2011

Effects of the National Institute for School Leadership's Executive Development Program on School Performance in Pennsylvania: 2006-2010 Pilot Cohort Results

John A. Nunnery, Ed.D.

Cherng-Jyh Yen, Ph.D.

Old Dominion University

Steven M. Ross, Ph.D.

Johns Hopkins University

March, 2011

OLD DOMINION UNIVERSITY
DARDEN COLLEGE OF EDUCATION
THE CENTER FOR EDUCATIONAL PARTNERSHIPS

Contents

EXECUTIVE SUMMARY

The National Institute of School Leadership's (NISL's) Executive Development Program (EDP) was established to provide professional development to school leaders to drive their schools to high performance. The program emphasizes the role of principals as strategic thinkers, instructional leaders, and creators of a just, fair, and caring culture in which all students meet high standards. Its primary goal is to ensure that the participating school leaders have the knowledge, skills, and tools to effectively set direction for teachers, support their staffs, and design an efficient organization. The curriculum is organized into four courses: *World-Class Schooling* (Principal as a Strategic Thinker and School Designer, Standards-Based Instruction), *Teaching and Learning*, *Developing Capacity and Commitment*, and *Driving for Results*. Training sessions are designed to be highly interactive through the use of simulations and assignment of "pre-work" and "homework" to participants.

This study examined the impact of EDP on student achievement in Pennsylvania schools from 2006-2010. It updates and extends a prior evaluation (Nunnery, Ross, & Yen, 2010a) study of this same cohort from 2006-2009, which found that elementary, middle, and high schools served by EDP principals had significantly larger gains in the percentages of students achieving proficiency in reading and mathematics.

Research Questions

The research questions addressed by the present study were:

1. How do the trends in school level performance in reading and English/Language Arts (ELA) differ between schools served by EDP-trained principals and matched comparison schools overall, and at the elementary, middle, and high school levels?

2. How do the trends in school level performance in mathematics differ between schools served by EDP-trained principals and matched comparison schools overall, and at the elementary, middle, and high school levels?

Summary Method

Participation by principals in the present EDP sample occurred through a multi-step process. Initially, the State Education Agency (SEA) publicized NISL to district superintendents. The latter, in turn, identified principals and assistant principals in their districts and encouraged them to apply. Actual applicants were selected by regional coordinators using an evaluation rubric. During the first two years of the program, participation was limited to principals or assistant principals in their first three years on the job. The highest weighting on the rubric evaluation was given to candidates from the lowest performing schools. All principals in the present sample completed EDP in 2008 or 2009.

On the basis of statistical analyses of school characteristics (e.g., percentage of students who are economically disadvantaged, special needs, LEP) and prior achievement in mathematics and reading/English Language Arts (ELA), 36 NISL elementary schools were individually matched to a highly similar comparison school in the same school district. Due to lacking appropriate within-district matches, an additional 32 elementary schools, all 19 NISL middle schools, and all 14 NISL high schools were matched to out-of-district comparison schools. Although this study was based on a rigorous *ex post facto* design with a carefully matched comparison group, selection effects cannot be ruled out given that the findings are not based on a fully randomized experiment.

Summary Results

Statistical analyses were conducted to compare the relationship between principals' participation in EDP and school-level performance in ELA and mathematics from 2006 (baseline)

through 2010[1].   In both subjects, schools led by EDP principals improved at a greater rate than

matched comparison schools. Specifically, all NISL schools improved in ELA performance at an

average rate of about +0.5% greater than the comparison schools *each year*. In other words, by the

end of the four years post-baseline, the typical NISL school had 2.16% more students achieving

proficiency in than otherwise would be expected. This translates to about 1,225 more students

between 2008 and 2010 who achieved reading/ELA proficiency in the NISL study schools that

would have been expected otherwise. Likewise, in mathematics, NISL schools also improved at an

average rate of about +0.5% proficient each year above the rate of improvement in comparison

schools. By the end of the four years post-baseline, the typical NISL school had about 1.92% more

students achieving proficiency than otherwise would be expected, which translates to a projected

increase of 1,089 students across all NISL schools included in the study.

EDP effects were smaller (though still statistically significant) in elementary schools than in

middle and high schools. Specifically, average annual increases in ELA proficiency over comparison

schools were 0.18%, 1.04%, and .81% for elementary, middle, and high schools, respectively. For

mathematics, the respective average annual increases were 0.17%, 0.51%, and 2.37%. Cumulative

four-year gains (2007-2010) in the percentage of students achieving proficiency in reading/English

language arts and mathematics are presented in Figure 1.

---

[1] Analyses consisted of factorial analysis of variance and two-level hierarchical linear modeling.
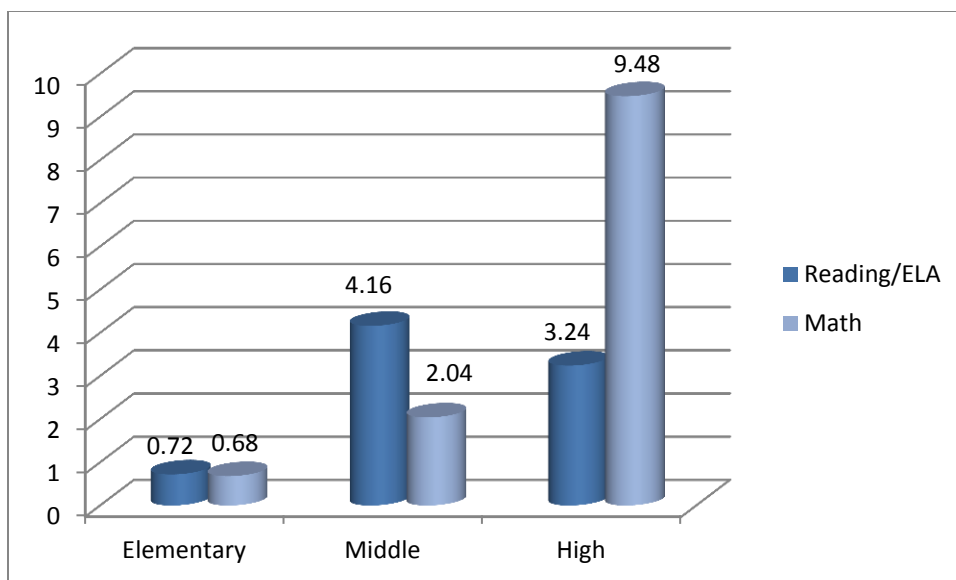
Figure 1. *Cumulative Four-Year Gains in the Percentage of Students Achieving Proficiency in Reading/English Language Arts and Math: Pennsylvania Executive Development Program Pilot Cohort 2007-2010.* Note. All gains statistically significant at *p* < .001.

Summary Findings

The results of this study show significantly larger gains by NISL schools relative to comparison schools in both reading/ELA and mathematics. In both subjects, NISL schools had lower percentages of students achieving proficiency in 2006 (the difference favoring comparison schools was statistically significant in mathematics). However, following principals' participation in EDP, the NSL schools were significantly superior in both subject areas in 2008, 2009, and 2010. Across all grade levels, the percentages of students achieving proficiency in NISL relative to comparison schools were +0.54 and +0.48 higher each year in reading/ELA and mathematics, respectively.

Compared to elementary schools, much stronger EDP effects were found in middle schools and high schools. This finding is particularly noteworthy, as there is little systematic evidence that any of the many high school reforms attempted to date have had a positive effect on student achievement (Fleischman & Heppen, 2009). In particular, very large gains in mathematics

performance were observed in EDP high schools. In 2006 (the baseline year), EDP high schools had only a +0.6% advantage in terms of percentage of students achieving proficiency in mathematics (46.3% versus 45.7%), whereas by 2010 EDP high schools had a 12% advantage (62.4% versus 50.5%). The greatest acceleration in mathematics performance in EDP schools occurred between 2009 and 2010, the first year following completion of the EDP program for all pilot cohort principals.

The results of this study represent highly promising evidence that the NISL Executive Development Program for school leaders results in statistically significant, substantial, and sustained improvements in student performance in reading and mathematics, particularly in the challenging context of secondary schools. This is particularly noteworthy given that the program is highly cost-effective, with current median participant costs of about $4,000. In the context of the current study, this amounts to only about $117 per additional student achieving proficiency: if the observed trends continued for one additional year, that cost would drop to about $69 per additional student achieving proficiency in either reading or mathematics.

Introduction

This study examined the impact of the National Institute of School Leadership's (NISL) Executive Development Program (EDP) on student achievement in Pennsylvania schools from 2006-2010. The current study used an *ex post facto* design in which schools served by the first Pennsylvania cohort of EDP principal participants were matched individually to control schools with similar school performance and demographic profiles in the pre-program year of 2006. This report is an update of the Nunnery, Ross, & Yen (2010a) study of this same cohort from 2006-2009, which found that elementary, middle, and high schools served by principals who had completed the EDP program had significantly larger gains in the percentages of students achieving proficiency in reading and mathematics over that time span. This study extends and improves upon the Nunnery, Ross, & Yen (2010a) study by adding an additional year of results, by utilizing a more sophisticated statistical model to control for non-linear secular trend in school performance trajectories, and by implementing school-level weights based on the number of students tested each year in order to get more precise error estimates for statistical hypothesis testing and effect size estimation.

The NISL Executive Development Program

The National Institute of School Leadership's (NISL's) Executive Development Program was established to train school leaders to drive their schools to high performance. The program emphasizes the role of principals as strategic thinkers, instructional leaders, and creators of a just, fair, and caring culture in which all students meet high standards. Its primary goal is to ensure that the participating school leaders have the knowledge, skills, and tools to effectively set direction for teachers, support their staffs, and design an efficient organization. The curriculum, which was

designed by experts on leadership training across a number of fields, was developed with an $11 million investment and five years of research and piloting.

Professional development goals of the EDP program are to provide high-quality instruction (both online and face-to-face), an advanced research-based curriculum, and an interactive approach to learning that includes simulations, case studies, school evaluations, and online activities. Key expectancies for NISL-trained principals include:

- Formulating a clear vision to inspire others in the school communities,

- Implementing fully-aligned, standards-based instructional systems,

- Building effective instructional programs in the core academic subjects, particularly math, language arts and science,

- Using data to produce continuous improvements in instruction and student achievement,

- Providing effective training programs to build a professional learning community for school faculty and staff, and

- Creating integrated school improvement plans that reflect strategic and systemic thinking.

The curriculum is organized into four courses: *World-Class Schooling* (Principal as a Strategic Thinker and School Designer, Standards-Based Instruction), *Teaching and Learning*, *Developing Capacity and Commitment*, and *Driving for Results*. Professional development sessions are designed to be highly interactive through the use of simulations and assignment of "pre-work" and "homework" to participants.

Methods

This study is an extension of the Nunnery, Ross, & Yen 2010 study of EDP effects in Pennsylvania. It extends the study by incorporating an additional year of school performance data (2010). It is based on a carefully matched comparison-group ex post facto design in which schools served by principals participating in the program were individually matched to control

schools with similar school performance and demographic profiles in the baseline (pre-program) year of 2006. It also improves upon the prior study by utilizing more advanced hierarchical linear modeling techniques to get more precise estimates of program impacts. Although this study employed a rigorous *ex post facto* design with a carefully matched comparison sample, there are potential confounds, or alternative explanations, for the observed results that arise because the findings are not based on a fully randomized experiment. Thus, it is possible that selection effects may account for some of the observed differences between NISL and comparison schools. Also, as is the case with any statistical model, it also is possible that unmeasured variables might account for some of the observed differences reported in this study.

Research Questions

1. *How do the trends in school level performance in reading and English/Language Arts (ELA) differ between schools served by NISL-trained principals and matched comparison schools overall, and at the elementary, middle, and high school levels?*

2. *How do the trends in school level performance in mathematics differ between schools served by NISL-trained principals and matched comparison schools overall, and at the elementary, middle, and high school levels?*

Sample

Program participant selection process. Participation by principals occurred through a multi-step process. Initially, the State Education Agency (SEA) publicized NISL to district superintendents. The latter, in turn, identified principals and assistant principals in their districts and encouraged them to apply. The potential applicants were characterized by the SEA and superintendents as mixed in leadership potential (experiences, accomplishments, and skills), with some demonstrating strong promise and others regarded as needing professional development support to improve instructional leadership skills. Actual applicants were selected by regional coordinators using an evaluation rubric (see Appendix B). During the first two years of the program, participation was limited to principals or assistant

principals in their first three years on the job. The highest weighting on the rubric evaluation was given to candidates from the lowest performing schools. Starting on January 1, 2008, a new state (Act 45 of 2007) policy requiring a principals' induction program went into effect. Based on the law's requirement that all school and system leaders meet approved continuing education requirements, all applicants henceforth were accepted to the program on a first-come-first-served basis. This change in recruitment practices, however, did not affect the present sample of NISL participants.

Study sample inclusion. Data from all Pennsylvania elementary schools with complete test score data from 2005-2006 through 2008-2009 were initially considered for inclusion in the analyses. There were a total of 70 NISL elementary schools, 19 NISL middle schools, and 12 NISL high schools. As explained below, 36 of the NISL elementary schools were included in a within-district matched samples analysis, and 32 were included in a separate set of analyses based on an out-of-district matched comparison sample. In the middle school sample, 19 of the NISL principals completed the NISL program in 2009, and 17 completed it in 2008.   In the middle school sample, 7 of the NISL principals completed the NISL program in 2009, and 12 completed it in 2008. In the high school sample, 6 of the NISL principals completed the NISL program in 2009, and 8 completed it in 2008. Thus, roughly half of NISL principals at any given grade level started the program in 2007 and completed in 2008, while half started in 2008 and 2009. The elementary within-district matched sample included an average of 4,565 students in comparison schools each year, and 5,898 students in NISL schools. Corresponding comparison and NISL average annual student sample sizes were 5,233 and 4,847 for elementary out-of-district matched samples, 8,916 and 7,498 for middle school out-of-district matched samples, and 3,017 and 2,552 for high school out-of-district matched samples.

Elementary school matching procedure. A principal components analysis was conducted using 2006 school performance index values in mathematics and reading, and the proportions of students who were economically disadvantaged, received special education services, or classified as having limited English proficiency (LEP). Results of the principal components analyses were used to construct a

regression-based factor score to use to identify matched pairs of schools for the analyses. Each NISL school was individually matched to a comparison school in the same school district that had the closest factor score. Matches were considered suitable only if the factor scores were within +/-0.25 standard deviation units. The matching process yielded a final sample for analysis of 36 NISL and 36 comparison schools at the elementary level that had matches within the same school district. Of the remaining 34 NISL elementary schools, two did not have test score data at all four time points. The remaining 32 were matched out-of-district, and these analyses were conducted separately. As shown in Table 1, this matching process yielded very closely matched samples. In 2006, the NISL elementary schools with within-district matches had a slightly lower percentage of economically disadvantaged students (15.8% versus 17.8%), slightly higher percentages of students with limited English proficiency (18.4% versus 14.8%), and a lower percentage of students who were proficient in math (78.0% versus 80.5%). For the out-of-district matches, all matching variables were within 0.1%, except percentage of limited English proficient students served (1.1% in comparison sites versus 0.6% in NISL sites).

Secondary school matching procedure. It was not possible to individually match middle and high school NISL schools to a comparison school within the same school district—in many cases, the NISL secondary school was the only school at that level within the district. At the middle and high school levels, an out-of-district match was made to each NISL school by matching the NISL school to a comparison school with the closest factor score. There were 19 NISL middle schools and 14 NISL high schools. As shown in Table 1, the matching process led to relatively well-matched samples for the middle school analyses, although the NISL sample was slightly more economically disadvantaged (27.6% versus 24.6%), and had somewhat lower initial scores in math (68.2% versus 72.0% proficient) and reading (69.9% versus 73.2%). Likewise, the high school samples were relatively well-matched, with comparison schools having somewhat higher percentages of economically disadvantaged students (27.9% versus 23.9%), higher proficiency rates in mathematics (47.8% versus 45.0%), and higher proficiency rates in Reading/ELA (63.8% versus 59.6%). NISL schools served higher percentages of special

education students (14.3% versus 10.9%). Independent samples *t*-tests conducted on school-level data showed that none of the NISL/comparison differences were statistically significant at *p* <.05. As shown in Table 1, proficiency levels at the high school level in 2006 were much lower than those observed at elementary or middle school, particularly in mathematics, where both NISL and comparison schools had less than 50% proficiency.

Table 1

Selected 2006 Characteristics of NISL and Comparison Schools

School Type	Economically Disadvantaged %	IEP %	LEP %	Math % Proficient	Reading % Proficient
Elementary[1]					
Comparison	17.8	6.41	14.8	80.49	69.20
NISL	15.8	6.61	18.4	78.04	68.10
Elementary[2]					
Comparison	17.0	7.6	1.1	79.0	68.7
NISL	16.9	7.5	0.6	78.9	68.5
Middle[2]					
Comparison	24.6	12.1	0.5	72.0	73.2
NISL	27.6	13.5	0.5	68.2	69.9
High[2]					
Comparison	27.9	10.9	0.2	47.8	63.8
NISL	23.9	14.3	0.2	45.0	59.6

[1]*N* = 36 NISL and comparison schools for elementary within-district matches. [2]*N* = 32 of each type for elementary out-of-district matches. *N* = 19 of each for middle schools. *N* = 14 of each for high schools.

Note. No NISL/comparison differences on matching variables were statistically significant.

Measures

The proportions of students in tested grade levels (3-8 and 11) who were economically disadvantaged, received special education services as evidenced by the existence of an individualized education plan (IEP), or who were classified as having limited English proficiency (LEP) were computed for each school to utilize in the school matching procedure, and to provide for descriptive comparisons between EDP and comparison schools. Summary indices of school performance were constructed for both reading/ELA and mathematics by computing the proportion of students at all tested grade levels (grades 3 to 8 and grade 11) that scored proficient or higher on the Pennsylvania State Assessment. Note that for high schools, test scores were available only for eleventh grade.

Analyses

Factorial analyses of variance. The first stage of the analysis was to estimate 2 (comparison versus EDP) X 5 (Years) factorial analyses of variance for both reading and mathematics performance indices to provide a straightforward, unadjusted perspective on the actual results obtained in comparison schools and schools led by EDP program completers. Preliminary diagnostics indicated no *program X school level* interactions, so these analyses were performed on the total data set including schools across all grade levels. Scheffe's test was employed as the post hoc multiple comparisons criterion to account for unequal school-level numbers of students. Where the *program X year* interaction effect was statistically significant, pairwise comparisons between comparison and EDP means were conducted within years using Holme's sequential Bonferroni technique to control for experimentwise alpha inflation.

Hierarchical linear models. Two-level multilevel models were estimated to determine whether there were statistically significant relationships between EDP participation status at level 2, and the linear and quadratic components of longitudinal school-level achievement trajectories. Achievement trajectories were indicated as the proportion of students scoring at or

above proficiency in reading or English/language arts and mathematics. Indicators were available for 2006 through 2010.

Level-1 model. The level one (outcome and year within schools) was specified as:

$$Y_{ti} = \pi_{0i} + \pi_{1i}(\text{Year-2008}) + \pi_{2i}(\text{Year-2008})^2 + \varepsilon_{ti}; \text{ where}$$

Y_{ti} is the proportion of students achieving at or above proficiency in year t at school i;

π_{0i} is the proportion of students achieving at or above proficiency at Year-2008 = 0;

π_{1i} is the linear coefficient indicating magnitude and direction change in the outcome per year in school i;

π_{2i} is the curvilinear rate of acceleration in the achievement trajectory curve or school i; and

ε_{ti} is a random level-1 error term assumed normally distributed with variance = 1.0.

The numbers of students tested in each subject each year were employed as level-1 weights.

Level 2 models. The level-2 (between schools) models were specified as:

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(\text{Program})_i + r_{1i}$$

$$\pi_{2i} = \beta_{20} + + r_{2i}$$

where

π_{0i} is the proportion of students achieving at or above proficiency at Year-2008 = 0 in year i;

β_{00} is the grand mean proportion of students achieving at or above proficiency at Year-2008 = 0;

r_{0i} is a random level 2 error term representing variation in mean proportions across years;

π_{1i} is the linear trend in performance in school i;

β_{10} is the grand mean linear trend in performance across schools;

β_{11} is the coefficient associated with the interaction of EDP program status and the linear trend in school performance;

r_{1i} is a random level 2 error term representing variation in linear slopes across schools;

π_{2i} is the quadratic trend in performance in school i;

β_{20} is the grand mean quadratic trend in performance across schools;

r_{2i} is a random level 2 error term representing variation in quadratic trends across schools.

Results

Reading: Factorial Analyses of Variance

Across all grade levels, 2 (Program) X 5 (Year) factorial analysis of variance revealed statistically significant effects for program ($F_{1,227566}$ = 1720.3, $p < .001$), year ($F_{1,227566}$ = 387.1, $p < .001$), and a statistically significant *program X year* interaction effect ($F_{4,227566}$ = 112.4, $p < .001$). Due to the interaction effect, a graphic profile was constructed and pairwise comparisons within years were conducted to determine the precise nature of the interaction (see Figure 2).
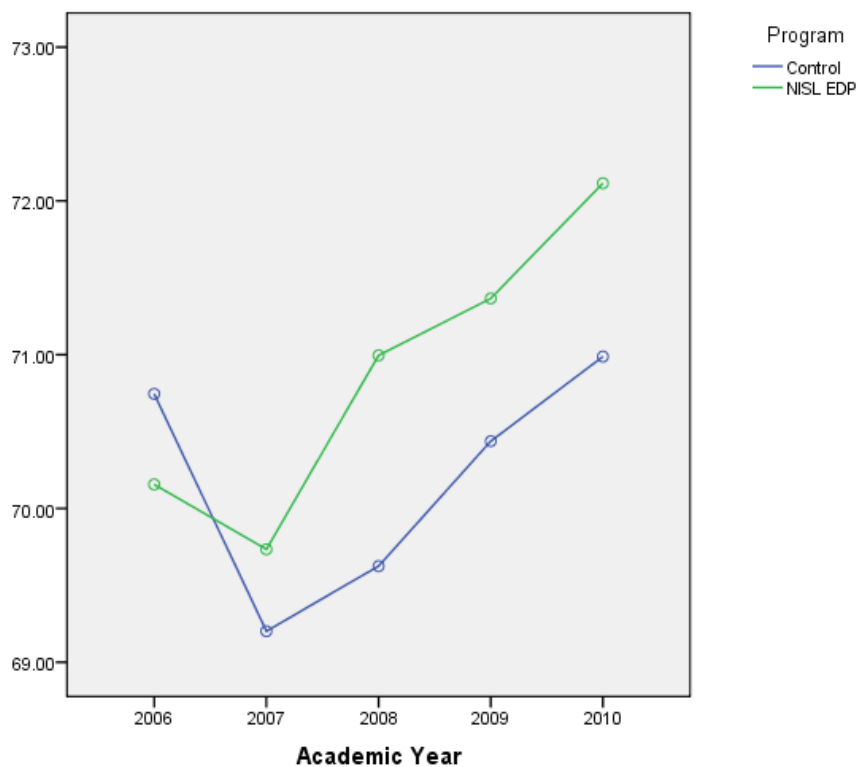

Figure 2. *Observed mean proportion of students achieving proficiency in reading/ English Language Arts by year: Control schools versus NISL EDP schools*. Note: Results include EDP schools that started in 2007 and in 2008.

Pairwise comparisons indicated no statistically significant differences between EDP and comparison schools in 2006 or 2007, whereas NISL schools had statistically significantly higher mean proportions of students reaching proficiency in 2007 (70.1% vs. 67.3%; $F_{1,43846} = 405.4$, $p < .001$), 2008 (70.9% vs. 68.6%; $F_{1,52555} = 345.1$, $p < .001$), 2009 (72.4% vs. 67.8%; $F_{1,33216} = 893.8$, $p < .001$), and 2010 (73.3% vs. 70.9%; $F_{1,53844} = 464.2$, $p < .001$).

Reading: Hierarchical Linear Model Results

The hierarchical linear model improves upon the analysis of variance model by accounting for the non-linear, or "curvy," aspects of school performance trajectories to better isolate program effects. Table 2 provides estimates of the fixed effects estimated in the hierarchical linear model for reading. A statistically significant fixed effect was observed for year ($\beta = +0.78$, $t = 78.72$, $df = 228016$, $p < .001$), indicating that, on average, all schools improved the percent proficient in reading/ELA by about 8/10ths of one percent each year. The year-squared (quadratic) effect was also statistically significant ($\beta = +0.40$, $t = 68.40$, $df = 228016$, $p < .001$), which is illustrated by the S-shaped NISL curve and U-shaped comparison school curve from 2006-2010 (see Figure 3). Finally, the *program X year* was statistically significant, ($\beta = +0.54$, $t = 38.59$, $df = 228016$, $p < .001$), which indicates that NISL EDP schools improved at an average rate of about +0.5% greater than the comparison schools *each year*. In other words, by the end of the four years post-baseline, the typical NISL EDP school had about 2.16% more students achieving proficiency in reading/English Language Arts than otherwise would be expected. This translates to about 1,225 more students achieving reading/ELA proficiency from 2008-2010 in the NISL study schools than would have been expected otherwise.

Table 2

Estimates of Fixed Effects: Reading/ELA

Parameter	β	s.e.	*df*	*t*	*p*
Intercept	69.09	1.41	200.0	48.95	.000
Program	0.56	1.99	199.9	-.27	.781
Year	.78	.009	228016.5	78.72	.000
Year-squared	.40	.006	228016.3	68.40	.000
Program X Year	0.54	.014	228016.5	38.59	.000

N = 228,221.

Cases weighted by Number Scored Reading
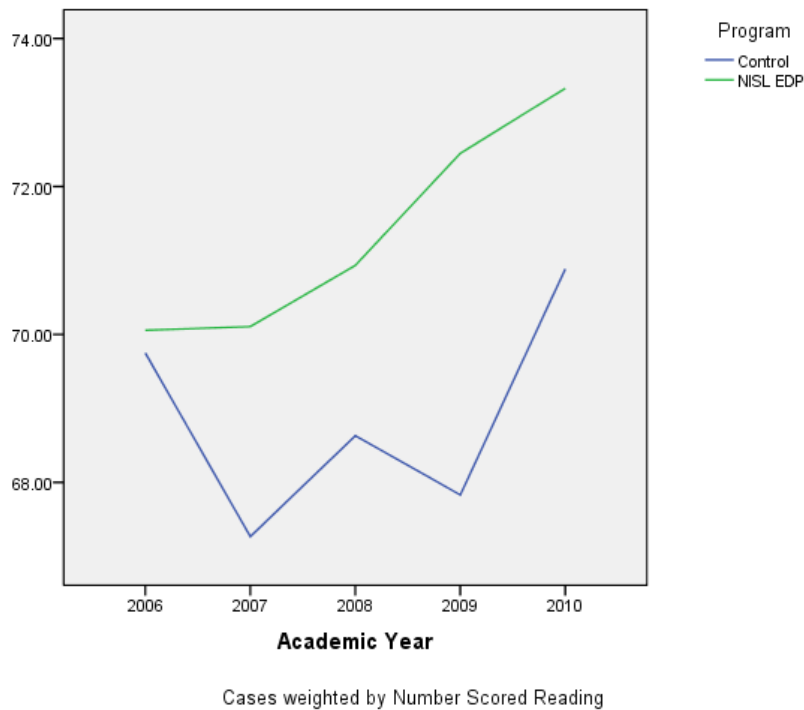
Figure 3. *Hierarchical linear modeling results in reading weighted by number of students at*

Level 1.

Mathematics: Factorial Analysis of Variance

A 2 (Program) X 5 (Year) factorial analysis of variance revealed statistically significant effects for program ($F_{1,228211}$ = 3839.7, p < .001), year ($F_{1,228211}$ = 940.1, p <.001), and a statistically significant *program X year* interaction effect ($F_{4,228211}$ = 125.0, p <.001). Due to the interaction effect, a graphic profile was constructed and pairwise comparisons within years were conducted to determine the precise nature of the interaction (see Figure 4).
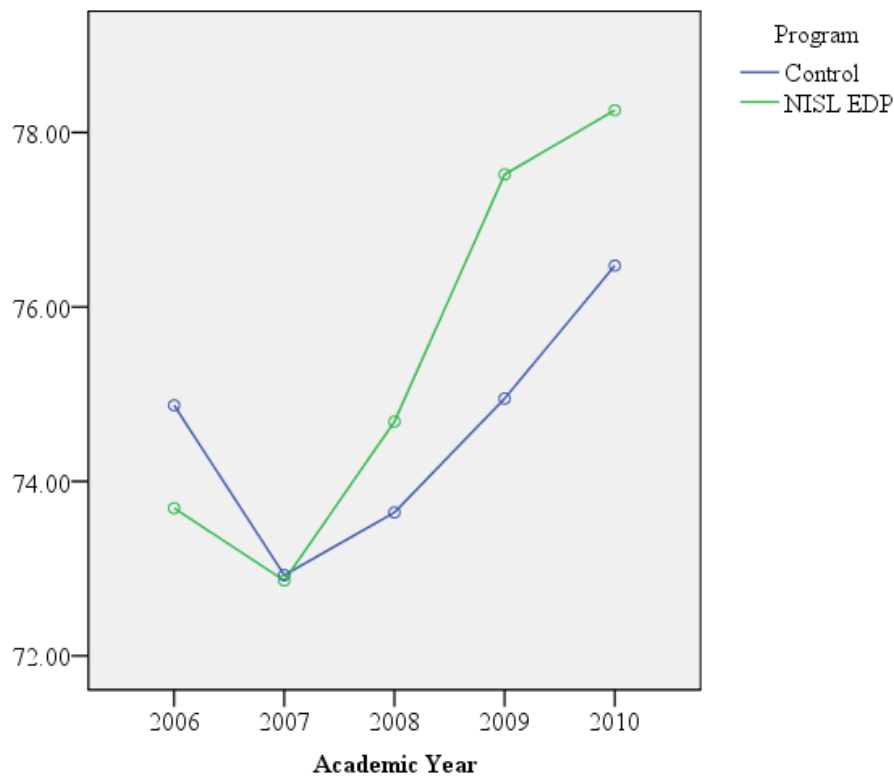

Figure 4. *Observed mean proportion of students achieving proficiency in mathematics by year: Control schools versus NISL EDP schools*. Note: EDP program implementation began in 2007 and was completed in 2009.

Pairwise comparisons indicated that comparison schools had a statistically higher mean proportion of students achieving proficiency in 2006 (74.9% versus 73.7%; $F_{1,44105} = 226.3$, p <.001) and no statistically significant difference in 2007. NISL schools had statistically significantly higher mean proportions of students reaching proficiency in mathematics in 2008 (76.2% vs. 71.7%; $F_{1,52555} = 978.7$, $p < .001$), 2009 (76.8% vs. 69.3%; $F_{1,33216} = 1598.6$, $p < .001$), and 2010 (79.6% vs. 75.9%; $F_{1,53844} = 821.5$, $p <.001$).

Mathematics: Hierarchical Linear Model Results for All Schools

Table 3 provides estimates of the fixed effects estimated in the hierarchical linear model for reading. A statistically significant fixed effect was observed for year ($\beta = +1.33$, $t = 126.39$, $df = 227371$, $p < .001$), indicating that, on average, all schools improved the percent proficient in mathematics by about 1.3% each year. The year-squared (quadratic) effect was also statistically significant ($\beta = +0.46$, $t = 74.07$, $df = 227371$, $p < .001$), which is illustrated by the accelerated growth in the NISL curve and U-shaped comparison school curve between 2006-2010 (see Figure 5). Finally, the *program X year* was statistically significant, ($\beta = +0.48$, $t = 32.83$, $df = 227371$, $p < .001$), which indicates that NISL EDP schools improved at an average rate of about +0.5% proficient each year over and above the rate of improvement in comparison schools. In other words, by the end of the four years post-baseline, the typical NISL EDP school had about 1.92% more students achieving proficiency in reading/English Language Arts than otherwise would be expected. This translates to about 1,089 more students achieving math proficiency from 2008-2010 in the NISL study schools than would have been expected otherwise.

Table 3

Estimates of Fixed Effects: Mathematics

Parameter	β	s.e.	df	t	p
Intercept	74.46	1.63	200.0	45.66	.000
Program	0.76	2.30	199.99	0.33	.744
Year	1.33	.011	227371.5	126.39	.000
Year-squared	0.46	.006	227371.3	74.07	.000
Program X Year	0.48	.015	227371.4	32.83	.000

N = 227,576.

Cases weighted by Number Scored Math
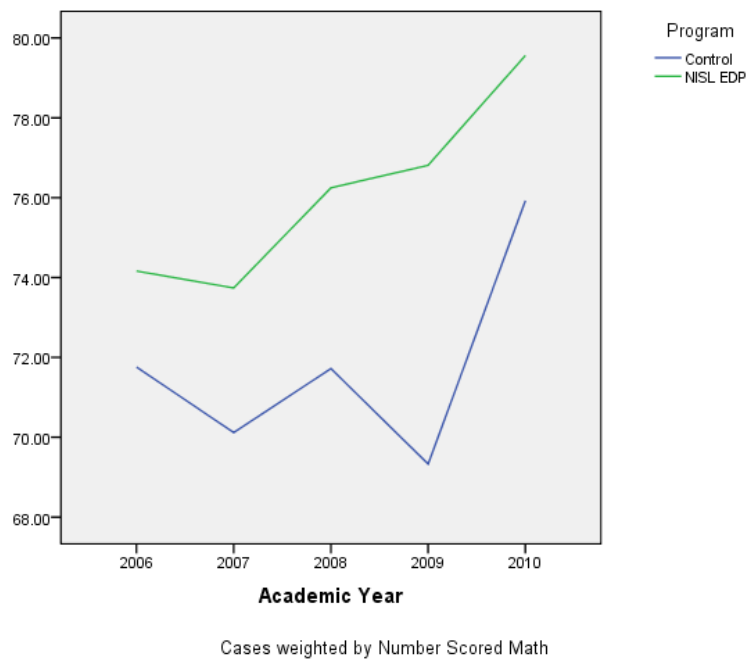
Figure 5. *Hierarchical Linear Modeling Results in Mathematics Weighted by Number of Students at Level 1.*

Hierarchical Linear Modeling Results by School Level in Reading and Mathematics

Table 4 provides tests of *program X year* interaction effects by school grade level configuration in both reading/ELA and mathematics. Statistically significant effects were observed in reading at all grade levels: elementary (β = +0.18; t = 8.48; df = 115,447; p <.001), middle (β = +1.04; t = 121.6; df = 82,659; p <.001); and high (β = +0.81; t = 21.83; df = 29,904; p <.001). Likewise, positive statistically significant program X year interaction effects were observed in mathematics at all grade levels: elementary (β = +0.17; t = 8.46; df = 115,191; p <.001), middle (β = +0.51; t = 23.55; df = 82,395; p <.001), and high (β = +2.37; t = 47.84; df = 29,779; p <.001). As shown in Figure 6, although all effects were statistically significant, much larger effects were observed for middle and high schools.
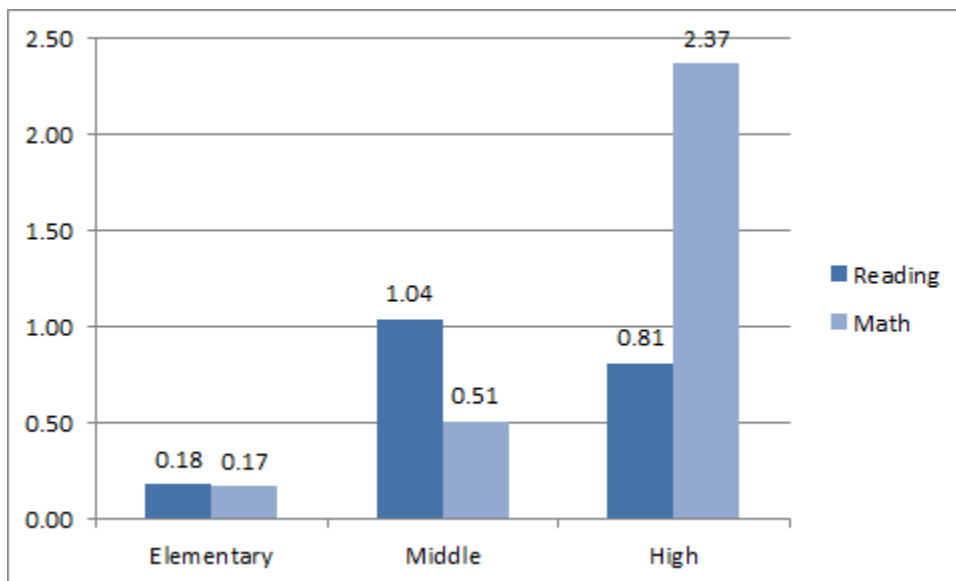

Figure 6. *Annual Acceleration in Percentage Proficient Gains in NISL versus Comparison Schools: Program X Year Interaction Effects by School Grade Level Configuration.*

Table 4. Hierarchical Linear Modeling *Program X Year* Interaction Effects by Grade Level.

Subject	Grade Level	β	*t*	*df*	*p*
Reading/ELA	Elementary	+0.18	8.48	115447	<.001
	Middle	+1.04	121.61	82659	<.001
	High	+0.81	21.83	29904	<.001
Mathematics	Elementary	+0.17	8.46	115191	<.001
	Middle	+0.51	23.55	82395	<.001
	High	+2.37	47.84	29779	<.001

Discussion

This study employed traditional analysis of variance to examine unadjusted longitudinal performance patterns in school performance in schools led by NISL EDP pilot cohort principals versus individually matched comparison schools. In both reading/ELA and mathematics, NISL schools had lower percentages of students achieving proficiency in these subject areas in 2006 (the difference favoring comparison schools was statistically significant in mathematics). As shown in Figures 5 and 6 above, NISL and comparison school performance trend lines "crossed over" between 2006 and 2008, with NISL schools having statistically significant higher levels of performance in both subject areas in 2008, 2009, and 2010. Although these simple models cannot account for non-linearity in trend lines and provide somewhat less precise estimates of annual program effects than the hierarchical linear models, the observed cross-over pattern on unadjusted outcomes is a powerful indication that implementation of the NISL program was associated with a substantial and sustained improvement in school performance during the study period from 2006-2010.

Hierarchical linear modeling analyses revealed statistically significant *program X year* interaction effects for both reading/ELA and mathematics. Across schools of all grade levels, NISL schools performance accelerated about +0.48 and +0.54 additional percent proficient each year in reading/ELA and mathematics, respectively. To put these statistics in perspective, Figure 7 illustrates the number of additional students estimated to have achieved proficiency in reading and mathematics between 2008 and 2010 in the NISL pilot cohort schools that met study inclusion criteria. These findings are consistent with previous state-wide studies conducted in Pennsylvania (Nunnery, Ross, Yen, 2010a), and Massachusetts (Nunnery, Ross, Yen,
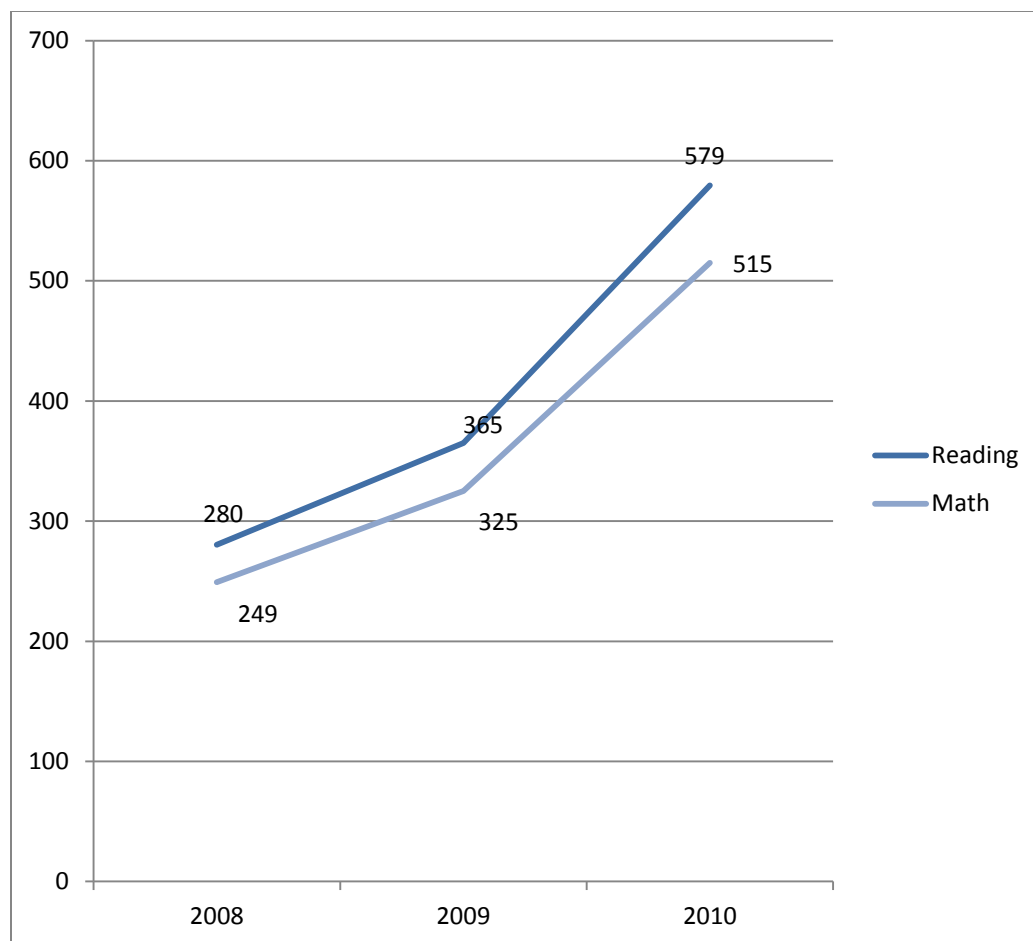
Figure 7. *Estimated Number of Additional Students Reaching Proficiency in Reading and Mathematics By Year in Pennsylvania NISL Schools: 2008-2010 Study Schools.*

2010b), which found positive associations between student achievement patterns and EDP program participation by school leaders. In Pennsylvania, NISL schools had statistically significantly higher proportions of students achieving proficiency in both reading and mathematics, with much larger proportions achieving proficiency in mathematics (see Figure 7 for Pennsylvania summary results from the previous study). As shown in Figure 8, schools led by EDP program completers achieved substantial and statistically significant gains in mathematics relative both to comparison schools ($d = +0.10$) and the Commonwealth as a whole ($d = +0.08$).
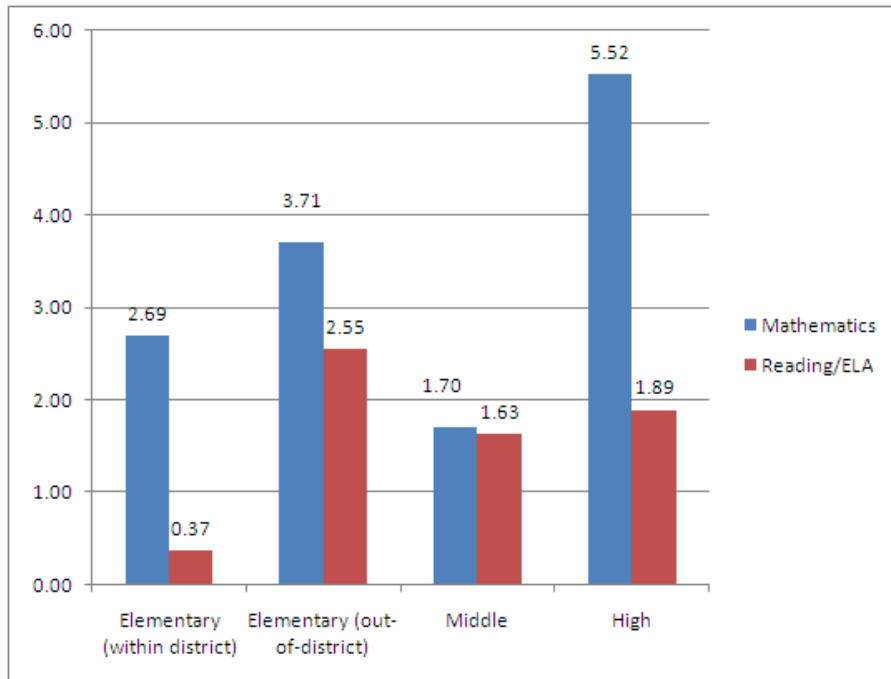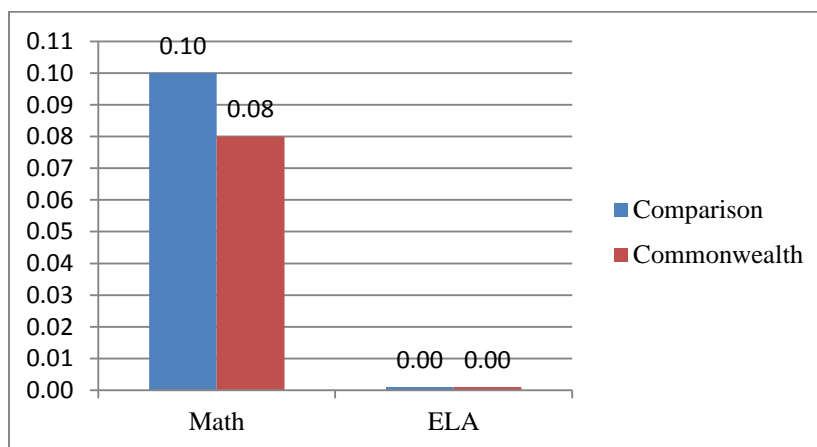
Figure 7. NISL Effects in Pennsylvania: Summary Results from Prior 2006-2009 Study of the Pennsylvania implementation of the EDP program: Differences in Percentages of Students Achieving Proficiency Relative to Matched Comparison Schools (from Nunnery, Ross, & Yen, 2010a).

Figure 8. NISL Effects in Massachusetts: Summary Results from the 2006-2009 Study of the Massachusetts implementation of the EDP program: Effect Size Estimates Expressed as Cohen's d. (from Nunnery, Ross, & Yen, 2010b).

Although statistically significant effects were associated with the NISL program at all grade levels and in both subject areas, much stronger improvements in school-level performance were observed in NISL middle schools and high schools. This finding is particularly noteworthy, as there is little systematic evidence that any of the many high school reforms attempted to date have had a positive effect on student achievement (Fleischman & Heppen, 2009). In particular, very large gains in mathematics performance were observed in NISL high schools. As shown in Figure 9, NISL high schools had only a +0.6% advantage in terms of percentage of students achieving proficiency in mathematics in 2006 (46.3% versus 45.7%), whereas by 2010 NISL high schools had a 12% advantage (62.4% versus 50.5%). Larger gains in high school mathematics for both NISL and comparison schools might have been expected given the low levels of initial performance in 2006 due to regression to the mean, but the difference in comparative gains is striking.
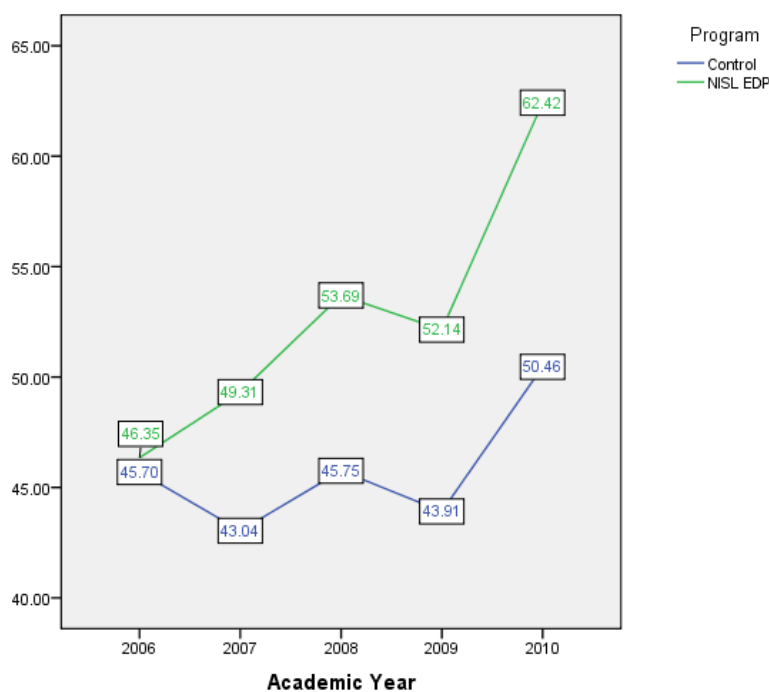

Figure 9. *Percentage Scoring Proficient in Mathematics from Pre-implementation Baseline (2006) through 2010 by Program: High Schools in Pennsylvania EDP Pilot Cohort Study.*

The greatest acceleration in math performance in NISL schools occurred between 2009 and 2010, the first year following completion of the EDP program for all pilot cohort principals. Given the staggered implementation (about half completing in 2008 and half completing in 2009), and the expectation of somewhat more modest effects during the two-year professional development sequence. NISL principals are expected to begin implementing strategies *during* the sequence, but full program effects would not be expected until post completion. The greater gains observed post-completion for all pilot cohort participants lends support to the notion that the effects are attributable to the program, although lack of randomization militates against a straightforward causal interpretation of the findings.

The results of this study represent highly promising evidence that the NISL Executive Development Program for school leaders may result in statistically significant, substantial, and sustained improvements in student performance in reading and mathematics, particularly in the challenging context of secondary schools. This is particularly noteworthy given that the program is highly cost-effective, with current median participant costs of about $4,000. In the context of the current study, this amounts to about $117 per additional student achieving proficiency already observed: if the observed trends continued for one additional year, that cost would drop to about $69 per additional student achieving proficiency in either reading or mathematics. Prior research has also indicated that the EDP program can be implemented with high fidelity in a cost-effective manner (Meristem Group, 2009).

References

Fleischman, S., & Heppen, J. (2009). Improving low-performing high schools: Searching for evidence of promise. *Future of Children*, *19*, 105-133.

Meristem Group (2009). *National Institute for School Leadership (NISL): Massachusetts program implementation 2005-2008*. Boston, MA: Meristem Group.

Nunnery, J.A., Ross, S.M., & Yen, C. (2010a). *The effect of the National Institute for School Leadership's Executive Development Program on school performance trends in Pennsylvania*. Norfolk, VA: The Center for Educational Partnerships at Old Dominion University.

Nunnery, J.A., Ross, S.M., & Yen, C. (2010b). *An examination of the effects of a pilot of the National Institute for School Leadership's Executive Development Program on school performance trends in Massachusetts*. Norfolk, VA: The Center for Educational Partnerships at Old Dominion University.