



Impact of the Expeditionary Learning Model on Student Academic Performance in Rochester, New York

July 2011



Executive Summary

The purpose of this study was to evaluate the impact of the Expeditionary Learning (EL) model on student academic performance in New York's Rochester City School District during the 2007-08 and 2008-09 school years. Using a quasi-experimental matched comparison group design, scores of two EL schools (one elementary and one middle) and twelve non-EL schools (four elementary and eight middle) were compared on the New York state English Language Arts and mathematics assessments.

Possible pre-existing differences between the EL and comparison schools were addressed in two ways. First, the comparison schools were chosen to match the EL schools on percentage of low-income students and English language learners. Second, differences in student gender, ethnicity, special education status, limited English proficiency status, disability status, and baseline test scores were accounted for using statistical methods.

The two primary outcome statistics were (1) the effect size of participating in an EL school on achieving a score of Proficient or higher on the state assessments, and (2) the percentage of students who would have shifted to a different Proficiency category based on this EL effect size.

The effect of being in an Expeditionary Learning school was positive, substantial, and significant ($p < .01$) for both years of elementary school mathematics, both years of elementary school English Language Arts, and both years of middle school English Language Arts. The difference between EL and non-EL schools on middle school mathematics in 2007-08 was not statistically significant, and non-EL schools showed a significant advantage ($p < .05$) in middle school mathematics in 2008-09.

The effect size of enrollment in an EL school suggested that, on the English Language Arts exam, 55% of elementary students and 39% of middle school students would have increased from the Non-proficient level to the Proficient level. On the mathematics exam, the EL effect size suggested that 65% of elementary students would have increased from the Non-proficient level to the Proficient level, and that the percentage of students scoring at the Non-proficient level on the Mathematics exam would have increased by 16%.

These strong and significant impacts should be of interest to those seeking school reform approaches that raise student scores on standardized assessments of elementary and middle school English Language Arts, as well as elementary mathematics. These findings should also raise concerns for those seeking to increase standardized assessment scores in middle school mathematics. Promising directions for future research include applying similar analyses to a larger sample of Expeditionary Learning schools, attempting to determine what aspects of the EL model are most responsible for observed impacts, and employing randomized designs that take advantage of Rochester's enrollment lottery.

Introduction

The purpose of this study was to evaluate the impact of the Expeditionary Learning (EL) model on student academic performance in New York's Rochester City School District (RCSD) during the 2007-08 and 2008-09 school years. Using a quasi-experimental matched comparison group design with statistical controls, scores of two EL schools and twelve matched non-EL schools were compared on the New York state English Language Arts and mathematics assessments.

Method¹

The RCSD provided student-level data for the five school years from 2004-05 through 2008-09. Data for each year included background variables (gender, ethnicity, limited English proficiency status, special education status, attendance, grade level, school attended), academic performance variables (scaled scores on the New York State mathematics and English language arts assessments), and an identification code that permitted tracking of individual students across years and data sets.

The study used a quasi-experimental matched comparison group design with statistical controls. The two Expeditionary Learning schools in the Rochester district, an elementary school and a middle school, served as the intervention group. The middle school was a school with grades 7-12, but Expeditionary Learning works only with grades 7 and 8 in the school. Since legal barriers prevented RCSD from providing student-level data on free/reduced lunch (FRL) status, socioeconomic data at the school level were used instead and served as the selection criterion for comparison schools.

The NYS Department of Education ranks all schools into need-level categories based on proportion of students with free-lunch eligibility and proportion of limited English proficiency (LEP) students. The RCSD comparison schools for this study were the 4 elementary schools that fell into the same need categories as the EL elementary school, and the 8 middle schools which both matched the EL middle school need category and were situated in secondary schools serving grades 7-12. The selection of comparison schools was further validated by comparing EL and comparison schools on school-level FRL statistics provided by RCSD. At the elementary level, 78% of EL students were FRL compared to 69% of comparison students. At the middle school level, 77% of EL students were FRL compared to 83% of comparison students.

Investigations of academic outcomes of EL vs. non-EL students at the elementary and middle school levels were then conducted for the 2007-08 and 2008-09 school years. Data provided by RCSD for earlier years were used to obtain baseline (6th grade) achievement scores for middle school students. For students who repeated 6th grade, the most recent 6th grade scores were used. Baseline achievement scores were not available for elementary students, because the first state assessment happens in 3rd grade, and by then students have already received substantial exposure to the EL model.

To obtain sample sizes that provided adequate power, regression analyses were conducted at the school level rather than at the grade level. Aggregating test scores at the school level first required standardizing scores for each grade level and school year to a common mean and standard deviation, because scores on RCSD academic assessments have different means and standard deviations across grade levels and school

¹ This study served as a reanalysis and extension of Albert (2010), and replicated many of that study's methods, but diverged where noted below.

years. Regression analyses were then conducted that controlled for gender, ethnicity, LEP status, and special education status as well as baseline (6th-grade) achievement scores for the middle school.

These analyses yielded eight regression coefficients -- one for each combination of school type (elementary or middle), school year (2007-08 or 2008-09), and test (ELA or math) -- that represented the impact of participation in an EL school. Because the scores had been standardized (to a mean of zero and standard deviation of one), the regression coefficients were in standard deviation units. For example, a regression coefficient of 0.78 would imply that EL students received test scores that were 0.78 standard deviations higher than non-EL students after taking into account the effects of gender, ethnicity, LEP status, SPED status, and baseline achievement scores.

Regression coefficients were then contextualized by calculating the magnitude of the shifts in Proficiency status that comparison group students would experience if their scores were increased or decreased by the amount implied by the relevant effect size. In a previous study of these data (Albert, 2010), these implied shifts were calculated using statistics that assumed normality of student test score distributions. Statistical guidelines are poorly defined regarding the extent of deviation from normality that can be tolerated without threatening the validity of statistical conclusions, and this is particularly true for samples as large as those in the Rochester analyses. However, investigation of student score distributions revealed sufficient deviations from normality that existing guidelines were clearly violated.

To address this concern, implied changes were instead calculated directly from the regression coefficients and the Rochester student data. First, the regression coefficient for each test and year was multiplied by the standard deviation of scores for the same test and year, yielding the predicted change in points attributable to participation in an EL school. This was done separately for each grade, because different grades had different standard deviations. For example, for a regression coefficient of 0.5 and a test with a standard deviation of 28, the regression model predicts that EL students would score $(28 * 0.5) = 14$ points higher than non-EL students on that test. The resulting number of points was subtracted from 650, which for every year and grade was the lowest possible Proficient score.

Continuing the example, $650 - 14 = 636$, so every non-EL student who scored between 636 and 649 would have moved into the Proficient category if they had received the 14-point benefit of participation in an EL school. Next, the number of non-EL students in each grade who scored in that range was divided by the total number of non-EL students in the grade. These results were averaged across the four elementary grades (3-6) or the two middle grades (7-8), weighted by the number of students in each grade, which yielded the percentage point change in Proficiency rates of non-EL students predicted by the regression model. Finally, the percentage of comparison group students who would have shifted Proficiency levels was calculated.

Findings

The total student population in RCSD (those who attended school for at least one day) was 15,109 in 2007-08 and 14,637 in 2008-09. The test scores of this entire group were used to calculate the means and standard deviations for each grade, year, and test used in the analyses.

Tables 1A and 1B show the total number of EL and non-EL students in each year's sample, as well as the number and percentage with valid data. To be included in the regression analyses, all students required valid current-year test scores. Middle school students also required valid baseline test scores. Elementary schools didn't have baseline scores, so the rows corresponding to baseline scores for elementary school students are blank.

The shaded cells indicate the final number and percentage of students who had complete data for each year and test and were therefore included in analyses. For example, Table 1A shows that there were 153 EL middle school students in 2007-08, that 92% of those students had valid math scores, and that 71% also had valid baseline math scores. Therefore, 71% of the EL students were included in analyses of the 2007-08 middle school math assessments.

Student Grouping	Elementary				Middle School			
	EL		non-EL		EL		non-EL	
	N	%	N	%	N	%	N	%
All Students	178	100	714	100	153	100	2806	100
Valid ELA	161	90	638	89	145	95	2419	86
Valid ELA & Baseline ELA	--	--	--	--	112	73	1904	68
Valid Math	160	90	636	89	141	92	2416	86
Valid Math & Baseline Math	--	--	--	--	108	71	1891	67

Student Grouping	Elementary				Middle School			
	EL		non-EL		EL		non-EL	
	N	%	N	%	N	%	N	%
All Students	160	100	650	100	163	100	2517	100
Valid ELA	156	98	593	91	151	93	2161	86
Valid ELA & Baseline ELA	--	--	--	--	127	78	1938	77
Valid Math	156	98	593	91	145	89	2168	86
Valid Math & Baseline Math	--	--	--	--	123	75	1956	78

Averaging across all math and ELA groups in both years, missing data resulted in the exclusion of 6% of EL elementary students, 10% of non-EL elementary students, 26% of EL middle school students, and 28% of non-EL middle school students.

For those who might compare these findings to those of Albert (2010), it is worth noting that the current study has higher exclusion rates for middle school students. Albert's assignment of baseline test scores did not account for grade-level retention, but 21% of students took three or more years to move from 6th grade to 8th grade, so the baseline score used for some students in that study was actually their 7th grade score. Those students were assigned different baseline score across the two studies, and more of them were excluded from analyses in the current study. Most often this was because their true baseline year was 2004-05, and RCSD didn't provide 6th-grade test scores for that year.

Table 2 shows the regression coefficient for each test (ELA and math), year (2007-08 and 2008-09), and school type (elementary and middle). Each regression coefficient indicates an effect size, in standard

deviation units, that is attributable to participation in an EL school, after accounting for gender, ethnicity, LEP status, special education status, and baseline achievement scores (for middle school only).

Table 2: Effect Size of EL Enrollment on RCSD Student Test Scores in English Language Arts and Mathematics		
	ELA Effect Size (std dev)	Math Effect Size (std dev)
Elementary 2008-09	0.58**	0.67**
Elementary 2007-08	0.40**	0.64**
Middle 2008-09	0.38**	-0.18*
Middle 2007-08	0.49**	-0.11, NS
* $p < .05$; ** $p < .01$; NS = not significant. Regression covariates include gender, ethnicity, LEP, SPED, and (for middle school only) baseline achievement scores.		

The effect of being in an Expeditionary Learning school was strong and significant ($p < .01$) for both years of elementary school ELA, both years of elementary school math, and both years of middle school ELA. The difference between EL and non-EL schools on middle school math in 2007-08 was not statistically significant, and non-EL schools showed a significant advantage ($p < .05$) in middle school math in 2008-09. This significant advantage for non-EL schools was smaller ($p < .05$) than the six categories in which EL participation was found to be a significant advantage ($p < .01$).

Tables 3A and 3B contextualize the effect sizes from Table 2 by providing the magnitude of the shifts in Proficiency status that comparison group students would experience if their scores were increased or decreased by the amount implied by the relevant effect size. The "Initial % Proficient" column shows the percentage of comparison students who scored at the Proficient level. The "Implied % Proficient" column shows the percentage of comparison students who would have scored at the Proficient level if the effect size of participating in an EL school had been applied to them. The "Implied EL Change" shows the number of percentage points that the comparison group would change based on the EL effect size. The final column shows the percentage of comparison group students who scored at the Non-proficient level who would have scored at the Proficient level if the effect size of participating in an EL school had been applied to them. In the two cases where the final column contains a negative number, it indicates the percentage increase in the number of students who would have scored at the Non-proficient level.

Level and Year	Effect Size (std dev)	Initial % Proficient	Implied % Proficient	Implied EL change (% points)	Implied % of students scoring Non-proficient who would shift to scoring Proficient
Elementary 2008-09	0.58**	75.2	90.7	15.5	62.6
Elementary 2007-08	0.40**	65.7	81.8	16.1	47.0
Middle 2008-09	0.38**	44.2	63.1	18.9	33.8
Middle 2007-08	0.49**	30.9	61.5	30.6	44.3

** p < .01. Regression covariates include gender, ethnicity, LEP, SPED, and (for middle school only) baseline achievement scores.

Averaging across the two years in Table 3A, applying the EL effect size would raise 55% of Elementary students and 39% of Middle School students from the Non-proficient level to the Proficient level on the English Language Arts exam.

	Effect Size (std dev)	Initial % Proficient	Implied % Proficient	Implied EL change (% points)	Implied % of students scoring Non-proficient who would shift to scoring Proficient
Elementary 2008-09	0.67**	81.6	94.8	13.2	71.6
Elementary 2007-08	0.64**	70.4	87.9	17.5	59.0
Middle 2008-09	-0.18*	50.9	41.1	-9.8	-20.0 [†]
Middle 2007-08	-0.11, NS	40.5	33.2	-7.3	-12.3 [†]

** p < .01; NS = not significant. [†]Negative numbers represent the percentage increase in the number of students who score at the Non-proficient level. Regression covariates include gender, ethnicity, LEP, SPED, and (for middle school only) baseline achievement scores.

Averaging across the two years in Table 3B, applying the EL effect size would raise 65% of elementary students from the Non-proficient level to the Proficient level on the Mathematics exam. At the middle school level, the percentage of students scoring at the Non-proficient level on the Mathematics exam would increase by 16%.

Conclusion

Participation in an Expeditionary Learning school resulted in substantial and statistically significant advantages for both years of elementary school ELA, both years of elementary school mathematics, and both years of middle school ELA. This level of positive impact should be of interest to those seeking school reform approaches that raise student scores on standardized assessments. Participation in

Rochester's EL middle school resulted in a substantial and statistically significant disadvantage in mathematics in 2008-09, a level of negative impact that should also be of interest to those seeking to raise student test scores.

Important questions remain for future research on the impact of the Expeditionary Learning model. First, what aspects of EL's model are responsible for the differences observed? Do these key aspects vary based on student backgrounds, grade levels, or other characteristics? What explains the very large ELA gain of 31 percentage points in 2007-08 compared to the large but notably lower gain of 19 percentage points the following year? Across those two years, were there identifiable changes in the ELA curriculum, staffing, student body, or other key factors that might explain these findings? And what explains the negative EL impact for 2008-09 middle school math? One hypothesis is that EL's inquiry-based math approach, which focuses on depth, understanding, and application of mathematical concepts, provides less of the breadth and repeated skill practice that are assessed on state exams.

The method used in this study was strong, controlling for many key variables known to influence student outcomes (i.e., gender, ethnicity, free/reduced price lunch status, special education status, English language learner status, and baseline test scores). However, the study was not able to account for factors that might influence which students enroll in a public school of choice such as the Expeditionary Learning schools in Rochester. It is typically assumed that such enrollment reflects higher levels of family investment in education, which could raise student test scores. However, some students and families are also attracted to schools of choice because they have not performed well in traditional schools, which could be predictive of lower test scores. Such factors are generally too numerous and complex to measure accurately, but a randomized or "lottery" study equalizes unmeasured factors across groups and would therefore permit even more accurate assessment of EL's effectiveness. Although the Rochester school district was unable to provide lottery data for this study, they do hold a lottery for their schools of choice, which offers strong opportunities for future research.

Finally, the two Expeditionary Learning schools in this study are part of a network of more than 160 EL schools across the country. Conducting similar analyses on a broader range of EL schools, and taking into account the extent of implementation of the EL model, are important and promising directions for future research.

Reference

Albert, D. (2010). Expeditionary Learning participation in the City School District of Rochester. Cambridge, MA: EduConsultant.