrolling large percentages (27-41%) of students who receive Bilingual or English as a Second language instruction.

In conclusion, the findings of this study support those from research conducted on other structured one-on-one tutoring programs, showing that the reading achievement of

elementary students at risk of reading failure can be improved through the use of supplemental, adult-instructed one-to-one reading interventions (Elbaum, Vaughn, Hughes & Moody, 2000; Pinnell, Lyons, Deford, Bryk, & Seltzer, 1994). Previously conducted self-evaluations by individual schools and districts (Capella, 2002; Price, 2002; Parman, 2001; Pugh, 2001), and other statistical analyses and research on the impact of Reading Rescue® have consistently noted positive impacts of the program (Britt, 2002; Gibson, 2002; Hoover, 1996; Hoover, 1999; Hoover & Sullivan, 1996). This study confirms the previous findings regarding the impact of Reading Rescue® tutoring, providing evidence based on a randomized experimental design. Although very difficult to implement in real-world educational settings, this design allows the strongest possible causal attributions to be made, and as noted by the *No Child Left Behind* legislation, provides the most scientifically based findings. Further research is needed to replicate these findings, and should include measures of program fidelity at both the school and teacher/tutor levels. Including a fidelity of treatment check will likely result in findings of greater effect sizes for those students receiving the one-on-one tutoring as intended by the Reading Rescue® program. In addition, research and evaluation is needed that examines the impact of other aspects of Reading Rescue® (e.g. effects of professional development on the quality of instruction in the regular classroom), as well as research that accounts for differences in program effect between students.

References

- Allington, R. L. (2004). Setting the record straight. *Educational Leadership*, 61(6), 22-25.
- Ambruster, B.B., Lehr, F., & Osborn, J. (2003). *Put Reading First: The Research Building Blocks for Teaching Children to Read*. Retrieved August 13, 2003
<http://www.ciera.org/library/products/others/putreadingfirst/>
- Britt, L. (2002). *Reading Rescue® in District Six Community Schools of the New York City Public School System: An Evaluation of Student Achievement in Year One of Program Adoption, 2001-2002*. Gainesville, FL: Report submitted to the Literacy Trust, Inc.
- Campbell, D.T. & Stanley, J.C. (1971). *Experimental and quasi-experimental design for research*. Chicago: Rand McNally.
- Capella, S. (2002). Invited testimony before the Council of the City of New York City, Committee on Education Oversight Hearing on Reading Instruction. New York City, New York. September 30, 2002.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Clay, M.M (1985). *The early detection of reading difficulties*. Exeter, NH: Heinemann.
- Elbaum, B., Vaughn, S., Hughes, M.T., & Moody, S.W. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology*, 92 (4), 605-619.
- Gay, L.R. & Airasian, P. (2000). *Educational Research: Competencies for analysis and application*. Upper Saddle River, NJ: Prentice-Hall, Inc.
- Gibson, N. (2002). *An evaluation of Reading Rescue programs in Georgia's Dodge, Fayette, Putnam and McDuffie County Schools, 2000-2001*. The Literacy Trust, Inc. Gainesville, Florida.
- Hiebert, E. H. (1994). Reading Recovery in the United States: What differences does it make to an age cohort? *Educational Researcher*, 23(9), 15-25.
- Hoover, N. L. (1996). *Primary teachers learn to teach at-risk students to read: Empowerment and improvement through a school-based inservice model*. Presented at the annual meeting of the American Evaluation Research Association. New York City, N.Y.
- Hoover, N.L. (1996). *Reading Rescue® at Interlachen Elementary School. Title I Newsletter*. University of Florida, Gainesville, Florida.

- Hoover, N. L. (1997). *Literacy – Realized or Rescued: The Primary School Mission*. *Education Times*. University of Florida, Gainesville, Florida.
- Hoover, N. (1999). *Preventing literacy failure through cost-effective tutoring: A statistical report on high-poverty & rural schools*. Gainesville, Florida: The Literacy Trust, Inc.
- Hoover, N. L., and Sullivan, L. (1996). *Teaching at-risk children to read*. Presented at the annual meeting of the American Evaluation Research Association. New York City, N.Y.
- Institute of Education Sciences. (2003). *Identifying and implementing educational practices supported by rigorous evidence*. Washington, DC: U.S. Department of Education. Available:
www.ed.gov/rschstat/research/pubs/rigorousetid/index.html
- Kennedy, M. M. and Others. (1986). *The effectiveness of Chapter 1 Services. Second interim report from the national assessment of Chapter 1*. Office of Educational Research and Improvement. Washington DC.
- Parman, L. (2001). *An analysis of Gates MacGinitie Reading Test Scores documenting achievement among Strom Lake, Iowa's Reading Rescue® students*. Presented at the Iowa Reading Conference. Des Moines, Iowa.
- Pedhazur, E. J. and Schmelkin, L. P. (1991). *Measurement, Design, and Analysis: An Integrated Approach*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Pinnel, G.S., Lyons, C.A., Deford, D.E., Bryk, A.S., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. *Reading Research Quarterly*, 29 (1), 9-39.
- Price, B. (2002). *Effectiveness of Reading Rescue® Programs*. Paper presented at the 47th Annual Convention of the International Reading Association. San Francisco, CA.
- Pugh, M. (2001). *Reading Rescue® in Walton County's Title I Schools: A report on success in helping the lowest performing students learn to read*. Presented at the annual meeting of the Florida Reading Association. Orlando, FL.
- Shadish, W.R., Cook, T.D., & D.T. Campbell (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin Company.
- Shanahan, T. (1998). On the effectiveness and limitations of tutoring. In P.D. Pearson & A. Iran-Nejad (Eds.), *Review of research in education*, 23 (pp. 217-234). Washington, DC: American Educational Research Association.

- Shanahan, T., & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk learners. *Reading Research Quarterly*, 30 (4), 958-996.
- Slavin, R. E. & Cheung, A. (2004). How do English language learners learn to read? *Educational Leadership*, 61(6), 52-57.
- Tabachnick, B. G., & Fidell, L.S. (1996). *Using Multivariate Statistics*. Third Edition. New York, NY: HarperCollins.
- Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology*, 40(1), 7-26.
- U.S. Department of Education, Office of Elementary and Secondary Education, *No Child Left Behind: A Desktop Reference*, Washington, D.C., 2002.
- U.S. Department of Education, National Center for Education Statistics, *Reading 2003 State snapshot report*, NCES 2004-456, Washington, DC: U.S. Government Printing Office, 2003. Last viewed on the internet 11/15/03
<http://nces.ed.gov/nationsreportcard/reading/results2003/>.
- Wasik, B.A., & Slavin, R.E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly*, 28, 178-200.