Journal of Education and Human Development
June 2016, Vol. 5, No. 2, pp. 62-73
ISSN: 2334-296X (Print), 2334-2978 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/jehd.v5n2a8
URL: https://doi.org/10.15640/jehd.v5n2a8

First Do No Harm: A Title I Natural Quasi-Experiment

Debra G. Porter, Ed. D.[1], Douglas Rush, Ph. D.[2], Jody Wood Ed. D.[3] & R. Larry Bohannon, Ph. D.[4]

Abstract

Supplemental reading services in three elementary schools were examined to determine its effectiveness in improving student reading achievement. Students were identified as in need of supplemental reading services at the end of their kindergarten grade. Due to more students requiring services than could be provided, additional services were only provided to students who were in most need during grade one. Surprisingly, it was discovered through a natural quasi-experiment that students who received additional supplemental reading support had lower reading achievement levels compared to those students who did not receive additional reading services. These findings suggest that the model used in providing supplemental reading services in all three schools, may not have been the best method to provide support to students struggling with learning to read.

Key Words: struggling reader, Title I, supplemental reading

1. Introduction

Title I of the Elementary and Secondary Education Act of 1965 (ESEA), was signed into law by President Lyndon Johnson on April 12, 1965. The purpose of the ESEA, as stated in Section 201 of Title I, was "to provide financial assistance...to local educational agencies serving areas with concentrations of children from low-income families to expand and improve their educational programs...to meet the educational needs of educationally deprived children." (Public Law 89-10-APR. 11, 1965). Title I has since grown to be the largest federal elementary and secondary education program consuming over 25% of the Department of Education budget (Promising Results, Continuing Challenges: Final Report of the National Assessment of Title I). In 2009-2010, Title I Part A funds provided services to over 21 million children in 56,000school districts at a cost exceeding $14,000,000,000 (Improving Basic Programs, 2009).In spite of this major financial investment, Title I funds have not appreciably improved the academic achievement of low income children and have not resulted in the narrowing of the academic achievement gap when comparing the reading and mathematics scores of low-income children to high-income children in elementary and secondary schools(Improving Basic Programs, 2009).

This article reports the findings of a natural quasi-experimental study that examined the reading achievement scores of first grade students who received Title I supplemental reading services and those students that did not receive supplemental services. The setting was at three elementary schools in a large suburban district in a mid-western state. As described more fully below, all students from the three schools were assessed at the end of kindergarten using universal screening to ascertain their reading potential and readiness skills for first grade. Students were then identified as in need of Title I supplemental reading services for first grade based on these results. More students were identified for supplemental reading services than could be placed.

[1] Southeast Missouri State University, USA. dporter@semo.edu, 314-541-2691
[2] Saint Louis University, USA.
[3] Saint Louis University, USA.
[4] Southeast Missouri State University, USA.

Therefore, some eligible students received supplemental reading services throughout first grade while others did not receive these services. These same students were assessed at the end of the first grade school year to determine reading progress. The investigators of this study matched 49 students who received Title I supplemental reading services with 49 students who did not receive the services using their kindergarten reading scores. Analysis of Covariance (ANCOVA) revealed a small, but significant negative effect on the end of the first grade reading assessment for those who received the supplemental reading services compared to the students who did not receive those services when additionally controlling for the initial reading score differences which may have existed after matching. In other words, it appears the students who did receive Title I pullout supplemental reading services did worse on the end of the year, first grade reading assessment than those students who did not receive the Title I services when adjusting for pre-intervention differences in reading ability. These results, while at first surprising, appear to be consistent with other studies which have found little or no positive impact of Title I supplemental reading services on reading achievement scores of first grade students (Improving Basic Programs, 2009)

1.2 A Brief History of Title I

Title I of the ESEA of 1965 provided direct federal funding to State Title I coordinators who then funneled the money to Local Educational Agencies (McClure, 2008).Weak federal guidelines on the use of the Title I funds allowed local school districts to divert the federal money to the entire school district rather than supporting programs for low-income children (McClure, 2008).In 1968, the Office of Education issued its Title I Program Guide Number 44. This guide mandated "Title I funds are not to be used to supplant state and local funds which are already being expended or which would be expended in the project areas or which would be expended in those areas if the services in those areas were comparable to those in non-project areas" (McClure, 2008, p15).In 1970, Congress amended Title I to include a comparability requirement (Section 105 (a) (3). In 1982, Congress enacted the Education Consolidation and Improvement Act, which renamed Title I as Chapter I and was reauthorized again in 1988 (Public Law 89-10-APR. 11, 1965). The Improving America's Schools Act of 1994 (IASA) reauthorized Title I but included three broad themes. First, states were charged with adopting performance base standards in reading and math as well as establish "adequate yearly progress" (AYP) for schools. Second, the 75% low-income threshold to adopt school-wide Title I funded programs was reduced to 50%. The third provided more local control over programs and allowed for a regulatory waiver provision (Promising Results, 2009).

1.3 Title I and Elementary School Student Achievement

Assessing the effectiveness of Title I programs has been challenging. Most Title I funds have been spent on school-wide initiatives. These school-wide funded programs provided benefits to both low-income and higher income students since the funds were not targeted at only low-income students. State assessments of student achievement are aligned to state curriculum and standards. Each state sets its own level of "proficiency" and uses different assessments to measure student achievement outcomes. The results on state assessments cannot be aggregated across all states to examine national trends. Several studies have attempted to determine whether Title I funding improves low-income student achievement at both the national, state and district level (Stullich, S., Eisner, E., & McCrary, J., 2007).

The 2007 National Assessment of Title I found both low-poverty (defined as 0 to 25 percent eligible for free or reduced price lunch) and high-poverty (defined as 70 percent to 100 percent eligible for free or reduced price lunch) school students had slightly improved scores on the fourth grade NAEP reading exam. However, the gap between the NAEP fourth grade reading achievement between low-poverty schools and high-poverty schools had widened between 1992 and 2005. In 1992, fourth grade students in high-poverty schools scored an average of 192 on the NAEP reading test as compared to an average of 225 on the NAEP reading test for fourth grade students in low-poverty schools, a 33 point gap. By 2005 the gap had widened by 36 points as high-poverty students scored 197 on the NAEP grade four reading test while low-poverty students scored 233 on the same exam (Stullich, S., Eisner, E., & McCrary, J., 2007). The 2007 National Assessment Report cautions assumptions about the effect of Title I or other federal programs due to the difficulty in assessing programs for the reason that change may reflect other state and local educational improvement efforts (p. xxii). These finding were consistent with the 1993 Final Report of the National Assessment of the Chapter I Program which found Chapter I (now Title I) participants did not improve their relative standing in fourth grade reading when comparing low-poverty and high-poverty school students (p. 11).

Furthermore, longitudinal data demonstrated Chapter I participants did not improve on standardized or criterion-referenced objectives when compared to non-participants with similar backgrounds and prior achievements (Stullich, S., Eisner, E., & McCrary, J., 2007).

Borman and D'Agostino (1996) conducted a meta-analysis of 17 studies of Title I reading programs and found a small, negative effect of reading achievement for first grade students who received Title I supplemental reading instruction when compared to similar students who did not receive Title I supplemental reading instruction. They also found a small but positive affect for students who received Title I supplemental reading instruction in second through twelfth grades when compared to students who did not receive such instruction. They concluded that Title I reading instruction may be more effective as programs matured and as students were exposed to the instruction for longer periods of time.

Slavin et al., (2009) conducted a systematic review of K-1 reading programs. They noted that children undergo "an extraordinary transition as readers" (p. 3) during kindergarten and the first grade. Children during that period learn the basic skills of turning print into meaning and learn the sound of letters and begin word formation. In later elementary grades, students build fluency, comprehension, and vocabulary. They further noted the difficulty in measuring reading achievement in the K-1 environment. Most K-1 studies use measures of phonemic awareness, which while precursors of reading, are not true measures of reading comprehension. They also noted that measures of reading comprehension and vocabulary tend to have floor effects in Kindergarten and the first grade (p.8).

2. Method

2.1 Setting

This study took place in a large school district located in a suburban county in a mid-western state. The racial makeup of the county, according to the American Fact Finder 2000 census, was 96% Caucasian, 2% African American, 2% Hispanic or Latino and 1% Asian. In 2006, the median family income for the county was $64,415 with 5.2% of the families considered to have an income below the poverty line.

The school district provided educational services to 17,765 students who resided in the largest city in the county. Data was collected from three of the district's elementary schools. School A and School B qualified for and provided supplemental reading services using federal Title I funding. School C did not qualify for federal Title I funding but adhered to the same Title I mandates and provided supplemental reading services using the same procedures as in School A and B. School A had a total enrollment of 494 students in grades K through 2 in 2007. During the period of the study, between 23.3% and 29.4% of School A's children qualified for free or reduced price lunch. School B had a total enrollment of 522 students in grades K through 5 in 2007. During the period of this study, school B had 37.1% students qualify for free or reduced price lunch. In 2007, school C had a total enrollment of 926 students in grades K through 5. School C had 10.9% students qualify for free or reduced price lunch during the period of the study (Missouri Department of Elementary and Secondary Education).

2.2 DIBELS: Identifying Students in Need of Supplemental Reading Services

During the 2007-2008 school year, the school district began a pilot program in Schools A, B and C to identify students in need of supplemental reading services. In January 2008, all kindergarten students in those targeted schools were assessed using the Dynamic Indicator of Basic Early Literacy Skills Universal Screening instrument (DIBELS) developed and managed by the University of Oregon. DIBELS is a criterion-referenced measure designed to benchmark reading skills across grade levels (Good, Kaminski, Simmons, & Kame'enui, 2001).

DIBELS assesses five components of beginning reading skills.

(1) Phonemic Awareness is measured by Initial Sounds Fluency (ISF) and Phoneme Segmentation Fluency (PSF);
(2) Alphabetic Principle is measured by Nonsense Word Fluency (NWF);
(3) Accuracy and Fluency with connected Text is measured by Oral Reading Fluency (ORF);
(4) Vocabulary is measured by Word Use Fluency (WUF); and
(5) Comprehension is measured through Retelling Fluency (RTF); (Good, Kaminski, Simmons, & Kame'enui 2001).

Protocol for administering DIBELS to kindergarten students includes assessing students' phonemic awareness and the alphabetic principle once all kindergarten students were assessed and scores were determined, students were placed into three categories: students at benchmark (benchmark), students in need of strategic supplemental reading intervention (strategic), and students in need of intensive supplemental reading intervention (intensive). A total of 167 kindergarten students in Schools A, B and C were identified as in either the strategic or the intensive categories and in need of Title I supplemental reading services.

Title I supplemental reading services were provided in Schools A, B and C by reading specialists who had received their state required reading specialist certification. All certified reading specialists were certified as an elementary education educator from the state in which the district is located as well as holding a Master's of Arts degree in Education with an emphasis in reading. School A had one reading specialist to provide services to students who qualified for services while Schools B and C had two reading specialists to provide services. Title I mandates the number of students to be serviced per reading specialists. Therefore, only 50 of the 167 students identified as being in need of supplemental reading services actually received those services during their first grade school year.

In order to service students who were most at risk, first grade classroom teachers in all three schools met with the reading specialists to determine which 167 eligible students would receive Title I supplemental reading services during the 2008 - 2009 school year. Based on additional criterion established by the district, 50 students were identified to receive Title I supplemental reading services the following school year. The remaining 117 students did not receive Title I supplemental reading services.

Race, sex, and socioeconomic factors were not used in the inclusion or exclusion selection process and were not considered as variables in this study. Attrition of subjects occurred during the course of the school year. Only 85 of the 117 students who were eligible, but did not receive supplemental reading services in the first grade, completed the school year and took the *Terra-Nova* assessment in April 2009. Similarly, 49 of the 50 students who were provided supplemental reading services completed the first grade school year and took the April 2009 *Terra-Nova* assessment. Attrition was primarily due to student mobility to other schools or school districts.

2.3 The Core Reading Curriculum in Schools A, B, and C

Prior to this study, the school district adopted a balanced literacy model of instruction with the intention to teach reading through small, differentiated, guided reading groups. A basal grogram was purchased that included a basal reader and a few trade books that were leveled as below grade level, on grade level, or above grade level. During the study, teachers in Schools A and B received professional development and coaching support in developing and implementing lessons through small, guided reading groups. School C began professional development and coaching the following school year. During the initial year of implementation, Schools A, B, and C had limited resources to support small group, guided reading instruction. Whole group instruction was utilized until the following school year when resources became available to support small, guided reading groups. Therefore, during this study, reading instruction in a small group was done to reteach skills not mastered in whole group setting. The sequence of word study skills coordinated with stories contained in the basal reading text. Because the majority of the stories in the basal were above grade level, these stories were read aloud either by the teacher or with the students as shared reading. School adopted a spelling curriculum that was in conjunction with the basal reader. The spelling assignments paralleled with the word study and vocabulary skills introduced in the basal.

2.4 Supplemental Reading Instruction

The supplemental reading instruction was accomplished using the "pull-out" model. Each of the 50 identified students was placed in a small group containing three to eight students. Groups were formed based on students' homeroom schedules by avoiding times when students were receiving direct instruction, recess, or any special class such as music, physical education, or art. Students in these groups were provided 30 minutes of supplemental reading instruction, five days a week during the entire first grade school year. The small groups remained constant during the school year and few students moved between groups. A common curriculum was not established for supplemental reading instruction in Schools A, B, and C. The reading specialists based their instruction on the reading process. The lessons incorporated semantic, orthographic, and syntactic cuing. The certified reading specialists used a variety of children's literature of various reading levels to develop comprehension and vocabulary skills.

When addressing phonics and word study, students used magnetic letters to build words, incorporated writing journals and practiced writing letters, words, sentences, and stories. In order to develop sight word knowledge, students read and reread books at their independent and instructional level. Formative assessment, to monitor student progress, was done weekly in order to adjust instruction and ensure student achievement.

2.5 Terra-Nova, the Second Edition Post-Test: Measuring Reading Ability

The School District had adopted *Terra-Nova, the Second Edition*, as a standardized achievement test to measure reading, language arts, and math skills of its first and second grade students. *Terra-Nova* is published by CTB/McGraw-Hill and the Reading/Language Arts component of the test integrates two communication skills areas, comprehension and phonics. It also is designed to test the content objectives of states and school districts as well as the conceptual framework of the National Assessment of Educational Progress [NAEP] (CTB/McGraw-Hill).

The *Terra-Nova* is scored using Item Response Theory (IRT) item-pattern scoring. The patterned IRT scoring method calculates scaled sores based on an overall pattern of correct responses rather than a simple calculation of correct responses. In theory, pattern IRT scoring provides a more accurate estimate of a student's performance level than a simple total correct response measure. The *Terra-Nova* Scale Score is the basic measure that reflects all information contained in each student's pattern IRT responses (CTB/McGraw-Hill).

2.6 Variables

The dependent variable in this study was the *Terra-Nova* test scores, which were obtained from the students at the end of their first grade school year. The independent variable was the grouping variable of whether the students were or were not receiving supplemental reading services during the first grade school year. The covariate in this study was the students' scores on the DIBELS pre-test.

2.7 Data Analysis

All data analysis was conducted using SPSS 19.The primary method of data analysis was Analysis of Covariance. ANCOVA is commonly used in education research where subjects cannot be randomly assigned to control and treatment groups necessary for true experimental research. ANCOVA allows a researcher to explore whether differences between groups on mean dependent variable scores are larger than to be expected by chance after removing the effect of one or more covariates. The use of ANCOVA is fraught with potential for error. Covariates must be correlated with the dependent variable, but uncorrelated with each other (Tabachnick and Fidell, p. 302).The purpose of the covariate in ANCOVA is to reduce the within-group error variation.

Successful use of ANCOVA requires that certain assumptions be met. First, the covariate must be measured before the treatment is administered and the dependent variables measured. This assumption is met in the study as the DIBELS pre-test covariate was administered during the students' kindergarten school year. The *Terra-Nova* dependent variable post-test was administered at the end of the students' first grade school year. The second assumption is there are enough subjects to provide sufficient statistical power to detect desired effects and the groups have equal sample sizes or that adjustments are made for unequal group sizes. A pre-study power analysis was not possible because the archived nature of the data limited the study to the number of students who were identified as in need of supplemental reading services and further limited to the number of students who were actually provided those services. According to Cohen (1988), a total of 49 students in each of the groups for an even balanced design should be sufficient to detect a medium to large effect.

The unequal number of students in the two groups presented two problems. First, the unequal sizes of the two groups create a non-orthogonal analysis. In addition, there was ambiguity regarding marginal means and possible overlapping of sums of squares to sources (Tabachnick and Fidell, p. 296). The second problem arose in that one of the factors considered in assigning students to the two groups was the DIBELS pre-test covariate. While the DIBELS score was not the sole determinate of assignment to the supplemental reading group, it did play a major role in the assignment. Therefore, the students who were not assigned to supplemental reading services generally had higher DIBELS scores than the students who were assigned to supplemental reading services. This potentially affects the independence of the DIBELS covariate from the supplemental reading services group independent variable.

Cook, Shadish, & Wong (2008) noted that using proxy pre-tests and using time and geographic matching might minimize bias caused by initial differences between the control and comparison groups. To resolve these issues, the study utilized a modified group matching technique.

The 49 students who were assigned to the supplemental reading group were matched on DIBELS scores with 49 students with similar DIBELS scores who were not assigned to supplemental reading services. This technique eliminated the unequal group design. Fields (2009) recommends matching the groups on covariate to reduce the potential for lack of independence between the covariate and the independent variable (p. 398).

The third assumption was there is an absence of multi-co linearity or singularity between covariates. In this study, there was only one covariate that eliminates this concern. The fourth assumption was there is normality of sampling distributions of means across the groups. Tabachnick and Fidell (2012) noted it is not possible to check this assumption without knowledge of population values. However, the central limit theorem provides with relatively equal group sizes robustness is to be expected within 20 degrees of freedom for error (p. 281).The final assumptions were homogeneity of variance which was tested with the Levene's test; linearity which was tested graphically using scatter plots; and homogeneity of regression slopes by using the ANCOVA custom model program in SPSS 19 (Pallant, 2005).

3. Results

3.1 Analysis of Assumptions for ANCOVA

Initially, an independent t-Test was conducted to test whether the covariate DIBELS score was independent of the matched grouping independent variable; supplemental reading services. The results indicated there was not a statistically significant difference in the mean DIBELS score when comparing the mean score of the group receiving supplemental reading with the group that did not receive supplemental reading services [M supp. service = 54.41, M no supp. service = 61.94, t (96) = 1.854, p (two tailed) = 0.067]. The non-significance finding indicated independence of the covariate DIBELS from the supplemental reading services independent variable. Even though the two matched groups did not statistically differ on the DIBELS proxy pretest, caution still needs to be exercised in comparing the post-test results due to a possible regression to the mean phenomenon (Linn, 1980).

A Hierarchical Multiple Regression (HRM) was then conducted to determine whether the DIBELS covariate was correlated with the *Terra-Nova* dependent variable. The covariate DIBELS score was entered as the first step in the model and the independent grouping variable was entered into the model in the second step. Entering the covariate as the first step in the regression presented the unique effect of the covariate DIBELS score on the *Terra-Nova* dependent variable to test for correlation between the covariate and the dependent variable. Entering the grouping independent variable into the model in the second step allowed for analysis of the unique effect of the grouping variable on the dependent variable after separating out the effect of the covariate. It allowed for an analysis of the entire regression model to determine the combined effect of the covariate and the independent variable of the dependent variable.

The HRM model revealed the covariate DIBELS was positively correlated with the *Terra-Nova* dependent variable [r = 0.370, n = 98, p < .001]. Tolerance (0.965) and VIF (1.036) scores indicated multi-collinearity was not present. Scatter plots and the Probability Plot (P-P of the Regression Standardized Residuals) indicated no deviation from normality. Analysis of the scatter plots and the Mahalanobis (Max. 9.378) distances did not reveal any outliers as being present in the data. The homogeneity of regression slopes assumption was tested using univariate analysis of variance to test the subject interaction effect between the covariate DIBELS and the grouping independent variable. The findings were not significant [F (1, 94) = 0.049, p = .826] indicating the homogeneity of the regression slopes and absence of interaction between the covariate and the independent grouping variable. This supported the earlier review of the scatter plots and the independence of the covariate from the independent variable as shown in the previous t-Test. A review of the HRM presented an interesting finding regarding the supplemental reading services independent variable. There was a negative correlation between receiving Title I supplemental reading services and the *Terra-Nova* dependent variable after removing or controlling for the effect of the DIBELS pre-test covariate [Part r = - 0.206, B = -13.388, SE = 6.233, t (97) = -2.219, p = 0.029]. Simply put, those students who received Title I reading services during their first grade school year scored lower on the *Terra-Nova* post-test than did similar students who were in need of but did not receive supplemental reading services after controlling for the effect of their pre-test DIEBLS scores. This finding was unexpected given the individual time and effort expended on providing supplemental reading services to those students.

This finding was further examined by running the ANCOVA using *Terra-Nova* as the dependent variable, DIBELS as the covariate, and using the grouping supplemental reading services as the independent variable. The Levene's test revealed no violation of the assumption for equality of error variance [F (1, 96) = 1.996, p. = .161]. The test of between subject effects demonstrated there was a statistically significant difference in the group means on the dependent variable *Terra-Nova* scores when comparing the scores for the group which received supplemental reading services with the group which did not receive supplemental reading services adjusting for or equalizing the effect of the pre-test DIBELS covariate [F (1, 95) = 4.924, p = .029, partial η^2 = .049. There was a statistically significant relationship between the DIBELS pre-test covariate and the *Terra-Nova* dependent variable [F (1, 95) = 12.209, p = < .001, partial η^2 = .114] when controlling for the effect of the grouping variable.

The adjusted group means for the study demonstrated the group which received supplemental reading services had lower adjusted group mean scores on the *Terra-Nova* post-test than did similar students who did not receive supplemental reading service [Adj M supp. services = 533.388, SE = 4.229, Adj M no supp. services = 566.775, SE = 4.229]. This again indicated the students who received supplemental reading services scored lower on the *Terra-Nova* post-test than did similar students who did not receive supplemental reading services after adjusting for the pre-test DIBELS score.

Tabbachnick and Fidell (2012) cautioned that adjusted group means must be interpreted with "great caution" (p. 280). In non-experimental work an assumption that all of the subjects start with equal pre-test scores on the covariate may be unrealistic. Other complicating issues arose because the variables in the model did not include all sources of possible variation in the dependent variable. Nevertheless, the findings of the ANCOVA do confirm participating in Title I supplemental reading services produced a small, but statistically significant, negative effect on students' post-test *Terra-Nova* scores.

4. Discussion

A speculation of four plausible explanations to elucidate these variations is proposed. These four themes reemerged throughout the literature and have been found to contribute to test score variance in other quantitative and qualitative studies. The themes are generalized and include a.) Teacher Quality; b.) Instructional conflict between classroom teachers and reading specialists; c.) Push-in vs. pull- out models, and e.) High stakes testing.

4.1 Teacher Quality

Although a plethora of various factors contribute to the success of children's learning, it is agreed a teacher's "...capacity to teach effectively is among the most important" (Darling-Hammond, 1999; Darling-Hammond and Youngs, 2002; Nye, Konstantopoulos, and Hedges, 2004; O'Connor et al., 2005). Furthermore, teacher education level was reported as an important factor in several studies (Miles, et. al., 2004; O'Connor et al., 2005; Hong and Radenbush, 2005; Kober, McMurrer and Silva, 2011) and was used as an indicator of student success (Darling-Hammond, 1999, Nye, Konstantopoulos, and Hedges, 2004) in several quantitative studies. Robust support for continued professional development (Adler and Fisher, 2001; Rea, Mclaughlin, and Walther-Thomas, 2002; O'Connor et al., 2005; O'Connor, Fulmer, Harty, and Bell, 2005; Garet et al., 2008) has been tied to student success in mathematics and reading programs in education with attention given to development of "knowledgeable teachers" (Adler and Fisher, 2001, p. 616). Garet (2008) and his colleagues affirm, "80 percent of elementary teachers reported participating in 24 hours of professional development or less on reading instruction during the 2003–2004 school year and experts have raised a concern that this level of professional development is not intensive enough to be effective" (2008, p. 1). Two Title I reading programs reported noteworthy gains in reading outcomes and contributed success in the programs to continuous professional development (Miles et al., 2004, O'Connor et al., 2005), Miles indicated, "the approach started with the professional development of our staff" (Miles et. al., 2004, pg. 321).

During this study, teachers from Schools A and B were provided professional development that focused on a balanced literacy model. The professional development provided was endorsed by the district's state education department and utilized a coaching model. Classroom teachers received 40 hours of professional development and observed model lessons presented by the professional trainer. As they became more knowledgeable on the structure of a balanced literacy model, classroom teachers taught lessons that were observed by the official trainer. With the trainer, the classroom teachers would reflect on the effectiveness of the lessons they taught and ways to improve their instruction.

Classroom teachers were shown how to assess students in order to guide their instruction. They were taught how to administer running records and then conduct a miscue analysis to determine each student's strengths and weaknesses. Using the assessments they administered, they learned how to form small, guided reading groups so they could scaffold student learning. While working with students in these small guided reading groups, classroom teachers were shown how to prompt and guide students through the reading process. They learned how to structure the school day and create a literate environment that promoted reading, writing, speaking and listening skills. School C did not receive any professional development during the year of the study but did the following year using the same balanced literacy initiative program endorsed by the state's education department.

4.2 Instructional Conflict between Classroom Teachers and Reading Specialists

A second prominent theme explaining test score variance among Title I reading programs that emerged in the literature was the conflict between the instructional styles of teachers and reading specialists. The theme generally consisted of a lack of coordination (Pugach and Wesson, 1995) between teachers and supplementary providers. Likewise, Shoho et al. (1997) note a lack of cooperation between special education and general education teachers. The general consensus was differing teaching techniques among various teachers, special education teachers, and supplementary reading specialists existed (Bean et al., 1991; Allington, 1983; Miles et al., 2004) and these differences result in "inconsistent instruction" (Miles et al. 2004, p. 321) and "poor reader call-outs" (Allington, 1983, p. 553) which could "...inhibit the development of student self-monitoring capabilities" (Allington, 1983, p. 553). To overcome this barrier, Winer and Ray (1994) suggest using a "collaborative consultation" approach first made popular by Bauwens, Hourcade and Friend (1989) that coordinates special educators and general educators and encourages them to teach in an inclusive setting. Miles et al. (2004) suggest being aware of the barrier and using a whole team approach which would include "...Title I reading specialists, instructional aides, and parents..." (p. 321) as partners in the student's reading program. They go as far as to suggest parental training initiatives should be implemented to ensure the success of the students' increased reading performances within the program while working at home (p. 321).

At the time of this study, the district did not incorporate professional learning communities and teachers were not provided time during the school year to collaborate and evaluate their instruction or student progress. Classroom teachers met monthly to discuss things such as field trips, parties, reporting of student progress, and other clerical items. Prior to the study when a classroom teacher had a student who needed to be considered for additional reading services, they would meet with the reading specialist and complete a reading criterion form. The reading specialist would then conduct an Individual Reading Assessment to determine whether the student qualified for additional reading services. If the student's score was within the range for qualification and there was a space available for an additional student, the student was provided services. Classroom teacher and reading specialist would briefly meet to discuss results but no collaboration was done to align instruction between the homeroom and supplemental reading instruction.

4.4 Push-in and Pull-out Models of Instruction

Supplemental academic support for students can be employed in two ways: A pull-out method in which students receive instruction in a location separate from their general education peers; and a push-in method in which reading specialists or supplementary providers work alongside the general education teacher to support the education of students in an inclusive setting. Researchers have argued the manner in which supplemental academic support is implemented can have adverse effects on student achievement.

The popularity of pull-out programs has waned for two considerable reasons despite the findings of Jenkins and Heinen (1989) claiming some students with learning disabilities preferred pull-out programs. First, it is argued pull-out programs cause fragmentation of the school day (Jenkins and Heinen, 1989; Bean et al., 1991; Allington, 1983; Rea et al., 2002) resulting in time off-task and arguably having a negative effect on test scores. Second, pull-out programs isolate students from peers, primary curriculum, and other classroom activities (Rea et al., 2002; Miles et al., 2004). In this case, it is argued the students experience segregation and continues to fall behind their peers as a result of the pull-out process.

The schools in this study used a pull-out method to provide supplemental reading services to students. Students left their homerooms for 30 minutes per day, five days a week. Students were placed with other students from their grade level and/or homeroom. Groups were limited to 10 students per group with most groups consisting of three to eight students. Classroom teachers were directed to avoid direct and explicit literacy instruction while students were provided supplemental reading services. The principals or other supervisors did not conduct a fidelity check to confirm this requirement. There was no district-wide approved curriculum and therefore, reading specialists relied on a variety of children's literature to teach comprehension, vocabulary, and fluency skills. Explicit instruction of phonics was taught using a variety of manipulative materials such as magnetic letters, dry erase boards and markers, big books, and poetry. Students sorted and classified letters and words, recognized and produced words with various onsets and rimes, identified and produced rhyming words, and practiced sight words by using leveled readers on their instructional or independent reading level. Students were instructed on using the cueing system to read. Students were guided to use syntax, semantics, and orthographic cues with the reading specialists scaffolding their instruction.

4.5 High Stakes Testing

Poor reading performance in primary grades has become a predictor of poor reading performance in later grades. As such, administrators have begun to identify and assess reading ability as a proscriptive approach to increasing reading ability. The argument has been proposed that Title I is problematic because it raises testing standards and consequently lowers children's test scores which results in an increased number of schools failing to meet NCLB standards (Wong et. al, 2009). Tobin and Calhoon (2009) argued there is a lack of empirical credibility in many of the reading programs put into use by school districts, further purporting heavy marketing and philosophical approaches often persuade school district administrators to employ such ineffective programs.

Interestingly, Borman and D'Agostino (1996) claimed "...in some cases teachers and administrators may vary pre- and post-test conditions in subtle ways in the hope of inflating the gains posted by the students in their programs" (p. 311) noting specifically the "summer effect" (p. 311) where students experience a loss in test score achievement gains over the summer; and an "unknown artifact" (p. 311) where administrators use the post-test from the previous year as the pre-test for the next year rather than practicing annual testing cycles. Overall, they concluded "...some testing and reporting effects may have influenced the norm-referenced outcomes..." (Borman and D'Agostino, 1996) and that "...Title I alone cannot be expected to serve as the great equalizer" (p. 324).

The school district, in which this study was conducted, formally assessed students during the fourth quarter of the academic school year. Grades three and above were assessed using the state assessment but the Terra-Nova standardized test was administered to first and second grade students. Several reading subtests are included in the Terra-Nova and results were provided the classroom teachers and parents at the end of the academic school year in which the test was taken. The district's purpose of administering the Terra-Nova standardized test was to determine effectiveness of instruction of reading and student knowledge of reading skills when compared to other students of the same grade level across the nation. Zemelman, Daniels, and Hyde (2005) warn that more standardized tests for students, teachers, and schools will lead to more systematic labeling and punishing of those that fail. They also are concerned that politicians, state and federal legislators, state education departments, and testing companies believe that through tighter control, more regulation, and the use of high-stakes standardized tests will improve schools (2005).

A major problem arises when a school district depends solely on standardized testing to assess curriculum, student performance, and teacher effectiveness. When a district incorporates ongoing formative assessment to guide instruction and monitor student progress, teachers can provide students with instruction that is meaningful and appropriate to their students' individual needs. In other words, teachers are not waiting until students fail before they provide the appropriate instruction students need to be successful. Teachers in the district for this study were required to conduct assessments, but these assessments were not ongoing to monitor student progress. These assessments were aligned with the basal series and used to determine students' knowledge of material after instruction. Assessments were completed in a whole group setting with the format of the tests being a combination of multiple-choice, fill-in-the- blank, matching, or constructive response. Students were also assessed on their oral reading fluency by reading aloud a specific grade level reading material. A combination of words per minute and number of errors determined if a student read at his or her given grade level.

This was conducted each quarter and results were reported on the quarterly report card as reading on grade level, above grade level, or below grade level. A running record and miscue analysis was not used by the classroom teacher nor was it used to determine students' reading levels for guided reading groups.

Reading assessment that occurs when teachers observe and interact with students as they read is much more affective in ensuring students' developing the skills needed to be proficient readers. By conducting an oral reading miscue analysis, teachers can understand what skills a student uses while reading and which skills need to be improved. Routman (2002) explains that one-to-one assessment conferences where the teacher prompts and scaffolds the student to problem solve as he/she reads aloud to the teacher is more effective than using a standardized test or basal series tests which does not emphasize the heart of reading. Since this district instructed through whole group, little or no one-to-one conferencing was done. Assessments consisted of tests provided by the basal series and was conducted at the end of each basal unit of instruction. Oral reading assessments that were conducted achieved one goal, to determine if the child was reading on grade level. A miscue analysis on the errors the student made was not required by the district and therefore, teachers had limited knowledge of their students' reading skills.

5. Conclusion

In conclusion, the researchers found while the influences may have been a result of either teacher quality, classroom teacher and Title I teacher collaboration, the pull-out model used, or testing changes, the students who received Title I services in this suburban school in a mid-west state did not score as high as students remaining in the regular classroom without the Title I intervention services. Unfortunately, it is unknown the exact cause which prevents further correction; however, it is incumbent upon the school district to investigate what changes might be in order to ensure this does not continue. The findings from this study, although disheartening, can be helpful to the district if they want to improve students' achievement, teacher effectiveness, and more specifically meet the needs of the struggling reader. One critical theme that surfaced again and again was the need for the district in this study to use research best practice for teaching all students. Although Schools A and B classroom teachers were being provided significant professional development that supported best practice, School C did not participate until the following school year. Research has proven that students who are provided reading instruction through small, guided reading groups are more likely to acquire the skills needed to be proficient readers. It is likely that once classroom teachers have put into practice a balanced literacy model, there will be more consistency between supplemental reading instruction and classroom instruction.

It is also clear that teachers who collaborate and meet regularly with each other to discuss student progress and seek professional development on best practice in teaching reading are more likely to provide instruction that is appropriate and effective for all students. Teachers who are assessing their students on an ongoing basis and monitoring progress can provide instruction that meets the needs of each individual student. Assessment that incorporates the opportunity for students to read aloud while the instructor is analyzing their errors and self-corrections to determine their guided reading level will be more successful in becoming a fluent reader. Finally, classroom teachers need to be proactive in differentiating their instruction to meet the needs of all their students. They should work with the reading specialists to provide the most effective instruction for students who are struggling the most with learning to read. If a pullout model is to be utilized, then only students who are most at risk should receive these services. These students should be placed in small groups with similar needs so the reading specialists can be more effective in providing instruction. It is important to remember that these services are in addition to the classroom instruction and never to be in place of the classroom instruction.

It is imperative to conduct further research on the progress of the participants in this study to determine if in fact as Slavin, et. al. (2009) suggested that children undergo "an extraordinary transition as readers" (p. 3) during kindergarten and the first grade. They learn the basic skills of turning print into meaning and learn the sound of letters and begin word formation. Not until later in elementary school do students build fluency, comprehension, and vocabulary (p.8). The students in this study will enter grade eight in the fall of 2016. At this point, the district will have fully implemented a balanced literacy model in all elementary schools. All classroom teachers will have been provided the required professional development and students will have been provided literacy instruction in their small, guided reading groups.

In conclusion, this study revealed the importance of educators using best practice. Education is a science and acknowledging the need for research that critically examines what works best to ensure student learning is essential. School districts that receive Federal funding have been entrusted to allocate these funds efficiently and effectively. Therefore, school districts are responsible in providing all students the opportunity to excel in school.

Works Cited

Adler, M. A., &Fisher, C. W. (2001). Early reading programs in high poverty schools: A case of beating the odds. *Reading Teacher,* 54 (6): 616-619

Allience for Excellent Education. (2003). *Reading for the 21st Century: Improving the Literacy of Children in Grades Four through 12.*

Allington, R. L. (1983). The reading instruction provided readers of differeing readin abilities. *The Elementary School Journal* 83(5), 548-559.

American FactFinder. (2007). Retrieved 2009 from U. S. Census Bureasu: http://factfinder.census.gov/servlet/SAFFPolulation

Bean, R. M., Cooley, W. W., Eichelberger, R. T., Lazar, M. K., & Zigmond, N. (1991). Inclass or Pullout: Effects of Setting on the Remedial Reading Program. *Journal of Literacy Research* 21, 445-464. doi: 10.1080/10862969109547753

Borman, G. D., & D'Agostino, J. V. (1996). Title I and Student Achievement: A Meta-Analysis of Federal Evaluation Results. *E-Journal of Educational Evaluation and Policy Analysis, 18(4), 309-326. Retrieved from http://eric.ed.gov/?id=EJ538585*

Brace, K. S. (2009). *SPSS for Psychologists, Fourth Edition.* London: Routledge Academics.

Bauwens, J., Hourcade, J. J., & Friend, M. (1989). Cooperative teaching: A model for general and special education integration. Remedial and Special Education, 10(2), 17-22. EJ 390 640

Cohen, J. (1988). *Statistical power analysis for the behavior sciences.* Hillsdale: Erlbaum.

Cook, T., Shadish, W. R., & Wong, V.C. (2008). Three conditions under which experiments and observational studies produce comparable casual estimates: New finding from within-study comparisons. Journal of Policy Analysis and Management, 27, 724-750.

CTB McGraw Hill. (n.d.). Retrieved 2009 from McGraw Hill Kids: http//www.mhkids,com/products/products

Darling-Hammond, L. (1999). *Teacher quality and student achievement: A review of state policy evidence.* Seattle: Center for the Study of Teaching and Policy.

Darling-Hammond, L. & Youngs, P. (2002). *Defining* "Highly Qualified Teachers": What Does "Scientifically-Based Research" Actually Tell Us*?. E-Journal of Educational Researcher,* 31(9), 13-25.

Fields, A. (2009). *Discovering Statistics Using SPSS*. Thousand Oaks, California: Sage Publishing Inc.

Garet, M.S., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., Uekawa, K., Falk, A., Bloom, H. s., Doolittle, F., Zhu, P., & Sztejnberg, L. (2008). The Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement. (NCEE 2008-4030). Retrieved from http://www.eric.ed.gov/?id=ED502700

Good, Kaminski, Simmons, & Kame'enui, (2002)*Center on Teaching and Learning, DIBELS Data System*. Retrieved from Dynamic Indicators of Basic Early Literacy Skilss: https://dibels.ioregon.edu/

Hong, G., & Raudenbush, S. W., (2005) Effects of Kindergarten Retention Policy on Children's Cognitive Growth in Reading and Mathematics. *Educational Evaluation and Policy Analysis, 27,* 205-224. Doi: 10.3102/01623737027003205

Improving Basic Programs Operated by Local Educational Agencies (Title I, Part A), Retrieved from http:// www2.ed.gov.

Jenkins, J. R. & Heinen, A. (1989). Students' preferences for service delivery: pull-out, in-class, or integrated models. *Exceptional Child 55(6*), 516-523.

Kaminski, R. A., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review* , 215-227.

Kober, N., McMurrer, J., Silva, M. R. (2011) *State Test Score Trends through 2008-09, Part 4: Is Achieving Improving and Are Gaps Narrowing for Title I Students?* (Center on Education Policy). Retrieved from Institute of Education Sciences website: http://www.ies.ed.gov

Linn, R.L. (1980). Regression toward the mean and the interval between test administrations. *New Directions for Testing and Measurement, 8,* 59-82.

McClure, P. (2008)*The History of Educational Comparability in Title I of the Elementary and Secondary Education Act of 1965* (Ensuring Equal Opportunity in Public Education, How Local School District Funding Practices Hurt Disadvantaged Students and What Federal Policy Can Do About It). Retrieved from Center of American Progress website: https://www.americanprogress.org/issues/education/report/2008/06/10/4567/ensuring-equal-opportunity-in-public-education/

Miles, P. A., Stegle, K. W., Hubbs, K. G., Henk, W. A., & Mallette, M. H. (2004). A Whole-Class Support Model for Early Literacy: The Anna Plan, *The Reading Teacher, 58(4), 318-327. doi: 10.1598/RT.58.4.1*

Missouri Department of Elementary and Secondary Education, Missouri Comprehensive Data System, Retrieved fromhttp://mcds.dese.mo.gov/quickfacts/SitePages/DistrictInfo.aspx.

Nye, B., Konstantopoulos, S., & Hedges, L.V. (2004). How Large Are Teacher Effects?. Educational Evaluation and Policy Analysis. *Educational Evaluation and Policy Analysis, 26,* 237-257. Doi:10.3102/01623737026003237

O'Connor, R. E. (2000). Increasing the intensity of intervention in kindergarten and first grade. *Learning Disabilties Research and Practice* , 43-54.

O'Connor, R. E., Fulmer, D., Harty, K. R., & Bell, K. M. (2005). Layers of reading intervention in kindergarten through third grade: changes in teaching and students outcomes. *Journal of Learning Disabilities, 38,* 532-538.

Pallant, J. (2005). *SPSS Surivial Manual.* London: Open Iniversity Press.

Promising Results, Continuing Challenges: Final Report of the National Assessment of Title I. (2009) National Center for Educational Evaluation and Regional Assitance (2009) Retrieved from www.IES.ed.gov.

Public Law 89-10-APR. 11 1965, (2009) Retrieved from U. S. Government Publishing Office.

Pugach, M. C., Wesson, C. L. (1995) Teachers' and Students' Views of Team Teaching of General Education and Learning-Disabled Students in Two Fifth-Grade Classes. *The Elementary School Journal 95*(3), 279-295.

Rea, P. J., McLaughlin, V. L., & Walther-Thomas, C. (2002) Outcomes for Students With Learning Disabilities in Inclusive and Pullout Programs. *Council of Exceptional Children. 68 (2) , 203-222.*

Routman, R. (2002) *Reading Essentials: The Specifics You Need to Teach Reading Well.* Heinemann Publishing, Portsmouth, NH.

Shoho, . R., KATIMS, D., & WILKS, D. (1997) Perceptions of alienation among students with learning disabilities in inclusive and resource settings. *The High School Journal, 81*(1). 28-36.

Slavin R., W. B. (n.d.). *Preventing early reading faliure with one-to-one tutoring: A review of five programs.* Retrieved 2009 йил 5-June from Success for All: http://www.successforall.net

Slavin, R. E., Lake, C., Davis, S., Madden, N. A. (2009) *Effective Programs for Struggling Readers: Best Evidence Synthesis*, Institute of Education Sciences, U. S. Department of Education

Stullich, S., Eisner, E., & McCrary, J. (2007) National Assessment of Title I Final Report. Retrieved from National Center for Educational Evaluation and Regional Assistance.

Tabachnick, B. G. (2001). *Using Multivariate Statistics.* Needham Heights: Pearson Education.

Tabachnick, B. G. & Fidell, L. S., (2012). *Using Multivariate Statistic, 6th Edition.* Needham Heights: Pearson Education.

Title I - Improving The Academic Achievement of the Disadvantaged (2002). U. S. Department of Education Retrieved from http://www.ed.gov

Tobin, K.G. & Calhoon, M. B. (2009) A comparison of two reading programs on the reading outcomes of first-grade students. *Journal of Direct Instruction,* 9, 35-46.

Winer, M. B., Rau, K. L. (1994) *Collaboration Handbook: Creating, Sustaining, and Enjoying the Journey.* St. Paul, MN: Amherst H. Wilder Foundation.

Wong, M., Cook, T.D., Steiner, P.M., 2009. No Child Left Behind: An interim evaluation of its effects on learning using two interrupted time series each with its own non-equivalent comparison series. Northwestern University Institute for Policy Research, Northwestern University, Working Paper Series WP-09-11.

Zemelman, S., Daniels, H. & Hyde, A. (2005). *Today's Standards for Teaching and Learning in America's Schools, third Edition:* Portsmouth, New Hampshire: Heinemann.