

Picturing Writing/Image-Making Integrated Curriculum Model: Evaluation of an Alternative Art-and-Literature-Based Approach to Literacy Learning

Authors:

Susan L. Frankel, Ph.D., RMC Research Corporation

Beth Olshansky, Center for the Advancement of Art-Based Literacy, University of New Hampshire

Jiaxiu Yang, Ph.D., RMC Research Corporation

April, 2015

Introduction

The Manchester School District (MANSD) in Manchester, New Hampshire received four years of federal funding (2006-2010) through a U.S. Department of Education Arts in Education Model Development and Dissemination (AEMDD) Grant to investigate the impact of Picturing Writing: Fostering Literacy Through Art (PW) and Image-Making Within the Writing Process (IM) on students' writing, visual literacy, and reading skills. In partnership with the Center for the Advancement of Art-Based Literacy at the University of New Hampshire and RMC Research Corporation, the goals of this initiative were: 1) to design and implement an effective, cohesive school-wide PW/IM integrated curriculum standards-based elementary school model, 2) to strengthen standards-based art and language arts instruction through enhanced professional development, 3) to determine through scientific evaluation the effect of the enhanced model on students' academic performance, and 4) to disseminate research findings and the enhanced model. This report summarizes the evaluation and the effects of PW/IM on students' writing achievement, students' ability to use the visual arts as a language for communicating their ideas, and students' reading achievement.

Background

Picturing Writing: Fostering Literacy Through Art and Image-Making Within the Writing Process are two dynamic art-and-literature-based approaches to literacy learning developed by Beth Olshansky at the University of New Hampshire. These innovative, multimodal instructional models have demonstrated their effectiveness in significantly improving student academic performance in writing as well as reading, particularly among at-risk learners, through the integration of standards-based art education into the core curricula.

PW and IM have established histories of proven results in improving student achievement through comprehensive, systematic integration of standards-based art education and standards-based language arts instruction (Frankel, 2011; O'Connor, 2010; Olshansky, 2007, 2008).

In 1993, based on research findings submitted to the U.S. Department of Education's Program Effectiveness Panel of the National Diffusion Network (NDN), Image-Making won validation as an "innovative and effective literacy program" and was awarded federal funding for national dissemination. A second research study, involving 555 first- and second-grade students from NH, HI and TX, was conducted from 1997-1998. This study was designed to document the impact of consistent use of PW/IM as a yearlong intervention. Findings demonstrated significant gains in writing achievement and

visual literacy of participating students as compared to students in demographically matched comparison groups. Particularly dramatic were the documented gains in writing skills made by at-risk students (Title I and Special Education); at-risk students kept close pace with their classmates.

In 1999-2000, the impact of PW on reading skills first became apparent. In a third-grade classroom in Fowler, CA, standardized test results from the John C. Fremont School revealed that 36% of students scored one to four years below grade level in reading comprehension before implementing PW (Siverolli Reading Assessment). After eight months of consistent use of PW, 97% scored at or above grade level with 75% scoring three to four years above grade level. Seventy-five percent (75%) of these students received Free and Reduced Lunch; 50% were English Language Learners.

A three-year Comprehensive School Reform Demonstration (CSR) Grant (1999-2002) for the Main Street School (grades K-2), Exeter, NH to integrate PW/IM into language arts, science, and art curriculum provided school-wide standardized test score data over a ten-year period (before, during and following the grant). Results from California Achievement Tests, the New Hampshire state language arts assessments (NHEIAP and later NECAP), and the Gates MacGinitie Reading Comprehension Tests documented that consistent use of PW and IM as a school-wide instructional model had a significant impact on students' acquisition of reading and writing skills, particularly for at-risk learners. For example, following school-wide adoption, students achieved and maintained a 10-point increase in California Achievement Test total language arts scores; on the NH state language arts assessment, Title I students consistently scored above the state average of all students in writing; and on the Gates MacGinitie Reading Comprehension Test, Title I and Special Education students scored above the national average of all students in reading. NH standardized test data over the next decade (NHEIAP and NECAP) continued to document a strong performance by Exeter's at-risk students in reading and writing as compared the performance of their at-risk peers across the state. Additionally, defying national trends, during the five-year period when disaggregated results on gender differences were made available, Exeter boys outscored girls across the state in reading for four out of five years (O'Connor, 2010).

These cumulative findings demonstrate that Picturing Writing and Image-Making serve to level the playing field by allowing all students to enter the reading/writing process from a position of strength and enthusiasm. This holds particularly profound implications for a variety of subgroups, including Title I students, Special Education students, boys, and English Language Learners (ELL) for whom pictures serve as universal language.

Picturing Writing/Image-Making Integrated Curriculum Model

Picturing Writing utilizes a variety of simple crayon resist-based painting techniques within a progression of carefully crafted art-and-literature-based mini-lessons designed to teach essential literacy skills through standards-based art and language arts instruction. Central to the process, students draft their ideas in pictures first. Through analyzing the work of professional artists and writers, learning the grammar and principles of each discipline through the study of quality picture books, and providing students access to quality art materials and instruction, students learn how to create pictures that tell a story and write words that paint pictures. An array of simple texturing techniques (e.g., watercolor and salt, scratching into wet watercolor, splatter painting, sponge painting, plastic wrap prints, and tissue

dabbing) allow students to think in more detail and therefore to access more descriptive language when writing to their pictures. Students carefully craft their own quality picture books as they apply the skills they have learned.

Image-Making utilizes collage made from hand-painted papers created by each student to provide diverse learners with concrete, visual and kinesthetic tools for constructing story. As students literally give shape to their ideas through moving cut-and-torn shapes across each page, they are able to rehearse, draft, and revise their stories long before setting pencil to paper. As they create a sequence of collage images to tell their story, students are able to orally rehearse their story-line through reading their pictures.

These dynamic models represent a major paradigm shift. They recognize that traditional straight verbal instructional methods used in classrooms across the nation are failing to meet the needs of our diverse learners and that many students who struggle with reading and writing are visual learners. PW/IM offer students a variety of visual tools for thinking and learning, thus eliminating the hidden bias within our schools that favors the verbal learner.

Facilitated within an “Artists/Writers Workshop” format, the PW/IM models treat words and pictures as complementary, parallel, and equal languages for learning. They provide diverse learners with concrete visual tools for developing and expressing their ideas throughout the writing process. Within Artists/Writers Workshop, daily lessons are facilitated in two complementary strands: an art strand, which focuses on the language of pictures, and a writing strand, which focuses on the language of words. The art strand always precedes the writing strand. Whether focusing on pictures or words, each lesson follows the same four-step format.

Artists/Writers Workshop

- Literature Share/Discussion
- Modeling
- Work Session
- Group Share

Through the use of quality picture books as mentor texts and the interweaving of literature-based art and writing lessons, PW and IM have demonstrated their ability to reach a wide range of learners. While aligned with English language arts and art standards, these art-infused models can be easily integrated across the curriculum.

As students engage in creating their own high quality, content-driven picture books, PW and IM provide a highly structured, enticing, alternative pathway into the reading/writing process that has proven to be an effective intervention, particularly for those who struggle with words. These comprehensive models are also aligned with many best practices including brain-based learning, differentiated instruction, project-based learning, self-efficacy theory, multimodal instruction, theory of multiple intelligences, and the zone of proximal development. Twenty-First Century Skills (the 4’Cs: critical thinking, creativity, collaboration, and communication) are naturally woven throughout each process. PW/IM also offer students the opportunity to experience the phenomenon of transmediation (the recasting or translating of meaning from one sign system to another) known to deepen students’ thinking, generate new ideas,

and create greater opportunities for reflection (Siegel, 1995). As students make meaning in pictures first and then translate that meaning into words, they are not only drawn more deeply into their picture, but also experience the image “as if they were there.” As they access descriptive language to describe what is happening in their picture, they experience their picture “come alive.”

Evaluation Approach

The evaluation was designed to test the impact of an enhanced PW/IM model on students' writing, visual literacy, and reading skills within a more diverse, high-poverty educational setting. While PW/IM have been widely adopted and studied in grades 1 and 2, this most recent evaluation targeted a wider range of grade levels (grades 1-4 and two English Language Learner Magnet Programs, grades 1-5) and included a much more diverse, urban population of students. Because the City of Manchester serves as a national refugee resettlement center, during the time of the study, over 70 languages were spoken within MANSD. Serving the highest poverty and the most diverse population in NH, the entire district was identified as "In Need of Improvement" under NCLB just prior to receiving funding for this initiative. This study of approximately 1500 students every year over three years was also the largest Picturing Writing/Image-Making study of its kind to date.

Study Design and Sample

A quasi-experimental design with a matched comparison group was used to assess the effectiveness of the enhanced PW/IM model on student reading, writing and use of art as a language for communicating ideas in grades 1-4. Three elementary schools in MANSD served as treatment schools and three demographically matched elementary schools in the same district (based on percent of students receiving free and reduced lunch) served as comparison schools. Two self-contained ELL Magnet programs within MANSD also served as treatment and comparison groups for evaluating the impact of the enhanced model on the academic achievement of this unique population of immigrant and refugee students new to the country. A large part of the enhanced model involved using proven PW/IM practices to design, field-test, and refine a school-wide spiraling PW/IM curriculum across participating grade levels that was aligned with NH's language arts and art standards and addressed MANSD's social studies and science curriculum. The developer, in conjunction with a small team of classroom coaches and grade-level teacher leaders, developed and refined a spiraling PW/IM curriculum that was implemented throughout the school year as well as across grade levels. Students who entered the study during the primary grades were able to participate in the enhanced model for several consecutive years. Depending on the grade level and the curriculum topic, the enhanced model involved the refinement or development of instructional units and materials.

During the first year of the four-year study, teachers in the treatment schools received initial training and then four months of classroom coaching while they piloted the model. These teachers were expected to implement Picturing Writing using an Artists/Writers Workshop format two-three times a week, each workshop session lasting for 60-75 minutes. Teachers in the comparison group implemented the writing program being used in their school. Beginning in Year 2 of the study, ongoing professional development occurred through an embedded coaching model. An annual external evaluation was conducted to examine the effectiveness of the PW/IM model.

The study sample consisted of about 350-400 students per grade in grades 1, 2, 3, and 4 and ELL Magnet grades 1-5 (a total of about 1500 students). On average, 77 teachers participated each year (ranging from 68-83) from seven elementary schools during the three years of data collection. The study in participating schools included all students in each classroom. An ELL Magnet Program in one treatment school was matched with an ELL Magnet Program in a seventh school that did not already serve as a school-wide comparison school.

Data Collection Methods and Procedures

To measure student art and writing outcomes, student pre-test and post-test art and writing samples were collected. Two separate scoring instruments that were originally approved by the Program Effectiveness Panel of the National Diffusion Network were revised, expanded, and tested to meet the needs of this enhanced PW/IM model for grades 1-4¹. One instrument focused on narrative writing skills (i.e., establishing a sense of setting, character development, plot development, use of descriptive language, and overall quality of writing, including voice and ability of words to paint pictures in the reader's mind). The second instrument was used to assess the use and quality of visual elements to communicate ideas (i.e., use of color, texture, shape, detail, composition, sequence, and overall quality of artwork to tell the story). Under the advisement of a panel of veteran elementary school art and writing educators, the scoring instruments were reviewed, revised, and piloted during the first year, and continued to be used in Years 2, 3, and 4 of the study. Special consideration was given to the NH State Department of Education Grade Level Expectations in Language Arts (GLEs) and the NH State Curriculum Frameworks for Language Arts. See Appendix A for a copy of the instruments. Independent trained scorers used the instruments to rate the quality of student art and writing.

Art and writing samples were collected in the fall (pre-test) and spring (post-test) of each year. In September, teachers in the treatment and comparison groups handed students a sheet of paper with a blank box at the top for drawing a picture and lines below for writing a story. Students were given 30 minutes to complete their picture and story. In May, teachers in the treatment and comparison schools collected stories written by students during the spring term. The post-test stories for the treatment students were written during the implementation of Picturing Writing or Image Making. The stories for the comparison group were written as part of the writing program implemented in those schools.

After each data collection period, all identifying information was removed from the stories. Each story was rated on two major areas: 1) text portion of the story, and 2) use of visual elements to express ideas. An additional blind-rating structure was used to score the writing portion of the spring samples since students' accompanying artwork would have identified students as belonging to either the treatment or comparison group. Student writing was separated from the art and typed in a uniform fashion, thus removing all identifying features. Independent raters with backgrounds in elementary education were trained to rate the text only (written) portion of the story as well as the picture/word relationship (i.e. whether and how the pictures conveyed what was written in the text). Pre-test and

¹ The PW/IM art and writing scoring instruments, originally developed in 1991 by PW/IM developer, Dr. Susan Frankel, and a team of veteran art and writing educators, during a study which resulted in the validation of Image-Making Within the Writing process as an innovative and effective literacy program by the U. S. Department of Education Program Effectiveness Panel of the National Diffusion Network. The scoring instruments have been refined over multiple studies, including this most recent one.

post-test samples were rated on a scale of 1 - 6 where “1” means text or visual quality is not present and “6” means a text or visual quality is extraordinary. Raters received 51 hours of training by the Project Director and Evaluator before achieving an inter-rater reliability of 92.5%. Each art/writing portion and text only portion of the story were scored by two trained raters. The average of the two scores was used as the student’s scores in the analyses.

To measure additional student academic outcomes, the study used State and District reading and language arts assessments, including student reading and writing data on the statewide New England Common Assessment Program (NECAP), student reading and writing scores on the district-wide Gates MacGinitie Reading Comprehension Tests, and student writing scores on district writing prompts. These tests were administered by all classroom teachers in the District; therefore, teachers participating in the evaluation used the same assessments as the other teachers in the District.

Statewide NECAP Test writing scores. Grade 5 NECAP Writing Assessment, administered in October of every year, was used to further assess writing achievement. (Grade 5 was the only grade level evaluated for writing on the state assessments.) Fifth-graders in the treatment schools were exposed to Picturing Writing in grade 4 during the previous school year. Student NECAP writing scores, aggregated at the school level, were compared between treatment and comparison schools. Each of the three treatment schools was compared with its matched comparison school based on school size and percent of students on free and reduced price lunches. Findings were summarized at the school level for students at the beginning of grade 5 in the fall of 2007, 2008, and 2010; Writing scores from tests taken in the fall of 2009 were not available as it was a “norming year.”

District writing scores. Manchester NH School District’s Writing Prompt assessment was used as yet another tool to measure student’s writing improvement. The assessment was administered three times a year. For this study, assessment scores from September (time 1) and May/June (time 3) were used as pre-test and post-test measures of writing proficiency. The assessment measures six areas: content and ideas, organization, conventions, word choice, voice, and presentation. Student writing was scored on a four-point rating scale with “1” meaning not proficient and “4” meaning highly proficient.

Statewide NECAP Test Reading scores. The NECAP assessment contains a reading component also administered in October of each year. Student NECAP reading scores, aggregated at the school level, were compared between treatment and comparison schools. As with the NECAP writing scores, each of the three treatment schools was compared with its matched comparison school based and percent of students on free and reduced price lunches.

District reading scores. The Gates MacGinitie Reading Comprehension Tests were used to measure *reading* achievement in participating students. The tests were administered in June of each year and data were available from students in both the treatment group and the comparison group. Scores were reported on a four-point scale; values of 3 and 4 were combined and considered as a measure of proficient level of reading.

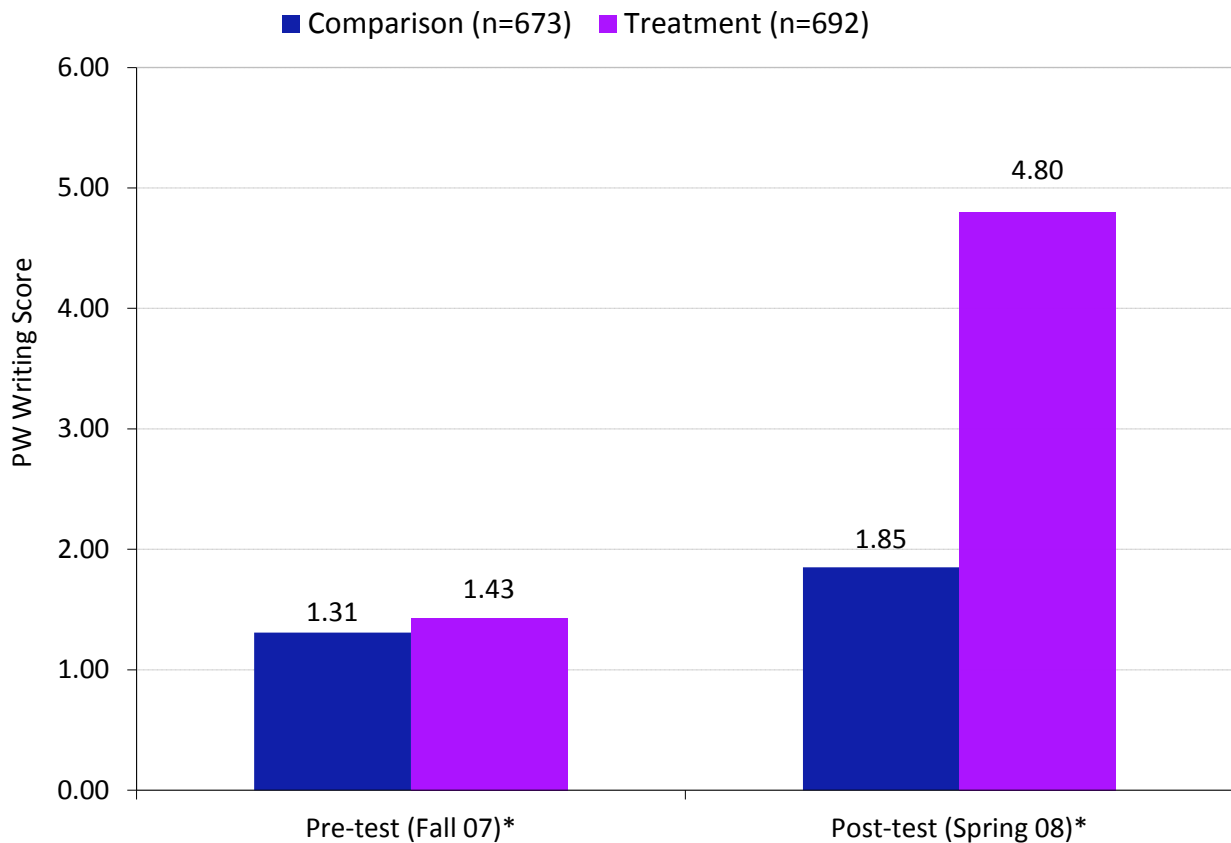
Data Analysis and Findings

Writing and Visual Literacy

Art and Writing. Pre-test and post-test scores from student art and writing samples were analyzed using analysis of variance (ANOVA) to test differences between treatment and comparison groups. Separate analyses were conducted on scores for each quality as measured by a scale or item from the scoring instruments, and for each grade and across all grades for treatment effect.

In the first year of the evaluation, statistically significant gains in writing were achieved by students in the treatment group as compared to the demographically matched comparison group across all grade levels. Both the comparison group and treatment group began in September with similar scores of 1.31 and 1.43, respectively, for the overall quality of student writing (scores that reflect less than minimal levels of achievement in basic writing skills), but by the spring dramatic differences in writing skills were apparent between the two groups. Although there was growth within the comparison group, much greater gains were made by students participating in Picturing Writing. The average score for the overall quality of student writing of the treatment group jumped from 1.43 to 4.80 (nearly “fully developed”). The average score of the comparison group increased slightly to 1.85, remaining a minimal score. See Exhibit 1. Findings were similar for each of the four grade levels.

Exhibit 1. Pre-Post Test Writing Scores, All Grade Levels, 2007-2008

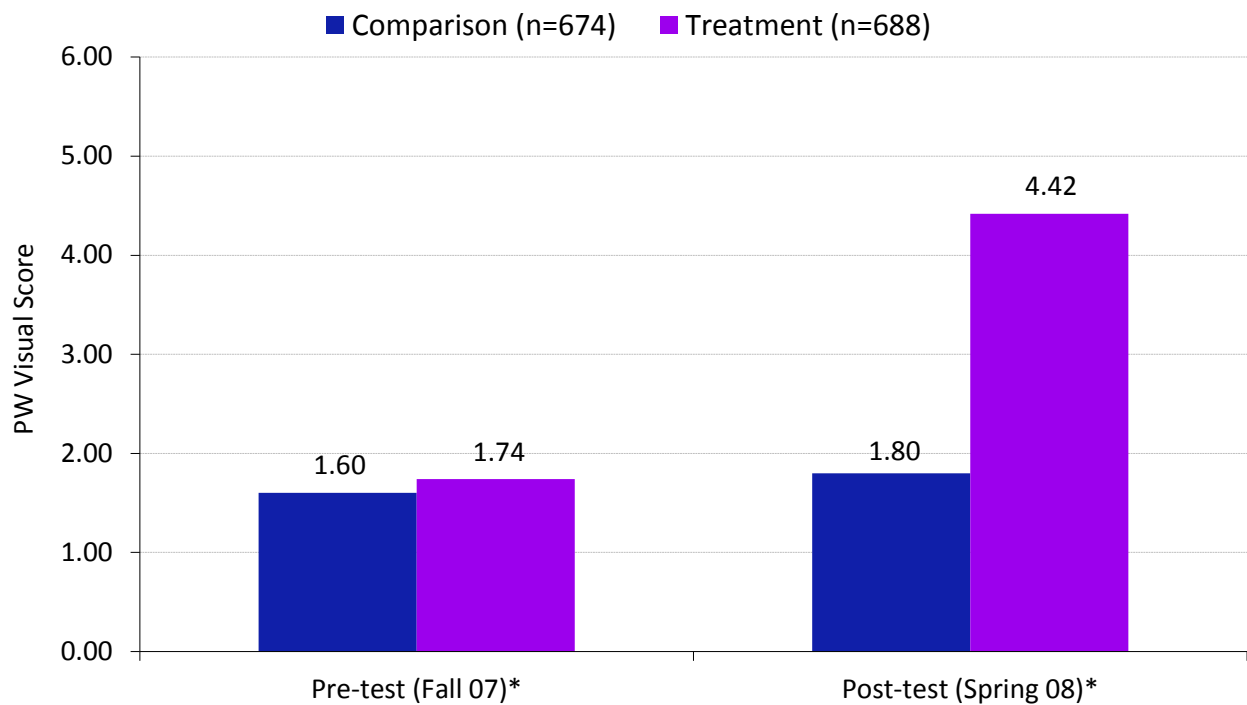


*The difference between comparison and treatment groups is statistically significant.

The findings in visual literacy showed a similar pattern in students’ ability to express their ideas through the use of visual elements; statistically significant gains were made by students participating in PW while only slight gains were made by students in the demographically matched comparison group. Students

from both groups began the school year with roughly the same level of skill (minimal scores of 1.60 for the comparison group and 1.74 for the treatment group) for overall use and quality of visual information. After participation in PW, the overall average score of the treatment students increased to 4.42 (display of trait is partially developed). The average score of students in the comparison group remained at the low end of the continuum with 1.80 (Exhibit 2).

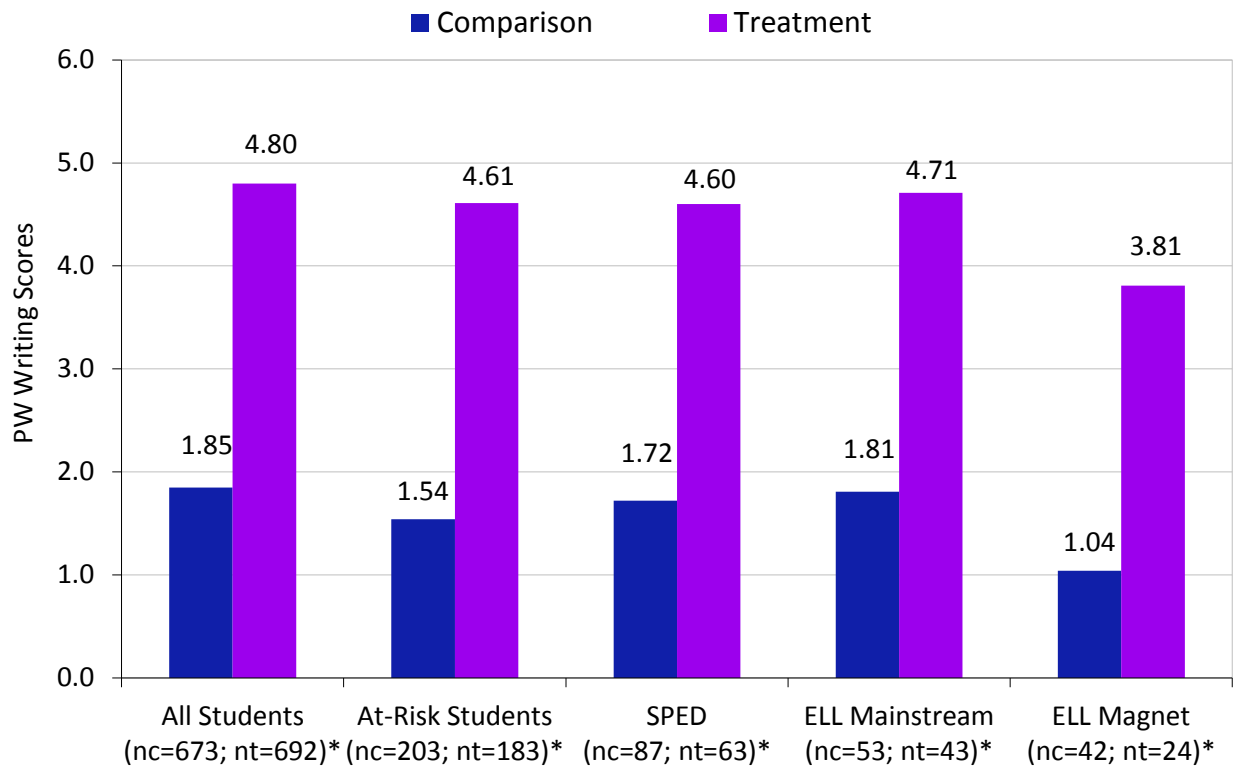
Exhibit 2. Pre- and Post-Test Visual Literacy Scores, All Grade levels, 2007-2008



*The difference between comparison and treatment groups is statistically significant.

This pattern also remained consistent for all subgroups of students targeted by No Child Left Behind – At-Risk Students (defined as any students who scored below benchmark in reading in the fall of that year), Special Education Students, ELL Mainstream Students, and ELL Magnet Students. All subgroups made statistically significant gains in Year One as compared to their demographically matched comparison groups, and kept close pace to “all students” in the treatment group. Note that even ELL Magnet Students (relatively new to the country) scored far greater than “all students” in the comparison group (Exhibit 3).

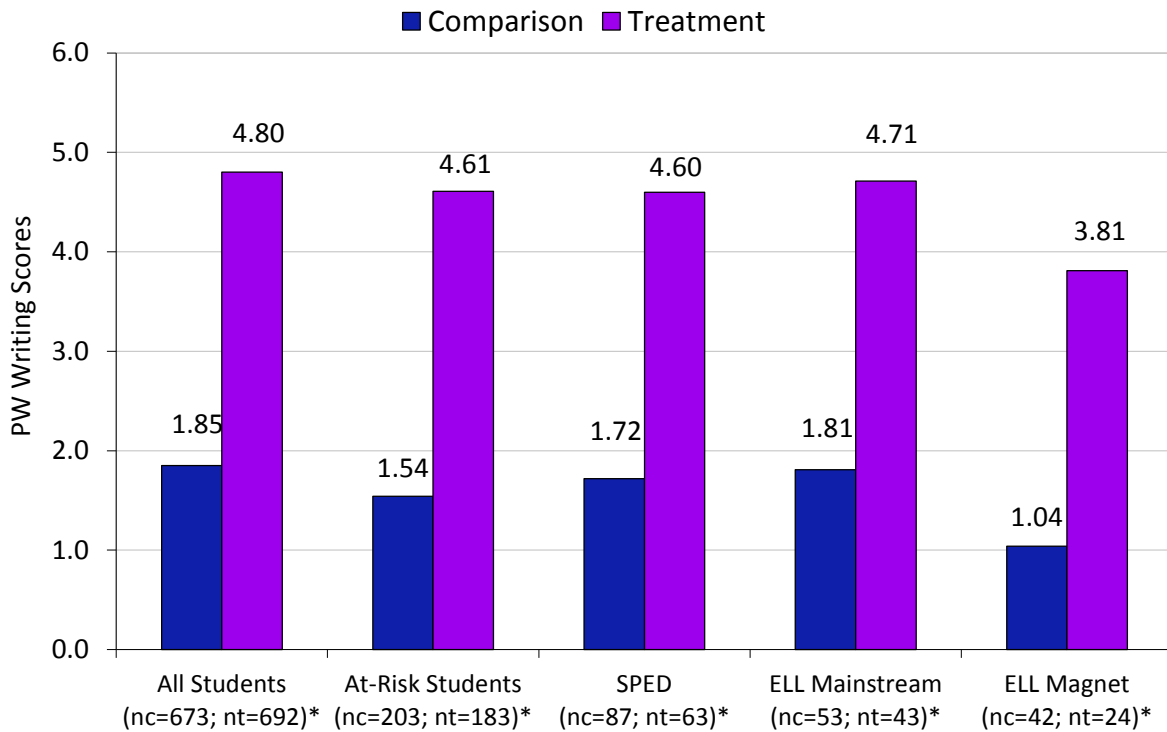
Exhibit 3. A Comparative Summary of PW Spring Writing Scores for Subgroups, Across Grade Levels, Spring 2008



*The difference between comparison and treatment groups is statistically significant.

Findings followed a similar pattern when looking at the Comparative Summary of Visual Scores for All Subgroups: Spring 2008 (Exhibit 4).

Exhibit 4. A Comparative Summary of PW Spring Visual Literacy Scores for Subgroups, Across Grade Levels, Spring 2008

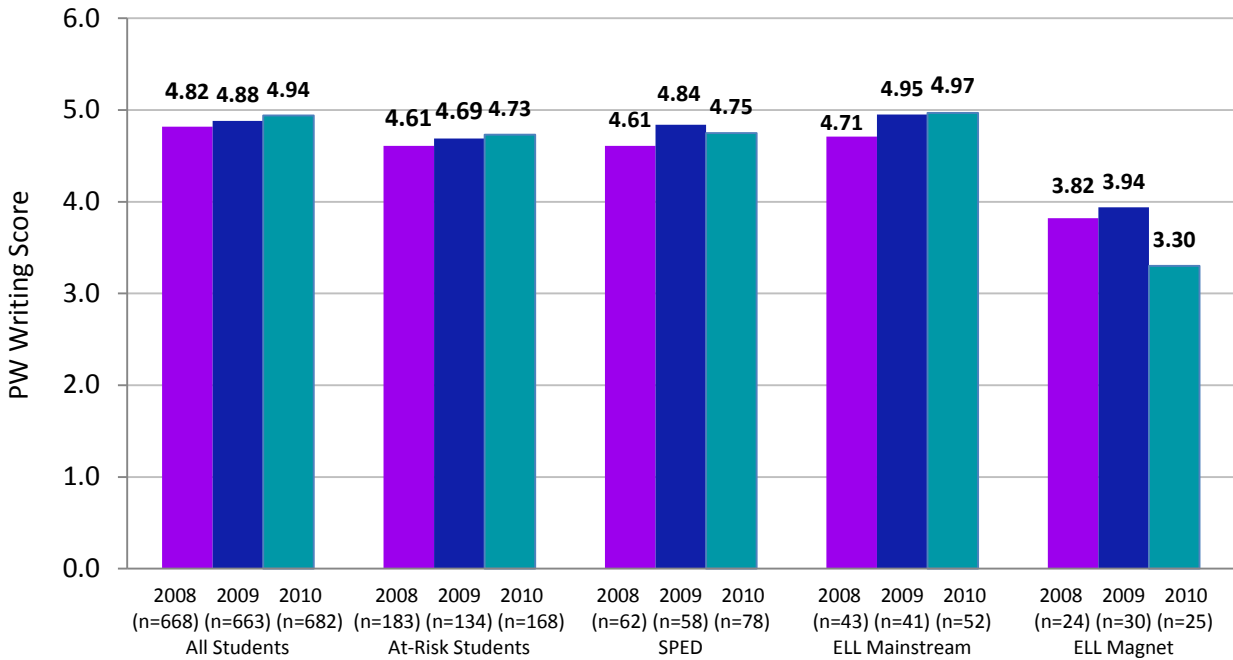


*The difference between comparison and treatment groups is statistically significant.

This pattern of findings remained consistent over the course of the three years of the evaluation. Using independently developed scoring instruments to evaluate students' story writing and use of visual elements to communicate ideas, significantly greater gains in writing and visual literacy were achieved by treatment students at all grade levels each year, as compared to the demographically matched comparison group.

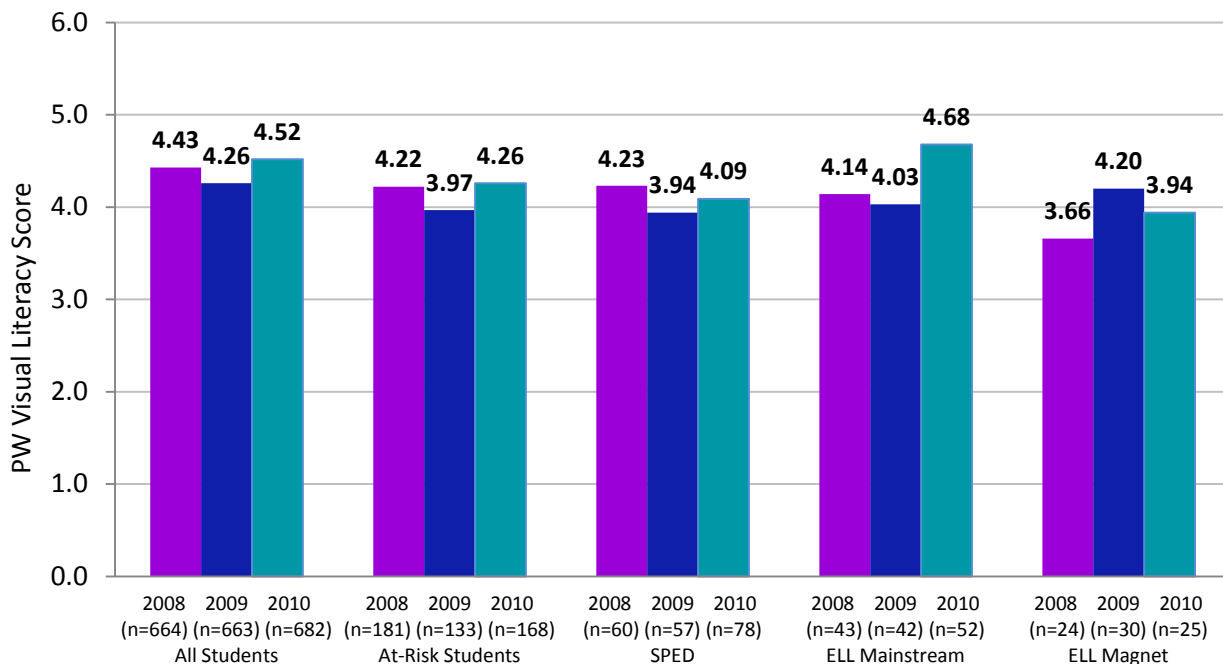
Exhibit 5 shows the spring writing scores each year for subgroups of students in the treatment group. Each year, after consistent participation in Artists/Writers Workshop (either PW or PW/IM models), students in various subgroups increased their overall writing scores, with minor variations seen among SPED and ELL Magnet populations dependent on the severity of these particular students learning needs, or in the case of the Magnet ELL students, whether or not they had any previous formal education. Within the treatment group, the overall spring writing scores for "all students" ranged between 4.82 and 4.94. Note that all subgroups except for those students brand new to the country kept close pace with "all students" by the spring of each year with the ELL Mainstream Students outscoring "all students" in writing in 2009 and 2010.

Exhibit 5. PW Spring Writing Scores, Treatment Students - Subgroups: 2008, 2009, 2010



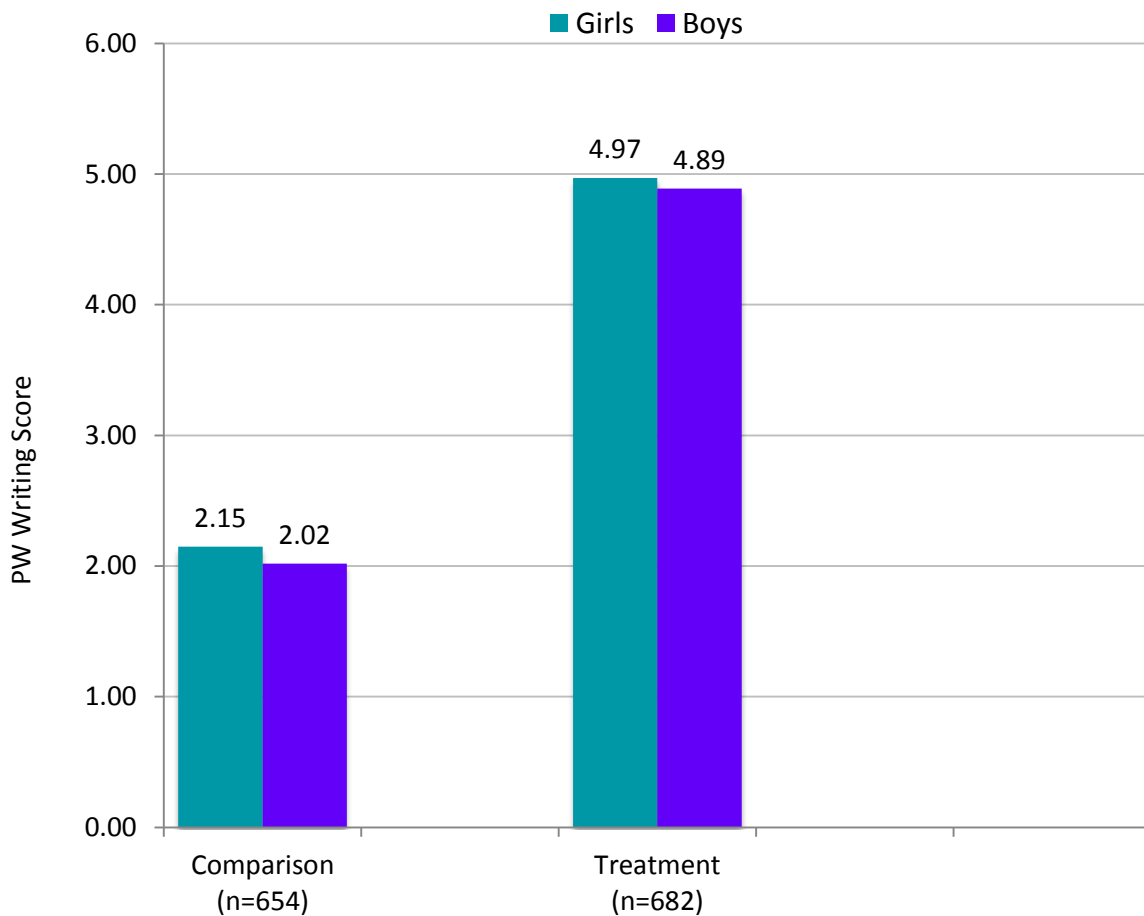
Spring scores on the use of visual elements for at-risk students in the treatment group across the three years of the study were consistently high compared to comparison groups though certainly displayed more variation year by year. As noted in our conclusion section, in school year 2008-2009 there was a radical turnover in teaching staff, thus many teachers during this year (except for the ELL Magnet teachers) would have been new to facilitating PW (Exhibit 6).

Exhibit 6. PW Spring Visual Literacy Scores, Treatment Students - Subgroups: 2008, 2009, 2010



Analysis by gender was conducted in the final year. Differences in writing between boys and girls were minimal. The average spring scores for the overall quality of writing for both boys and girls participating in PW/IM were far greater than the scores for boys and girls in the comparison groups. The scores for boys and girls in the comparison group increased marginally to 2.02 and 2.15, respectively. The scores for boys and girls in the treatment group increased to 4.89 and 4.97, respectively. These findings suggest that boys, who generally underperform in the area of language arts, are more successful with hands-on visual and kinesthetic language arts instruction (Exhibit 7).

**Exhibit 7. PW Spring Writing Scores Treatment vs. Comparison Students, Girls and Boys:
All Grade Levels 2009-2010**



*The difference between comparison and treatment groups is statistically significant, $p < .000$.

**The difference between females and males is statistically significant, $p < .05$.

Writing

Writing Scores on State Assessment (NECAP). Table 1 shows results of Grade 5 NECAP Writing Test. Findings are summarized at the school level for students at the beginning of grade 5 in the fall of 2007, 2008, and 2010; students in the treatment schools participated in Picturing Writing in grade 4 during the prior year, although fifth graders taking the NECAPs in the fall of 2007 only participated in PW beginning in January of that year. Writing scores were not available for the fall of 2009 because 2009-2010 was a statewide “norming” year for the writing assessment.

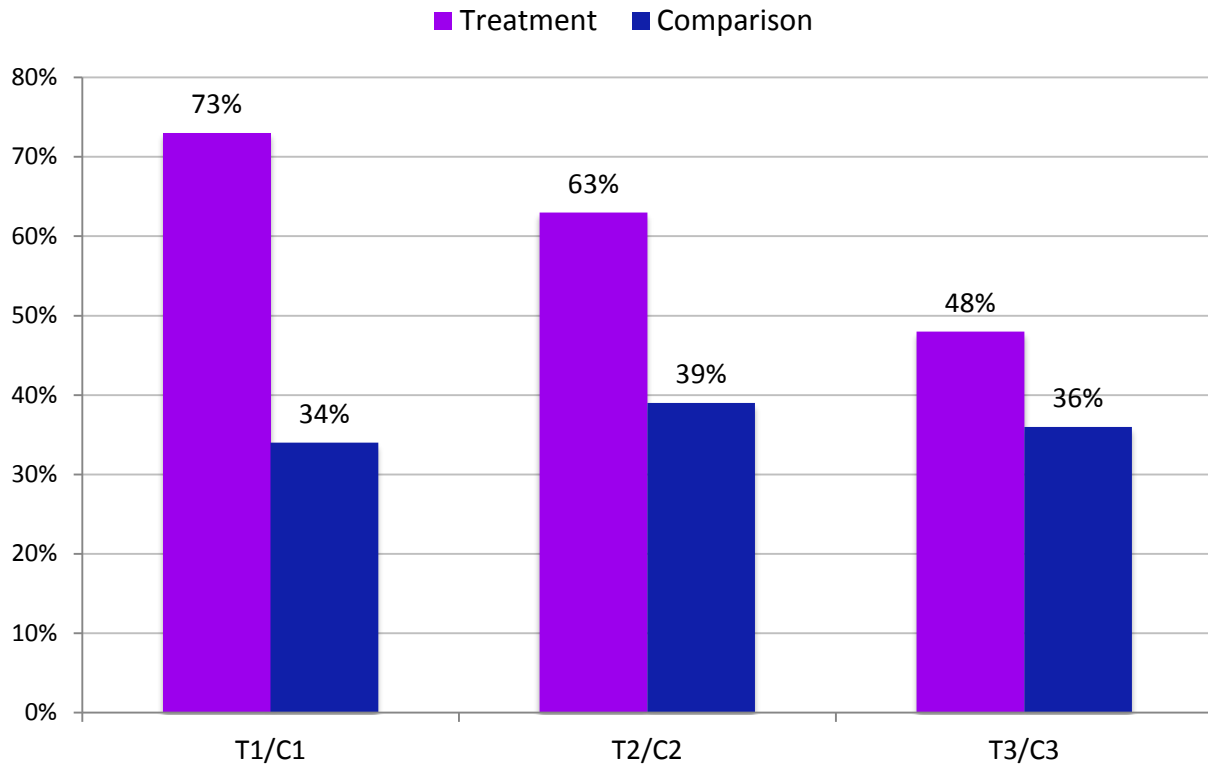
Table 1. Comparison of Students Scoring at Proficiency or Above on 2008 and 2007 NECAP Grade 5 Writing Test Between Comparison and Treatment Schools

	Fall 2007			Fall 2008			Fall 2010		
	Comparison Schools	Picturing Writing Schools	District	Comparison Schools	Picturing Writing Schools	District	Comparison Schools	Picturing Writing Schools	District
Number of students in grade 5 at beginning of Fall who took NECAP writing test	320	192	1,100	286	202	1,106	279	151	1,020
Number of students who scored at proficient or above proficient levels	101	71	343	96	104	419	93	93	424
Percent of students who scored at proficient or above proficient levels	32%	37%	31%	33%	51%	38%	33%	62%	42%

The percent of students proficient on the Grade 5 NECAP Writing Test was higher for schools that implemented PW than for the comparison schools, and the district overall. In the fall of 2008 NECAP Writing results, 51% of students in the treatment group scored proficient compared to 33% of students in the demographically matched comparison group whereas in the fall of 2007 only 37% of students in the treatment group scored proficient compared to 32% of students in the demographically match comparison group. In 2010, the final year of our study, 62% of treatment students scored proficient while 33% of comparison students scored proficient.

When looking at the test results of PW treatment schools as compared to demographically match comparison schools, some interesting findings emerge. While one might assume that the school with the strongest test results was the treatment school with the lowest percentage of students on free and reduced lunch, in this case, the T1/C1 partner schools had the highest percentage of students on free and reduced lunch (57% and 52% respectively). Closer analysis revealed that this treatment school was the only treatment school with consistent administrative support throughout the four-year research period (Exhibit 8). The other two schools experienced turnover in administration, resulting in a lack of strong leadership regarding PW implementation.

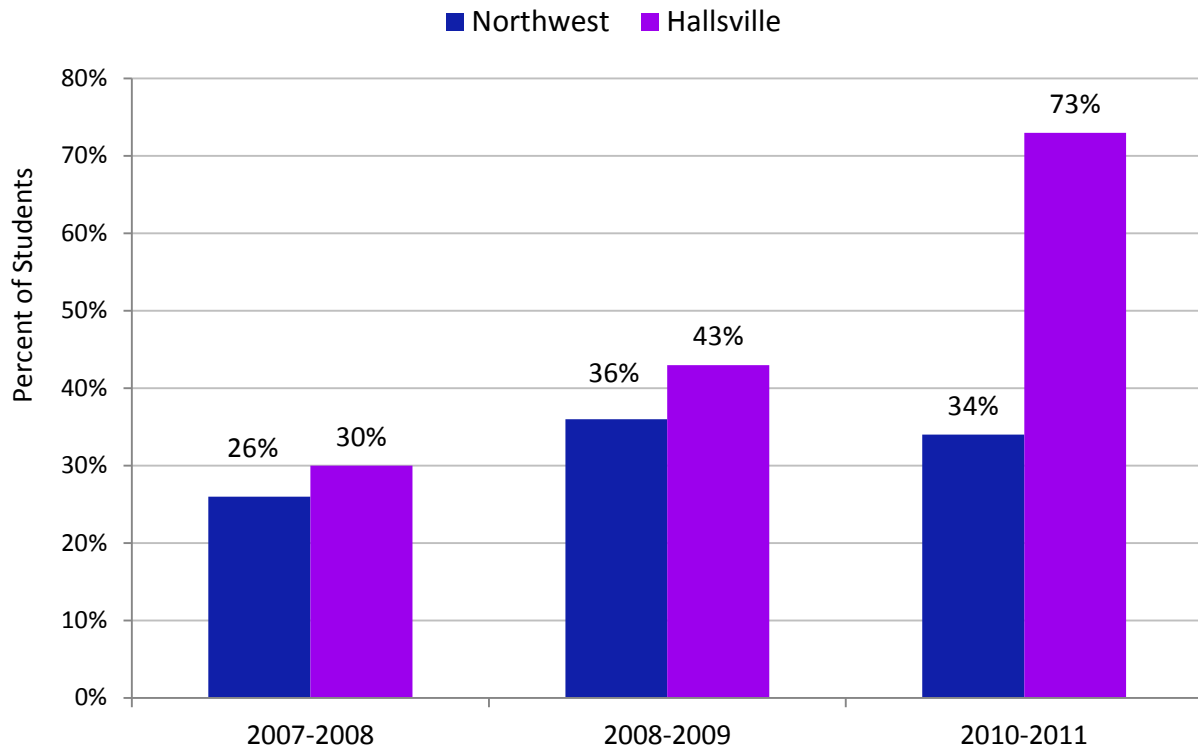
Exhibit 8. Percentage of Fifth Graders Scoring Proficient or Above on NECAP Writing 2010 by Treatment/Comparison Schools



Defining T1 as the “high fidelity school” based on its consistent administrative support, it is informative to drill down deeper to look at the potential of PW when its art-based writing process is fully supported and implemented.

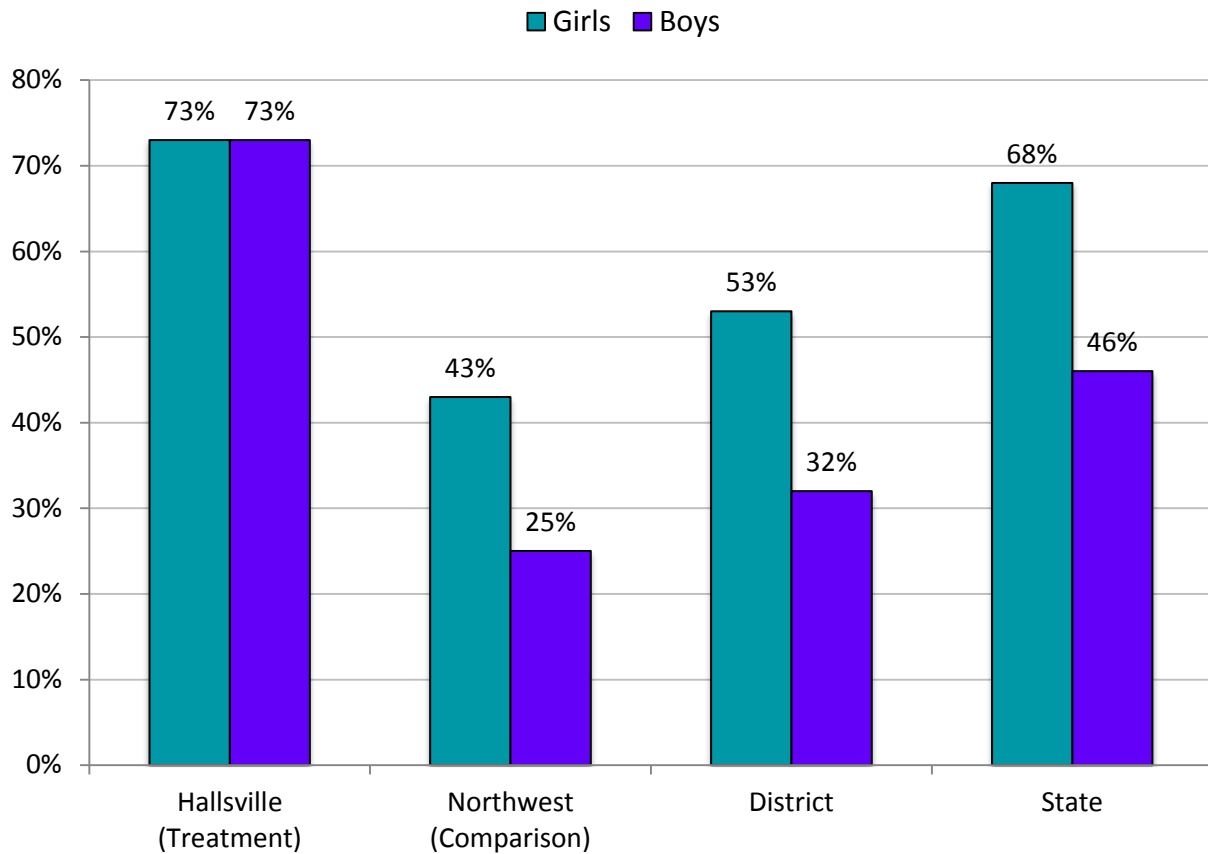
Exhibit 9 shows the writing gains over time of the high fidelity school as compared to its demographically matched comparison school.

Exhibit 9. Percent of Students Proficient and Above on Grade 5 NECAP Writing Test



Most interesting are the findings of boys versus girls on the NECAP Writing Assessment. Note that boys and girls scored identically and identically well in writing. This did not happen in the comparison school, anywhere in the District, nor anywhere in the State (Exhibit 10). This is a particularly interesting finding given that boys typically lag behind girls in reading/writing across the nation (National Center for Education Statistics, 2000).

Exhibit 10. Grade 5 NECAP Writing Test 2010-2011 Percent Proficient and Above by Gender



This finding is also interesting in the context of the brain research conducted by Michael Gurian, who has observed physiological differences between the brains of boys and girls. Given that Gurian discovered that boys have more cortical areas of the brain dedicated to spatial functioning while girls use more cortical area of their brains for verbal functioning, this finding regarding the performance of boys participating in a visual/spatial approach to writing is of particular interest (Gurian, 2005).

Also of interest are the comparative findings for Title I students (Exhibit 11) and Economically disadvantaged students (Exhibit 12).

Exhibit 11. Grade 5 NECAP Writing Test 2010-2011 Percent Scoring Proficient and Above: Title I Data

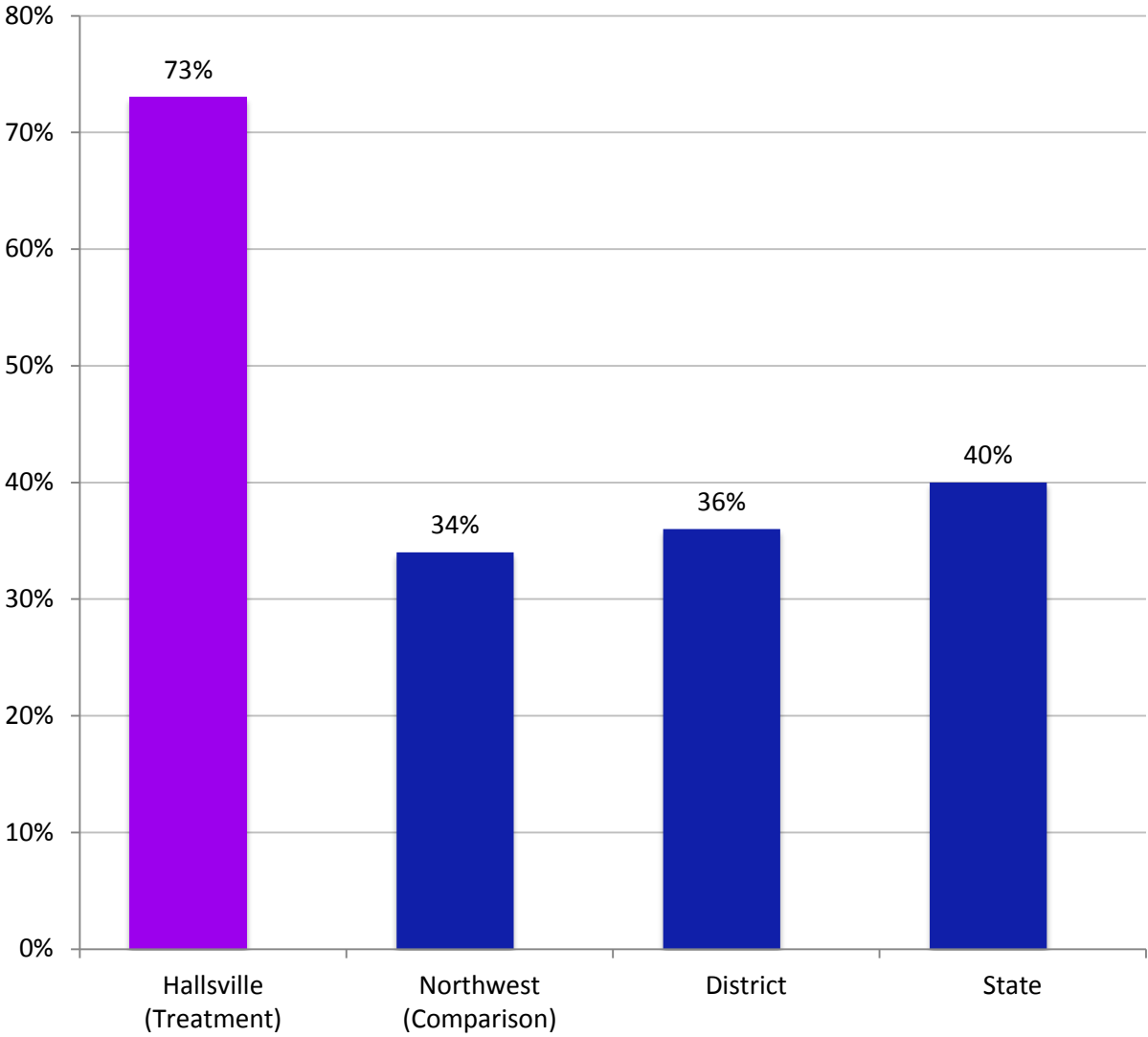
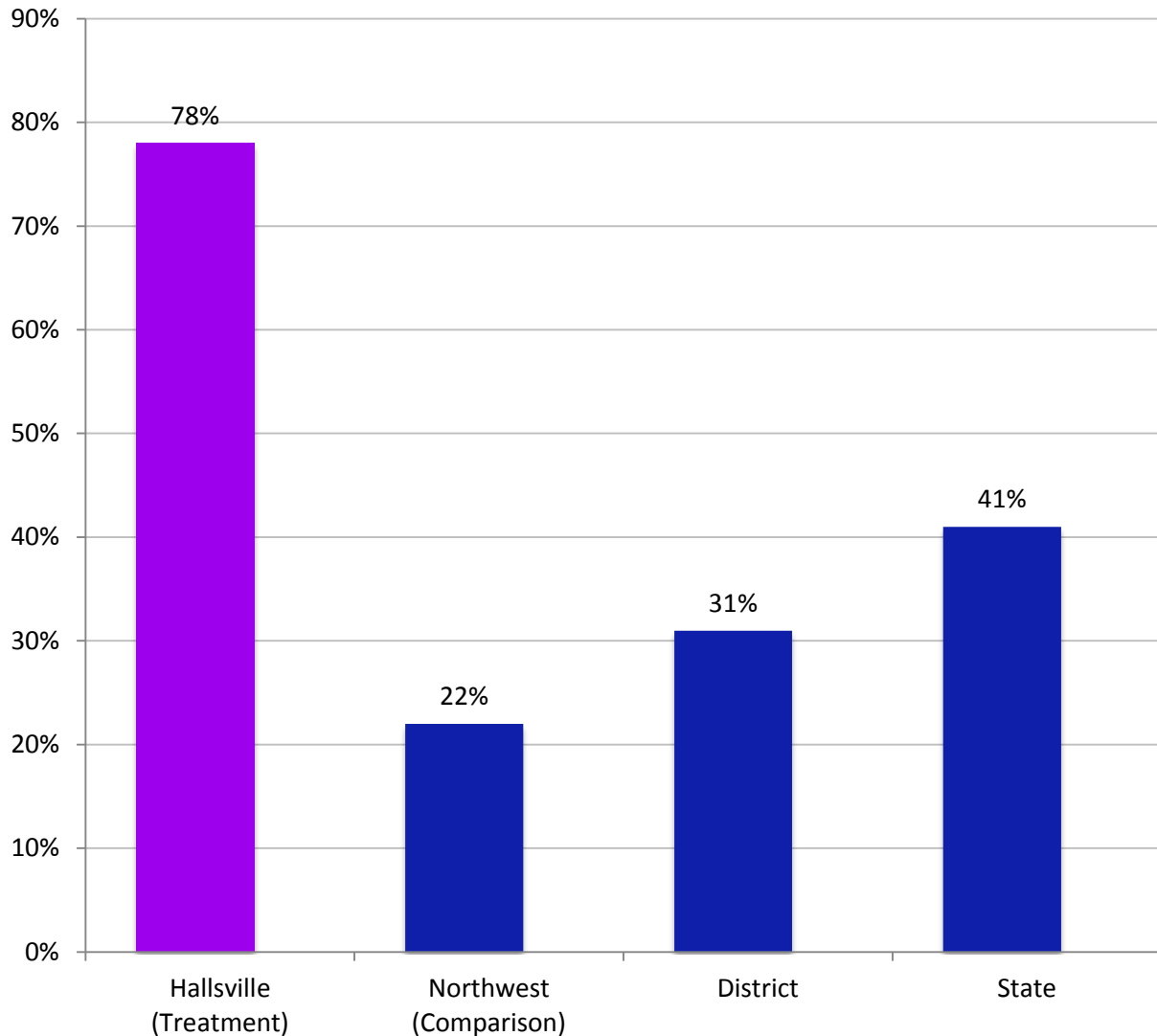


Exhibit 12. Grade 5 NECAP Writing Test 2010-2011 Percent Scoring Proficient and Above: Economically Disadvantaged Data



District Writing. The Manchester NH School District’s Writing Prompt assessment was also used to measure student’s writing improvement. Student writing was scored on a four-point rating scale with “1” being not proficient and “4” being highly proficient. The assessment measures six areas: content and ideas, organization, conventions, word choice, voice, and presentation.

For each year of the research study, pre-post matched data were available from students in the treatment and the comparison group. Across the three years, on most of the writing areas, students participating in Picturing Writing scored higher than their comparison peers. One consistent finding across the years was that students participating in Picturing Writing scored higher than students in the comparison group in the writing area of voice, and the difference across all grade levels was statistically significant each year. Tables 2, 3 and 4 show the findings for all students by year and grade level.

Table 2. Table Student Performance on School District’s Writing Prompt by Treatment Group – All Students (Grades 1-4), 2007-2008

Trait of Writing	Comparison (n=706)			Treatment (n=705)			F ₁	F ₂	Effect Size
	Mean	Std. Dev	Adj. Mean	Mean	Std. Dev	Adj. Mean			
Content									
Fall 2007	2.06	.82		2.05	.80		.04		
Spring 2008	2.77	.82	2.70	2.90	.83	2.90		10.81*	.13
Organization									
Fall 2007	1.80	.77		1.80	.78		.01		
Spring 2008	2.64	.89	2.64	2.71	.90	2.70		2.04	.06
Conventions									
Fall 2007	2.19	.89		2.06	.85		8.35*		
Spring 2008	2.81	.84	2.78	2.83	.86	2.86		3.98	.16
Word Choice									
Fall 2007	1.84	.74		1.80	.71		1.08		
Spring 2008	2.49	.77	2.49	2.59	.81	2.59		8.10*	.15
Voice									
Fall 2007	2.01	.82		2.00	.82		.11		
Spring 2008	2.67	.81	2.66	2.88	.83	2.88		27.43**	.21
Presentation									
Fall 2007	2.26	.94		2.16	.87		4.12		
Spring 2008	2.94	.87	2.92	3.05	.87	3.76		13.78**	.21

F₁ tests the difference on pretest (fall 2007) between comparison and treatment groups.

F₂ tests the difference on posttest (spring 2008) between comparison and treatment groups after adjusting pretest differences.

*p<.01; **p<.000.

Table 3. Table Student Performance on School District’s Writing Prompt by Treatment Group – All Students (Grades 1-4), 2008-2009

Trait of Writing	Comparison (n=717)			Treatment (n=665)			F ₁	F ₂
	Mean	Std. Dev	Adj. Mean	Mean	Std. Dev	Adj. Mean		
Content								
Fall 2008	1.98	.85		2.16	.86		15.21***	
Spring 2009	2.79	.82	2.83	2.90	.82	2.86		.97
Organization								
Fall 2008	1.80	.82		1.91	.84		5.24*	
Spring 2009	2.65	.84	2.67	2.74	.84	2.72		1.33
Conventions								
Fall 2008	2.04	.88		2.13	.92		3.81	
Spring 2009	2.76	.84	2.78	2.84	.90	2.82		.67
Word Choice								
Fall 2008	1.79	.77		1.89	.76		5.86*	
Spring 2009	2.53	.79	2.55	2.64	.81	2.61		2.79
Voice								
Fall 2008	1.88	.85		2.18	.87		42.12***	
Spring 2009	2.72	.82	2.77	2.92	.78	2.86		5.12*
Presentation								
Fall 2008	2.12	.95		2.28	.91		10.53**	
Spring 2009	2.97	.82	3.01	3.06	.84	3.02		.08

F₁ tests the difference on pretest (fall 2008) between comparison and treatment groups.

F₂ tests the difference on posttest (spring 2009) between comparison and treatment groups after adjusting pretest differences.

*p<.05; **p<.001; ***p<.000.

Table 4. Table Student Performance on School District’s Writing Prompt by Treatment Group – All Students (Grades 1-4), 2009-2010

Trait of Writing	Comparison (n=667)			Treatment (n=680)			F ₁	F ₂
	Mean	Std. Dev	Adj. Mean	Mean	Std. Dev	Adj. Mean		
Content								
Fall 2009	1.98	.89		2.09	.85		5.76*	
Spring 2010	2.81	.82	2.82	2.95	.79	2.93		6.83**
Organization								
Fall 2009	1.77	.85		1.87	.86		5.32*	
Spring 2010	2.64	.84	2.66	2.77	.86	2.75		4.34*
Conventions								
Fall 2009	2.03	.96		2.11	.94		2.39	
Spring 2010	2.78	.83	2.79	2.90	.83	2.89		5.36*
Word Choice								
Fall 2009	1.78	.81		1.84	.77		2.33	
Spring 2010	2.54	.81	2.55	2.65	.79	2.70		5.84*
Voice								
Fall 2009	1.93	.89		2.05	.87		6.27*	
Spring 2010	2.73	.79	2.75	2.98	.76	2.96		28.56***
Presentation								
Fall 2009	2.17	1.02		2.28	1.00		4.24*	
Spring 2010	2.94	.83	2.96	3.08	.84	3.06		5.75*

F₁ tests the difference on pretest (fall 2009) between comparison and treatment groups.

F₂ tests the difference on posttest (spring 2010) between comparison and treatment groups after adjusting pretest differences.

*p<.05; **p<.01; ***p<.000.

Further analysis showed that in 2008-09 and 2009-10, at-risk students across all grade levels scored higher than their peers in the comparison group in all six writing areas. In 2007-08 at-risk students participating in Picturing Writing scored higher than students in the comparison group in five of the six writing areas. Tables 5, 6, and 7 show the findings for at-risk students by year and across grade levels.

Table 5. Table Student Performance on School District’s Writing Prompt by Treatment Group – At-risk Students (Grades 1-4), 2007-2008

Measure	Comparison				Treatment				F ₁	F ₂
	Mean	Std. Dev	n	Adj. Mean	Mean	Std. Dev	n	Adj. Mean		
Writing - Content										
Fall 2007	1.66	.71	187		1.67	.72	201		.01	
Spring 2008	2.24	.72	187	2.24	2.44	.80	201	2.44		7.04*
Writing - Organization										
Fall 2007	1.46	.61	187		1.40	.67	201		.91	
Spring 2008	2.05	.75	187	2.04	2.21	.84	201	2.22		5.30
Writing – Conventions										
Fall 2007	1.69	.77	187		1.57	.65	201		2.88	
Spring 2008	2.19	.76	187	2.17	2.20	.80	210	2.23		.66
Writing - Choice										
Fall 2007	1.50	.60	187		1.43	.54	201		1.22	
Spring 2008	2.03	.63	187	2.02	2.05	.69	201	2.06		.42
Writing - Voice										
Fall 2007	1.64	.70	187		1.57	.67	201		1.00	
Spring 2008	2.20	.72	187	2.19	2.38	.80	201	2.39		7.57*
Writing - Presentation										
Fall 2007	1.79	.83	187		1.66	.74	201		2.65	
Spring 2008	2.45	.86	187	2.43	2.50	.90	201	2.53		1.45

F₁ tests the difference on pretest (fall 2007) between comparison and treatment groups.

F₂ tests the difference on posttest (spring 2008) between comparison and treatment groups after adjusting pretest differences.

*p<.01; **p<.000.

Table 6. Table Student Performance on School District’s Writing Prompt by Treatment Group – At-risk Students (Grades 1-4), 2008-2009

Trait of Writing	Comparison (n=204)			Treatment (n=136)			F ₁	F ₂
	Mean	Std. Dev	Adj. Mean	Mean	Std. Dev	Adj. Mean		
Content								
Fall 2008	1.50	.74		1.62	.71		1.96	
Spring 2009	2.25	.75	2.26	2.36	.80	2.34		.89
Organization								
Fall 2008	1.40	.68		1.38	.63		.09	
Spring 2009	2.08	.75	2.08	2.23	.81	2.23		3.58
Conventions								
Fall 2008	1.53	.78		1.51	.74		.05	
Spring 2009	2.14	.83	2.13	2.16	.80	2.17		.14
Word Choice								
Fall 2008	1.35	.64		1.38	.61		.24	
Spring 2009	1.98	.64	1.99	2.04	.65	2.03		.45
Voice								
Fall 2008	1.39	.68		1.60	.72		7.25*	
Spring 2009	2.19	.77	2.21	2.38	.75	2.34		2.34
Presentation								
Fall 2008	1.57	.83		1.70	.85		1.84	
Spring 2009	2.49	.86	2.52	2.45	.88	2.42		1.01

F₁ tests the difference on pretest (fall 2008) between comparison and treatment groups.

F₂ tests the difference on posttest (spring 2009) between comparison and treatment groups after adjusting pretest differences.

*p<.01.

Table 7. Table Student Performance on School District’s Writing Prompt by Treatment Group – At Risk Students (Grades 1-4), 2009-2010

Trait of Writing	Comparison (n=165)			Treatment (n=171)			F ₁	F ₂
	Mean	Std. Dev	Adj. Mean	Mean	Std. Dev	Adj. Mean		
Content								
Fall 2009	1.47	.82		1.61	.75		2.76	
Spring 2010	2.28	.81	2.31	2.42	.71	2.40		1.29
Organization								
Fall 2009	1.27	.72		1.44	.70		4.63*	
Spring 2010	2.13	.82	2.16	2.22	.75	2.19		.11
Conventions								
Fall 2009	1.50	.90		1.52	.77		.07	
Spring 2010	2.12	.76	2.13	2.23	.78	2.22		1.59
Word Choice								
Fall 2009	1.34	.71		1.36	.64		.06	
Spring 2010	2.03	.75	2.03	2.09	.67	2.09		.61
Voice								
Fall 2009	1.46	.79		1.58	.77		1.94	
Spring 2010	2.22	.77	2.24	2.50	.80	2.48		8.55**
Presentation								
Fall 2009	1.61	.92		1.75	.93		2.16	
Spring 2010	2.38	.85	2.40	2.51	.88	2.48		.85

F₁ tests the difference on pretest (fall 2009) between comparison and treatment groups.

F₂ tests the difference on posttest (spring 2010) between comparison and treatment groups after adjusting pretest differences.

*p<.05; **p<.005.

Reading

While Picturing Writing and Image-Making are instructional models used to teach writing, because the reading and writing processes are so closely linked, our study tracked State and District reading scores.

Reading Scores on State Assessment (NECAP). Table 8 presents the results of the NECAP Reading Test over three years of the evaluation. Findings are summarized at the school level for students at the beginning of grades 3-5 in the fall of 2007, 2008 and 2009. In the fall of 2008 and 2009, students in the treatment schools participated in Picturing Writing in grades 2-4 during the prior year. In the fall of 2007, however, students in the treatment schools participated in Picturing Writing for less than 6 months in grades 2-4 in the previous year.

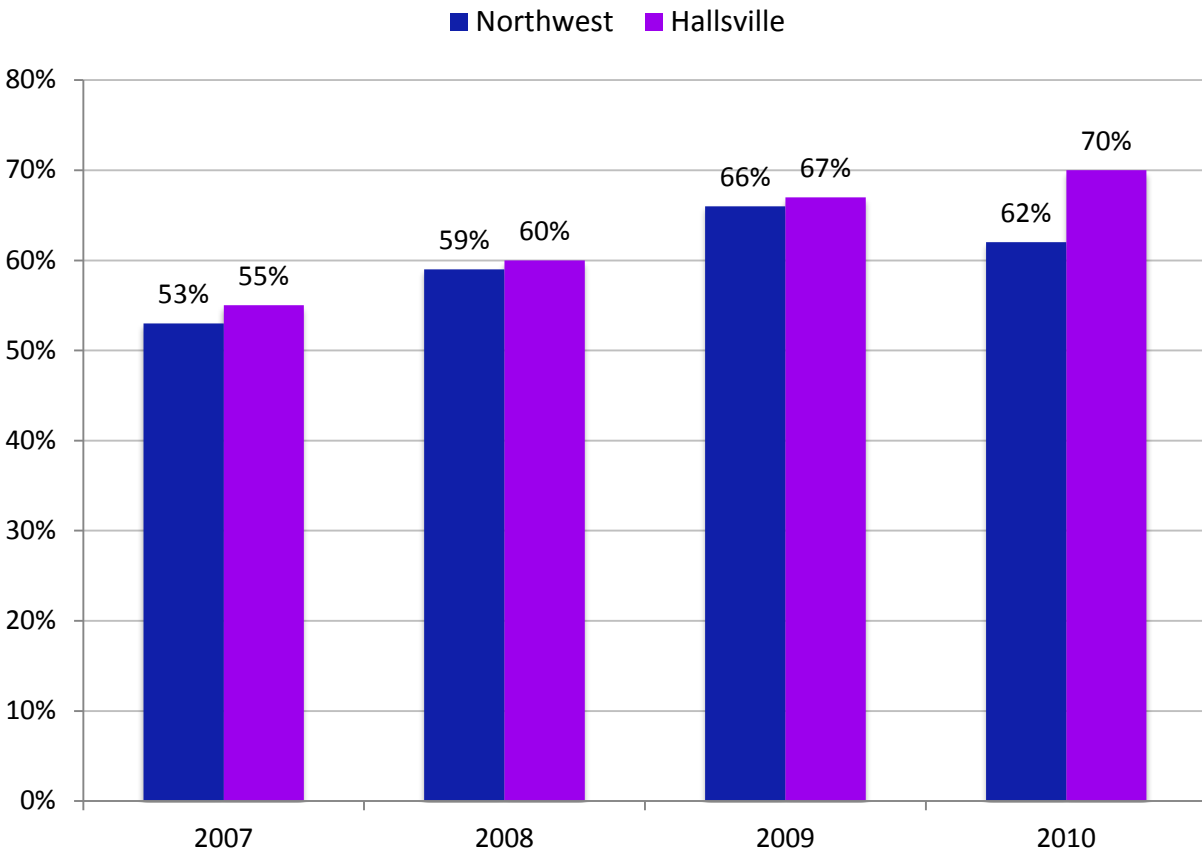
Table 8. A 3-Year Comparison of Students Scoring at Proficiency or Above on NECAP Reading Tests Between Comparison and Treatment Groups

	Fall 2007		Fall 2008		Fall 2009		Fall 2010	
	Comparison Schools	Picturing Writing Schools	Comparison Schools	Picturing Writing Schools	Comparison Schools	Picturing Writing Schools	Comparison Schools	Picturing Writing Schools
Number of students in grades 3-5 at beginning of Fall who took NECAP reading tests	918	591	872	555	847	542	866	553
Number of students who scored at or above proficient levels	554	356	526	359	561	369	548	358
Percent of students who scored at or above proficient levels	60%	60%	60%	65%	66%	68%	63%	65%

Findings show that in the fall of 2007, the percent of students who were proficient on the NECAP Reading Test was the same for the treatment and comparison schools, 60%, indicating that at the beginning of the research study, students in the demographically matched schools had similar percentages of students reading at or above proficient levels. In subsequent years, the percent of students reaching proficiency was slightly higher for schools implementing PW than for comparison schools.

When looking at the NECAP reading results for the “high fidelity” school, the school with consistent administrative support, we see a stronger pattern emerge. From 2008-2010, the treatment school’s NECAP reading scores show consistent growth in reading by Grade 5 whereas the comparison school’s scores show erratic results (Exhibit 13).

Exhibit 13. Percent of Student Proficient and Above on Grade 5 NECAP Reading Test



District Reading. The Gates MacGinitie Reading Comprehension Tests were used to measure reading achievement. The tests were annually administered in June. Scores were reported on a four-point scale, values of 3 and 4 indicate that students were proficient at reading.

Table 9 presents the percent of proficient students in both treatment and comparison groups for each year of the research study and by grade level. Across all three years, the percent of students scoring proficient in grade 4 was higher for the treatment group than for the comparison group. For grades 1-3, the findings differed by grade level and by year. In 2009, the difference between the percent of grade 1 students scoring proficient in the treatment and comparison groups was statistically significant, with students participating in Picturing Writing scoring higher than their comparison peers.

Table 9. Student Performance on Gates-MacGinitie Reading Tests By Comparison and Treatment Groups and Grade – 2008-2010

Measure/Grade	Comparison		Treatment	
	% Proficient	n	% Proficient	n
Grade 1				
2008	75%	133	70%	125
2009**	68%	146	83%	144
2010	76%	185	73%	188
Grade 2				
2008	71%	142	74%	82
2009	76%	120	73%	125
2010	72%	176	79%	193
Grade 3				
2008	74%	138	78%	152
2009	76%	127	66%	105
2010	74%	184	66%	185
Grade 4				
2008	65%	121	70%	139
2009	70%	147	76%	145
2010	71%	160	77%	147

X² tests the difference between comparison and Picturing Writing groups.

*p<.01; **p<.001

Overall Summary of Findings and Conclusions

Summary of Findings

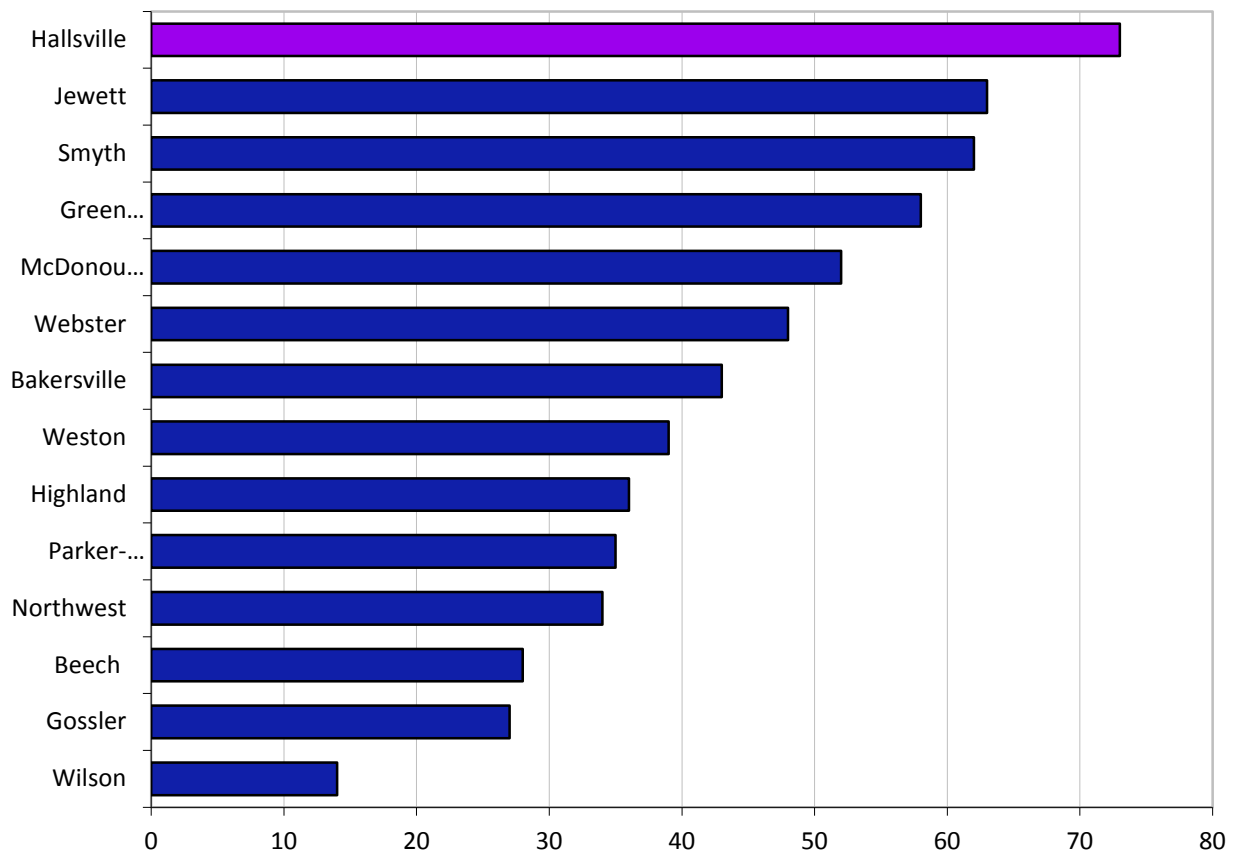
Overall outcomes from the evaluation findings were as follows:

- Within the independent study, statistically significant greater gains were documented each year in writing for treatment students at all grade levels as compared to the demographically matched comparison group. This held true for all at-risk subgroups as well, including “Below Benchmark Readers,” Special Education Students, ELL Mainstream Students, and ELL Magnet Students.
- In that same study, statistically significant greater gains were documented each year in visual literacy for treatment students at all grade levels as compared to the demographically matched comparison group. This held true for all at-risk subgroups as well, such as “Below Benchmark Readers,” Special Education Students, ELL Mainstream Students, and ELL Magnet Students.
- The performance of boys in the treatment group was comparable to the performance of girls in the treatment group. These findings suggest that boys, who tend to lag substantially behind girls in the area of writing across the nation, are more successful when participating in the hands-on visual and kinesthetic approaches to writing available within the Picturing Writing and Image-Making processes.
- The percent of student who scored proficient or above on the NECAP Grade 5 Writing Test was higher for schools that implemented PW/IM compared with the comparison schools, the district, and the state overall. In the fall of 2010 NECAP Writing results, which reflects the last year of student participation in PW/IM within the grant period, 60% of the students in the treatment group scored proficient or above as compared to 36% of students in the

demographically matched comparison group. This compared to 43% of the students across the district scoring proficient or above and 56% across the state.

- While differences among the three treatment school scores were apparent, the high fidelity school (the school with consistent administrative support) demonstrated the strongest results in writing, and particularly for those subgroups that typically perform poorly on standardized assessments. On the 2010 NECAP Fifth Grade Writing, boys in the high fidelity school scored equally well as the girls; Title I students and economically disadvantaged students also demonstrated impressive gains over their Title I and economically disadvantaged peers in the comparison school, across the district, and across the state.
- Exhibit 14 shows 2010 comparative writing results for fifth graders across the district, with our high fidelity school scoring top in the district despite the fact that 8 of the lower performing schools had a smaller percentage of students on free and reduced lunch.

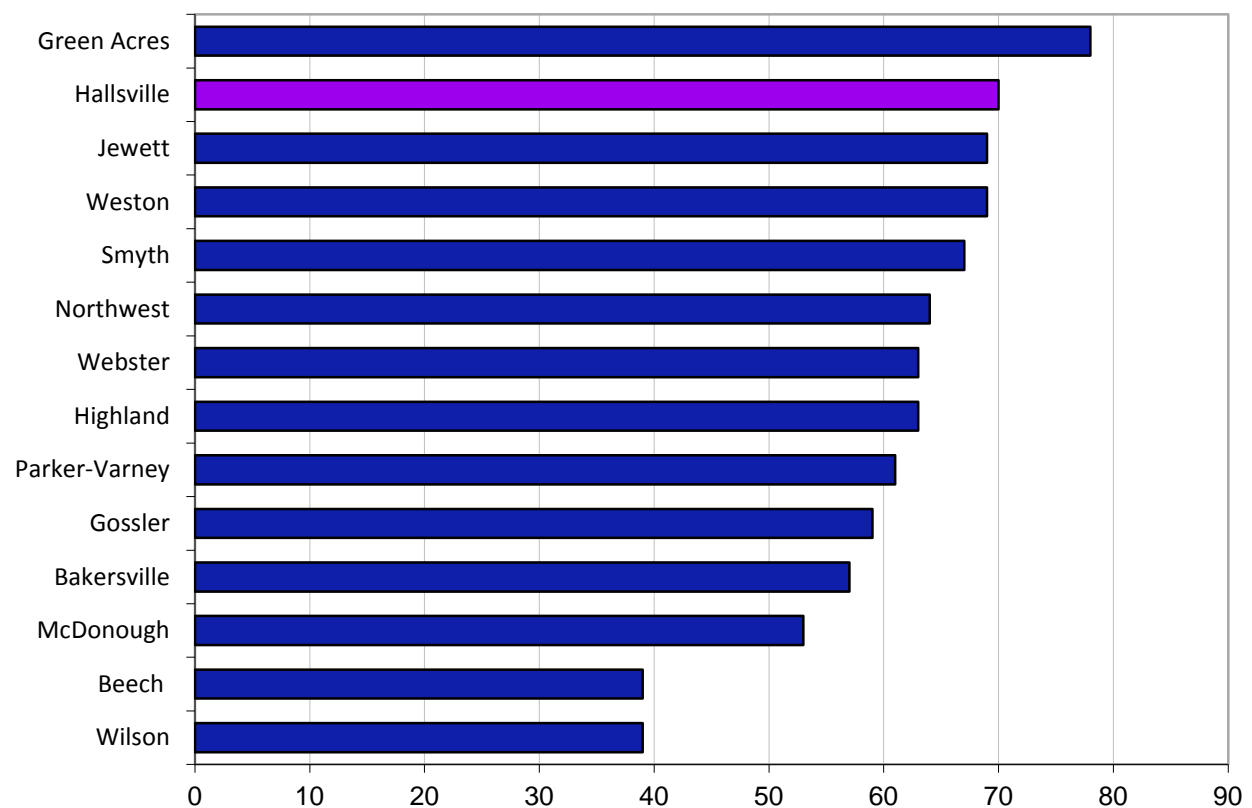
Exhibit 14. Grade 5 NECAP Writing Test 2010-2011, Percent Scoring Proficient and Above Across the District



District measures in writing also showed overall positive trends. The findings differed by grade levels. One consistent finding across all grades over the three years was that the treatment students scored higher than their comparison peers in the writing area of voice; the difference was statistically significant each year.

Findings from standardized reading tests (Gates McGinitie) varied grade level to grade level and test-by-test and, as a result, the reading data collected did not produce clear patterns in findings. Overall, the reading assessment data (Gates McGinitie and the state NECAP Reading Assessment) did not show statistically significant differences between student’s scores in the treatment classrooms and student’s scores in the comparison classrooms. However, strong gains in reading on the NECAP Assessments were apparent for the high fidelity treatment school. It scored second in the district in reading on the 2010 NECAP Reading Assessment. The top scoring school served the most privileged population in the district (only 15% of students on free and reduced lunch) while our high fidelity treatment school served a much more needy population (57% of students on free and reduced lunch).

Exhibit 15: NECAP Reading 2010-2011 Percentage Scoring Proficient and Above: Grade 5



Conclusion

Our data reveal that students in grades 1-4 benefit greatly from participating in the enhanced Picturing Writing/Image-Making art-and-literature-based literacy model in the areas of writing and visual literacy. As a result of this AEMDD-funded research study, we have found that when students are given PW/IM’s highly structured visual-verbal approach to literacy learning and thus two “languages” (the language of pictures and the language of words) to think, develop, and record their ideas, they display deeper levels of thinking and greater power of expression (both in pictures and in words). Through the power of transmediation (in this case, translating or recasting meaning from pictures to words), it is clear that encouraging students to think in detailed pictures first before they write helps to deepen their thinking and provides access to greater detail and description in their writing. We found that the highly

structured visual interventions provided by PW and IM are particularly helpful to at-risk learners such as “Below Benchmark Readers,” Special Education students, and English Language Learners for whom pictures serve as a universal language for thinking and expressing their ideas. Students who, for one reason or another, struggle with writing, have demonstrated that when they are able to utilize their visual abilities, their verbal skills are strengthened. Given the pressure to have all students perform well on standardized tests, these data will be both useful and important.

These positive outcomes bore out in the district and state writing data collected and analyzed as well. On the final NECAP Writing Assessment (2010), 24% more fifth graders who participated in the treatment prior to fifth grade reached proficiency or above in writing than their comparison peers. Additionally NECAP Writing data regarding traditionally underserved populations (boys, Title I, and Economically Disadvantaged) offer great promise for those students who have repeatedly demonstrated that they do not learn effectively through traditional straight verbal approaches.

With its documented positive results for at-risk populations, who tend to be less verbal and more visual in their orientation, the enhanced PW/IM model offers a viable alternative instructional model. This is particularly relevant, given the number of schools and districts across the nation that have been identified as low performing schools and the importance of writing as an essential literacy skill.

Additionally, with the growing number of immigrant children in our schools, documented success with English Language Learners will provide important data in determining future policy and instructional practices by giving teachers critical new tools for reaching all students, even those with very limited English proficiency.

There also has been a great deal of discussion among educators across the nation regarding how to close the achievement gap for boys in the area of language arts. Hands-on, art-based literacy methods have proven to be a good match for boys and align well with research on brain differences between girls and boys (Gurian, 2005).

Anecdotally, all students have loved participating in these highly structured, comprehensive art-and-literature-based literacy methods. Their level of engagement and motivation, particularly among those students who are not generally engaged by traditional verbal methods of instruction, was outstanding.

Students developed a new appreciation for picture books as mentor texts, both for their artwork and their writing. All students showed a great deal of pride in the published picture books they artfully created and that motivated them to want to read their own and each other’s books. This fostered an enthusiasm for reading, especially among those students for whom reading did not come easily. Seeing the success and enthusiasm of their children, parents were likewise very enthusiastic.

We encountered many unanticipated challenges that affected the impact of the treatment. Because Manchester School District had been identified as a “District in Need of Improvement” just prior to the beginning of our work in the schools, MANSD teachers had been given many additional mandated programs to implement. This left little time for teachers to fully embrace and implement the treatment. The reading portion of the treatment was never fully implemented which may have resulted in mixed findings on the district and state reading assessments.

We also experienced tremendous turnover in administration and staff. From the time of our early meetings with the district when we first received a commitment to move forward with grant writing to the end of the funded research period, Manchester School District had three superintendents. During the four-year funded period, within our three treatment schools, there were a total of six principals (three principals in one of the treatment schools, two principals in another, and one principal in the third). Incoming superintendents and principals were not as invested in supporting implementation of the model as those original administrators who first made the commitment to embrace this alternative approach to literacy learning. Additionally, changes to the NH State Law regarding teachers' retirement benefits resulted in an enormous turnover in teaching staff. During the course of the study, we witnessed approximately a 50% turnover in teaching staff within our treatment schools. Despite efforts to train new teachers as they arrived at each treatment site, teacher fidelity was greatly impacted by a combination of the turnover and lack of administrative support in two schools. Ironically, our highest performing treatment school in both reading and writing ended up being the one school with the highest percentage of students on free and reduced lunch but which had consistent administrative support. This same "high fidelity" school demonstrated impressive gains for boys as well as Title I and economically disadvantaged students.

Reflecting on these findings, we learned about the universality of appeal and effectiveness of this approach in teaching writing to a wide range of learners, including recent immigrant and refugee student populations. We also learned about the critical nature of administrative support on both the district and school building level for long-term school-wide adoption. We strongly recommend securing administrative buy-in before commencing a school-wide adoption. Strong, ongoing administrative support is essential to the success of ongoing implementation of the treatment.

References

- The 2000 National Assessment of Educational Progress finds boys one and one-half years behind girls in reading/writing (National Center for Education Statistics, 2000).
- Frankel, S. (2011). Picturing writing: Arts in education model development and dissemination (AEMDD) research findings, Retrieved from <http://www.picturingwriting.org/pdf/AEMDDFindings.pdf>.
- Gurian, M. (2005). *The Minds of Boys*. San Francisco: Jossey-Bass/John Wiley.
- National Center for Education Statistics. (2000). *National Assessment of Educational Progress: The nation's report card*. Washington, DC: U.S. Department of Education.
- O'Connor, S. (2010). Picturing Writing/Image-Making standardized test score data Exeter, New Hampshire, Retrieved from http://www.picturingwriting.org/pdf/ExeterData2010_update.pdf
- Olshansky, B. (2007). Impact of an art-and-literature-based approach on writing and reading skills of at-risk learners: sixteen years of evidence. *Arts & Learning Research Journal*, 23 (1), 1-30.
- Olshansky, B. (2008). *The power of pictures: Creating pathways to literacy through art*. San Francisco, CA: Jossey-Bass.
- Siegel, M. (1995). More than words: The generative power of transmediation for learning. *Canadian Journal of Education*, 20(4), 455-475.

Appendix A:

Instruments

Text Only Scoring Instrument

1. Lead sentences		
Sense of setting (time of day, weather, season, place)	_____	_____
Hook	_____	
2. Character traits or development		
Describes what character is doing	_____	_____
Describes (shows or tells) how character is feeling	_____	
3. Story development/Plot		
problem or anticipated event/drama or suspense	_____	_____
solution or sense of relief	_____	
resolution or sense of conclusion	_____	
sense of sequence/cohesiveness	_____	
4. Descriptive language		
strong adjectives	_____	_____
strong verbs and adverbs	_____	
sensory component	_____	
inclusion of detail (including similes, personification...)	_____	
transitional phrases	_____	
5. Overall quality		
Do words paint a picture in reader's mind?	_____	_____
Do words establish mood?	_____	
Does the piece have voice?	_____	
6. Overall use and quality of written information to tell the story		
Additional Characteristics:	_____	_____

1= none, 2 = minimal, traces, 3= abbreviated
4 = partially developed, 5 = fully developed, 6 = extraordinary

This scoring instrument was used to evaluate the quality of students' story writing.

Use of Visual Information to Communicate Ideas Scoring Instrument

1. Color

Color accurately represents setting elements _____

Color used to enhance mood _____

2. Texture

Texture used to represent story elements _____

Texture used to enhance mood _____

3. Shape

Shape/contour used to represent setting _____

Shape attempts to accurately represent key objects _____

4. Detail

Details used to enhance setting _____

Details used to depict character _____

5. Composition

Foreground/background provide information about setting _____

Placement/size creates sense of depth _____

Placement/size used to enhance sense of importance _____

Picture composition conveys meaning of text _____

6. Sequence

Sequence reflects changes in time of day/weather, etc. _____

Changes in perspective enhance the story _____

Picture sequence tells the story _____

7. Overall use and quality of visual information to tell the story _____

1= none, 2 = minimal, traces, 3= abbreviated

4 = partially developed, 5 = fully developed, 6 = extraordinary

This visual instrument was used to evaluate students' use of visual elements to make meaning and communicate their ideas.