Richmond, A. S., Aberasturi, S., Abernathy, T., Delvecchio, T., & Aberasturi, R. (2009). One school district's examination of least restrictive environments: The effectiveness of pull-out programs and inclusion classrooms. The Researcher, 21(1), 53-66.

One School District's Examination of Least Restrictive Environments: The Effectiveness of Pull-out and Inclusion Instruction

Aaron S. Richmond[1], Suzzane Aberasturi[2], Tammy Abernathy[2], Ruth Aberasturi[3], & Toni DelVecchio[3]

Metropolitan State College of Denver[1] *University of Nevada-Reno[2]*

Carson City School District[3]

The purpose of this study was to determine if the Least Restrictive Environment (LRE) policy of a local school district leads to positive learning outcomes for elementary students who have reading and math learning disabilities (LD). We were interested in determining whether students with LD benefit more from math and reading instruction in different placement options (pull-out vs. inclusion programs). We randomly sampled elementary students who were either reading or math LD and received instruction in either inclusion or pull-out programs. We also sampled typically achieving students. In a six-month pre-post design, we compared scores from the Woodcock-Johnson Tests of Achievement. The results found that both reading pull-out and inclusion programs showed an increase in reading comprehension achievement scores but there was not a difference between the two programs. However, LD students in the pull-out program made significantly more gains than the inclusion and the control group students in letter-word identification. There were no other differences detected. The results are discussed and applied to the ongoing debate in regards to placement options along the LRE continuum for LD students.

***Keywords**: Least Restrictive Environment, Math and Reading Learning Disabilities, Elementary Education*

For long as students with disabilities have been educated in public schools, general education reform movements have had a significant impact on these students. Conventional conversations consider general and special education as separate systems whose policies overlap only rarely. This history goes back to the initial poor participation of general education teachers and parents in the initial development of the Education for All Handicapped Childrens Act of 1975. In actuality, any policy decision effecting children in schools, impacts all children in schools. What remains unclear is the effectiveness of particular policy shifts on particular groups of children. That work is consistently given to researchers to answer. The inclusion of students with disabilities into general education classrooms represents a key case in point.

Inclusion as we think of it today has strong ties to the 1983 National Commission on Excellence in Education Report. Using the effective schools research (Hawley &

Contact information: Aaron S. Richmond, Department of Psychology, Metropolitan State College of Denver, Campus Box 54, P. O. Box 173362, Denver, CO, 80217-3362; E-mail: arichmo3@mscd.edu

Rosenholtz, 1984; Purkey & Smith, 1983, 1985) and its premise that environmental characteristics impact achievement, many schools took the research and the admonishment of the National Commission on Excellence as an opportunity to raise achievement scores by focusing on mean outcome scores and bypassing the achievement levels of individuals (Eubanks & Levine, 1983).

While this movement initially circumvented the needs of children with disabilities and did not mention them either, savvy scholars (Gartner & Kipsky, 1984; Gerber & Semmel, 1985; Lilly, 1987; Sapon-Shevin, 1987; Wang, Reyonds, & Walherb, 1988) discovered the flaw in those schools subscribing to the effective schools research. Simply emphasizing high standards, teachers catered to students who learned easily and rewarded teachers for pushing ahead rather than working to teach all students. As a result, student achievement levels in classrooms could become more variable while overall achievement scores increased. This caused students who struggled to learn (i.e., students with learning disabilities) to become marginalized and fall further behind their classmates.

Simultaneously, policymakers such as Will (1986), and others (Bickel & Bickel, 1986; Edmonds 1979) espoused that if the effective schools research could improve educational outcomes for general education students, it could benefit all students. Further, Will (1986) proposed that special education services for students with disabilities would be obsolete and that all students could thrive in the general education environment, if the effective schools research was implemented correctly. Numerous advocates of integration subscribed to this position (Gartner & Lipsky, 1984; Lilly, 1988; Pugach, 1988; Stainback & Stainback, 1989). Many of these advocates perceived the original Education for All Handicapped Children Act of 1975 (IDEA), to be a mandate to integrate all children with disabilities into general education classrooms.

Debate raged between scholars finding flaws in the effective schools research and it's over emphasis on academic outcomes at the school level while ignoring outcomes at the individual level. The debate persisted as there were not any data related to efficacy to validate the need for special education placement options, such as resource rooms, or full-time general education placement.

Regardless of the lack of empirical evidence, inclusion increased 95% between 1987 and 1993 (Lerner, 1997). Inclusion, often referred to as a "moral imperative" (Cook, Semmel & Gerber, 1999, p. 199), flourished and continues to do so despite that fact that little data to support improved outcomes for students with mild disabilities exists (Baker & Zigmond, 1995; Ochoa & Olivarez, 1995). This shift in placement for students with disabilities suggests that the general education teachers have become increasingly responsible for the performance outcomes of both general and special education students.

PULL-OUT PROGRAMS

Originally, the philosophy behind pull-out programs was for students with disabilities to receive specialized or intensive instruction in an identified area of need in a small-group environment with a specially trained teacher. Reintegration of the student into the general education classroom when the student achieved the ability to perform independently was the goal of this type of program (Zigmond & Baker, 1996). However, Denton, Vaughn, and Fletcher (2003) found most studies of this nature indicated special education placement tended to stabilize growth, particularly in the area of reading, rather

than accelerate it. The authors cited a paper by Hanushek, Kain, and Rivkin (1998) that found students in special education pull-out programs experienced a change of only a slight increase in reading ability. In fact, most students, once classified as LD, remained under the special education umbrella and were placed in a resource room for the remainder of their years (Denton et al., 2003).

Several studies (Bentum & Aaron, 2003; Gartner & Lipsky, 1984; McKinney & Feagans, 1984) have examined the inability of pull-out programs to produce any significant increases in achievement. First, Bentum and Aaron (2003) found significant declines in achievement in spelling when students received instruction in a resource room environment as compared to students in an inclusion program. However, the researchers found no differences in the areas of reading and math achievement. Second, in Wang and Baker's (2001) meta-analysis, they found that students with learning disabilities in inclusion programs experienced statistically higher achievement gains when compared to students with learning disabilities in pull-out programs. However, these results are suspect as only 3% of the participants were actually classified as LD. The majority of the participants were mentally retarded or hearing impaired. Therefore, it is difficult to generalize these results to students with learning disabilities.

As is evident by prior research, there are holes and inconsistencies in the current research as to the efficacy of pull-out programs for students with learning disabilities.

INCLUSION PROGRAMS

As with the research on the effects pull-out programs have on students, there are inconsistencies in the literature on the effect inclusion programs have on students with learning disabilities. For example, Rea, McLaughlin &, Walther-Thomas (2002) found that when they compared special education students in inclusion versus pull-out programs, students in the inclusion programs had higher grades and achieved higher or comparable scores on standardized tests. Additionally, Daniel and King (1997) found that students who received their instruction in inclusion programs had higher grades for reading but not for math.

In a study that was inconsistent with Rea et al. (2002) and Daniels and King (1997), Manset and Semmel (1997) found that inclusion programs worked well for some students, but not for all, and that each situation needs to be evaluated on a case-by-case basis. These studies suggest that decisions regarding placement options for students with learning disabilities should not be wholesale policy decisions, but rather individual decisions as mandated by IDEA (Individual with Disabilities Education Act).

While some of the research purports the effectiveness of inclusion programs (Rea et al., 2002; Daniels & King, 1997) there is some research that highlights the shortcomings of inclusion programs. Zigmond and Baker (1996) examined education programs and teachers' opinions about full inclusion programs for LD students. Zigmond and Baker report that although most teachers incorporated techniques to include all students (e.g., reading tests out-loud and allowing students to pick test items), they felt students with learning disabilities still needed more individualized instruction. In their study, many of the teachers resorted to small group instruction conducted by the special education staff in the hallway. One fifth grade teacher expressed the following concern, "We're not teaching them how to read. I think we're just doing accommodation. Nobody has time to teach these kids [fifth graders] how to read back

at their second grade level" (Zigmond & Baker, 1996, p. 31).

Similar to the research on pull-out programs, the inconsistencies in the literature necessitate a clear need to assess the effectiveness of these two types of instruction for students with LD.

Purpose and Rationale of Study

Despite pronouncements of the moral good of inclusion or need for smaller instructional environments that pull-out programs provide for some children, the LRE and Individualized Education Plan for each child has once again become part of federal law IDEA (2004). Educators have the option and responsibility to make sensible placement decisions in the best interest of the students they serve. Research is one tool educators can use to make these decisions, but with a literature base in its infancy that cannot keep up with the policy shifts, it remains in the hands of site personnel to make professional decisions. Studies espousing the benefits of inclusion (Daniel & King, 1997; Rea et al., 2002) are as numerous as those suggesting students with disabilities or a reading difficulty need specialized, small-group, intensive instruction (Denton et al., 2003; Marston, 1996; Whinnery & King, 1995; Zigmond & Baker, 1996).

Even though much of the research has discussed flaws with both types of placement options the question remains whether students with LD experience the same gains in achievement as their non-disabled peers and which program, inclusion versus pull-out, accomplishes this goal. Based on the lack of and inconsistent prior research, this study sought to answer the following questions:

1. Is there a difference in academic reading achievement among students receiving instruction in a pull-out program or an inclusion program or their non-LD cohorts in a general education program?
2. Is there a difference in academic math achievement among students receiving instruction in a pull-out program or an inclusion program or their non-LD peers in a general education program?

This study differs from other studies in that the comparisons are made using randomly selected students from several classrooms, both special education and general education, and from six schools. This design controls for the impact that one good teacher might have on the results. In addition, students were not selected for pull-out or inclusion based on functioning. Instead several of the schools in this district had a policy of either doing inclusion or pull-out regardless of functioning. Because the students with learning disabilities were all enrolled in the same district, consistent district policies were used in the identification of the disability and the eligibility process for each student. Summarily, every effort was made to minimize the effects of confounding variables and still maintain a ecologically valid assessment of student achievement.

Method

Participants

Participants consisted of five groups of students in grades 1-6. These groups were: (1) students without disabilities, (2) students with LD in math receiving inclusion instruction, (3) students with LD in math receiving pull-out instruction, (4) students with LD in reading receiving inclusion instruction, and (5) students with LD in reading receiving pull-out instruction. Tables 1 describe the demographics of the students in the three groups for both the reading and math studies.

For both math and reading LD, the inclusion group was comprised of students

identified as LD (in their respective area) who received all of their instruction in the general education classroom with the special education teacher providing classroom support. The pull-out students were identified as having LD and were pulled out of the general education class to be taught by a special education teacher in a resource setting. Students in the control condition were typically developing students without disabilities who received their instruction in the general education classroom.

Table 1. *Percentages and Means for Demographic Data*

Reading LD			
Type of Classroom	Grade Level	Gender	Average Age
Inclusion	2nd -17.2%	Males-72.4%	Males-9.9
(*n* = 25)	3rd -27.6%	Females-27.6%	Females-9.7
	4th -17.2%		
	5th -37.9%		
Pull-out	2nd -5.6%	Males-69.4%	Males-10.0
(*n* = 33)	3rd -22.2%	Females-30.6%	Females-9.9
	4th -27.8%		
	5th -44.4%		
Control	2nd -2.9%	Males-52.9%	Males-10.1
(*n* = 34)	3rd -32.4%	Females-47.1%	Females-9.7
	4th -23.5%		
	5th -41.2%		
Math LD			
Inclusion	2nd -9.5%	Males-38.1%	Males-10.4
(*n* = 21)	3rd -28.6%	Females-61.9%	Females-10.1
	4th -14.3%		
	5th -47.6%		
Pull-out	2nd -4.3%	Males-34.8%	Males-9.7
(*n* = 23)	3rd -21.7%	Females-65.2%	Females-10.4
	4th -30.4%		
	5th -43.5%		
Control	2nd -3.3%	Males-56.7%	Males-10.1
(*n* = 30)	3rd -26.7%	Females-43.3%	Females-9.9
	4th -23.3%		
	5th -46.7%		

Students with LD were identified using the more traditional discrepancy model indicating a significant difference between their ability and achievement in the area of math and/or reading as set by the state statutes. This discrepancy model was based on a regression formula with more than one standard deviation difference necessary as the IQ increases and less as it decreases. Eligibility for special education placement was determined separately from this study by licensed school psychologists in the school district. The students included in this study were randomly selected from a pool of either pull-out or inclusion students within all the six elementary schools in the district (*N* = 8,576 children). Students were excluded if they were identified as other

health impaired, mentally retarded, autistic, English as a Second Language, emotionally handicapped or physically handicapped. If a student was identified as primarily LD yet met the criteria for these other categories they were also excluded. This was done on a case-by-case basis with the Special Education Director and the individual special education teachers in the district. These decisions were made to ensure that our sample only included students that represent the LD population.

Students without LD were randomly selected from the schools represented in this study. Students were selected within certain grade levels to match the grade level of students with disabilities assigned to groups (e.g., if 10 students participated from one school, then 10 non-LD students were also sampled from that same school and grade level).

SAMPLE AND POPULATION DESCRIPTION

There are several distinct factors regarding the school district that should be explained. First, for all schools in this study, the school district had the policy of using inclusion or pull-out regardless of the student's degree of functioning. Choosing to use inclusion or pull-out exclusively was a decision made by the individual schools. Second, the school district had a total 12.6% indentified disabled students. Third, each of the six elementary schools conducts a 90-minute literacy instruction block each day. All elementary schools in the state are under the 90-minute literacy block mandate as well. Reading is the major part of the block which includes; comprehension, writing, listening, and speaking. Fourth, schools were able to select reading programs from the district adopted list. Additionally, the schools in this project participate in a highly prescriptive reading program.

Lastly, the quality of classroom instruction was not assessed or observed. It was assumed that the schools and classrooms complied with the district and state literacy and math standards.

MEASURES

The Woodcock-Johnson Tests of Achievement 3rd edition (WJ-III; Woodcock, McGrew, & Mather, 2001) was used as the measure of achievement. Only the six subtests that compose the reading and math scales were used (i.e., the Letter-Word Identification, Passage Comprehension, Reading Fluency, Math Fluency, Calculation, and Applied Problems subtests). The WJ-III has acceptable levels of reliability (median reliability coefficient alphas range from .81 to .94) (Woodcock et al., 2001).

Reading achievement measures. The Passage Comprehension subtest gives a measure of a student's ability to understand what they read. The Reading Fluency subtest gives a measure of the speeding rate of the student and the Letter-Word Identification subtest measures the student's ability to pronounce words or phonetic ability (Woodcock et al., 2001).

Math achievement measures. The three sub-tests of the WJ-III that were used to asses math achievement were the Math Fluency, Calculation, and Applied Problems (Woodcock, et al., 2001). The Math fluency subtest essentially measures the efficiency of completing simple math tasks in 3 minute intervals. The Calculation subtest requires students to solve simple arithmetic computations. The Applied Problems subtest assesses a students' ability to solve oral, mathematical word problems (Woodcock et al., 2001)

To control for a testing effect, the WJ-III includes both an A and B version. Students received version A during the pre-assessment and version B during the post-assessment (Woodcock, et al., 2001).

PROCEDURE

After informed consent was obtained, all participants were administered version A of the WJ-III (Woodcock et al., 2001). This administration took place during the beginning of the second half of the school year. Participants received either the math subtests of the WJ-III or the reading subtests of the WJ-III. On average, licensed school psychologists took 30 minutes to administer the assessment.

During the first day of testing all students were administered a basic demographic questionnaire.

Four months later, all participants were administered their respected subtests of the WJ-III B version.

RESULTS

For purposes of clarity, we divided the results into two studies: reading and math.

READING STUDY

A single factor multivariate analysis of variance (MANOVA) was conducted to determine the overall effect of placement on the LRE continuum on the four dependent variables: Letter-Word Identification, Reading Fluency, Passage Comprehension, and Broad Reading scores. The score used in the analysis was the difference between the post-score and the pre-score on each of the subtests. Following the procedures of Mertler and Vanatta (2002), data were initially screened for significant outliers, violations of assumed normality and homoscedasticity. Two cases were deleted as they were considered significant outliers by calculating the Maholanobis' distance (Mertler & Vannatta, 2002; Stevens, 1996). The Leven's test for Homoscedasticity was not significant (p>.05) indicating variability among the dependent variables (DVs) were similar (Tabachnick & Fidell, 1996). Using both the Box's M Test and the Kolmogorov-Smirnov Test, the four DVs were tested for normality. Results indicate that all DVs were normally distributed, $p>.05$. Since the assumption of normality was meet, Wilks' Lambda Wilks' Λ) was used as the test statistic for the MANOVA (Mertler & Vannata, 2002). MANOVA results indicated that class type (Wilks' $\Lambda=.78$, $F(2, 89) = 2.82$, $p<.01$, $\eta^2 = .12$) significantly affects the combined DV of Letter-Word Identification, Reading Fluency, Passage Comprehension, and Broad Reading scores (see Table 2 for means and standard deviations). This was a medium to large effect. Green, Salkind, and Akey (2000) report .01 as a small effect, .06 as a medium effect, and .14 as a large effect.

Four single factor univariate analyses of variance (ANOVAs) were conducted as follow-up tests. In order to control for Type I Errors on the four ANOVAs, a Holm's Sequential Bonferroni method was used to lower alpha levels (Green et al., 2000). A Levene's test of homogeneity indicated that all DVs exhibited similar variances ($p> .05$). ANOVA results indicated that classroom type significantly differs for the DVs letter-word identification ($F(2, 89) = 3.50$, $p<.05$, partial $\eta^2 =.07$), Passage Comprehension ($F(2, 89) = 4.48$, $p<.05$, partial $\eta^2 =.09$), and Broad Reading score ($F(2, 89) = 5.75$, $p<.01$, partial $\eta^2 =.11$). Refer to Figure 1 for the Letter-Word Identification, Passage Comprehension, and Broad Reading scores. However, there was not a significant effect for Reading Fluency ($F(2, 89)= 0.72$, $p>.05$, partial $\eta^2=.02$). The calculated effect size for Reading Fluency indicates that a small proportion of Reading Fluency is accounted for by class type.

Scheffé post-hoc tests were conducted for the Letter-Word Identification, Passage Comprehension, and Broad Reading scores on each of the three significant DVs (see Table 3). For Letter-Word Identification the Scheffé post-hoc results indicated that pull-out students significantly differed from the control students. For the Letter-Word Identification subtest, the participants in

pull-out programs increased their score more from pre- to post-test than the participants in the control condition. Control students' scores declined. The inclusion students did not significantly differ from either control students or the pull-out participants.

The Scheffé post-hoc results for the Passage Comprehension and the Broad Reading subtests indicated that the pull-out and inclusion students differed significantly from Control students (see Table 3). Both pull-out and the inclusion students' scores increased from pre- to post-test on the Passage Comprehension and Broad Reading subtests, while the control students' scores decreased over time.

Table 2. *Pre and Post-Test Scores for the Reading Dependent Variables Sorted by Classroom Type*

DV	Classroom Type	Pre-Test		Post-Test	
		M	SD	*M*	SD
Letter-Word Identification	Inclusion	81.54	9.69	80.46	9.38
	Pull-out	81.37	7.17	82.89	9.11
	Control	101.74	7.56	100.03	7.18
Reading Fluency	Inclusion	84.44	9.11	87.20	9.19
	Pull-out	86.30	8.38	87.50	6.19
	Control	100.26	7.39	101.79	8.4
Passage Comprehension	Inclusion	7938	8.03	82.08	6.91
	Pull-out	80.40	8.40	82.46	7.28
	Control	97.12	6.19	93.94	8.44
Broad Reading	Inclusion	78.42	9.75	80.46	9.09
	Pull-out	79.49	8.19	81.37	7.18
	Control	99.85	7.01	99.06	6.91
ORF	Inclusion	44.96	25.79	56.50	25.00
	Pull-out	48.65	20.55	60.06	22.30
	Control	104.38	23.76	117.30	23.05

Figure 1. Estimated Marginal Mean Difference Scores of Letter-Word Identification, Passage Completion, & Broad Reading by Classroom Type

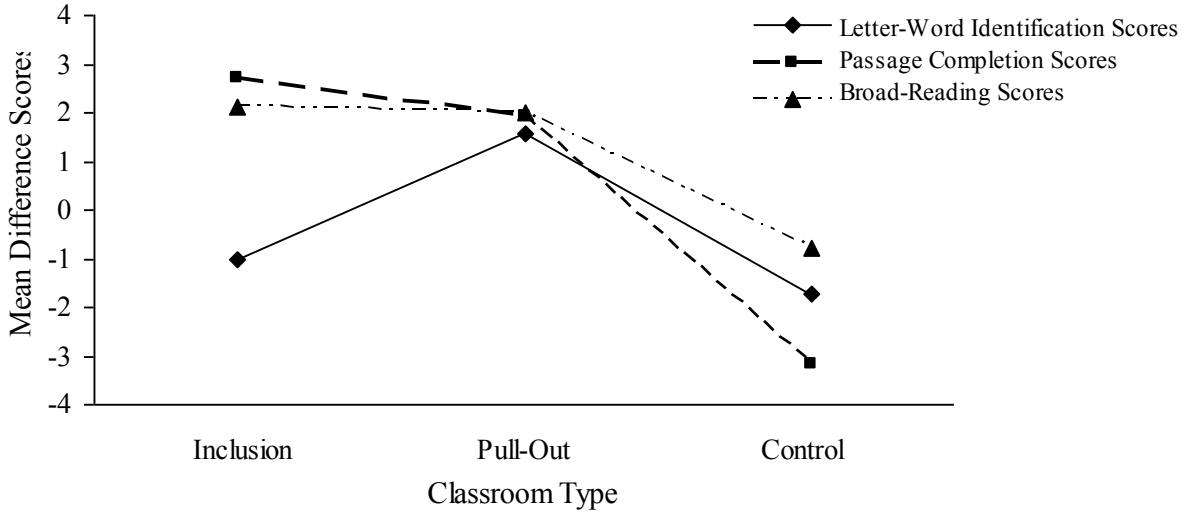

Table 3. *Mean Difference Reading Scores Among the Three Treatments*

Letter-Word Identification			
Conditions	Inclusion	Pull-out	Control
Inclusion	—	2.58	.71
Pull-Out		—	-3.28*
Control			—
Passage Completion			
Conditions	Inclusion	Pull-out	Control
Inclusion	—	.81	5.90*
Pull-Out		—	5.09*
Control			—
Broad-Range Reading			
Conditions	Inclusion	Pull-out	Control
Inclusion	—	.12	2.91*
Pull-Out		—	2.80*
Control			—

Note: (*) Denotes $p<.05$.

Table 4. *Pre and Post-Test Scores for the Math Dependent Variables Sorted by Classroom Type*

DV	Classroom Type	Pre-Test		Post-Test	
		M	SD	*M*	SD
Broad Math	Inclusion	88.24	8.29	87.76	7.82
	Pull-out	85.52	9.94	84.43	11.80
	Control	105.16	6.49	107.33	7.87
Math Calculation	Inclusion	90.71	8.64	91.42	10.28
	Pull-out	86.69	10.96	86.34	9.64
	Control	104.46	7.51	106.16	8.41
Math Fluency	Inclusion	90.61	10.95	87.19	7.97
	Pull-out	90.08	11.45	85.34	13.50
	Control	102.93	11.95	105.33	7.25
Applied Problems	Inclusion	88.19	8.00	91.76	10.59
	Pull-out	86.26	10.30	86.95	13.41
	Control	104.53	7.53	106.10	12.43

Note. Inclusion (n = 21), Pull-out (n = 23), Control (n = 30)

In addition, an analysis of the Broad Reading scores was undertaken to determine if the three groups scores changed significantly from pre- to post-test. The scores in Table 3 represent the mean difference between pre and post assessments for each the three groups. The results found that students in both the inclusion (t = 3.08, n = 26, p < .01) and the pull-out groups (t = 2.92, n = 35, p <.01) increased their scores significantly from pre- to post assessment. There was no difference between the pre and post assessments for the control students on the Broad Reading score (t = 1.19, n = 34, p > .05).

MATH STUDY

Difference scores were calculated on the pre-post scores of math subtests of the WJ-III (Woodcock et al., 2001). A MANOVA was conducted to determine the effect of instructional setting on the four dependent variables of Math Calculation, Math Fluency, Applied Problems, and Broad Math. See Table 4 for means and standard deviations. All pre-data procedures suggested by Mertler and Vanatta (2002) were conducted. Three participants from the control group were deleted because their scores were deemed to be significant outliers. All other assumptions (i.e., normality, linearity, and homoscedasticity) were met. MANOVA results indicate that instruction type (Wilks' Λ=.815, $F(2, 136) = 1.83$, $p = .077$, $\eta^2 = .10$) did not affect the combined dependent variables of Math Calculation, Math Fluency, Applied Problems, and Broad Math. Therefore, there were no differences in students' performances in the different instructional settings. The average difference between the pre- and post-scores ranged from -.24 to .78, which indicates the students did not show large increases or decreases in scores from pre to post assessment regardless of type of classroom.

DISCUSSION

The results of this study found that for reading comprehension and for reading overall, it did not seem to make a difference whether students with LD were pulled-out of the general education classroom for instruction or whether instruction was conducted in the general education classroom (i.e., inclusion). However, the results of this study did suggest that for Letter-Word Identification, which offers a measure of phonemic ability, the students that were pulled-out improved their scores significantly more than students without disabilities in general education classrooms (control group). This suggests that pull-out may be the more appropriate placement option for building specific skills as opposed to conducting specialized instruction in the larger classroom. Schools needing to enhance specific skill sets for students, for example phonics, may find as this study suggests that pull-out models continue to be viable instructional models for certain tasks and specific students. This result did not hold true for math instruction as there were no differences between students in pull-out, or inclusion settings, or students without disabilities in the control group on the difference scores between pre- and post tests.

Interestingly, special education students, regardless of whether they were instructed in pull-out or inclusion settings, showed more improvement than students without disabilities on Letter-Word Identification and Broad Reading. Of concern was the lower post-test score for students without disabilities. Unfortunately, this finding fuels the continuous debate about the impact of struggling students on their typically developing peers. While students with disabilities were able to succeed in both the pull-out and inclusion setting at the same rate, students without disabilities were failing to make progress One possible explanation for the success of students with LD in inclusion settings was the level of support they received from special education teachers working in the general education classroom. Further, in this particular district, unlike others across the country, retention of qualified special educators has not been a critical problem. Special education teachers at these school sites averaged six years experience. Students with LD who received pull-out services were often present in the general education classroom for the 90 minute literacy block and then received additional specialized reading instruction later in the

day, thus giving these students additional instructional time.

More concerning than whether students with LD thrive in pull-out or inclusion settings in terms of their reading performance is the poor performance of students without LD. Critics of inclusion models may be quick to say that instruction is "dummied down" for students who struggle, thus deflating the scores. Cook et al. (1997) refer to this as "leveling". The shrinking standard deviation from the pre and post-test scores (Letter Word Recognition, Broad Reading, ORF) may lend some weight to this argument. Another reason for the flat or declining scores may be the nature of the test. Students with LD are socialized to assessment instruments such as the WJ-III, while students without disabilities are not used to these norm-referenced assessments with their out of context items. Further, students with disabilities are used the assessment process and participate in it regularly. The students without disabilities who participated in this studied had recently completed state and county assessments and may not have thought this follow-up assessment was important enough to give it their full attention. Finally, the short time span between the pre and post-test (one semester) may be too tight of a timeline for this type of study. However, typically, in investigations where data are collected in a shorten time frame, it is usually the results of students with LD that is compressed rather than the scores of the students without disabilities.

Another result of this study shows that although both groups of students with LD (pull-out and inclusion) made significant increases from pre- to post-testing, they still remained in the below average range of functioning. The average post-test score on Broad Reading for the inclusion group was 80.46 and the average post-test score for the pull-out group was 81.69. Regardless of how much progress these students made, they are still falling over one standard deviation below the mean in their achievement scores in the area of reading. Future research conducted in these schools needs to monitor the rate of progress between students with LD and students without disabilitiesAs long as students with disabilities are maintaining a projected growth rate and maintaining or closing the gap between themselves and their peers then participation in the general education classroom is successful.

A significant contribution of this project is the generalizability of the results. The participants for this study represent students from six schools within a school district where as the majority of previously conducted studies examined only one or two classrooms. Therefore, this study may have better generalizability than previous studies (Bentum & Aaron, 2003; Gartner & Lipsky, 1984; McKinney & Feagans, 1984) because it represents a wider range of teaching styles and functioning. In addition, students were not selected for the pull-out or inclusion groups based on functioning. Instead several of the schools in this district had a policy of either doing inclusion or pull-out regardless of functioning. The homogeneity of Woodcock-Johnson scores between the two special education groups reflect that the students possessed relatively equal ability at least in the area of achievement. The size of the study also highlights its weakness. Given the district's policy of adopting multiple literacy programs, not all students included in this study participated in the same literacy program. Further, the study did not control for quality of teaching or adherence to the 90 minute literacy block rule.

CONCLUSION

In conclusion, given typical reading instruction as outlined by this district,

students with LD remained behind their nondisabled peers, yet they did not lose ground. Students with LD progress at the same rate regardless of whether they were being instructed in an inclusion setting with support or in a pull-out program, thus making the placement choice more dependent on other more personal variables associated with the child. What remains clear from this study is that decisions about the LRE for a specific child are as idiosyncratic as ever, (Semmel & Manset, 1997) and students with LD do not appear to be harmed by thoughtful decisions made on their behalf. As directed by IDEA placement decisions should be made on a case-by-case basis. As caseworkers make placement decisions, it is critical that instructional deficits be carefully considered in the decisions and that the best environment for teaching the instructional skill deficits of the child be considered. Further, it is essential that school personnel begin to think of placement options in a more fluid manner. For example, students may need pull-out services for an area of need, but only temporarily. One year later, the child may need pull-out services again.

REFERENCES

Baker, J. M., & Zigmond, D. N. (1995). The meaning and practices of inclusion for students with learning disabilities: Implications from the five cases. *Journal of Special Education, 29*, 163-180.

Bentum, K. E., & Aaron, P. G. (2003). Does reading instruction in learning disability resource rooms really work? A longitudinal study. *Reading Psychology, 24*, 341-382.

Bickel, W. E., & Bickel, D. D. (1986). Effective schools, classrooms, and instruction: Implications for special education. *Exceptional Children, 52*, 489-500.

Cook, B. G., Gerber, M. M. & Semmel, M. I. (1997). Are effective schools reforms effective for all students? The implications of joint outcome production for school reform. *Exceptionality, 7*(2) 77-95.

Daniel, L. G., & King, D. A. (1997). Impact on inclusion education on academic achievement, student behavior and self-esteem, and parental attitudes. *Journal of Educational Research, 19(2)*, 67-80.

Denton, C. A., Vaughn, S., & Fletcher, J. M. (2003). Bringing research-based practice in reading intervention to scale. *Learning Disabilities Research & Practice, 18*(3), 201-211.

Edmonds, R. R. (1979). Effective schools for the urban poor. *Educational Leadership, 37*(1), 15-24.

Education for All Handicapped Children Act of 1975, 20 U.S.C., 1400 *et seq.* (1975).

Eubanks, E. E., & Levine, D. U. (1983). A first look at effective schools projects in New York City and Milwaukee. *Phi Delta Kappan, 64*, 697-702.

Gartner, A., & Lipsky, D. K. (1984). Beyond special education: Toward a quality system for all students. *Harvard Educational Review, 57*, 367-395.

Gerber, M. M., & Semmel, M. I. (1985). The microeconomics of referral and reintegration: A paradigm for evaluation of special education. *Studies in Educational Evaluation, 11*, 13-29.

Green, S. B., Salkind, N. J., & Akey, T. M. (2000). *Using SPSS for windows: Analyzing and understanding data.* (2nd Ed.) Upper Saddle River, NY: Prentice Hall.

Hanushek, E. A., Kain, J. F., & Rivkin, S. K. (1998). *Does special education raise academic achievement for students with disabilities?* (National Bureau of Economic Research, Working Paper No. 6690). Cambridge, MA. Available at http://www.nber.org/papers/w6690.

Hawley, W. D., & Rosenholtz, S. J. (1984). Good schools: What research says about improving student achievement. *Peabody Journal of Education, 61*(4), 1-178.

Individuals with Disabilities Education Act (IDEA) Amendments of 1997. 11 stat 37 (20 USC 1401 [26]).

Lerner, J. (1997). *Learning disabilities: Theories, diagnosis and teaching strategies* (7th ed.), Boston: Houghton Mifflin.

Lilly, S. (1987). Lack of focus on special education students in literature on educational reform. *Exceptional Children, 53*, 325-326.

Lilly, S. (1988). The regular education initiative: A force for change in general and special education. *Education and Training in Mental Retardation, 23*, 253-260.

Manset, G., & Semmel, M. I. (1997). Are inclusive programs for students with mild disabilities effective? A comparative review of model programs. *Journal of Special Education, 31(2)*, 155-180.

Marston, D. (1996). A comparison of inclusion only, pull-out only and combined service models for students with mild disabilities. *Journal of Special Education, 30*(2), 121-132.

Mertler, C. A., & Vannatta, R. A. (2002). *Advanced and multivariate statistical methods* (2nd ed.). Los Angeles: Pyrczak.

McKinney, J. D., & Feagans, L. (1984). Longitudinal studies of learning disabled and average achievers: Academic and behavioral characteristics. *Learning Disability Quarterly, 7*, 251-265.

National Commission on Excellence in Education. (1983). *The nation at risk: The imperative for educational reform.* Washington, DC: U.S. Department of Education.

Ochoa, S., & Olivarez, A. (1995). A meta-analysis of peer tutoring socio-metric studies of pupils with learning disabilities, *Journal of Special Education, 29,* 1-19.

Pugach, M. (1988). The consulting teacher in the context of educational reform. *Exceptional Children, 55*, 266-277.

Purkey, S. C., & Smith, M. S. (1983). Effective schools: A review. *The Elementary School Journal*, *83,* 427-452.

Purkey, S.C., & Smith, M.S. (1985). School reform: The district policy implications of the effective schools literature. *The Elementary School Journal, 85*, 353-389.

Rea, P. J., McLaughlin, V. L., & Walther-Thomas, C. (2002). Outcomes for students with learning disabilities in inclusive and pullout programs. *Exceptional Children, 68*(2), 203-233.

Sapon-Shevin, M. (1987). The national education reports and special education: Implications for students. *Exceptional Children, 53*, 391-398.

Stainback, W., & Stainback, S., (1989). Common concerns regarding merger. In S. Stainback, W., Stainback, & M. Forest (Eds.), *Educating all students in the mainstream of regular education* (pp. 255-274). Baltimore, MD: Brookes.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences (2nd ed.).* Hillsdale, NJ: Lawrence Erlbaum Associates.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics (3rd ed.).* New York: Harper Collins.

Wang, M. C. & Baker, E. T. (2001). Mainstreaming programs: Design features and effects. *Journal of Special Education, 19*(4), 503-521.

Wang, M. C., Reynolds, M. C., & Walberg, H .J. (1988). Integrating the children of the second system. *Phi Delta Kappan, 70*, 248-251.

Whinnery, K. W. & King, M. (1995). Perceptions of students with learning disabilities. *Preventing School Failure, 40*(1), 5-10.

Will, M.C. (1986). *Educating students with learning problems: A shared responsibility.* Washington, DC: U.S. Department of education, Office of Special Education and Rehabilitative Services.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside.

Zigmond, N., & Baker, J. M. (1996). Full inclusion for students with learning disabilities: Too much of a good thing? *Theory into Practice, 35*(1), 26-34.