

III. Evaluation Findings

The following sections present evaluation findings pertinent to each local project benchmark for Year Four (2009-2010) and, for student objectives, over the full course of program implementation (2007-2010).

A. Year Four (2009-2010) Findings

In this section, findings related to Year Four (2009-2010) of the HSV initiative are presented by individual goal and benchmark.

Goal 1: Students will increase their oral and written communication skills and will deepen their appreciation for the arts.

Benchmark 1.1: In each year of project implementation, treatment EP students' gains in literacy achievement will exceed those of their similarly situated peers in control schools, as measured by their gains on the NYS ELA and ISAT assessments.

As described above, repeated measures ANOVAs were conducted to assess change in EP students' NYS ELA and ISAT Reading scale scores from 2009 (Year Three) to 2010 (Year Four) and to test whether patterns of change differed between students in the treatment and the control groups.

New York City

Table 14
EP Achievement Test Results – NYS ELA
Means and Standard Deviations by Group – NYC (2009-2010)

		NYC ^a	
		Treatment (N=204)	Control (N=199)
Spring 2009	Mean	655.60	653.50
	Standard Deviation	14.742	15.661
Spring 2010	Mean	649.05	641.99
	Standard Deviation	646.00	16.464

^a Repeated measures ANOVAs include students for whom data were available at all time points included in the analysis.

Table 15
EP Achievement Test Results – NYS ELA
Results of Repeated Measures ANOVA – NYC (2009-2010)

Source	NYC		
	F	p	Cohen's d
Time (Pre, Post)	159.974	<.001*	1.34*
Group (Treatment, Control)	8.838	.003*	0.31
Time*Group	12.041	.001*	0.37*

The tables above display 2009 and 2010 mean scale scores on the NYS ELA for EP students in the treatment and control groups in NYC as well as results of the repeated measures ANOVA. While mean

NYS ELA scale scores declined significantly over time from 2009 to 2010 for both groups, **results of the repeated measures ANOVA revealed that the mean scale score for EP students in the control group declined significantly more from Year Three to Year Four than that of the treatment group (see Table 15).** Calculated effect sizes provided evidence that this difference was educationally meaningful as well.

Benchmark 1.5: *Treatment EP and ELL students will demonstrate improvement in the quality of their writing and performing skills in each project year, as measured by an analysis of pre/post gains on locally developed authentic assessment rubrics.*

Student Writing Samples

Student writing samples collected from treatment schools in NYC and Chicago in fall 2009 and spring 2010 were blinded by the evaluation team (i.e., all identifiers for students' name, school, city, and pre/post date were hidden in order to remove any potential bias before scoring) and scored by a non-participating teaching artist in summer 2010. Matched pre- and post-assessments were scored for 72 NYC students and 54 Chicago students and paired *t*-tests were conducted on total pre-post scores for the matched samples within each city. Means and *t*-statistics are presented in Table 32. **Results of the paired *t*-tests indicated that, in Chicago, treatment students' overall writing scores increased significantly from pre-assessment (fall 2009) to post-assessment (spring 2010). In NYC, however, there was no significant change from pre- to post-assessment.**

Table 32
Student Writing Samples – Paired T-Test Results
Year Four (2009-2010)

	Fall 2009	Spring 2010	Test Statistics
NYC (<i>N</i> =72)	9.47 (<i>SD</i> = 2.08)	9.29 (<i>SD</i> = 2.85)	$t_{71} = 0.560$ $p = .577$
Chicago (<i>N</i> =54)	8.28 (<i>SD</i> = 2.75)	9.70 (<i>SD</i> = 2.56)	$t_{53} = 2.958$ $p = .005^*$

Performance Rubrics

Student performance rubrics were completed for each treatment student by teaching artists in the participating treatment schools in spring of each implementation year.¹ Matched pre- and post-assessments were scored for 137 NYC students and 31 Chicago students and paired samples *t*-tests were conducted on pre/post scores for all three sections of the rubric for the matched samples within each city. Means and *t*-statistics are presented in Table 33. **Results of the paired samples *t*-tests indicated that, in NYC, treatment students made significant gains from spring 2009 (end of Year Three) to spring 2010 (end of Year Four) on the performance section of the rubric but that scores on the Interpretation and Poem sections did not change significantly over time. In Chicago, treatment students made significant gains from spring 2009 (end of Year Three) to spring 2010 (end of Year Four) on the Performance and Interpretation sections of the rubric, but scores on the Poem section did not change significantly over time.**

Table 33

¹ Performance Rubrics were not completed for NYC HSV students in Year Two, spring 2008.

**Performance Rubric Section Scores by City –
Year Three to Year Four Change - Paired Samples *t*-Test Results**

		NYC (N=137)			Chicago (N=31)		
		Year Three (Spring 2009)	Year Four (Spring 2010)	Test Statistics	Year Three (Spring 2009)	Year Four (Spring 2010)	Test Statistics
Performance	Mean	2.44	2.56	$t_{136} = 2.110$ $p=.037^*$	1.93	2.45	$t_{30} = 5.508$ $p<.001^*$
	<i>SD</i>	<i>0.66</i>	<i>0.64</i>		<i>0.47</i>	<i>0.42</i>	
Interpretation	Mean	2.27	2.37	$t_{136} = 1.611$ $p=.109$	1.44	2.16	$t_{30} = 8.038$ $p<.001^*$
	<i>SD</i>	<i>0.73</i>	<i>0.73</i>		<i>0.35</i>	<i>0.39</i>	
Poem	Mean	2.52	2.57	$t_{136} = 1.014$ $p=.313$	2.34	2.36	$t_{30} = 0.242$ $p=.810$
	<i>SD</i>	<i>0.65</i>	<i>0.60</i>		<i>0.41</i>	<i>0.43</i>	

IV. Summary and Conclusions

Overall, the HSV project was successful in all our years of its implementation. This report describes results for the fourth and final year of implementation and across all implementation years, where possible. Across both NYC and Chicago, 485 students participated in the treatment in Year Four (2009-2010) and more than 300 of these students participated in the full three years of the project (2007-2010).

Program Impacts on Students – Reading Achievement:

- **Year Four:** In NYC, the NYS ELA scale scores declined significantly over time from 2009 to 2010 for both treatment and control groups; however, control group students declined significantly more from Year Three to Year Four than treatment group students. In Chicago, ISAT Reading scale scores increased significantly from 2009 to 2010 for both the treatment and control groups, however the control group posted greater gains than the treatment group.
- **Full Implementation:** In NYC, NYS ELA scale scores increased over time from baseline (2006) to the final year of implementation (2010); however, this growth was inconsistent across the years (see the Year 4 finding above). In Chicago, mean ISAT Reading scale scores have increased significantly over time from pre-implementation (2006) to post-implementation Year Four (2010) for both the treatment and control groups; however, differences in the pattern of growth were identified. While the treatment group mean scores increased consistently over time (from baseline to 2010), the control group mean scores stabilized from 2008 to 2009 and then sharply increased from 2009 to 2010.

Program Impacts on Students - Attitudes Toward Literacy:

- **Year Four:** There were very few between group differences in gains in students' attitudes towards literacy identified during Year Four (fall 2009 to spring 2010), with only one area identified for NYC students (treatment student gains exceeded control student gains in agreement with the statement: "I plan on going to college") and no areas identified for Chicago students.
- **Full Implementation:** Between group differences were identified in gains in NYC students' attitudes across the three years of program implementation [fall 2007 (pre-implementation baseline) to spring 2010 (post-implementation)] as treatment student gains exceeded control

student gains in agreement with the statements: “*I like going to school,*” “*I enjoy learning about new things,*” “*I like to write stories,*” and “*I follow school and classroom rules.*”

Program Impacts on Students - Motivation:

- **Year Four:** Changes in students’ motivation toward learning during Year Four (fall 2009 to spring 2010) varied by city. In general, NYC students reported lower motivation at the end of the school year (spring 2010) than they did at the beginning of the school year (fall 2009) on two of the four motivation subscales “Task Involvement” and “Effort.” But this rate of decline did not vary by study group. In Chicago, no significant changes over time or treatment/control differences were identified for any scale.
- **Full Implementation:** Across the three years of program implementation [fall 2007 (pre-implementation baseline) to spring 2010 (post-implementation)], several significant changes in NYC students’ motivation toward learning were found, with both treatment and control groups declining over time on: “Effort,” “Competition,” and “Social Concern.” Notably, on “Effort” and “Social Concern,” the control group declined more than the treatment group over the course of program implementation.

Program Impacts on Students - Writing and Performance:

- **Year Four:** In Chicago, treatment students’ writing improved from fall 2009 to spring 2010 during Year Four of the project; however, no similar improvements were found for NYC students. In NYC, treatment students’ improved from Year Three to Year Four on the Performance section of the Performance Rubric, while Chicago students improved on the Performance and Interpretation sections.
- **Full Implementation:** Chicago treatment students demonstrated significant gains from Year Two to Year Four on all three sections (performance, interpretation, poem) of the Performance Rubric. Results for NYC are not available.