

March 2006

Project RAISE, Reading and Arts Integrated for Student Excellence, Final Report

Nancy Andrzejczak

Follow this and additional works at: <http://digitalcommons.unl.edu/cehsgpirw>



Part of the [Education Commons](#)

Andrzejczak, Nancy, "Project RAISE, Reading and Arts Integrated for Student Excellence, Final Report" (2006). *Research and Evaluation in Literacy and Technology*. 3.

<http://digitalcommons.unl.edu/cehsgpirw/3>

This Article is brought to you for free and open access by the Education and Human Sciences, College of (CEHS) at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Research and Evaluation in Literacy and Technology by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Project RAISE,

*Reading and Arts Integrated for
Student Excellence*

Final Report

Nancy Andrzejczak
*Lake Elsinore USD
Project Director*

Dr. Guy Trainin
*University of Nebraska, Lincoln
External Evaluator*

Table of Contents

A. Summary of Study Results

- A.1 Project Overview
- A.2 Main Findings: Visual Arts and Writing
- A.3 Main Findings: Readers Theatre and Reading

B. Impact on Teachers

- B.1 Visual Arts and Writing
 - Visual Art Professional Development Plan
 - What we actually did
 - How we measured growth
 - Data Collection Methods
 - Results of Professional Development Methods
- B.2 Readers Theatre and Reading
 - Readers Theatre Professional Development Plan
 - What we actually did
 - How we measured growth
 - Data Collection Methods
 - Results of Professional Development Methods

C. Impact on Students—Visual Arts

- C.1 Assessment Plan
- C.2 How we measured objectives
- C.3 Year 1
- C.4 Year 2
- C.5 Years 3 & 4
- C.6 Summary/conclusion

D. Impact on Students—Readers Theatre

- D.1 Assessment Plan/Theoretical Framework
- D.2 Pilot Study
- D.3 Year 1
- D.4 Year 2
- D.5 Years 3 & 4
- D.6 Summary/conclusion

E. Implications for the Field and Recommendations for Policymakers

A: Summary of Study Results

A.1: Project Overview

What is Project RAISE?

Project RAISE is a federally funded Arts Education Model Program Development and Dissemination Grant (AEMPDD). The proposal was granted funding for three years from 2001-2002 until 2003-2004. The proposal had two strands, one investigating the impact of a visual art strategy and the second examining the impact of Readers Theatre on reading comprehension. The proposal was a project of the Lake Elsinore USD. The AEMPDD required grantees to work with schools with Kindergarten-8th student populations over 35% free/reduced lunch.

The original concept of RAISE was to improve arts (visual and theatre) achievement through instruction provided by the generalist teacher. We understood that in the context of budgetary constraints and a focus on the three R's, the concept of arts specialists or artists in residence was not sustainable. Arts instruction would be sustainable only if provided by the generalist teachers in elementary school settings. In this way, the arts become integral parts of literacy instruction. The goal was to ensure that the arts were part of everyday curriculum while leaving room for the integration of additional resources such as specialists and artists-in-residence as they become available.

This original plan called for improving the quality of arts instruction in the generalist classroom, through staff development on instructional strategies, arts education issues and standards based instruction and assessment. The project sought to increase the capacity of generalist teachers to teach the arts. The main vehicle was the creation and field-testing of staff development models. An additional goal was to improve the quality of arts education to the students, through artist in residence, partnerships with museums and art providers, and the creation of standards based curriculum.

What are the questions addressed?

- The AEMPDD asked all grantees to investigate three questions:
- What impact does arts integration have on the academic performance of students?
 - What impact does arts integration have on the artistic performance of students?
 - Does the impact of arts integration have special impact on the performance of at-risk students?

What were the main questions of Project RAISE?

- **Project RAISE** asked what impact does visual art as a prewriting strategy have on the academic and artistic performance of students, specifically in the realm of literacy and writing?

- **Project RAISE** asked what impact does Readers Theatre as an instructional strategy have on the academic and artistic performance of students, specifically in the realm of literacy and reading?
- **Project RAISE** asked were generalist teachers able to provide and to sustain quality arts education when provided with professional development?

A.2: Main Findings: Visual Arts

What did we learn about the impact of visual arts on the writing of students?

1. Students whose teachers used the program with high fidelity outperformed low fidelity classrooms in writing performance, aesthetic response writing, and the California Standards Test (CST) in English Language Arts.
2. Teachers who adapted the program to their own context maintained involvement and their students tended to out perform other teachers who did not contextualize the program.
3. Teachers reported that the program increased student motivation and improved writing and this was supported in the data.

From our study and research of how the student used visual art in the writing process four important themes emerged.

- Visual Art was inherently motivating and engaging for all students, especially English Language Learners and economically disadvantaged students, who could experience success on a level playing field.
- The student’s aesthetic perception was positively impacted through visual art instruction linked with writing.
- The student was able to elaborate on their ideas and communicate their thoughts and concepts holistically by creating visual art first.
- Both student and teacher could use the art as a basis for the compositional event. The student could reflect on the image finding ideas or sensory details. Emerging writers were able to “store” ideas in their paintings and then could use more of their working memory for the writing process. The teacher could use the art as a basis for the writer’s conference, thus increasing student and teacher clear communication.

What are the implications of these findings?

The implications of this study are that teachers should incorporate visual art teaching routinely into their language arts lessons, especially with younger students and students who have difficulty with English. The common practice of writing and illustrating a story does not have the value added that the art first approach does. The findings support this supposition. The use of text only as a methodology to teach language and writing does not take advantage of the natural use of interwoven symbols that children routinely use drawing and writing a story. By using art first students can

learn about visual art, which is inherently valuable for students. Furthermore, through the enhancement of verbal skills, the student gains not only academically but also culturally. Orr (2003) studied how the impact of student wealth (SES) on achievement is mediated largely by cultural capital. Indeed cultural capital becomes an important predictor of academic success and helps explain the persistence of achievement gaps. Integrated with this idea of cultural capital are factors of verbal facility including vocabulary (English, 2002). In fact, English (2002) hypothesizes that unless we find a way to bridge this cultural gap, the achievement gap will persist. This program provides student support for all three areas: cultural capital, verbal facility and vocabulary.

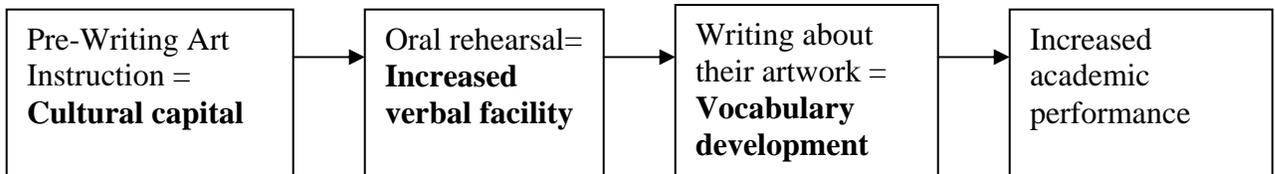


Figure 1: Alignment of the VIEW critical elements

How can these findings be used to improve student achievement?

Teachers should be encouraged to link visual images, museum images or art created by the student, to language. Students should be actively engaged in the writing process. Students who write on a regular basis will gain fluency and coherence in their writing.

Districts should be encouraged to support the arts in the classroom not only because they provide engagement for the students, but also because as this study demonstrates connections between text and image help students in understanding and using language.

A.3: Main Findings—Readers Theatre

What did we learn about the impact of Readers Theatre on the reading/literacy of students?

- Readers Theatre does have a positive impact on student’s reading achievement in terms of comprehension and fluency.
- Readers Theatre was difficult to implement in classrooms due to time constraints and preparation needed. This impacted the fidelity and sustainability.

What are the implications of these findings?

- Readers Theatre has some promise as an instructional strategy for reading. While the results were significant, the time, effort, and costs exceeded the benefits.
- Readers Theatre was a motivating and engaging strategy for students K-8. Middle school students were equally engaged as younger students. Middle school students reported that Readers Theatre helped them in word meaning and reading.

Some middle school teachers reported used the strategy with success for content area comprehension. The increased engagement impacted test scores.

How can these findings be used to improve student achievement?

- Readers Theatre can be used as a motivational strategy to engage students in reading. We believe that Readers Theatre used as periodic strategy will have significant impact on the classroom achievement.
- Readers Theatre can provide a model of prosody and fluency for struggling readers as they hear competent oral reading modeled by peers.

B: Impact on Teachers

B.1: Visual Arts and Writing

The initial concept was to evaluate a program supported by the University of New Hampshire, Picturing Writing (Olshansky, 1990, 1994, 1995, 1998). This scripted program appeared to be an excellent choice for evaluation and expansion of the program. The project used Picturing Writing for the first two years. At the end of the second year, some troubling issues began to surface. The first was the quality of the visual art was not increasing as predicted. The art instruction lacked any depth and was often repetitive and formulaic in nature. It seldom used correct art vocabulary and did not meet California Standards. An attempt was made to add supplementary lessons, but it proved difficult to add this overlay to the scripted format. Even more troubling was the increasing problem of teacher sustainability. The limited units available and their inability to meet content and grade standards had begun to take its toll. Teachers reported that while they saw success in student writing, they could not afford the time necessary to do the scripted lessons on top of their already rigid requirements.

At this point, the project leadership reassessed the Picturing Writing program. It was felt that while we had seen student success in writing measures that its limitations of scripted format, lack of depth and breadth in visual arts, and limited relationship to classroom standards in the content areas indicated that we needed to develop an alternative. We began work on creating a program based on research that followed a structure but allowed total teacher flexibility to adapt the structure to whatever standards or level they taught. This new program **VIEW**, Visual Integration to Enhance Writing also took on a new partner, the Smithsonian American Art Museum. This museum partnership was an effort to add further content depth to the visual art instruction. This partnership sought to infuse stronger visual art instruction impacting students cultural knowledge. Thus student data from the third year came from classrooms that were both trained in Picturing Writing and VIEW. However from the teacher survey, the number of teachers who adapted the program to fit their own needs, it would appear the vast majority of high fidelity teachers had moved into the **VIEW** format.

Primary Visual Art Professional Development plan:

- Expand and strengthen arts instruction in elementary & middle grades through staff development for teachers from throughout the district, but focus on one elementary school
- Train mentor teachers
- Develop, field test and disseminate staff development modules
- Develop and provide resources, materials and expertise to classroom teachers

- Provide professional development & practice using performance-based assessment and critical friends roundtable to provide feedback through mentor coaching model
- Provide opportunities for pilot teachers to attend and present at professional conferences & symposia
- Develop partnerships with museums to improve standards based instruction

What we actually did

To highlight the achievements of the project, we have listed the activities by goal and year. This will allow the reader to follow the development and formative nature of the professional development activities.

Goal 1: *Expand and strengthen arts instruction in elementary & middle grades through staff development for teachers from throughout the district, but focus on one elementary school*

- Year 1*
 - Trained 110 elementary and middle school classroom teachers in Picturing (PW) & Image Making (4 sessions)
- Year 2*
 - Provided 8 additional trainings in PW units (n=158 teachers) including all of staff at Butterfield Elementary (LEUSD) (Some teachers from year 1 repeated training)
 - Provided 3 after school refresher sessions, 4 meetings (LEUSD)
- Year 3 & 4*
 - Provided Teacher Action Research class. Provided training for elementary teachers at target site, district wide, and other districts. Total participants at all trainings, (n=429)

Goal 2: *Train mentor teachers*

- Year 1*
 - Attended two week Trainer training in PW (4 teachers)
 - Completed Project Zero training (Project Director & Co-Director Visual Arts)
- Year 3*
 - Informal training with key teachers to lead staff development
 - Developed units (Elements, Narrative Writing and Holiday Art) and presented at pre service days and state art conference.

Goal 3: *Develop, field test and disseminate staff development modules*

- Year 3*
- Collaboration & creation of VIEW (Visual Integration to Enhance Writing) using research from a variety of sources to meet California standards (6 lead teachers),
 - Provided: Introduction to VIEW (3 sessions 1 1/2 hours each: before school year) (n= 100 teachers) (LEUSD)
 - Provided one (before school) + one (during school) session of VIEW Elements & Applications (6 hours) (n= 76 teachers) (LEUSD)
 - Provided two Narrative Writing (12 hour) sessions (n= 41 teachers) (LEUSD)
 - Provided VIEW: E&A training to 50 teachers in Belmont-Redwood Shores School District
 - Provided training with SAAM in museum online resources (n= 8 teachers)
- Year 4*
- Expanded VIEW into Holiday Art (standards based visual arts instruction centered around holiday) (n= 5 lead teachers)
 - Provided Elements & Application (6 hours) –(n= 15 teachers LEUSD)
 - Provided Holiday Art (1 1/2 hours) – (n = 101 teachers LEUSD)
 - Provided training to 30 teachers in Belmont-Redwood Shores School District

Goal 4: *Develop and provide resources, materials and expertise to classroom teachers*

- Years 1-4*
- Developed library, provided for bulk ordering of materials, trained Artist-in-Residence
 - Provided release time for Project Co-Director to assist and support teachers at Butterfield
 - Developed video tapes to complement VIEW manuals in Year 3 & 4
 - Developed and disseminated Newsletters

Goal 5: *Provide professional development & practice using performance-based assessment and critical friends roundtable to provide feedback through mentor coaching model*

- Years 2-4*
- Developed professional development on performance-based assessment. In conjunction with the site administrator, we organized grade-level assessment in Year 2 and Year 3 of student

pre-writing and pre-art assessments as teachers began their annual planning

- Co-sponsored grade-level days, which allowed collegial planning and sharing. Teachers also visited other classrooms.
- Co-director and trainers acted as coaches and provided classroom support.

Goal 6: *Provide opportunities for pilot teachers to attend and present at professional conferences & symposia*

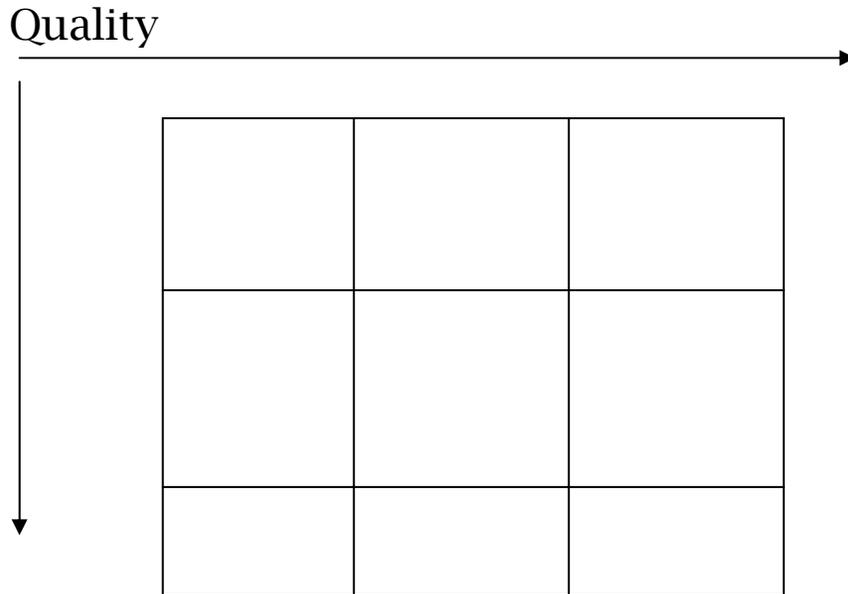
- Year 1*
 - Project directors and/or teachers participated at Arts Education Model Program Development & Dissemination (AEPDD) meetings
- Year 2*
 - Attended Arts Education Partnership (AEP)
 - Project directors and/or teachers participated at Arts Education Model Program Development & Dissemination (AEPDD) meetings
 - Presented results at California Art Education Association (CAEA)
 - Shared results at California Model Arts Program Network (MAP) meetings & ArtsWork Conference,
 - Presented at National Association of Elementary School Principals (NAESP),
 - Arts Education Partnership (AEP)
- Years 3 & 4*
 - Attended National Art Education Association (NAEA).
 - Project directors and/or teachers participated at Arts Education Model Program Development & Dissemination (AEPDD) meetings
 - Presented results at California Art Education Association (CAEA)
 - Shared results at California Model Arts Program Network (MAP) meetings & ArtsWork Conference
 - Presented results at American Education Research Association (AERA)
 - Presented results at Arts Education Partnership (AEP)
 - Presented results at National Art Education Association (NAEA)

How we measured teacher change & growth

Fidelity Matrix

The fidelity matrix was developed to give an accurate and reliable measure of teacher practice. We analyzed program fidelity on two dimensions: adherence to the critical elements of the program and time spent on program units as presented in Figure 2. The determination of teacher fidelity was based on Teacher Logs, Interviews, Surveys, Observations and Student Work Samples, representing a triangulation of different

sources. Inter-rater reliability for the teacher fidelity measure was calculated using Cohen’s Kappa, a statistic that takes into account chance agreement between raters. The final agreement was .91, which indicated a very high inter-rater reliability.



Time

Figure 2: *The Fidelity matrix based on Time and Quality*

Data Collection Methods

Surveys

We used both formative (year 1 & 2) and summative (year 3) teacher surveys. In the formative surveys, we focused on how teachers used the Picturing Writing program in their classroom. The Year 2 formative survey also contained teacher self-efficacy items. It was apparent that this survey alone would inflate actual fidelity of project teachers. It was at this that we developed the fidelity matrix. The Year 3 summative survey asked teachers to rate their continued usage of the visual/writing, training levels, and impact on practice. We were able to match this survey to classroom observations, teacher logs and student samples to a sample group of teachers. This time the survey results coincided with other measures giving a level of reliability and validity to the survey.

Interviews

We primarily performed informal interviews, which were designed to be less intimidating and more helpful in nature. A formal interview was held at the end of Year 3 for Diane Austin with teacher representative from the target schools.

Classroom Observation

The classroom/teachers were observed three times a year. The Co-director, Director, External Evaluator, Site Administration, and Diane Austin from the Office of Innovation and Improvement Project Management performed these observations. A form

was developed so we could easily record the classroom environment and teaching strategies. In some cases the observations were random and unannounced.

Teacher logs

The teacher logs were monthly calendars that recorded time spent and activity. They also had comments or needs section. The on-site Co-Director collected and quantified the results. In some cases the log was an accurate gauge of activity and in others it under-represented the activity.

Student Work Samples

During the first year, teachers voluntarily submitted work samples of all stages of the writing/art process. We performed a case study of one teacher's class, collecting a portfolio from each of the students.

In the second and third years, the project focused on the target sites, Butterfield Elementary and Terra Cotta Middle Schools. To make the collection process manageable, we asked teachers to pick 6 students (academic rating: 2 highs, 2 middles, and 2 lows). We collected both writing and art samples from these students. In year 3, we digitally collected the samples rather than color copies.

Results of professional development efforts

The results presented here draw mainly on the summative evaluation. We specifically address the main questions that were outlined in the proposal which generalize to many similar projects. The first question that will be addressed is insuring sustainability of project after funding is gone. Following the issue of sustainability we will discuss the changes in teacher practice that resulted from this project and items we have identified that still need attention. Finally we will discuss the importance of a supportive administrative environment.

In June 2004, all teachers who had participated in training at some level were surveyed. Out of the 196 teachers still working in the district, 141 responded. Teachers were asked questions dealing with continued use of the program, training attended, and impact on teaching practice. The teachers were given an opportunity to respond with written comments.

a. Insuring sustainability

The results of the summative survey show very clearly that the ability to adapt the curriculum to teachers' own use was a critical element in ensuring sustainability two and three years after initial training. Close to half of the teachers (46%) who adapted VIEW to their own use, use it regularly. Only 18% of those who did not adapt VIEW to their own use were still using it regularly after the end of the project. Over 90% of teacher who adapted VIEW to their own use were still using occasionally or regularly after the end of the project. In Figure 3, the average number of training days the high use group attended was nearly double the number of days. The Pearson Correlation between usage level and training days is .38 ($p < .01$).

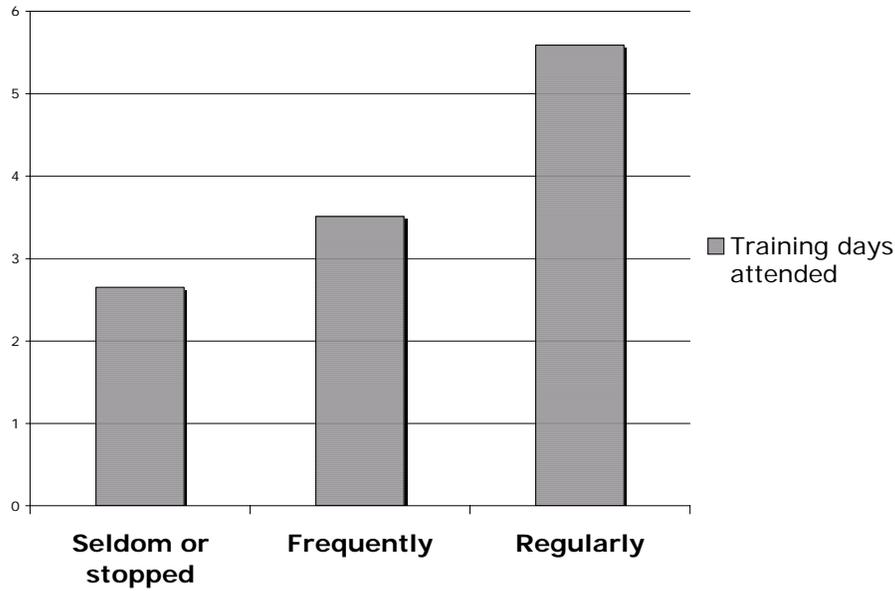


Figure 3: *Average Training days by usage group*

To ensure program fidelity, the project included professional support that we deemed necessary to insure short-term fidelity and long-term sustainability. We provided numerous training opportunities, on-site support, and afternoon refresher sessions, newsletters with concrete ideas and samples, artist-in-residence teacher training, teacher field trips to museums and manuals with videotape instructional support. The average number of training days attended for all teachers was 4. The teachers who attended greatest number of training days used VIEW/ Picturing Writing the most regularly. Only 19% of teachers of the teachers who use VIEW/PW frequently/regularly attended only a few training days ($n < 3$). While we are not suggesting causality, we are suggesting that to ensure sustainability the likelihood of that happening is increased with multiple training events.

b. Changes in teacher practice

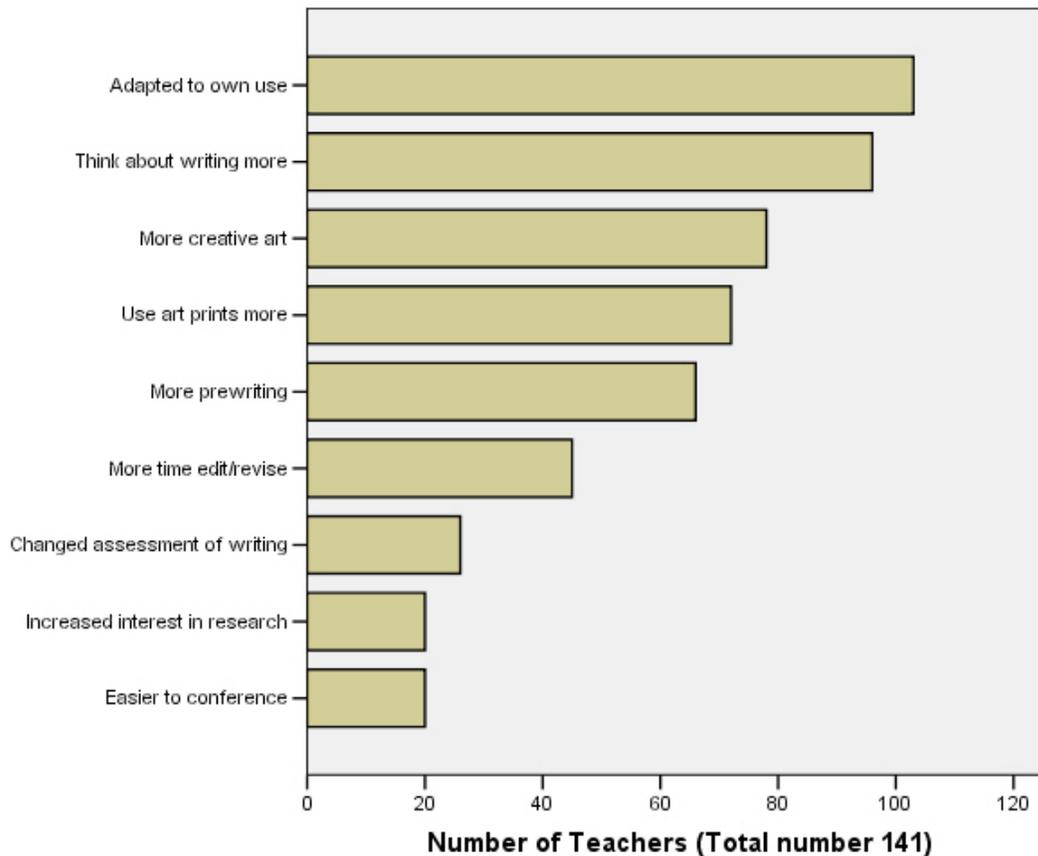


Figure 4: *Changes in teacher practice by numbers of responses*

The changes in teacher practice are divided into three levels: high, moderate, and low impact. In the high impact band on Figure 4, teachers indicated that they adapted VIEW to their own use, thought about using writing more, and used more creative art. The teachers who adapted the process were able to internalize it and frame it within their practice. The moderate band included the increased use of art prints, pre-writing strategies, and edit/revision with students. This middle band demonstrated that teachers not only thought about their writing instruction, but also were beginning to alter distinct portions of their instructional strategies. The lowest band dealt with more difficult issues such as embedded assessment, student conferencing, and teacher research. These topics received limited professional development during the project.

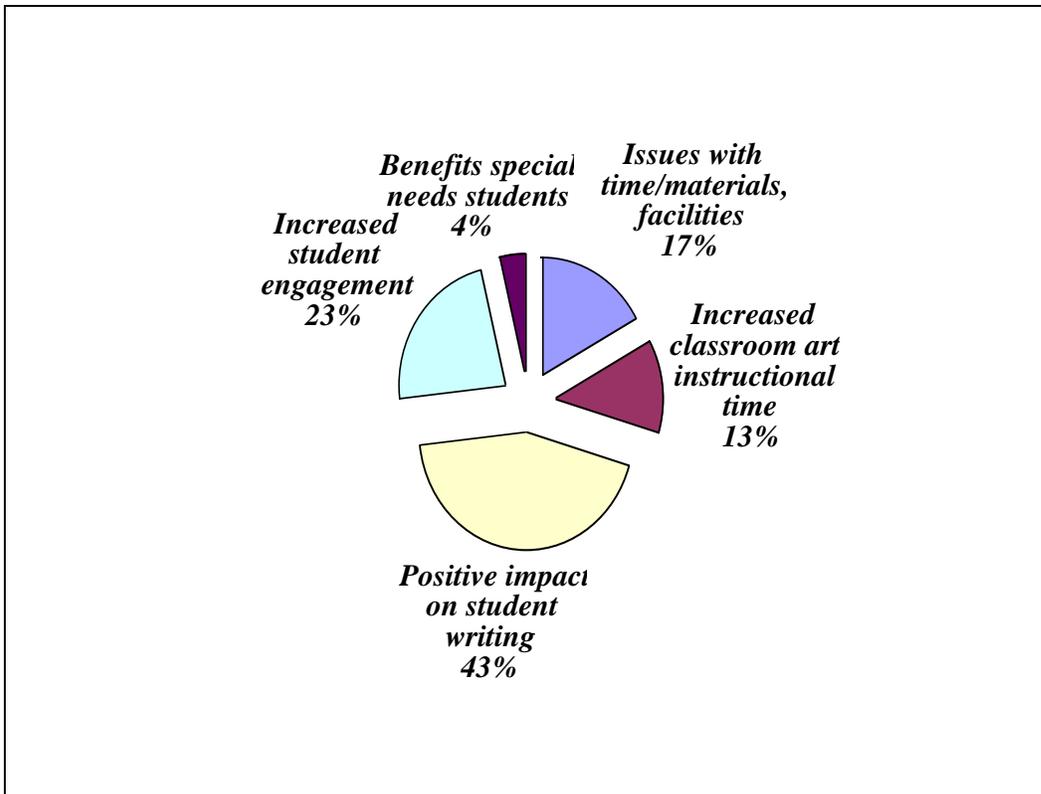


Figure 5: Analysis of teacher written comments on summative survey.

Of the 141 survey respondents, 82% (n=115) wrote written comments. The comments were coded and then grouped by similarities. The majority of comments (51%, n=95) reflected the impact of the program on teaching and student achievement. The subcategories included: student writing is improved, student vocabulary is more increased/detailed, useful in other content areas, and assistance in teaching special needs students. Motivational issues for students as well as the teacher was the next most frequent response category with the observation that students were more engaged or motivated in writing instruction, the single most frequent response overall (n=38) dealt with the motivation of students.

The goal of increasing or at the very least maintaining the use of visual art in the elementary classroom was met. Since 85% of the respondents used VIEW frequently or regularly, and the VIEW strategy employs visual art as the strategy, we can infer that visual art instruction is happening in those classrooms. When asked if they used more creative art in their classrooms, 55% responded yes. The other 45% indicated they didn't use more. Inferring from the 85% usage figure, there were at least 30% whose creative art time stayed fairly constant. When asked if they were using more art prints, 52% indicated they were using more prints. Lastly, 10% (n= 13) of the respondents wrote that they were providing more art instruction in their classrooms. For the group that didn't use VIEW, facilities issues such as access to water became a barrier to doing art. Upon comparison there were numbers of teachers who were able to overcome that barrier.

Of the teachers who reported that they did VIEW or PW infrequently or had stopped, time issues (n= 17) were the most frequent problem. Less than 5% (n = 7) reported that the approach was uncomfortable or did not coincide with their style of

teaching. In general, the approach we had taken over the 3 years had mediated potential problems in the area of materials supplies and training frequency and support. Time was the one area we could not solve. While training had tried to address ways to adapt VIEW into the grade level content areas, not all teachers who had been trained in years 1 and 2 had attended the VIEW trainings, which emphasized more adaptability and flexibility of approach.

c. Issues still needing attention

In terms of professional development, there are some issues that still are unresolved. Funding precluded the extensive use of substitute days for teachers to observe each other or for extensive grade level planning. In looking back, the limited grade level planning was very worthwhile and allowed the exchange of ideas and strengthened implementation. It was however, costly with each grade level-planning day costing nearly \$1000 per day and \$6000 for all grade levels to meet.

Further professional development is needed to two areas. First in linking assessment to instruction, we found that too often teachers would not use assessment as a roadmap to guide instructional modifications. Teachers tended to use assessment as solely the ending activity and then move on. This professional development would be best handled by having small groups of teachers plan instruction and then analyze student work as embedded assessment and discuss modifications in instruction.

Secondly, the art instruction was still an issue. The videotapes helped dramatically in the reinforcing some art content vocabulary usage. However, continued reinforcement of visual art vocabulary and skills is needed. The majority of teachers have had little or no training in visual art, however a desire for further training is indicated in the survey. Manuals and videotapes must continue to reinforce the elements and principles of art as well as media techniques. Teachers often would develop competency in a particular project, but not be able to transfer the skills to other projects. An example of this was the Holiday Art project, *Birches in the Winter*. This project, due to its high student success factor, is incredibly popular. Student examples of *Birches* were seen over and over across the district kindergarten through 8th grade. However, we did not see any projects that used those specific techniques with another subject matter or project. This appears to indicate that we need to further provide more generic art professional development with multiple possible applications to eliminate the “district-wide” approach of everyone doing the same thing.

d. Administrative Support

The ability to maintain the program and provide needed support at the target site was a direct reflection of the site administrator’s attitudes and behaviors. At sites where trained teachers were not supported through materials, time or attitude, teachers either decreased involvement or resorted to covert behaviors.

At the target site, the principal attended the on-site trainings and work sessions. Her active participation in evaluating student work created a positive collegial atmosphere. She would note positive implementation in classroom and teaching

observations. This reinforced the teachers' beliefs that they could adapt and use the strategy to support their individual classroom needs.

This support was most noteworthy the first year, when the school fell into school improvement status. Pressure from outside the district to modify the program could have adversely impact the research and program implementation. She served as a buffer and provided a safe environment for the strategy to take root.

She co-sponsored key grade level planning and program training days. We were able to use these days for student work evaluation as well as implementation planning. We saw teacher confidence increase as their ideas were positively reinforced by principal comments.

District wide, support was also provided by access to staff development funding and time. The program played an active role in staff development pre-service and in-service days throughout the year. Access to the program allowed the program to provide training and materials to almost 200 more teachers than just the target site. This broadening of the impact has led to a policy implications including the use of Title I funds to support materials for VIEW at some sites and the choice of district writing curriculum based on its compatibility with VIEW pre-writing.

B.2: Readers Theatre and Reading

Unlike the Visual Arts and Writing strand of Project RAISE, Readers Theatre did not begin with a previously developed program and professional development. The first year was a pilot and field-testing year. As the program moved into Year 2, the staff development manuals began to be developed and tested. By Year 3, there was a completed manual for both the one and two day workshops. The project had created two videotapes on Staging for Readers Theatre and Scripting for Readers Theatre. T

Primary Readers Theatre Professional Development plan:

- Expand and strengthen arts instruction in elementary & middle grades through staff development for teachers from throughout the district.
- Train mentor teachers
- Develop, field test and disseminate staff development modules
- Develop and provide resources, materials and expertise to classroom teachers
- Provide professional development & practice using performance-based assessment and critical friends roundtable to provide feedback through mentor coaching model.

What we actually did

To highlight the achievements of the project, we have listed the activities by goal and year. This will allow the reader to follow the development and formative nature of the professional development activities.

Goal 1: *Expand and strengthen arts instruction in elementary & middle grades through staff development for teachers from throughout the district, but focus on one elementary school*

- | | |
|---------------|--|
| <i>Year 1</i> | <ul style="list-style-type: none">• Trained 37 elementary and middle school classroom teachers in Readers Theatre. |
| <i>Year 2</i> | <ul style="list-style-type: none">• Trained 150 teachers. Provided in depth training for field test group of teachers. (n=5) |
| <i>Year 3</i> | <ul style="list-style-type: none">• Provided training for study participants at two sites and neighboring districts. (n= 88) |

Goal 2: *Provide training for mentor teachers.*

- | | |
|---------------|--|
| <i>Year 2</i> | <ul style="list-style-type: none">• Provided training for study mentor teachers at the Readers |
|---------------|--|

Theatre Institute.

Dr. William Adams, Professor, and founder of The Institute trained 37 teachers in two Readers Theatre workshops. The first workshop provided an introduction to RT principles and classroom use and the second supported teacher experiences as well as increased teachers understanding.

A new cohort of teachers (n=40) was trained in the instructional strategies of Readers Theatre. Training and dissemination occurred for neighboring districts and the Rockford project site. A one-day and two-day staff development seminars had been developed with teacher resources including videotapes for teacher/classroom use. The seminar has been piloted with positive results. The participating teachers last year helped create Readers Theatre scripts for use in the fourth grade based on basal text stories. These scripts are now available for any trained teacher in the district. (Total trained teachers 150) At the recent seminar, scripts were developed for other grade levels.

The participating teachers last year helped create Readers' Theatre scripts for use in the fourth grade based on basal text stories. These scripts are now available for any trained teacher in the district. (Total trained teachers 150) At the recent seminar, scripts were developed for other grade levels.

A new cohort of teachers (n = 88) was trained in the instructional strategies of Readers Theatre. Training and dissemination occurred for neighboring districts. A one-day and two-day staff development seminars had been developed with teacher resource manuals including videotapes for teacher/classroom use. Teachers developed classroom applications of Readers Theaters using basal text stories in 2-5th grade.

Goal 3: *Develop, field test staff development modules*

Goal 4: *Develop and provide resources, materials and expertise to classroom teachers*

Goal 5: *Provide professional development & practice using performance-based assessment and critical friends roundtable to provide feedback through mentor coaching model*

How we measured teacher change & growth

Fidelity Matrix

We found that the fidelity matrix as developed for the Visual Art/Writing strand worked equally well with teachers in this part of the program. Time spent doing Readers Theatre and the quality of instruction were valid measures of teacher change and level of implementation.

Data Collection

Unlike visual art, Readers Theatre performances tended to be more informal in nature and often were more ephemeral. We did collect some videotaped performances. The primary data collection tools for teacher change consisted of teacher surveys, student surveys, and classroom observation.

Results of professional development efforts

C. Impact on Students—Visual Arts

The main goal of the project was to demonstrate the effectiveness of the strategy, visual arts with writing, on student academic performance as well as skills in Visual Arts. The teacher level of fidelity to the program was a predicator of student achievement. It became extremely important to ascertain the level of implementation. Teachers were rated high, middle and low fidelity based on: evidence of critical elements: in student work and instructional practice, teacher logs, classroom observation, and survey responses. This evidence was used to place the teachers on a fidelity matrix, which accounted for both time and quality of implementation. Then using the matrix placement as an internal control group we were able then to perform a series of statistical studies comparing various student measures. The unit of study was the student and student nested in classrooms.

C.1: The Visual Arts and Writing Strategy assessment plan (Picturing Writing and VIEW):

- Assessing student growth in creating and responding to the arts using multiple measures.
- Assess the impact of the Visual Arts strategy on language arts achievement.
- Assess the motivational entry point aspect of the visual arts strategy.
- Assess if there is a correlation between academic growth and artistic growth.

C.2: How we measured the objectives

Pre-Post Writing Samples:

The writing samples were collected in Year 1 only at the end of the year. With funding for the project beginning mid-year, pre test samples were not available. In years 2 & 3, pre/post writing tests were administered. The writing test for each grade level used a similar prompt. Each student was provided a graphic organizer. Every classroom used a large color print of a work of art. Each grade level used the sample painting. The paintings were chosen for their subject matter and relationship to the program. Initially, we picked landscapes to reflect the instruction of Picturing Writing. It was observed that a landscape without people or animals did not provide enough action for detailed descriptions using active verbs and sensory details. Thus an effort was made for Years 2 and 3 to use more active paintings. Since each grade level was given a general writing prompt and an aesthetic response prompt, the print was also chosen based on the art elements in the aesthetic question.

The scoring guide was based on the work of Dr. Calfee, UCR. Initially in Year 1 the guide was 7-point analytic guide, adjectives, verb use, vocabulary, coherence, and grammar. The scoring guide was aligned with the California English Language Arts

framework. In years two and three, the guide was expanded to a 9 point guide. This expansion was done to increase variability and decrease the ceiling effect.

Aesthetic Response Section Pre/Post:

This response section was designed to generate some indication of art knowledge and vocabulary development and at the same time “piggy-back” on to the writing assessment so as to lessen the number of testing times needed. This piggyback approach had some major problems. First a number of teachers skipped this section. Even though it was clearly indicated in the administration instructions, teachers either ignored the instructions or didn’t want to devote the time needed. Secondly, students would answer this section briefly in what would appear to be a rushed manner. Nevertheless, the limited time for testing would probably make this piggyback approach necessary. Perhaps by changing the order of items and writing the aesthetic response first, it would increase the numbers of responses and quality. Additional training for the teachers in test administration might also solve this problem.

The other issue that surfaced immediately was that the students seldom used art vocabulary. This led to rather flat low scores on the aesthetic response section during the 1st and 2nd years. This was one of the contributing factors to the creation of VIEW. VIEW emphasized the elements of art and had an increased use of correct art vocabulary and skills. The videotapes were also an effort to improve art education instruction for the students and to provide reinforcement of skills, vocabulary and knowledge for the teacher as well.

The aesthetic responses in Year 3 appeared to show increased quality and a small group of second graders were scored. The scoring was based on a 7 point scoring guide. The guide was developed using anchor papers to isolate criteria through examination of selected student work samples.

Art Production Test:

In attempt to follow the NAEP test in addition to the aesthetic response, a creative response was attempted. Students were asked to paint a landscape. In the second and third years the same person in the multipurpose room administered this test. We had felt that the controlling of the administration of this test was important. However, the time and facility for the test created a rather false atmosphere and student work was often rushed and did not reflect the quality of work in their portfolios.

The artwork was scored originally on a 7pt rubric 0-6. Like the writing guide, the 6 points did not allow for enough variance and the scores tended to be mostly in the low half. So for the following year a 9point analytic scoring guide based on California Visual Art content standards was developed. A reliability and validity study was performed on the instrument in 2005. Up until the reliability study was performed scores seemed fairly stable as when supervised the graders provided a fair amount of reliability. However with minimal supervision the rater-reliability dropped. A scoring guide with illustrations was developed to ensure higher rater reliability. It was retested and found with the addition of visual and text the raters were able to increase reliability.

Standardized Measures used:

The standardized measures used in Year 1 included the SAT9 and the California Star test. We looked at student scores in English Language Arts and Reading. We had originally hoped to see impact on reading scores, but no evidence has been seen in the standardized scores. However, there was consistent impact on the English Language Arts scores of the Star test, which tests student proficiency in terms of California Standards. We also saw some impact on the standardized tests SAT9 and CAT6.

In year 1 we had access to the California 4th grade writing test and saw some impact on that instrument. While students from the treatment group outperformed the control, the increase was not statistically significant.

Student Work Collection:

It was determined to systemically collect student work in years 2&3. Six students were chosen from each classroom, 2 highs, 2 middles and 2 lows. Teachers were asked to submit student work each trimester. Student work was copied or digitally stored and the originals were returned. The major problem that surfaced was that teachers would forget and send work home. Even with monthly reminders, student work was lost. It would appear that it might be easier to collect everything in a hanging file box, copy at the trimester, and then send home. Remembering to cull out the work of six proved to be for many a task that was easily omitted. Nevertheless, we have a large amount of student work. A limited number of students have a 3-year portfolio demonstrating growth.

Compounding the collection of student work is that the target site has a high student turnover rate. Thus the number of students with complete portfolios for 2 or 3 years is much smaller than it would be in a school with a stable population.

C.3: Year 1: District Wide Field Test

Participants

The data from the experimental fifth grade classrooms was dropped to prevent a strong bias in this measure that is highly correlated with age. As a result the final sample for the analysis of writing products was 269. The fifth grade classrooms were included in the analysis of standardized test results, yielding an N=342, in this case 46 students did not have scores from both the current and previous years. The full sample was 52% male. Ethnic distribution was 46% Caucasian, 42% Hispanic, 8% African-American, and 4% other.

Methodology

Data Collection

Students were assessed in the spring. All students were presented with a visual prompt in the form of a piece of art and asked to write. Students were asked describe the scene presented and the story it depicted. Students were given time to use a pre-writing scaffold that helped generate connected vocabulary. Students in all classrooms were given 15 minutes for the pre-writing activity and 30 minutes to write their responses. The

written products were analyzed in three domains: quantity, quality, and the use of color vocabulary.

Writing quantity was assessed by counting the number of words. Inter-rater agreement on word counts was .98 all disagreements were resolved by a third rater.

Writing quality of samples was scored using analytic rubrics based on work by Calfee and his colleagues (2003). Writing was assessed on five discreet traits in reading: coherence, grammar, vocabulary, adjectives, and verb use the actual rubrics are presented in the appendix. The inter-rater agreement for scoring the samples was .89. While the individual traits are useful for feedback and instructional strategies a factor analysis found a strong one factor solution representing the data. We therefore created a single writing score by adding all scores. The reliability of this score was satisfactory $\alpha=.90$.

The use of color words was used as a proxy for sensitivity to artistic features, albeit a very rudimentary measure. Two measures were used: frequencies of using basic color words i.e. blue, red, and frequencies of using more advanced color words and analogies such as maroon, velvet black.

State-wide standardized testing (SAT9) were also analyzed to examine transfer. The scores for reading and writing are based on the augmented test that included items beyond those of the national sample to fit the California Standards. The reading test included items focused on word analysis and vocabulary development (52%), reading comprehension (34%), and literary response and analysis (14%). The scores for reading and writing are based on the augmented test that included items beyond those of the national sample to fit the California Standards. The reading test included items focused on word analysis and vocabulary development (52%), reading comprehension (34%), and literary response and analysis (14%). The language test includes items addressing vocabulary, semantic and grammatical structures.

Treatment fidelity is essential for establishing the internal validity of any study. When conducting research in schools it becomes an essential part in interpreting the results. The fidelity matrix presented in figure 1 illustrates the two main components of fidelity: time of implementation and the use of critical program elements. The four critical elements in Picturing Writing were: exposure to both art and writing on the topic of choice, art creation before writing, oral rehearsal, planning the written product, and connecting the written product to the art. Time was defined by curricular units i.e. full implementation of *Time of Day*.

The data used in establishing the time and the use of critical elements included observations, teacher logs, and student products. Observations were conducted by project coordinators and focused on visible products and teacher debriefing. Teacher surveys were administered twice, once at the middle and again at the end of the school year. Teacher logs were handed in monthly. Phone interviews were conducted in cases when other evidence was lacking. All reports were triangulated by analyzing student products. Inter-rater reliability for the teacher fidelity measure was calculated using Cohen's Kappa, a statistic that takes into account chance agreement between raters (in this case 33%), was .94 indicating a very high inter-rater reliability.

Results

An ANOVA was conducted to assess the impact of the picturing writing program on the quantitative measure of writing length. Adjusted mean results are in Table 2. There was no interaction effect between grade level and experimental group. Both main effects were significant. The main effect of age was significant as expected $F(2,266)=25.1$, $p<.0001$, $MSE=822.2$. Experimental group was also significant $F(1,266)=9.4$, $p<.005$, $MSE=$ with an effect size $d=.37$.

An second ANOVA was conducted to assess the impact of the picturing writing program on the qualitative measure of writing (combined score). Adjusted mean results are in Table 2. Here too there was no interaction effect between grade level and experimental group. Both main effects were significant. The main effect of age was significant as expected $F(2,266)=19.3$, $p<.0001$. $MSE=10.1$ Experimental group was also significant $F(1,266)=9.1$, $p<.005$, with an effect size $d=.35$.

In writing about art students participating in the project used significantly larger vocabulary when discussing when discussing a painting than control. An example of a robust response is “The sky looks like dark green army material. The hills look like dark brown chocolate milk” (Eric, 3rd Grade). Project students used complex descriptions of colors significantly more than controls $F(1,284)=24.0$, $p<.0001$, $d=.58$. Overall project students wrote longer texts than controls. The overall effect was $F(1,284)=18.7$, $p<.0001$, $d=.51$. The effect was especially large in the fourth grade where students in the project wrote on average 50% more than their classmates.

SAT9 results were analyzed using a hierarchical linear model. The use of the hierarchical model was necessary since students in the program were nested within classrooms. Spring Language scores were modeled at two levels. On the student level scores were predicted by the previous years’ language scores. On the classroom level program participation was used as well as a random classroom effect. The fixed effect for program participation was significant $F(1, 321)=4.21$, $p<.05$.

A similar analysis was conducted to examine the impact on standardized reading scores. The effect here was even larger $F(1, 321)=6.29$, $p<.05$. Figure 2 presents the results converted to mean percentile scores (analyses were conducted using NCE).

Year 1: Project students performed better on the state-wide writing assessment $F(1,385)=7.77$, $p=.006$, $d=.28$

Transfer to standardized measures of language is expected to be smaller than impact on direct writing measures. Despite this project PW students performed considerably better than control students and the district as a whole on the language part of the SAT9 (Figure 2) even after controlling for language scores from the previous year $F(1,1016)=6.32$, $p=.012$, with a relatively small effect sized $d=.05$.

C.4: Year 2: Target School Study

Qualitative Case Study of Student Writing with Art.

Methodology and results

In the beginning of year 2, a small case study was performed on two students. Students were observed and interviewed on videotape. They were asked about their work

and their processes. The observer also questioned the parents and teachers about student work, attitudes, and behaviors. The responses were coded and important themes emerged. This study will be included in the appendix.

This study looks at the benefits of integrating visual art creation and the writing process. The qualitative inquiry uses student, parent, and teacher interviews coupled with field observation, and artifact analysis. Emergent coding based on grounded theory clearly shows that visual-art creation enhances the writing process. Students used more time for thought elaboration, generated strong descriptions, and developed concrete vocabulary. The advantages of using production of art and artwork in the pre-writing process provided a motivational entry point, a way to develop and elaborate on a scene or a narrative. This study shows that the benefits of a rich visual art experience can enhance thought and writing in response to the finished artwork.

Quantitative Study

Methodology

Participants

The study was conducted in an elementary school in a rural Southern California school district. The study included 3rd – 5th grade teachers (N=24, 7 male teachers). A sample of 397 students nested in the 24 classrooms was drawn. In the sample 62% were male. Ethnic distribution was 48% Caucasian, 39% Hispanic, 5% African-American, and 8% other.

Data Sources

Teacher fidelity data included observations, surveys, phone interviews, teacher logs, and student products. Observations were conducted by project coordinators and an external evaluator and focused on visible products and teacher debriefing. Teacher surveys were administered twice, once at the middle and again at the end of the school year. Teacher logs were handed in monthly. Phone interviews were conducted in cases when other evidence was lacking. All reports were triangulated by analyzing student products.

Student achievement data in this analysis were writing sample scores. Writing samples were scored using analytic rubrics based on work by Calfee and his colleagues (2003). Standardized writing prompts were used in fall and spring of the school year. Visual art was evaluated using a standard based rubric created by the researchers an example is the scoring

Procedure

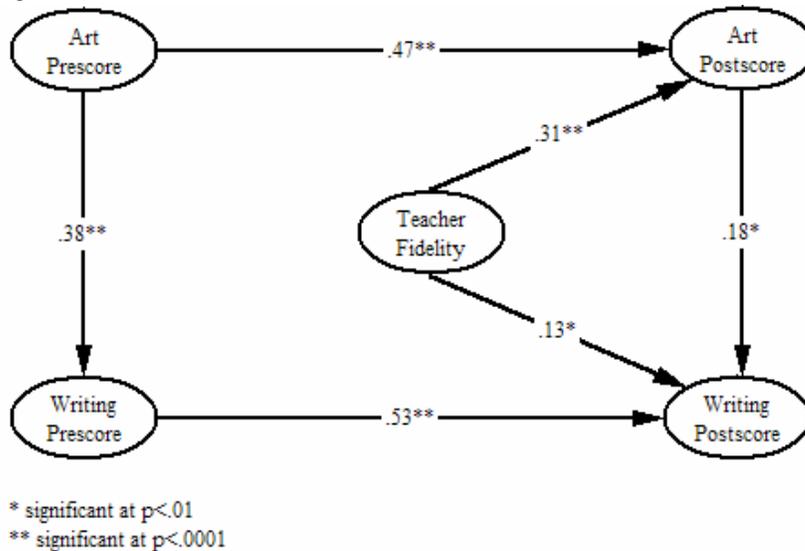
Program fidelity data was analyzed on two dimensions: adherence to the core ideas of the program, and time spent on program units as presented part B. To provide criterion related validity, student achievement was analyzed using mixed linear analysis with grade fidelity as fixed factors, a random teacher intercept as a random factor, and writing pre-scores as a covariate.

Results and Conclusions

Inter-rater reliability for the teacher fidelity measure was calculated using Cohen's Kappa, a statistic that takes into account chance agreement between raters (in this case 33%). The result was .94, indicating a very high inter-rater reliability. The

mixed linear model analysis used an unstructured covariance structure. All fixed effects were highly significant: Fidelity $F(1,384)=19.74$ $p<.0001$; Grade $F(1,384)=38.34$ $p<.0001$; as well as the interaction of Grade and Fidelity $F(1,384)=18.85$ $p<.0001$. The main effect for grade was expected.

Figure 1



Results

The main effect for teacher fidelity was followed by inspection of the coefficients and provided support to the validity of the fidelity measures showing that student achievement was increased as fidelity increased. The interaction effect showed that differences in teacher fidelity were significantly more important in 2nd grade than 4th grade. It is hypothesized that the effects of this particular program are greater in the earlier grades as the art help students manage cognitive resources as they write. The cognitive load during writing is decidedly heavier for younger students who possess less working memory capacity and less fluency in writing mechanics.

The data was further analyzed in a latent variable model. The model in Figure 1 explores the relationship between art production and writing with teacher fidelity as a contributing factor. All paths were significant, the Goodness of Fit Index was .91, CFI=.95, RMSEA=.60 with a confidence interval between .51 and .68 all indicating a reasonable fit between model and data.

The model clearly shows that teacher fidelity was a significant predictor of student achievement in both art and literacy.

Educational Importance of the Study

For research-based programs to be truly scalable, research and evaluation must provide valid and reliable ways to measure fidelity. Researchers and evaluators move on the continuum from the controlled lab environment to controlled school sites, design experiments (Calfee et al, 2001), and evaluation studies. This study has shown that as professionals guide educators, they can and should explore the issue of fidelity to ensure

the validity of their causal argument. Causality is not the only concern, however, an exploratory approach, as presented here, performed in early stages of scaling up can show which aspects of the program are the key to its success. An additional contribution of the approach presented here is the creation of fidelity tools that will allow administrators and educators to assess the efficacy of the future implementation of educational programs.

C.5: Year 3 & 4

Methodology

Participants This study focused on the entire second grade (n = 146) of a school in Southern California. The school has over 70% of the students on free/reduced lunch and approximately 25% of the students are English Language Learners. The high student mobility presented some difficulties in maintaining sample size over the school year (Students present for both pre/post tests n=76). The teachers were participants in a longer 2-year study and had attended between 3-14 days of training in the use of art with writing. The teachers provided all classroom instruction.

Procedure: Teachers were asked to intertwine the visual art instruction with the regular district adopted writing curriculum. They were asked to follow the critical elements of the process, (1) Visual art lesson and production, first; (2) Oral discussion of art work; (3) writing and editing; and (4) provide informal or formal publication.

Data Sources

Instructional delivery: Teachers were asked to provide monthly time logs and samples of student work. In addition, classroom observations and both formal/informal interviews were conducted.

Student Achievement: (1) Students were pre/post tested in art production. Each student was asked to paint a landscape. All students were provided crayons, watercolors, and various texture producing items (toothpicks, tissue, salt, etc). Student work was scored on a 9-point analytical scoring guide, examining color, space, composition, and media skill. The scale was developed using state standards. (2) Students were pre/post tested in two writing tasks (Trainin et al., 2005). The first task involved a narrative or descriptive response to the same print. The second task involved a specific prompt designed to solicit an aesthetic valuing response using the same print. Student writing was scored on a 9 point analytical scoring guide, adjective choice and usage, verb choice and usage, general vocabulary, coherence, and grammar/writing conventions. Scale was based on writing scales developed by Dr. Calfee, University of California Riverside and aligned with the state English Language Arts standards. The aesthetic response was scored using an 8-point holistic scale focused on the use of visual arts vocabulary and the ability to analyze a work of art.

Results and Conclusions

Based on the teacher logs, interviews, surveys, classroom observations and student

work samples, teachers were divided into two groups: (1) high fidelity—those teachers who followed the critical elements and provided regular art/writing instruction throughout the entire year and (2) low fidelity—those teachers who missed components of the critical elements and did the program infrequently or seldom.

Table 1: Comparison of Pre and Post Art Writing Scores

Fidelity	Pre-Art Writing		Post-Art Writing	
	M	SD	M	SD
High (n=39)	2.31	1.13	3.77	.1.16
Low (n=37)	2.30	.88	2.32	1.30

Using the aesthetic rubric, a sample scoring was done of 76 students from 7 classes, four high fidelity and three low fidelity. The pretest means of all classes were virtually the same with a minimal .01 difference (SD 1.13, .88) (refer to Table 1). The mean score of the high fidelity group grew to nearly 4 (M=3.77). A repeated measures analysis of the art writing scores was performed and found to be significant ($F(1,74) = 14.07, p < .001, r^2 = .16$.) This is not surprising considering the low fidelity group provided limited instruction in visual art concepts. From the results for the creative art production pre/post samples, we found that within subject contrasts revealed that students in the experimental group had superior performance to those in the low fidelity group $F(1,78)=23.93, p<.001, r^2 = .24$. The important consideration is that this same group outperformed the other group on the more traditional writing prompt as well. In both writing length and coherence the experimental group outperformed the lower fidelity group. (Coherence: $F(1,110) = 4.42 p < .05, r^2 = .04$; Quantity $F(1,110) = 10.86, p < .001, r^2 = .09$) Thus the high fidelity group was able to outperform the low fidelity group in the overall score as well in these critical areas and gain a level of cultural and aesthetic awareness.

In the second grade group, there is a statistically significant correlation between the students' high writing quality and quantity scores and the high aesthetic response. This certainly does not indicate a causal relationship. It does, however, suggest that students who spend time developing the skills and vocabulary to respond to an aesthetic question do not then show a decline in their overall academic performance in writing. It may also suggest some of the impact of cultural capital as suggested by Orr (2003) and English (2002) on academic achievement.

Table 2: Correlation coefficients for Relations Among The Post Writing Quality and Quantity Scores and Aesthetic Response Writing Scores

Measure	Aesth. Post Writing	Post Writing	Post Writing Quant.
Aesthetic Post Writing	--	.38**	.40**
Post Writing (Avg.)		--	.60**
Post Writing Quant.			--

** $p < .01$

The impact was tied to teacher fidelity in that the strongest results came from classrooms with the best implementation. It is difficult however in this study to say whether there is a causal relationship. Since fidelity played such an important role, do

these high fidelity teachers provide a richer classroom environment? Is the difference in scores due to the overall teacher quality as opposed to the intervention? Based on the overwhelming teacher survey response, the teacher's responses (district survey n=141) would seem to indicate that for the majority of teachers the visual arts and writing program did play a role in increased student achievement. Nevertheless, the data would support a future study that would isolate the components of the study in randomized trials.

C.6: Summary/Conclusion

In terms of academic impact, in all three years the study found a positive impact of the visual art on student language and writing. The impact was tied to teacher fidelity in that the strongest results came from classrooms with the best implementation. It is difficult however in this study to say whether there is a causal relationship. Since fidelity played such an important role, do these high fidelity teachers provide a richer classroom environment? Is the difference in scores due to the overall teacher quality as opposed to the intervention? The overwhelming teacher survey response would seem to indicate that for the majority of teachers the program did play a role in increased student achievement. Nevertheless, the data would support a future study that would isolate the components of the study in randomized trials.

However in general these common threads reach across the various results. The students found it easier to write when they had time to work through their thinking while painting. The painting in front of them served as a stimulus as well as point of reference for the writer's conference. This point of reference is particularly important with English Language Learners who might lack the words but can create the image. The image then serves as a bridge between student and teacher. There was both engagement and an increased sense of personal investment in the writing process. Students would be more likely to search the thesaurus for just the right word for their own painting. Teachers repeatedly wrote or spoke about the increased vocabulary and specificity of word choice. Lastly children became eager and most importantly more fluent writers.

The sacrifice of the discipline on the altar of integration is an important issue in the field of art education. Eisner(2000) often warns that art becomes the handmaiden of the other content area. The quality of the art and aesthetic responses in the first two years did improve somewhat. Since the original program was not an art program it seldom used specific art vocabulary, historical/cultural context, and aesthetic valuing. It did stimulate more creative production into the classroom, but after an initial boost the visual art began to stall. The creation of VIEW and subsequently Distant VIEW worked to provide stronger instruction in visual art. Distant VIEW in particular with the museum connection brings children into a world they have never seen. Abigail Housen's (2001) study of aesthetic responses of children showed that younger students benefited more than older children, our experience parallels her experience that the earlier this aesthetic work begins the stronger the impact. The addition of videotapes and expansion of media also increased the diversity and quality of creative products that were observed. Thus in terms of the second question, integration can be achieved that does not sacrifice artistic achievement, but special supports must be provided for the generalist teacher.

Lastly, did the program show a benefit for at-risk student populations? In general the program had a positive effect. Both anecdotally and in the statistical comparisons, the ELL subgroup appeared to benefit from this approach. This would be congruent with the theories concerning the impact of cultural capital on student achievement and the instruction of second language learners. Cummins(2000) in his study of bilingual students proposed that certain instructional strategies would facilitate bilingual learning. One such support is the use of pictures, thus the finding of a positive impact on bilingual students would be supported by studies on bilingual instruction. Further study of the program on bilingual populations is warranted. This strategy is easily implemented and does show promise as mediator of language in the classroom.

Therefore, the project did show positive findings for all three questions. Further questions are raised and there is a need to perform randomized studies to show causality. Nevertheless, we can state conclusively, that the doing visual art was not detrimental, it provided a natural bridge to literacy. As human beings we are a visual species; images are powerful communicators. To capitalize on the synergy between verbal and visual symbols to engage children in language is a powerful tool. The value added component makes this approach even more important for our at-risk students.

“We do not describe the world we see: we see the world we are able to describe.” Senge (2002).

D. Impact on Students—Readers Theatre

The objectives of this report are: (1) to discuss the efficacy of Readers Theatre as an instructional component in the elementary and middle school classroom and (2) to discuss the challenges of scaling up intervention research. In a series of designed experiments, we analyzed the impact of Readers Theatre as an instructional strategy to improve oral reading fluency and comprehension. Study research designs progressed from a pullout intervention with a single expert delivering instruction to a scaled-up model with multiple grade levels and teachers. Throughout the four studies, Readers Theatre demonstrated a significant impact on student achievement and self-efficacy.

The scaling up process not only investigated the efficacy of the approach, but also began to come to grips with the ever-increasing issues of professional development and implementation fidelity. These issues of scalability are essential for the translation of research into practice.

D.1: Pilot Study

Methodology

Participants

One hundred and eighty six fifth-grade students (52% Female 48% Male) from three schools in the Lake Elsinore school district participated in the pilot study. The sample included a high percentage of low SES (65%) and English Language Learners (23%).

Procedure

Students in each classroom were divided into tertiles (low, middle and high achievement) and then randomly assigned to treatment or control groups. A poetry selection and a narrative text were selected from district basal text. Classroom teachers taught these texts only when the experimental children were not in the classroom. The arts specialist provided two-1 -hour sessions a week of instruction on Readers Theatre techniques, dramatic reading of poetry selections, and scripting, repeated practice, and performance.

Data sources

Student achievement was measured using three measures. The first was the San Diego Quick, which is a set of graded word lists used to determine word recognition (Ekwall & Shanker, 1988), The second was a computer adaptive reading comprehension assessment tool developed by Renaissance Learning, Inc (2001). These two were administered as pre/post assessments. The third measure was the basal unit test, Boston: Houghton Mifflin, Reading, California, 5th grade, 1999.

Results and Conclusions

The Readers Theatre group had significantly higher assessment scores in all measures. In reading recognition (San Diego Quick) analysis of covariance (controlling for initial achievement) showed a significant effect $F(1,185)=4.38, p<.05$, with an effect size (Cohen's d) $d=.31$. For the Star comprehension test, analysis of covariance of the reading comprehension showed $F(1,185)=4.1, p<.05$, and an effect size $d = .30$. An analysis of covariance on the unit test using the Star pre-test a covariate showed $F(1,185) =13.01, p<.001$, and a moderate effect size $d = .53$.

The conclusion was that a highly controlled application of Readers Theatre for a limited time produced moderate effects in reading recognition and comprehension. The effects were larger for curriculum-based assessment (unit test) and smaller for standardized assessments. The study showed a clear advantage for the use of Readers Theatre over standard reading instruction.

D.2: Year 1~Field Test in Middle School

Methodology

This field test was conducted with a group of Middle School Students who were enrolled voluntarily in an elective. The 7th and 8th grade elective class ($n= 25$) focused on both Picturing Writing and Readers Theatre. The goal of this field trip was to test the potential unit items and to see how middle school students' especially low readers would respond to this reading intervention. Art specialist teaming with the regular classroom elective teacher provided instruction.

Procedure

Students were tested using a computer adaptive reading comprehension assessment tool developed by Renaissance Learning, Inc (2001). Students were to be tested in October, March and then in June. A computer failure destroyed the June scores. The students participated in a survey asking their self-evaluation of the Readers Theatre class.

Results

In the survey of middle school participants, 68% felt Readers Theatre helped their reading comprehension by being able to chunk phrases into comprehensible units. Half of the students felt that Readers Theatre motivated them to read more, developed their vocabulary and provided a sense of ownership of the materials they read.

The work with the middle school students developed a curriculum to be used with that age group that was high interest and addressed both reading fluency and comprehension needs. Reading assessment results showed a small advantage for students who participated in Reader's theatre activities.

D.3: Year 2

Methodology

Participants

Year-two participants were 146 fourth grade students from four schools in Lake Elsinore school district participated in the study. Two of the schools had less than 17% free/reduced lunch (High SES) and the other two had greater than 70% of free/reduced

lunch students (low SES). A sub-sample of 9-11 students was selected from each classroom representing low, middle, and high achieving students.

Procedure: In this study, the arts specialist trained the teachers in Reader Theatre techniques. The training involved two days of training on general Readers Theatre concepts and delivery. The study teachers then attended a specialized professional development session on the study procedures. Participating teachers chose the two focus basal narrative texts, Reading –California Grade 4, Boston: Houghton Mifflin (2001). They learned the scripting techniques developed by Adams (2003), which allows the teacher (or students) to adapt the basal text into a Readers Theatre script without modifying a single word in the original text. These fourth grade teachers then provided the classroom instruction in this scripting technique to the students. The teachers were randomly assigned to an experimental block in an alternate block design.

Data sources

There are two data sources, a curriculum-based measurement of Oral Reading Fluency (ORF) with grade level text and unit tests (Houghton Mifflin, 2001) for the two separate narrative texts. Each unit test includes vocabulary, explicit and implicit comprehension questions.

Results and Conclusions

The group using Readers Theatre in each block had significantly higher results. We found that in the first block, the experimental group had superior performance. The first covariate was the pre-Oral Fluency Test and the second block covariate was the first unit test. The first block analysis of covariance showed $F(1,45) = 48.4, p < .0001, d = 2.0$. For the second block, the analysis of covariance showed $F(1,185) = 6.51, p < .001, d = .37$.

In highly controlled applications of Readers Theatre, the curriculum-based results were highly significant. The large effect size in the first block is a result of a ceiling effect in the outcome measure, i.e. many students in the experimental group scored 100% correct on the test. The relatively moderate effect size in the second block is more in line with results of the first study, but also may reflect carryover effects from the first block. Analysis of the oral fluency results showed gains for all students for the two blocks with benefits for all ethnic and language groups.

This study showed that teachers could effectively Readers Theatre components as part of their regular reading class instruction. It further strengthens the conclusion that this approach helps all students not just struggling readers. It is important to note as in Study 1, the benefits are not limited to reading fluency but extend to comprehension and vocabulary.

D.4: Year 3

Methodology

Participants

This study included 570 students from 2nd -7th grades (males, $n = 297$) from four schools in Lake Elsinore school district. Experimental teachers ($n = 15$) were selected from a volunteer pool. Control teachers ($n = 11$) were solicited from the schools. Experimental group included 317 students.

Procedure

The experimental volunteer teachers attended a two-day Readers Theatre training and one day pre-experiment orientation session. The experiment group further attended a second day during the middle of the study to review early results and to focus implementation. The experimental teachers provided 1 hour a week of Readers Theatre activities for the study period. They were encouraged to script stories as well as to engage in choral reading and prepared script performance. Teachers kept a monthly time log and were observed teaching Readers' Theatre in the classroom. Both experimental and control groups were to use the district mandated reading program by Houghton Mifflin (2001).

Data Sources

An outside assessment team administered all instruments three times during the academic year (fall, winter and spring). The instruments used included a curriculum-based measurement of Oral Reading Fluency (ORF) with grade level text, DIBELS (Good & Kaminsky, 2002). The Interactive Reading Assessment System (Calfee and Calfee, 1978) vocabulary subset was used (2-5th grades only).

Fidelity to the treatment included direct observation of student practice and teacher logs, which reported activities and instructional minutes.

Results and Conclusions

From the results for the oral reading fluency, we found that within subject contrasts revealed that students in the experimental group had superior performance to those in the control group $F(1,564)=4.72, p<.05$, with a small effect size $d = .2$. There was also an interaction between grade and experimental status. Multiple comparisons showed that differences were more significant in third and seventh grade.

In conclusion, scaling-up diluted the impact of the Readers Theatre in the classroom despite the increased amount of time that teachers were able to use the Readers Theatre strategy. While there appeared to potentially more time, the looser controls allowed a diluted implementation. We found that the instruction tended to be sporadic, instead of the sustained regular implementation of Readers Theatre in the first two studies. The small effect size reflects diminished implementation and the reduced novelty effect.

D.5: Summary

While Readers Theater has often been cited as a repeated reading strategy (e.g. Rasinski, 2003) few studies have shown its actual contribution to students reading performance. These series of studies has shown that a carefully conducted Readers Theater component can increase student performance in both fluency and comprehension. The evidence from the three studies shows not only that Readers Theater is an effective strategy when used by experts but also that it can be implemented successfully in the classroom.

One caveat emerges here: the Readers Theater component used in our studies was NOT based on ready-made scripts written by others, but rather on classroom existing curriculum. This feature cannot be underestimated as it connects content and domain knowledge with the practice of repeated readings. The use of scripting techniques by teachers and students led to longer reading periods and immersion with the texts that increased both fluency and comprehension.

The series of studies presented the dilemma of scaling up of validated instructional components. It is not simple or straightforward transferring practice from carefully controlled experiments to wide scale application. The main dangers are demands on teacher and classroom time, teacher fidelity to the core ideas of the instructional component, and long-term commitment. The result of these challenges was smaller but still significant effects in large-scale implementation.

In considering future scaled-up experiments, we feel that the professional development and teacher support must allow for incremental increases in teacher self-efficacy and classroom practice. The scripting technique with its time demands often resulted in severe dilemmas for the classroom teacher and researcher alike. Based on discussions with participants, teachers would have benefited from pre-scripted basal stories for the first semester. This would have allowed easier classroom implementation with student benefits of improved prosody, fluency, and comprehension. Teachers would build their own implementation skills while seeing more immediate student results with less initial effort and instructional time. Increased self-efficacy should result in stronger fidelity to the implementation during the latter part of the experiment. This process of professional support has wide application in a variety of implementation studies.