PARTNERSHIPS IN ARTS INTEGRATION RESEARCH (PAIR) PROJECT

Final Reports

(Updated 12.12.2012)

Comprehensive Report

Part I: Analysis of Teacher Impact

Analysis of Teacher Professional Development Outcomes and Their Possible Relationship with Student Learning Outcomes
(*Dr. Gail Burnaford, September 2010*]

Part II: Analysis of Student Academic Performance Impact

A Comparative Analysis of Control-Treatment School Student Standardized Academic Test Learning Outcomes by All School Types and Student Academic Classifications
(*Dr. Lawrence Scripp, June 2011*]

Part III: Analysis of PAIR Student Arts Integration Assessments and the Intersections Between Teacher and Student Performance Outcomes

PAIR Student Learning and Survey Results and a Comprehensive Statistical Analysis of their Links with Teacher Professional Development and
Student Standardized Test Outcomes by all Demographic Factors
(*Dr. Lawrence Scripp, April 2012*]

Addendum A: A Brief Summary of Findings from the PAIR Report

Meeting the burden of proof: Statistical evidence for the impact of arts integration based on causal links among teacher professional development, arts learning and academic outcomes.

(*Dr. Lawrence Scripp, April 2012 and Laura Tan Paradis, December 2012*)

Addendum B: A Powerpoint Summary of Findings from the PAIR Report

Arts Learning Assessment Practices and Analytic Processes in the Partnership for Arts Integration Research (PAIR) Project

(*Dr. Lawrence Scripp, April 2012 and Laura Tan Paradis, December 2012*)

PAIR Project Executive Summary

The PAIR (Partnerships for Arts Integration Research) COMPELETE FINAL REPORT is an evaluation of a four year, federal Department of Education funded Arts in Education Model Development and Dissemination (AEMDD) project administered by the Chicago Arts Partnerships in Education (CAPE) in partnership with the Chicago Public Schools. This project brought together 3 pairings of school populations (a world languages focused magnet cluster school with a fine-arts focused magnet cluster school; a math and science focused magnet cluster school with a fine arts focused magnet cluster school; and a literature and writing focused magnet cluster school with a fine arts magnet cluster school) to work with teaching artists in 4th, 5th, and 6th grade classrooms. Results from the six schools were compared with six control schools of similar status, resources, student population, demographic factors, and comparable levels of academic achievement prior to the start of the PAIR project.

The PAIR research and evaluation focuses extensively on teacher impact and student achievement. Two principal investigators noted for their work in the fields of teacher education, student learning, and arts in education teaching and learning practices engaged in this research: Dr. Gail Burnaford, School of Education faculty at Florida Atlantic University, who examined the impact of PAIR on classroom teachers, and Dr. Lawrence Scripp, Director of the Center for Music-In-Education, Inc, who analyzed student arts integration and academic learning outcomes and their relation to PAIR teacher professional development outcomes and controlled for student demographic factors. Burnaford's and Scripp's cumulative findings on the impact of PAIR on teacher professional development, student learning and the intersections between teacher and student outcomes over the three-year time period of the project are presented in the three-part comprehensive report.

Lawrence Scripp and Laura Tan Paradis (PAIR research coordinator) provide a brief summary of the project findings as an addendum to the comprehensive three-part PAIR Report.

PARTNERSHIPS IN ARTS INTEGRATION RESEARCH (PAIR)

FINAL REPORT

Part 1: Teacher Impact
(September, 2010)

Gail E. Burnaford, Ph.D.
Principal Investigator
Professor, Curriculum, Culture, and Educational Inquiry
Florida Atlantic University
777 Glades Road
Boca Raton, FL 33486
561 706-8228 burnafor@fau.edu
http://www.coe.fau.edu/faculty/burnafor/

With Thanks to Olga Vazquez, Research Assistant and Doctoral Student,
And Laura Tan Paradis, CAPE Research Associate

The PAIR project was funded through a grant to Chicago Public Schools from the US Department of Education's AEMDD Program. Chicago Arts Partnerships in Education (CAPE) contracted with Chicago Public Schools for this project. CAPE then subcontracted with Dr. Larry Scripp to serve as co-principal investigator for the project in collaboration with Dr. Gail Burnaford, Principal Investigator, Florida Atlantic University.

PAIR Report Part 1 of III: Table of Contents

PAIR Report Part One: Teacher Impact 1

Part 1Table of Contents 2

Part 1 List of Figures and Tables 4

Introduction 6

Teacher Learning Goals and Results for 2009-2010 8

Data Collection and Analyses for Year Three 9

- Quantitative Data: Indicators of PAIR's Impact on Participating Teachers 9
- Qualitative Data: Indicators of PAIR's Impact on Participating Teachers 10

Goal 1 Collaboration and Professional Development 12

- Goal 1A - 60% of PAIR teachers communicate and collaborate, in and across schools. 12
- Goal 1B - 90% of PAIR teachers develop partnerships with teaching artists and external professional development resources. 14
- Goal 1C - 50% of PAIR teachers provide professional development for their schools. 17

Goal 2 Curriculum Documentation, Inquiry and Reflection 19

- 2 A – 50% of PAIR teachers develop inquiry questions and demonstrate reflection. 19
- 2 B - 90% of teachers develop and document at least one curriculum plan representing arts integration and content learning aligned with Illinois Learning Standards. 22
- 2 C - 90% of PAIR teachers document implementation and assessment of arts integrative curriculum in the PAIR project. 25

Goal 3 Improved Instruction through Arts Integration 29

- 3A- 90% of PAIR teachers use and document research-based effective teaching practices in the PAIR unit. 29

Summary of Teacher Goals 34

- Professional Development and PAIR 35
- Collaboration/Planning and PAIR Professional Development 37
- Documentation and PAIR Professional Development 41
- Peer-to-Peer Sharing and PAIR Professional Development 45
- Extending and Deepening the Curriculum and PAIR Professional Development 46
- Attendance at PAIR Professional Development Sessions 47

Portfolio Conference Categories of Reflective Practice and Professional Development Attendance 51
Four Teacher Profile Effects and PAIR Program Impact 52
The Content Expertise Effect: Pedagogical Content Knowledge Matters 52
The Documenting to Learn Effect: Collecting Student Work Inspires Reflection 55
The Fourth Grade Effect: Designers Have Ownership 57
The Healy Effect: Initiatives Build on Other Initiatives 62
Considerations for Making Connections: 64
Teacher Impact Effects and Student Learning Outcomes 64
Conclusions 66
References 72

Part I: List of Tables and Figures

Table One	Teachers' Speech Occurrence Categories During Portfolio Conferences: Relationships to Goal Areas	11
Table Two	Reflective Practice Portfolio Conference Collective Rankings of PAIR Teachers Based on Number of Occurrences in Categories	11
Table Three	Mean Scores by Grade Levels for Items 6 and 9: Year-End Curriculum and Teaching Survey	13
Table Four	PAIR Teachers and Collaboration: Survey Responses and Portfolio Conference Comments	14
Table Five	Mean Scores by Grade Levels for Items 1, 2, 3, 4, 5, and 12: Year-End Curriculum and Teaching Survey	16
Table Six	Mean Scores by Grade Levels for Items 15, 16: Year-End Curriculum and Teaching Survey	18
Table Seven	Means Scores by Grade Levels: Items 10, 11, 20, 24, 40 - Year -End Curriculum Survey	20
Table Eight	Grade Level Means for Items 25 and 26: Year-End Curriculum and Teaching Survey	26
Table Nine	Impact of PAIR on Students' Learning Teacher Comments from Portfolio Conferences	27
Table Ten	HAL Students and Assessing Learning Teacher Comments During Portfolio Conferences	28
Table Eleven	Grade Level Means for Items 21, 22, 23: Year-End Curriculum Survey	30
Table Twelve	Effective Teaching Practices Documented on PAIR Student Work Labels	32
Table Thirteen	Teacher Progress and Learning Over Time in PAIR Project Portfolio Conference Comments	33
Table Fourteen	Teacher Goal Outcomes Based on Qualitative and Quantitative Data	35
Table Fifteen	Level of Use of An Innovation (Hall and Hord, 2006)	37
Table Sixteen	Notes from a Professional Development Session	38
Table Seventeen	Sharing Data on Collaboration: Notes from a Professional Development Session	38
Table Eighteen	Professional Development Session Survey Means Across Four Sessions for Section 1: PAIR Collaboration and Planning	40
Table Nineteen	Sharing Results from the Research with PAIR Participants: Notes from a Professional Development Session	41

Table Twenty	Collaborative Reflection on PAIR Documentation Panels: Notes from Professional Development Discussion	43
Table Twenty-One	Professional Development Session Survey Means Across Four Sessions: Section 4 Documentation of Student Learning	45
Table Twenty-Two	Professional Development Session Survey Means Across Four Sessions: Section 3 Peer to Peer Sharing	46
Table Twenty-Three	Professional Development Session Survey Means Across Four Sessions: Section 2: Extending and Deepening the Curriculum	47
Table Twenty-Four	Attendance at Professional Development Sessions Year Three of PAIR Project	48
Table Twenty-Five	Attendance at Professional Development Sessions Across Three Years of PAIR Project	49
Table Twenty-Six	Teacher Descriptions of Practice After PAIR Project Ends	51
Table Twenty-Seven	Portfolio Conference Coded Categories of Reflective Practice: High and High/Middle	52
Table Twenty-Eight	The Content Expertise Effect: Pedagogical Content Knowledge Matters	54
Table Twenty-Nine	The Documenting to Learn Effect: Collecting Student Work Inspires Reflection	55
Table Thirty	The Documentation Effect: Teachers' Responses on Year-End Survey Impact of PAIR Documentation of Learning by Students	56
Table Thirty-One	The Documentation Effect: Teachers' Responses on Year-End Survey – PAIR Documentation of Student Learning by Teachers	57
Table Thirty-Two	Six Teachers and Significant Differences from all PAIR Teachers Grades 4, 5, 6 on Specific Indicators as Self-Reported on Year-End Curriculum and Teaching Survey	60
Table Thirty-Three	The Fourth Grade Effect: Designers Have Ownership	61
Table Thirty-Four	The Four Teacher Impact Effects in PAIR Project	64
Table Thirty-Five	Data Quality Campaign: Advances in State Student Data Systems	71
Figures A, B, C	Curriculum Maps From PAIR Professional Development Sessions	23
Figures D, E, F	Documentation Panels Constructed During PAIR Professional Development	42

* * *

Introduction

Forty years ago, there was widespread belief that teachers and schools had little influence on students' achievement independent of their socioeconomic background and context (Jencks, 1972, Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld, and York, 1966). More recent studies of teacher effects at the classroom level, however, such as those using the Tennessee Value-Added Assessment System, have found that differential teacher effectiveness is a strong determinant of differences in student learning, far outweighing the effects of differences in class size and heterogeneity (Sanders & Rivers, 1996; Wright, Horn, & Sanders, 1997).

Students who are assigned to several ineffective teachers in a row have significantly lower achievement and gains in achievement than those who are assigned to several highly effective teachers in sequence (Sanders & Rivers, 1996). Teacher effects appear to be additive and cumulative, and generally not compensatory. These issues have been the topic of much other research over the last 50 years (Darling-Hammond, 1999). More and more research is conducted with teacher practice and professional development as part of the context for investigating student outcomes. That is what the PAIR project has done during this research initiative.

The Partnerships in Arts Integration Research (PAIR) project was a three-year initiative focused on the intersections between arts and non-arts content learning in two mathematics and science, two world languages and two writing Magnet Cluster Schools in Chicago. This section of the final report will focus on the impact of the project on the teachers, with particular attention to the third year of the project in which documentation was more intentional and systematic in each school. The 6 PAIR schools were matched with 6 control schools also in the Arts Magnet Cluster Schools program in Chicago Public Schools. A Year-End Curriculum and Teaching Survey was administered to 4th, 5th, and 6th grade teachers in all twelve schools during Year Three of the project. Other data were also collected from the teachers in the 6 PAIR schools, including professional development session surveys and attendance figures, portfolio conference transcribed comments, student work and teacher practice labels and documentation from work completed at professional development sessions (documentation panels and curriculum maps).

The design of the program, in which the fourth grade teachers participated in Year One of PAIR and contributed to the initial planning process for arts integration, followed by the addition of 5th grade

teachers in the second year and 6th grade teachers in the third year, will be discussed. PAIR teachers had access to two teaching artists each year in two art forms. The seven goals related to teacher impact in the project were related to the nature and degree of collaboration fostered by PAIR, the curriculum development process and products related to arts integrated content-focused units, and the ways in which documentation and assessment informed teaching, learning and research within the project.

PAIR builds upon the outcomes of a larger three-year Department of Education Professional Development Grant project that ended in 2008 titled Building Learning Communities through Culture, Leadership and the Arts (BCCLA). In that project, 59 Fine and Performing Arts Magnet Cluster Program Schools experienced professional development that supported the development of arts integrated curriculum, leadership, and community. The BCCLA project focused intensive professional development on arts specialists in the Fine and Performing Arts Magnet Schools (Burnaford, 2009). The PAIR project also intended to build community across magnet cluster schools in and through arts integrated curriculum and was designed to engage non-arts teachers in such innovations. The PAIR project involved classroom teachers in six magnet cluster schools, two each in the areas of Writing, World Language, and Mathematics in order to determine how engagement with multiple arts experiences over three years not only develops arts learning but also influences non-arts content learning in students. Consistent with the research, the PAIR project research team investigated what classroom teachers learned in and through PAIR, as per Ball and Cohen's "pedagogy of professional development" (1999), in which there is always a dynamic interaction between teachers, students, curriculum content and the school environment. Research suggests that when teachers' knowledge, skills, and strategic actions are seen as resources, student achievement rises significantly (Phillips, 2003). Any consideration of students' learning within an initiative such as PAIR must also accommodate the investigation of how their teachers are seeing the initiative, accommodating the new practices, and developing their expertise as a result of the intervention.

To that end**,** the PAIR research team collected both quantitative and qualitative data over three years to address seven teacher learning goals. This report will first discuss the quantitative and qualitative data for each of those goals, noting significant differences with any or all of the control group school data collected during Year Three (See Reports for Years One and Two for treatment/control group goal data for those years).

Teacher Learning Goals and Results for 2009-2010

Goals for teacher learning were developed by the PAIR/CAPE/Chicago Public School project directors in collaboration with the researchers. The goals have remained consistent through the three years of the project, although they have been reordered and synthesized in order to improve data collection procedures and align instrumentation as the program was implemented and revised each year of the project. It should be noted that the implementation goals are extraordinarily high for this project, with 3 of the 7 goals expecting 90% implementation levels for teachers. This speaks to the investment that the partner arts organization, CAPE, and the district have in the project.

Goal 1 Collaboration and Professional Development

Goal 1 A - 60% of PAIR teachers communicate and collaborate, in and across schools.

Goal 1 B - 90% of PAIR teachers develop partnerships with teaching artists and external professional development resources.

Goal 1C- 50% of PAIR teachers provide professional development for their schools.

Goal 2 Curriculum Documentation, Inquiry and Reflection

2 A – 50% of PAIR teachers develop inquiry questions and demonstrate reflection.

2 B- 90% of teachers develop and document at least one curriculum plan representing arts integration and content learning aligned with Illinois Learning Standards.

2 C 90% of PAIR teachers document implementation and assessment of arts integrative curriculum in the PAIR project.

Goal 3 Improved Instruction through Arts Integration

3 A- 90% of PAIR teachers use and document research-based effective teaching practices in the PAIR unit.

* * *

Data Collection and Analyses For Year Three

Quantitative Data: Indicators of PAIR's Impact on Participating Teachers

A. *Year-End Curriculum and Teaching Survey*

- Frequencies for PAIR Teachers were calculated for each item. The Likert Scale for the Year-End Curriculum and Teaching Survey 2009-2010 was as follows:

1 = never happened
2 = rarely happened (one or twice during the unit)
3 = happened sometimes (at least 3 times during the unit)
4 = happened about often (at least 5 times during the unit)
5 = frequently happened (at least 7 times during the unit)
6 = always happened (every week during the unit)

Occurrences at the level of 3 or higher were counted in frequencies reported for each goal below.

- Tests for Significant Variation - Treatment and Control Group Teachers (1 of 6 control group schools, Taylor, did not report scores for 2009-2010)
- Tests for Significant Variation - Treatment Teachers and Grade Levels (4,5,6)
- Tests for Significant Variation – Individual Teachers and All PAIR Teachers (4,5,6)

B. *Professional Development Session Survey* (administered 4 times in Year Three)

- Mean scores across 4 sessions by teacher
- Mean scores across 4 sessions by grade level (4,5,6)

C. *Attendance at Professional Development Sessions* (Years One, Two, Three)

- Frequency by teacher

D. *Effective Teaching Practices Frequency Reported on Student Work Labels*

- Identified practices by teacher

Qualitative Data: Indicators of PAIR's Impact on Participating Teachers

The research team collected a variety of data each year of the PAIR project with the most consistent and thorough collection occurring in Year Three. If there were certain teachers who demonstrated impact, across a variety of variables, then perhaps those teachers' students' outcomes could be explored to see if there is a relationship between impact at the teacher level and student learning. These kinds of relationships are being explored in the larger educational field, particularly with respect to professional development. The PAIR project contributes to this discussion with the results from Year Three. It must be noted that the sample sizes for these tests were always small. The number of possible teachers at each grade level in both treatment and control schools limits the power of the findings.

The qualitative data collected from PAIR teachers are as follows:

A. *Open-Ended Responses in Year-End Curriculum and Teaching Surveys*

B. Student Work Label Effective Teaching Practices

C. *Portfolio Conference*[1] *Comments (from Transcriptions)*

Each transcript was coded and the number of occurrences of these speech categories was counted. Teachers were then grouped according to the number of Occurrences as High, High/Middle, Middle/Low and Low number of Occurrences. The Portfolio Conference Comment by PAIR teachers were coded according to emergent categories of reflection on the meaning of PAIR experiences for themselves and their classrooms (See Table One). These six categories seemed to indicate teacher learning related to specific PAIR Goals. Six teachers did not articulate comments in any of the six emergent categories; it is interesting to note that four of those six are from the same school. Teachers' and school names have been assigned a numerical code to ensure confidentiality. The first number in each code refers to the PAIR treatment school. The summarized data from the portfolio conferences in Tables One and Two below were used to address goals, and the project's impact on specific teachers and will be discussed in Goal sections.

[1] Portfolio Conferences protocols devised by Dr. Scripp are described and analyzed in detail in Part 3 of this report.

Table One: Teachers' Speech Occurrences in Categories During Portfolio Conferences: Relationships to Goal Areas

Reflective Practice In Portfolio Conference Teacher Comments	*Described the impact of PAIR on their non-arts curriculum (Goal 2A)*	*Described the impact of PAIR on students' learning in non-arts content (Goal 2C)*	*Described progress across multiple years in PAIR (Goal 2A)*	*Described what they would do with PAIR learning after project was over (Goal 2A)*	*Assessed student learning evidenced in portfolio conferences (Goal 2C)*	*Described collaborations with other teachers at their school for PAIR project (Goal 1A)*
3 or more occurrences during conferences	1.20 3.20 5.10 5.30 5.40 7.10 9.40 11.10 11.20 11.30 11.60	1.40 1.50 3.20 5.10 5.20 5.60 7.10 7.30 7.40 9.60 11.10 11.20 11.40	1.10 3.20 5.10 7.10 7.40 11.10	1.10 3.20 5.20 7.30 11.30	1.50 7.10 7.30 7.40 7.60	3.30 9.30 9.40 9.50
2 occurrences						
1 occurrence						

No speech occurrences in these six categories were identified for: 3.10, 3.40, 3.50, 3.60, 7.50, 11.50.

Table Two: Reflective Practice Portfolio Conference Collective Rankings of PAIR Teachers Based on Number of Occurrences in Categories (See Table One)

Yellow (High)	Green (High/Mid)	Purple (Mid/Low)	White (Low)
3.20 5.10 7.10 7.30 11.10 11.40	1.10 1.50 5.20 7.40 11.20 11.30	1.20 3.30 5.30 5.40 7.60 9.30 9.40 9.50 11.60	3.10 3.40 3.50 3.60 7.50 11.50

Goal 1 Collaboration and Professional Development

Goal 1A - 60% of PAIR teachers communicate and collaborate, in and across schools. MET IN 2008-2009

In a recent review of the literature regarding professional learning communities, researchers noted that there are few studies that move beyond self-reports of positive impact (Vescio, Ross and Adams, 2008). In their review, Vescio et al noted that the collective results of these studies suggest that well-developed learning communities have positive impact on both teaching practice and student achievement. The PAIR project also relies on self-report of teachers and, while the PAIR design did not specifically target the development of formal professional learning communities, the project was from the beginning focused on the variety of collaborations and the ways in which teachers could learn through such structures during the project. As the project progressed, it became clear that PAIR teachers participated in collaborations of several distinct types: 1) collaborations with two partner teaching artists assigned to their schools; 2) collaborations with PAIR teachers on their grade level; 3) collaborations with PAIR teachers on other grade levels; 4) collaborations with PAIR teachers in the partner magnet cluster school (writing, mathematics or language/culture); 5) collaborations with all PAIR teachers and artists in the initiative.

Year Three's Year-End Curriculum and Teaching Survey was redesigned to determine more fully the specific nature of collaboration that the PAIR project engendered among participating teachers. For Goal 1 A, 65.7% reported *discussing the project with teachers from other grade levels* (#6 *item on the survey*) and 84.4% reported *collaborating with at least one other teacher in the school to extend and deepen the curriculum* (#9 *item on the survey*). On the basis of these reported frequencies, **Goal 1A was met for Year Three as they are above the 60% level.**

Other data contribute to this discussion of goal implementation as well. Mean scores by grade levels for these two indicators of collaboration reveal more collaboration on the 4th grade level than in grades 5 and 6 (See Table Three), though not at the level of significance, with less reported collaboration among the 6th grade teachers in the project. This is somewhat predictable as the 6th grade teachers joined the project in Year Three, the last year of the project.

Table Three: Mean Scores by Grade Levels for Items 6 and 9:
Year-End Curriculum and Teaching Survey

grade		plan6	curr9
4.00	**Mean**	**4.4444**	**4.4444**
	N	**9**	**9**
	Std. Deviation	**.88192**	**1.42400**
5.00	Mean	2.8000	4.0000
	N	10	10
	Std. Deviation	1.61933	1.15470
6.00	Mean	3.1000	3.9000
	N	10	10
	Std. Deviation	1.28668	.99443
Total	Mean	3.4138	4.1034
	N	29	29
	Std. Deviation	1.45202	1.17549

grade		plan6	curr9
4.00	**Mean**	**4.4444**	**4.4444**
5.00	Mean	2.8000	4.0000
6.00	Mean	3.1000	3.9000
Total	Mean	3.4138	4.1034

Based on the Year-End Curriculum and Teaching Survey, 2009-2010, there was no significant difference between treatment and control group teachers as groups regarding the *degree of collaboration between classroom teachers*. Neither were there significant differences in the *degree of collaboration between classroom teachers among the matched pairs* of control and treatment schools. This may be due to the magnet school cluster norms of collaboration in both sets of schools.

In the open-ended response section of the Year-End Survey, PAIR teachers were more able to describe the challenges and new conceptions of collaboration that they had gained during the project (See Table Four). The responses reveal some of the central insights with respect to the structure of PAIR and the nature of collaboration over three years with a grade level being added each year.

First, as some indicated, it was the 4th grade teachers who were most intimately involved with the design of the project because they were partners from the first year. Some 5th and 6th grade teachers perceived themselves as implementing what someone else designed. Also, some 5th and 6th grade teachers had the impression that the ideas for PAIR work came primarily from the artists. It was probably the case that the artists served as the liaison for teachers across grades and that the PAIR projects were in fact designed by the 4th grade teachers with the artists. Teachers did recognize the

opportunity embedded in the PAIR project for genuine exchange of ideas among teachers. That was a value expressed in writing on this survey as well as in the portfolio conferences.

Table Four PAIR Teachers and Collaboration:

Survey Responses and Portfolio Conference Comments

This form of instruction is very beneficial for students. It provides them with two different interpretations or styles of teaching. The teachers also grow by learning to accommodate different views. However, the artist seemed pretty set on one idea from the beginning. Further collaboration may have been needed (7.60).
Collaborative teaching is important because it helps teachers prepare lessons with a richer background than if using only their own knowledge. When teachers collaborate, it helps build consistency among classes and grade levels. Also, the exchange of ideas often creates new directions to take on teaching (11.60).
It enables colleagues to share ideas. Sometimes another's idea will trigger additional ideas for others. There was not effective collaboration between the artist and myself. I was not able to clearly present ideas for the project. The artist did come up with ideas (5.30).
Well, it worked out nicely with me because I collaborated with our Spanish teacher. And the different things that we did, we made sure that we were on the same page so that she was there to support us as far as...especially the first unit that we did, the literature unit we did with the story was *Esperanza Rising*, so there were a lot of Hispanic culture and a lot of Spanish words that we were able to use and learn in that way, so it worked out just fine (3.30)
Well, first, summarizing the project, we decided [to share]. This year we took a book, which was *In the Time of the Drums*, and we applied it to the arts and we applied it to writing. And our goals basically was we wanted – since we did it last year, we wanted to see if we could get better at doing it as teachers, Miss G and I, get better at doing it as teachers with our students. So we did some of the same things. Some things we changed. We came up with some new ideas. She did some different things. Miss G did some different things and I did some different things, but we both shared ideas (9.40).

Goal 1B - 90% of PAIR teachers develop partnerships with teaching artists and external professional development resources. MET IN 2008-2009

For Goal 1 B, there were a variety of indicators assessed:

81.3% of PAIR teachers reported *brainstorming as part of collaboration and planning with teaching artists (#1 item on the survey);*

94.4% reported *coming to consensus at planning meetings with teaching artists (#2item on the survey)*;

90.7% reported *modeling discussion and feedback with the teaching artists in the classroom when students were present* (with no teachers reporting that this never or rarely happened) (#3 *item on the survey*);
68.8% reported *planning outside of classroom time (#4 item on the survey)*;
50.1% reported *meeting with teaching artists outside of the school* (#5 *item on the survey*).
78.2% reported that they were actively *integrating the arts when the teaching artist was not there (#12 item on the survey).*

Mean scores for these indicators were also compared across grade levels (See Table Five). Ranges for total means were from 2.7931 (regarding the frequency with which *teachers met with artists outside of school*) to 4.9655 (regarding the frequency with which teachers *modeled discussion and feedback with artists when students were present*). These figures would suggest that teachers felt they did a great deal of co-teaching and were very much involved when the teaching artists were in their classrooms. This phenomenon might suggest that teachers did not feel the need to meet with teachers outside of schools in order to effectively co-teach their PAIR units. Co-teaching represents the most complex and sophisticated measure of collaboration, beyond the instrumental logistics of meetings to plan activities that are typically associated with collaboration and that are investigated in 4 of the 5 items related to Goal 1B.

Qualitative data, including portfolio conference comments indicate that many teachers found ways to communicate with artists that suited them and that, for some teachers, this need to communicate decreased as they became more comfortable with the goals of the project for their own classrooms.

(Tables on next page)

Table Five: Mean Scores by Grade Levels for Items 1, 2, 3, 4, 5, and 12: Year-End Curriculum and Teaching Survey

grade		plan1	plan2	plan3	plan4	plan5	curr12
4.00	Mean	4.0000	4.6667	4.7778	3.4444	3.6667	4.2222
	N	9	9	9	9	9	9
	Std. Deviation	.86603	1.22474	1.09291	1.01379	.86603	1.56347
5.00	Mean	4.7000	4.8000	5.1000	3.5556	2.7000	4.7000
	N	10	10	10	9	10	10
	Std. Deviation	1.05935	1.39841	.87560	1.23603	1.15950	.94868
6.00	Mean	3.8000	4.7000	5.0000	3.5000	2.1000	4.0000
	N	10	10	10	10	10	9
	Std. Deviation	1.81353	1.70294	1.05409	1.71594	1.72884	1.41421
Total	Mean	4.1724	4.7241	4.9655	3.5000	2.7931	4.3214
	N	29	29	29	28	29	28
	Std. Deviation	1.33815	1.41160	.98135	1.31937	1.42376	1.30678

grade		plan1	plan2	plan3	plan4	plan5	curr12
4.00	Mean	4.0000	4.6667	4.7778	3.4444	3.6667	4.2222
5.00	Mean	4.7000	4.8000	5.1000	3.5556	2.7000	4.7000
6.00	Mean	3.8000	4.7000	5.0000	3.5000	2.1000	4.0000
Total	Mean	4.1724	4.7241	4.9655	3.5000	2.7931	4.3214

Based on the Year-End Curriculum and Teaching Survey, 2009-2010, there were significant differences between treatment and control group teachers as groups regarding the relationships between teachers and teaching artists *brainstorming to plan* (t=2.230, df=55, significance <. 05 = .030), *coming to consensus in planning (t= 2.74, df = 54, significance.<.05 = .008),* and *modeling discussion and feedback in front of students in the classroom (t= 3.928, df = 55, significance <.05=.000).* These results suggest that the project had considerable, if predictable impact on teachers' planning processes with external partners.

There were also significant differences among matched pairs of treatment and control group teachers as follows:

Healy/Holden: There were significant differences (with more frequent occurrences reported in the treatment school teachers) in the frequency of *teacher brainstorming to contribute to teacher/artist planning (t=.739, df=8, significance <.05=.001)*, the frequency of *coming to consensus during planning meetings between artists and teachers(t=.950, df = 7, approaching significance <.05=.085)*, and the frequency of *modeling discussion and feedback in front of students in the classroom (t=1.144, df = 8, significance <.05-.047).*

Swift/Ebinger: There were significant differences (with more frequent occurrences reported in the treatment school teachers) in the frequency of *modeling discussion and feedback in front of students in the classroom (t=.839, df = 6, significance <.05=.024),* and the frequency of *meetings with artists outside of school (t=1.769, df=6, significance <.05 = .005)*.

Ward/Carson: There were significant differences (with more frequent occurrences reported in the treatment school teachers) in *modeling discussion and feedback in front of students in the classroom (t=8.781, df=11, significance <.05 = .000),* and the frequency of *meetings with artists outside of school*, *including email, phone or in-person meeting times (t =4.366, df=11, significance .<05 = .044, t= 3.813, df=11, significance <.05 = .001)* .

The most interesting thing about these comparative data between the PAIR schools and the control group schools is that three of the six PAIR schools showed significant differences with respect to what happens in their classrooms while working with teaching artists. Healy, Ward and Swift teachers all indicated co-teaching behaviors: this suggests something about the nature of their classrooms in this initiative that moved beyond the traditional "one teacher at a time" model.

Based on these data, **GOAL 1B WAS MET FOR 2009-2010.**

Goal 1C- 50% of PAIR teachers provide professional development for their schools.
NOT MET IN 2008-2009

Based on the Year-End Curriculum and Teaching Survey, 2009-2010, there was no significant difference between treatment and control group teachers as groups regarding the degree to which

teachers report receiving feedback either from teachers in their own schools or in other schools regarding their curricular units.

This year's Year-End Curriculum and Teaching Survey also asked teachers whether they had *shared their PAIR units with other teachers in their own school* (item # 15 on the survey and 65.7% reported they did) and *with teachers outside of their school* (item # 16 on the survey and 37.5% reported they did).

Mean scores are again reported by grade levels for these indicators (See Table Six). The averaged mean scores across three grade levels for these two indicators of project dissemination were 3.3793 (item #15) and 2.4138 (Item #16), indicating that such sharing had happened *sometimes (at least 3 times during the arts integration unit) or rarely (once or twice during the arts integration unit).* There were no significant differences across the grade levels in this regard; 4th, 5th, and 6th grade teachers alike reported that this dissemination did not happen often.

Table Six: **Mean Scores by Grade Levels for Items 15, 16: Year-End Curriculum and Teaching Survey**

Grade		share15	share16
4.00	Mean	3.8889	2.8889
	N	9	9
	Std. Deviation	1.45297	1.26930
5.00	Mean	3.1000	2.0000
	N	10	10
	Std. Deviation	1.66333	.81650
6.00	Mean	3.2000	2.4000
	N	10	10
	Std. Deviation	1.31656	1.07497
Total	Mean	3.3793	2.4138
	N	29	29
	Std. Deviation	1.47391	1.08619

grade		share15	share16
4.00	Mean	3.8889	2.8889
5.00	Mean	3.1000	2.0000
6.00	Mean	3.2000	2.4000
Total	Mean	3.3793	2.4138

The Project Directors felt that PAIR was not structured to provide the necessary support for teachers to formally "provide professional development for their schools" as Goal 1C states. This may be a useful follow up to the PAIR project during 2010-2011. It should also be noted that the PAIR Project Directors from CAPE and the Magnet School Cluster Administration will offer a project dissemination day in November of 2010 during which the PAIR projects will be shared.

Based on these data and acknowledging the opportunity for teachers to share their work as a culminating event in November, **GOAL 1 C has been met in 2009-2010.**

Goal 2 Curriculum Documentation, Inquiry and Reflection

2A – 50% of PAIR teachers develop inquiry questions and demonstrate reflection.
MET IN 2008-2009

Based on the Year-End Curriculum and Teaching Survey, 2009-2010, there were no significant differences between treatment and control group teachers as groups regarding the degree to which *teachers report innovating approaches or processes in their own teaching* (#10) or the *development of inquiry questions to guide the curriculum (#11).*

Similarly, there were no significant differences in the treatment and control group teachers as groups with respect to *reporting their success at predicting the quality of student work over time (#20),* except in the case of the matched pair of Healy and Holden, in which *Healy teachers did report increasing success at predicting quality as the project continued (t=.688, df=8, significance = <.05=.018).* There were no significant differences in the treatment and control group teachers as groups with respect to the *degree to which they reported reflecting on their teaching* (#24).

On the Year-End Curriculum and Teaching Survey, PAIR teachers reported occurrences as follows:

87.5% reported *thinking of innovative arts integration approaches within their content areas;*
75.1% reported *developing inquiry questions* to explore the PAIR curriculum;
84.5% reported *becoming more successful at predicting the quality of student work* during the PAIR unit;
87.6% reported *reflecting on their practice* during the PAIR unit;
78.1% reported *participating as a student* and learning while the teaching artist was in the classroom.

Table Seven: Means Scores by Grade Levels

Items 10, 11, 20, 24, 40 - Year -End Curriculum Survey

Grade		curr10	curr11	doc20	doc24	art40
4.00	Mean	4.4444	3.7778	4.4444	4.4444	4.3333
	N	9	9	9	9	9
	td. Deviation	1.42400	1.30171	1.33333	1.33333	1.87083
5.00	Mean	4.5000	4.1111	4.1000	4.8000	4.1000
	N	10	9	10	10	10
	Std. Deviation	1.17851	1.36423	.56765	.63246	1.44914
6.00	Mean	4.1000	3.4000	4.5000	5.1000	4.9000
	N	10	10	10	10	10
	Std. Deviation	1.28668	.96609	1.50923	1.10050	1.19722
Total	Mean	4.3448	3.7500	4.3448	4.7931	4.4483
	N	29	28	29	29	29
	Std. Deviation	1.26140	1.20570	1.17339	1.04810	1.50205

Grade		curr10	curr11	doc20	doc24	art40
4.00	Mean	4.4444	3.7778	4.4444	4.4444	4.3333
5.00	Mean	4.5000	4.1111	4.1000	4.8000	4.1000
6.00	Mean	4.1000	3.4000	4.5000	5.1000	4.9000
Total	Mean	4.3448	3.7500	4.3448	4.7931	4.4483

One element of investigation with respect to teacher learning related to practice is the importance of reflection and learning from mistakes, missteps, and discomfort with new ways of teaching. Much of CAPE's work is founded on the principle of risk-taking and questioning that is necessary if one is to move beyond current practice. Projects such as PAIR allow researchers to follow that process of what Lee Shulman calls "error, success and refinement or – in a word, teacher-knowledge growth" (1987, p. 4) for those teachers who choose to seize the innovation. As Shulman says, "The neophyte's stumble is the scholar's window" (p. 4). What can the research team learn then from the teachers' reflections on PAIR?

Reflection is a difficult aptitude to measure and sometimes may be a challenge to recognize. Table One exhibits teachers' rankings in six basic categories of reflection, based solely on their transcribed comments related to the categories of teacher impact identified by the researchers in the Portfolio Conferences. Table Two exhibits the overall rankings as High Number of Occurrences, High/Middle Level of Occurrences, Middle/Low Level of Occurrences, and Low Level of Occurrences. These comments address several goals for teachers in the PAIR project including the goal to encourage reflection about their teaching and about what the PAIR project might do to enhance their teaching. Teachers in Year One had to rethink their assumptions about such projects with external partners that are focused on classroom time with students. They had to adjust to the notion that the research team was also interested in teachers' learning, not just the impact that the project has on students.

Now, at the end of the project, teachers, such as the "neophytes" to arts integration referenced below, can articulate what they have learned:

> *Rather than look for their simple grammatical errors and sentence structures and development of the story, I kind of put that to the side now—and I don't look at their mistakes in spelling. I look at the ideas and things developed and the process that they took in writing their stories. That's a big difference (11.10).*
>
> *I was able to trace students' thoughts, their difficulties and then plan to go about this work. This helps me to know if they are going in the right direction. Are they learning the strategies and skills? I am able to understand the process students are planning at this level of involvement (5.60).*
>
> *This (PAIR project) gave me lots more ideas. Go beyond the concept (3.20).*

This year, in the portfolio conferences, 7 teachers reported the adoption of inquiry questions in their PAIR units. In addition, teachers demonstrated reflective practice in the coded categories of the conferences. Coded comments revealed categories of 'reflective practice' that the research team then explored with other qualitative variables to discern whether certain teachers had more characteristics that could indicate PAIR impact on practice. While these categories do not suggest a definitive ranking of teachers as "Highly reflective" or "Not reflective" by any means, they suggest a way of considering what teachers said during the portfolio conferences about PAIR and the impact it has had on their teaching and their students. It must also be noted that the Conferences, which followed a protocol, were not strictly scripted and the conversations did have some appropriate variability across

the teachers and their students in the 6 treatment schools. In addition, 6th grade teachers in their first year of the project were understandably less able to discuss their progress during the PAIR project over time due to the design of graduated implementation across grade levels.

Based on these qualitative and quantitative data, GOAL 2 A WAS MET for 2009-2010.

2B- 90% of teachers develop and document at least one curriculum plan representing arts integration and content learning aligned with Illinois Learning Standards. MET IN 2008-2009

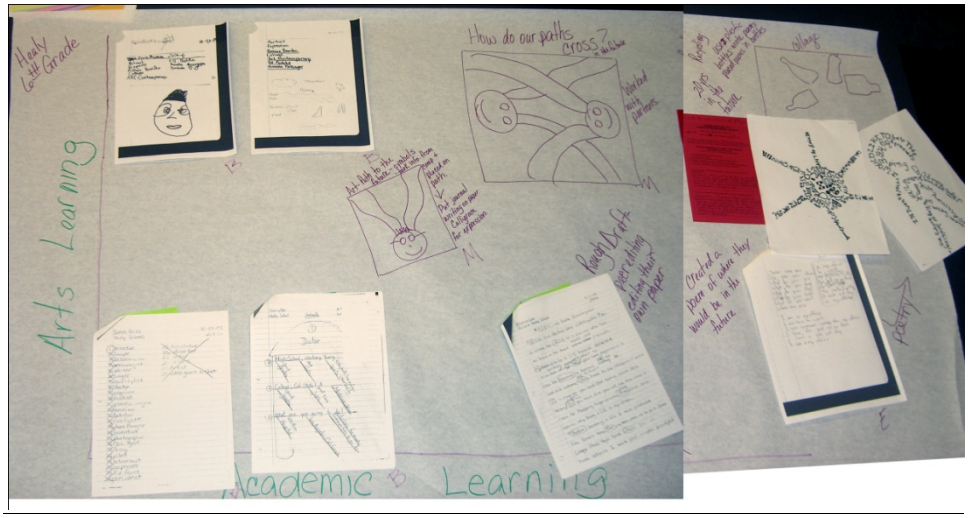
Curriculum development is discussed in the professional development research literature as a central aspect of teacher learning and growth (Gordon, 2004; King & Lawler, 2003). When teachers analyze student work, they can begin to "see" how students are thinking and reflect on implications for their teaching practices from that standpoint (Fickel, 2002). Similarly, teachers can learn much from shared lesson and curriculum planning in a common grade level or for a common unit, as the PAIR project suggests.

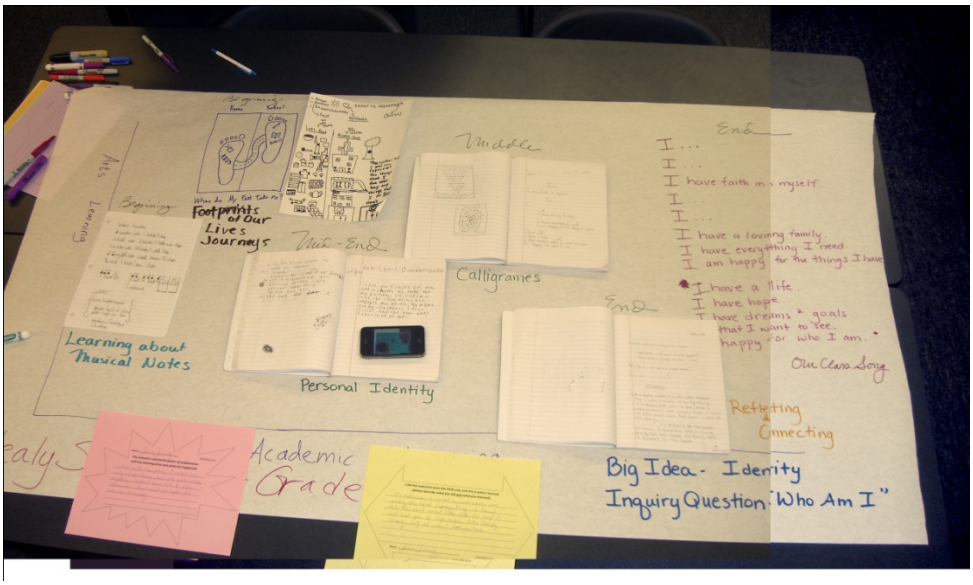
Arts teachers and teaching artists, however, are seldom expected to develop long-term, high quality arts curriculum that is relevant, related to non-arts learning, and open to ongoing collaboration and dialogue as replicated over time, nor are they accustomed to seeing exhibitions and performances of 'student work' as tools for pedagogical or curricular improvement. Especially in funded projects, there are few useful curriculum unit plans available for replication in the field. Further, most "written" curriculum guides in schools are not actively used and referred to by teachers. There are few curriculum guides available for dissemination that reflect collaboration explicitly. Finally, most curriculum plans developed by teachers have few or no examples of student work to demonstrate rich learning formatively or summatively.

The challenge then in PAIR was formidable. The PAIR units developed in Year One did appear to remain fairly consistent during Years Two and Three. The professional development offered to teachers and artists in Year One was in part aimed to help teachers learn to document those curriculum units more fully in order to improve them, learn from them, and share them with the larger magnet cluster group of schools in the district. The teams in the schools increased the level of their documentation of the units over the three years and, in Year Three, many seemed to understand the value of such a practice. As in the BCCLA project two years earlier, the notion of curriculum planning,

while a required compliance element for teachers, is not considered a crucial part of teacher practice, even for the most effective teachers. The PAIR/CAPE leadership for professional development determined that curriculum maps that incorporated the activities of both teaching artists and the grade level teams of teachers was the most effect way of documenting the curriculum plans. These maps were used as tools to guide the initiative and were informally created during professional development sessions in Years One and Two. In Year Three, the maps, coupled with documentation panels, were designed to show the Arts and Academic Learning Outcomes embedded in the PAIR activities (See Figures A, B, C).

Figures A, B, C: Curriculum Maps from PAIR Professional Development Sessions





Based on the Year-End Curriculum and Teaching Survey, 2009-2010, there were areas in which the teachers indicated that the arts integration initiatives in PAIR contributed to students' understandings in the focal non arts content areas. Part of the curriculum mapping/unit process was to determine the intersections in students' learning areas among art forms and the non-arts content that was the focus of the magnet cluster partner PAIR schools.

In the mathematics magnet cluster school matched pair, Swift and Ebinger, there were significant differences regarding *reported implementation of movement/dance to help students understand mathematics,* with more frequent occurrences reported in the treatment school teachers. (*t=1.899, df=6, significance = <.05 = .035*). Swift worked with a dance/movement artist as well as a sculptor/visual artist; the trend toward understanding through movement and dance is therefore encouraging in the treatment school.

In both PAIR mathematics magnet cluster schools, Swift and Thorp, there were significant differences with their matched control group magnet cluster schools with respect to the degree of implementation of the use of music to help students understand math concepts or math problems (*Swift/Ebinger: t = 1.604, df=6, significance = <.05=.048, Thorp/Hale: t= 1.332, df=10, significance = <.05=.021).* While this might initially seem surprising given that the artists working with Swift and Thorp over the three years were not music teaching artists, music was an integral part of the movement/dance initiatives on the part of one of the teaching artists. Some of the teachers continued to express discomfort with the dance and movement from a personal experience standpoint, but it's possible that

the introduction of music to their classrooms was both beneficial and less risky than the movement itself.

Based on the Year-End Curriculum and Teaching Survey, 2009-2010, there was approaching significant difference in the language arts/writing magnet cluster school matched pair, Eberhart and Pirie, with respect to *reported implementation of one particular art form – images/visual art – to help students understand books, stories, writing or vocabulary (t=1.852, df=7, approaching significance = <.05=.074).* Again this trend is intriguing in that the teaching artists at Eberhart were not visual artists but rather focused on music and drama integration with writing. Based on the documentation in the PAIR Pails, the teachers did use more visual imagery, graphic organizers, and visual symbols more frequently with these art forms to help students relate their writing to the art forms, thereby building upon the artists' contributions to the curriculum.

This area was also a significant difference for another matched pair in the language and culture magnet cluster schools, Healy and Holden, in which the Healy teachers *reported more use of images/visual arts to help students understand books, stories, writing or vocabulary (t=1.352, df=8, significance = <.05=.000).* This might suggest that even though the emphasis content area for Healy in the PAIR project was culture, the clear integration with writing, language, and cross-cultural communication through the arts was evident to the teachers.

2C 90% of PAIR teachers document implementation and assessment of arts integrative curriculum in the PAIR project. Formerly Goal 3A NOT MET IN 2008-2009

Consistent with what has been reported for the Goals above, there is much richer documentation and more data sources to address this goal in 2009-2010 than in the previous two years of the project. First, on the Survey, PAIR teachers reported as follows:

> 84.4% reported that they and/or their students were documenting and assessing during the PAIR unit;
>
> 68.8% reported that they used pre and post assessments to see what students were learning through arts integration.

Based on the Year-End Curriculum and Teaching Survey, 2009-2010, there were no significant differences between treatment and control group teachers as groups regarding *the degree to which*

students documented and assessed their own learning (#25 item on the survey), nor were there significant differences in *the degree to which teachers used pre and post assessments to ascertain student learning (#26 item on the survey).*

There was a significance difference in one matched pair, Healy and Holden, with respect to *the degree to which students documented and assessed their own learning*, in which Healy teachers reported that students documented and assessed their learning more often *(t=.717, df=8, significance = <.05=.013).*

Grade level means were compared for these two items on the Year-End Curriculum and Teaching Survey (See Table Eight). Across the grade levels, it seemed that teachers reported more *student involvement in documenting and assessing* than in the *teachers using pre/post assessments* for the PAIR project. Although pre/post assessment was featured in professional development sessions for this project, it became apparent that teachers rarely used pre/post measures in their teaching, nor did they see that approach as meaningful for an arts integration project such as PAIR. Teachers reported assessing informally, through observation, or through the actual projects that students produced, in order to determine whether PAIR had been effective. If specific pre/post measures are important for future projects, it would be necessary to work with teachers to discover how to deign and use them as part of their teaching practice.

Table Eight: Grade Level Means for Items 25 and 26: Year-End Curriculum and Teaching Survey

grade		doc25	doc26
4.00	Mean	4.2222	3.6667
	N	9	9
	Std. Deviation	1.39443	1.50000
5.00	Mean	4.6000	3.9000
	N	10	10
	Std. Deviation	1.26491	1.28668
6.00	Mean	4.6000	3.6000
	N	10	10
	Std. Deviation	1.26491	1.83787
Total	Mean	4.4828	3.7241
	N	29	29
	Std. Deviation	1.27113	1.50941

This goal addresses teachers' perception of how much their students learned through the PAIR project. Even though they did not offer too many specific assessment instruments or rubrics associated with the project, although encouraged to do so in the PAIR PAILS, they definitely had anecdotal views on how the project had affected their students' learning in their comments during the Portfolio Conferences (See Table Nine). The Magnet Cluster Schools might draw on these experiences to develop more tangible ways for teachers to provide evidence of their sense that students did learn specific skills and concepts as they attest to in these comments from this arts integration initiative in a non-arts content area.

Table Nine: Impact of PAIR on Students' Learning

Teacher Comments from Portfolio Conferences

Just their overall taking a test, and their writing has really improved. Some of their Young Author stories, I see, reading through them, they really came up with three problems that a character came across in their narrative stories. (11.40)
And I find that it appeals to different intelligences. They actually get up, get out of their seats. They get around and move. It's not just visual and listening skills. I think a lot of students, in this class especially, are kinesthetic. They need to get up and move and do things with their hands. I think it helped them to develop a better understanding of the material. As we read, we tried to connect to a novel in reading. I think all the students were engaged. I think they were very excited about learning. It was kind of fun to be in the classroom. And I think it...I'm *hoping* it helped develop higher level thinking skills (11.40).
I think it bring a lot of creativities in my classroom and it really gives them a lot more vocabulary that they don't use daily. It kind of like gave them another way to express themselves and to learn in a different way other than the way that we teach traditionally, you know, through papers and through the board (3.20).
I would...I see that now, when we're writing and we do writer's workshop almost every day, and their characters are better defined, there is more detail about their characters. We're learning...I really think I've seen a progression, and even for the time that, you know, what I had in their portfolios and now, months later, it's a better awareness of how you write a story, that there has to be a story. You just can't write then, then, then, then, then, that there's characters in their stories – and this goes for my whole class, not just these three students – that they're really getting the idea of how to put a story together. But what I see is each piece gets better and better. (11.20).

The Portfolio Conference afforded teachers the opportunity to view the "HAL" students in the project (that is, the High Achieving, Average Achieving and Low Achieving students) as they interacted with the researcher regarding their work in the PAIR project. The interviewer then asked them to comment on what they saw, with particular attention to how and whether these categories of achievement (as

measured on Reading ISAT scores) were evident in this particular approach to assessment (See Table Ten). For some teachers, it was clear that they viewed the conference as a unique form of assessment that was distinctly different from more traditional forms of assessment.

Table Ten: HAL Students and Assessing Learning

Teacher Comments During Portfolio Conferences

The fact that they understood the story so well. You know, I assess them by giving them a little matching vocabulary test that they really got the true meaning of all these different complexities in the story that we read. They really got the meanings of the stories. (11.40) Well, I mean, like I said, the fractions. I mean, I don't know if I'm saying what you want me, like, but, like they understand the fractions. I know that. I mean, I can tell that from what they said that for the most part they understand the parts of a whole. They understand the sides of shapes (5.10)
(Interviewer): But what was it that they said that convinced you? **Teacher:** Well, they were able to identify a third, a sixth, a half. They were able to identify that that's a hexagon because it has six sides. From what we saw them do, I think they understand symmetrical and asymmetrical (5.10).
Now that they're not here, I can speak freely about their growth and development. They're language learners, English Language Learners, so during this year I didn't see, or didn't get the growth in the writing as I expected to see in my previous classes, especially the first year with the project in a mono English classroom. This year they were still developing the writing skills. It's the last step that they develop as bilingual students. However, their confidence and ability to speak to others using the English language has grown significantly, and I think it's because of the fact that they can express their knowledge through other means rather than the simple paper and pencil. Because paper and pencil basically means they're going to be graded as far as what they write and how they spell things. But when they have the opportunity to say it to you, they know that they can express their ideas, what they know and what they've learned, a whole lot more freely than they can on paper and pencil (11.10).
(Interviewer): You think people watching the tape would know that they're low, medium and high? I don't think so. That's why I'm very glad, because all three were answering, all three had something to say, and it wasn't intense. They were really putting things together and everything else. And I don't think...this is part of the program where I see that they're integrated, and you can't really tell the difference (7.10).
I don't think if you just looked at them you could say like, oh, well, this one really struggles with math and this one – I mean, I don't think you can really tell across the projects. (5.40).

These teachers were able to assess understanding differently in the conference context and thereby viewed the HAL students' capacities in new ways.

In summary, even though the goal regarding documentation and assessment was not met at the high frequency level of implementation projected (90% of teachers), the amount of documentation has increased steadily each year of the three-year project and this year's work contributed to a richer portrait of the work on the part of students and teachers. The portfolio conference process is promising as a systemic part of integration initiatives in the future. The conference depends on documentation and could influence what teachers collect from students and how they interact with students themselves in teacher or artist-led conferences.

Experience with other CAPE projects suggests that authentic documentation procedures that are useful for assessment and contribute to curriculum improvement take time and skill to develop. There were teachers in the project who excelled at documenting and collected data that appeared to be beneficial for their own practice and not merely for the purposes of the project team (See "Documenting to Learn Effect" below). Such efforts should be highlighted and further research conducted on how/whether documenting to learn can be scaled up in schools and partnerships such that it impacts teacher practice and student achievement.

Based on these data, GOAL 2 C WAS NOT MET FOR 2009-2010.

Goal 3 Improved Instruction through Arts Integration

3A- 90% of PAIR teachers use and document research-based effective teaching practices in the PAIR unit. Formerly 2C NOT MET IN 2008-2009

There were three items (# 21, 22, 23) on the Year-End Curriculum and Teaching Survey that addressed Goal 3 A in addition to other data sources collected this year.

On the survey PAIR teachers reported as follows:

78.2% reported that they make use of their students' feedback to co-create curriculum;

84.4% reported that they conduct critique sessions during the PAIR units.

90.7% reported that they look at student work in order to revise their curriculum.

These elements were essential components of the CAPE/PAIR methodology and were consistent sources of discussion in planning meetings and professional development meetings. It is reassuring to see the progress on these indicators from Year Two to Year Three. The mean scores on these items are consistently high on these three items across grade levels, indicating that teachers did feel an investment in working on their curriculum from the perspective of what students were actually learning, articulating and producing (See Table Eleven). That was a major goal for PAIR this year.

Table Eleven: Grade Level Means for Items 21, 22, 23:
Year-End Curriculum Survey

Grade		doc21	doc22	doc23
4.00	Mean	4.2222	4.2222	4.0000
	N	9	9	9
	Std. Deviation	1.30171	1.39443	.86603
5.00	Mean	4.0000	4.4000	4.6000
	N	10	10	10
	Std. Deviation	1.24722	.96609	.84327
6.00	Mean	3.9000	4.2000	4.8000
	N	10	10	10
	Std. Deviation	1.28668	1.31656	1.13529
Total	Mean	4.0345	4.2759	4.4828
	N	29	29	29
	Std. Deviation	1.23874	1.19213	.98636

grade		doc21	doc22	doc23
4.00	Mean	4.2222	4.2222	4.0000
5.00	Mean	4.0000	4.4000	4.6000
6.00	Mean	3.9000	4.2000	4.8000
Total	Mean	4.0345	4.2759	4.4828

Based on the Year-End Curriculum and Teaching Survey 2009-2010, there were no significant differences between treatment and control group teachers as groups regarding the *degree to which students co-created of curriculum* or *participated in critique sessions*. Teachers in both groups did not report significant differences in *the degree to which they used student work to analyze and revise their curriculum*. Neither were there significant differences *between classroom teachers among the matched pairs* of control and treatment schools in these areas (See Table Eleven).

PAIR teachers submitted student work for the PAIR project and attached student work labels to describe the teaching practices associated with that work. The labels offered a practices menu for the teachers derived from the Effective Teaching Practices Survey that CAPE has been adapting for

different projects for the past 7 years. The Practices are based on the standards from the Center for Research on Excellence and Diversity in Education (www.crede.org). Project resources did not permit direct, regular and systematic observation of all teachers in the project. However, the research team did collect student work and teachers' Effective Teaching labels on that work in which they described the teaching and learning processes through which that work was produced.

The research team coded the labels to determine which teachers made use of the opportunity to reflect on their practice during PAIR and the subsequent student work that emerged. A few teachers submitted no labels. Many teachers identified very similar teaching practices that they employed during the PAIR unit. However, there were seven areas that seemed to distinguish teachers' practice as self-reported on the student work labels were:

1. *Assessments of any kind used during PAIR (including exhibitions, performances, tests, conferences, writing samples).*
2. *Documentation of students' co-creation of curriculum (including planning, brainstorming, developing new ideas, making choices, working together, improvising).*
3. *Documentation of students teaching other students.*
4. *Documentation of students teaching the teacher.*
5. *Student participation in documenting and assessing during PAIR (including critique, naming learning strategies they were using)*
6. *Students writing their own inquiry questions to guide PAIR work.*
7. *Students using new arts vocabulary.*

Table Twelve illustrates the results from the student work labels by teacher. Half (50%) of the teachers in the project reported that *students taught other students* in the PAIR project. This practice is one of the research-based Effective Teaching Practices derived from the CREDE standards and a goal practice for arts integration initiatives through CAPE. Twenty-eight percent reported the *documentation of assessment* and *the co-creation of curriculum with students*. Twenty-two percent of responding teachers reported *students' participation in documentation and/or assessment, writing inquiry questions,* and *using new arts vocabulary on the student work labels.* The most unusual Effective Teaching practice from the label menus, *students taught the teacher* was noted by 2 of the 32 teachers.

Table Twelve: Effective Teaching Practices Documented on PAIR Student Work Labels

Evidence of PAIR Teacher Impact based on Student Work Labels	Documented assessment	Documented students' co-creation of curriculum	Documented students teaching other students	Documented students teaching the teacher	Students participated in documentation and/or assessment	Students wrote inquiry questions	Students used new arts vocabulary
	1.50	1.20	1.20	11.20	1.10	3.30	1.40
	3.10	1.50	1.40	11.40	1.20	5.20	3.20
	3.20	3.10	1.50		1.50	5.40	7.10
	3.50	3.20	3.10		3.10	5.60	9.30
	5.20	3.30	3.50		3.50	9.50	9.50
	7.10	3.40	3.60		5.20	11.20	11.20
	7.30	3.50	5.20		5.40	11.40	11.40
	7.50	3.60	5.60		5.60		11.60
	9.40	5.10	7.10		7.10		
	9.50	5.20	7.30		7.30		
	11.20	5.30	7.50		7.50		
	11.30	5.40	7.60		9.30		
	11.40	5.60	9.40		9.50		
		7.10	11.20		11.40		
		7.30	11.30				
		7.40	11.40				
		7.50					
		7.60					
		9.30					
		9.40					
		9.50					
		11.20					
		11.30					
		11.40					
		11.60					

In addition, categorized comments by teachers in the portfolio conferences represent triangulated data that suggest the project's impact over time on their practice as mathematics, world language and writing teachers (See Table Thirteen). The teachers who articulated their growth and understanding over time during PAIR are understandably 4th or 5th grade teachers, due to the design of the program.

Table Thirteen: Teacher Progress and Learning Over Time in PAIR Project Portfolio Conference Comments

I think that the teaching artists also learned the curriculum more and they kind of know, "Oh, you guys do this, don't you, here," and we kind of, you know, that was really helpful. And, you know, I just think like with anything, the longer you know someone, you just kind of...everything runs smoother (5.10)
Well, the original goal was to see how the arts itself affected the students' abilities as far as their writing, and how the different arts, for example, music, because we had a percussionist and a playwright, would enhance the child's development of characters and stories. The first year was pretty new. We still had to develop some skills as far as what the artists were teaching the children and my own understanding of what it was. This year I think it's been...I've seen...to get the kids to actually think more critically. Last year it was okay, okay, it's going through the machinations, for the most part. This is what I'm supposed to do, this is what we did (11.10).
Well, I think these students are much better at relating it to math than last year. I think they were much more...they expressed themselves much better mathematically, which is because when I was in the room, we really stressed this is not only art. I mean, we talked about it before, but we really stressed it this time because I didn't want it to happen again where they didn't see the value of the math in the project (7.30).

Based on these multiple data sets, **Goal 3A was met in 2009-2010.**

Variability Between Partner PAIR Schools

PAIR schools were paired according to magnet cluster school focus:

Thorp and Swift – Mathematics

Eberhart and Lee – Writing

Healy and Ward – World Language

Based on the Year-End Curriculum Survey, there were no significant differences in the Thorp and Swift teachers' implementation within the goal areas. In other words, the teachers at these two schools who engaged with teaching artists in the teaching of mathematics perceived the project similarly in terms of the nature of *collaboration*, the *extending and deepening of the curriculum*, their *sharing with peers* at their schools, *documentation and learning through the arts forms* offered by the two teaching artists.

There were a slight difference in the Eberhart/Lee pairing for 2009-2010 with Eberhart teachers reporting more success at *predicting the quality of student work* from the PAIR unit over time (*t=2.279,*

df=15, significance = <.05=.015). This is encouraging and seems to suggest, if in only one small way, that Eberhart teachers were understanding the project more fully and what it could mean for their students as they examined arts integrated student work samples to enhance student learning.

Healy and Ward demonstrated several areas of significant difference: Healy teachers reported more *collaboration* at their school: (*t=1.143, df = 20, significance = <.05=.049*). Healy teachers reported more *innovation in arts integration* in their classrooms: (*t=1.491, df = 20, significance = <.05=011*). Healy teachers reported more *feedback on units among peers at their school*: (*t=1.674, df=20, significance <.05=.012*) (See "The Healy Effect" below).

Summary of Teacher Goals

PAIR teachers increased in the level of implementation for all seven goals in the project during Year Three. The summary, with comparisons between Year Two and Year Three, based on quantitative and qualitative data, appears in Table Thirteen. Three goals were not met in 2008-2009 and only one goal was not met in 2009-2010.

These goals were assessed with respect to teachers' self-reported responses on targeted items on the Year-End Curriculum Survey for Year Three of the project and results were also compared to paired control group teachers in each goal area. Grade level means were also reported.

During Year Three, the research team isolated the specific indicator related to documentation of Effective Teaching Practices, formerly included in Goal 2C as part of documentation in general and made it a separate goal, 3 A, for Year Three. During Year Three of the project, the level of documentation, including PAIR PAIL student labels and curriculum maps, coupled with the complete set of portfolio conferences that included each PAIR teacher, provided increased evidence of implementation of all goals above from qualitative sources.

The one goal not met in 2009-2010 was set at a 90% implementation level that, in hindsight, was difficult to achieve with limited designated resources focused on technical assistance for documentation and intensive professional development in assessment practices in this project. The district might consider ways in which to advance more research-based authentic assessments that

address the curricular designs of integration while accommodating content area learning in the Magnet Cluster Schools.

Table Fourteen: Teacher Goal Outcomes Based on Quantitative and Qualitative Data

	2008- 2009	**2009-2010**
Goal 1 Collaboration and Professional Development		
Goal 1 A - 60% of PAIR teachers communicate and collaborate, in and across schools.	**Met**	**Met**
Goal 1 B - 90% of PAIR teachers develop partnerships with teaching artists and external professional development resources.	**Met**	**Met**
Goal 1C- 50% of PAIR teachers provide professional development for their schools.	**Not Met (50% level)**	**Met (Pending November Dissemination Session)**
Goal 2 Curriculum Documentation, Inquiry and Reflection		
2 A – 50% of PAIR teachers develop inquiry questions and demonstrate reflection.	**Met**	**Met**
2 B- 90% of teachers develop and document at least one curriculum plan representing arts integration and content learning aligned with Illinois Learning Standards.	**Met**	**Met**
2 C 90% of PAIR teachers document implementation and assessment of arts integrative curriculum in the PAIR project.	**Not Met (80% level)**	**Not Met (90% level)**
Goal 3 Improved Instruction through Arts Integration		
3 A- 90% of PAIR teachers use and document research-based effective teaching practices in the PAIR unit	**Not Met (80% level)**	**Met**

Professional Development and PAIR

Over the past 25 years, professional development for teachers has gone from a choice to a mandate (Lieberman and Wilkins, 2006). Further, professional development is increasingly cited as a key mechanism for improving schools (Lieberman and Wilkins, 2006; Frechtling and Killeen, 2003; Elmore, 2002; Newmann, King and Youngs, 2000). This means that districts must learn from small-

scale projects such as PAIR in order to scale up professional development that is meaningful for a larger set of schools with common characteristics.

There are many models for professional development in educational literature that share common features, including grade level, content, or team professional development coupled with individual experiences, focus on content, attention to inquiry and analyzing student work. The emphasis on content has been a central feature of professional development research for decades (Grossman, Schoenfeld with Lee, 2005, Invarson, Meiers and Beavis, 2005; Guskey, 1985; Joyce and Showers, 1982). A focus on content and reported impact on practice is clearly evident in this research. PAIR's design to intentionally focus on mathematics in two schools, culture and language in two schools, and writing in two schools addresses this criterion clearly. The results of this research indicate that content expertise is a variable in the discussion of the degree to which teachers implement an innovation in arts integration with external partners.

Data collection with respect to professional development focused on artifacts constructed during the sessions the Professional Development Session Surveys completed for each session, and attendance data (See Maps in Goal 2B above, Session Discussions in Tables Fifteen and Sixteen and Documentation Panels in Figure D, E, F below). The four sections of the Professional Development Session Survey correspond to Teacher Goals (see above sections) and are parallel with sections of the Year-End Curriculum Survey. Means for items in the four sections of the Session Survey (*Collaboration and Planning*, *Peer to Peer Sharing* and *Extending/Deepening the Curriculum*, and *Documentation of Student Learning*) are reported below, with sample documentation as they were collected during the professional development sessions.

The Professional Development Session Survey's scale was based on the research by Hall and Hord (2006) regarding the measurement of the *level of use* of an innovation (See Table Fifteen). The Survey is compelling in that it asks teachers to not only comment on the experiences in the professional development session, but also asks them, in the same Likert scale, to assess the degree to which the indicator is happening in their own setting and context. In other words, the teachers were asked to comment on the degree to which the professional development goals and the actual teaching occurring are parallel. Current research in this area is very much focused on how to assess the connections between what teachers experience in professional development and what they are doing in their classrooms.

Table Fifteen: Level of Use of An Innovation (Hall and Hord, 2006)

(*How do we know when an innovation has taken hold?)*

Users • Renewal • Integration • Refinement • Routine • Mechanical Use
Nonusers • Preparation • Orientation • Nonuse

This survey incorporating Hall and Hord's levels of use with respect to arts integration innovations needs more development and testing, but holds promise as a means of discerning the intersections between professional development and implementation of effective teaching practices.

Collaboration/Planning and PAIR Professional Development

Just as the research team asked teachers to document, so we also documented the professional development session discussions in order to study the coherence between session topics and levels of use in teachers' practice. The professional development session documentation of participant discussions for Year Three appears in Tables Sixteen, Seventeen, Nineteen and Twenty.

In October of 2009, the PAIR teachers and artists conducted a brainstorming sessions focused on collaboration (See Table Sixteen).

[continued on next page]

Table Sixteen: Pair Project Brainstorming on Teacher/Artist Collaboration

Notes from a Professional Development Session

What does co-teaching with artists and teachers look like?

- *Planned session in which artist takes the lead, then the lead shifts to the teacher, then back again… supporting actor vs. lead actor – planned in terms of minutes…(what could artist do when teacher is the lead actor?)*
- *Teacher restates*
- *Teacher reminds*
- *Teacher has a visual in the classroom to refer to/add to/document on as the t.a. teaches*
- *Teacher actively DOES the project with the students – becomes a part of a group*
- *Teacher leads the 'debrief' by interviewing the artist about the session they just completed*
- *Teacher documents – takes photos, takes notes, interviews groups or students as they are working – and makes that documentation visible for students and others*
- *Teachers takes two small groups; artist takes two small groups – they work parallel on the task presented by the artist*

What does it mean to reinforce or respond to work with what the teaching artist does when the artist is not there?

- *Teaching artist and teacher plan for 'BEFORE THE VISIT' and 'AFTER THE VISIT' intentionally and explicitly*
- *Teacher has students journal/write respond to content before or after artist visit*
- *Teacher maps content connections with arts experience with students*
- *Teacher focuses on arts and non arts content vocabulary and links them visually in the room*
- *Teacher and artist create common assessments – including paper/pencil assessments, student critiques, conference sessions*
- *Teacher names and reinforces content standards being addressed in and through arts integration/artist visits*

In that same Professional Development session, the research team shared some of the results from Year Two of the PAIR project regarding collaboration (See Table Sixteen).

Table Seventeen

Sharing Data on Collaboration: Notes from a Professional Development Session

An impressive 73.3% of PAIR teachers reported on this survey at the end of Year Two that they believed they have been participating in a learning community with goals aligned to the magnet cluster initiative, their schools, and Chicago Public Schools. Further, almost half of the PAIR teachers (46.7%) noted that they felt they did receive the district support necessary to collaborate across schools with community partners.

- *There is evidence that explicit collaboration among PAIR teachers across the PAIR network has been challenging again during Year Two of the Project. Only 14.2% of PAIR teachers reported that this collaboration was occurring.*

- *The Post-PAIR Annual Professional Development Survey (self-report) revealed that 73.4% of participating teachers now feel somewhat or very successful collaboratively planning arts integration projects with colleagues.*
- *On the Annual Professional Development Survey administered to PAIR teachers, the 4th grade teachers, who have just completed their second year in the program, reported they were feeling significantly more successful in collaboratively planning arts integration projects with their colleagues than 5th grade teachers, who were new to the program this year, did (t=2.530, df = 8, significance <.05 = .035). These results indicate that the PAIR program is having an effect over time and has promoted cross-school curricular sharing and reflection, which can only be fostered in multi-year initiatives.*
- *When asked about specific activities that contribute to cross-school sharing, teachers did not appear to be too active:*
  - *7.2% report that they helped to drive effective and real collaboration as part of PAIR with multiple external partners and community resources.*
  - *17.8% report that as part of the PAIR project they exhibited leadership in getting artwork displayed or performed outside of school, finding an audience during the year.*
- *PAIR TEACHERS*
  - *10.7% report that as part of the PAIR project, they assisted in scheduling for arts integration co-planning and co-teaching time with teaching artists.*
- *PAIR TEACHERS:*
  - *32.1% reported that they reinforced or built on arts activities when the teaching artists were not there.*
  - *17.9% reported that they taught with teaching artists, integrating teaching and learning in arts and non-arts content areas.*
- *The 4th grade teachers (who have been in the project for two years) reported significantly more frequent participation in driving effective and real collaboration than the 5th grade teachers, new to the project this year, did(t = -2.185, df=14, significance < .05=.046)*
- *PAIR TEACHERS:*
  - *21.4% reported, through PAIR, being part of a leadership team to plan for the arts in their schools.*
  - *25% reported that they shared the PAIR unit with their faculty peers who were not involved in the project.*
- *It is interesting to note the extent to which PAIR teachers felt that the professional development provided by the project enabled them to plan, implement and assess the arts integration/content units.*
  - *60% reported that they participated in PD occasionally or frequently that helped them learn how to plan for curriculum integration.*
  - *53.3% reported that they occasionally or frequently participated in PD that helped them learn how to implement curriculum integration lessons.*
  - *46.7% reported feeling somewhat successful at assessing students' learning in the content field in the PAIR project. No teachers reported feeling very successful at this assessment.*
- *PAIR teachers generally reported feeling somewhat successful with using new instructional approaches learned through arts integration in PAIR project (66.7%), although no project teachers reported that they felt very successful in this endeavor.*
- *Nearly half of PAIR teachers in Year Two noted that students were beginning to participate in documenting their learning in these projects (42.8%) and 17.9% of the PAIR teachers reported planning with TAs to document student work that shows arts and non-arts content learning.*

On the PAIR Professional Development Session Survey, Section One asked the teachers to comment on various elements of the collaboration they were engaged in with teaching artists and other teachers (items 1 – 6). Scores on these items across the four professional development sessions ranged from 3.3816 to 4.5132, indicated that teachers acknowledged that *some elements of collaboration were not currently happening in their site, but they received some preparation for those elements in the PD* while others reported that *elements of collaboration were happening in their sites and that they were making progress toward making them routine, thanks to this PD* (italics = language of the Likert Scale on the survey).

Table Eighteen

Professional Development Session Survey Means Across Four Sessions for Section 1: PAIR Collaboration and Planning

PD		Q1	Q2	Q3	Q4	Q5	Q6
Session 1	Mean	4.5000	4.5000	3.8846	4.2692	3.9615	4.6154
	N	26	26	26	26	26	26
	Std. Deviation	1.06771	1.24097	1.70474	1.53773	1.79957	1.57675
Session 2	Mean	3.7308	4.4615	4.1923	4.1154	2.8077	4.1923
	N	26	26	26	26	26	26
	Std. Deviation	1.75631	1.20767	1.47022	1.42343	1.81150	1.52366
Session 3	Mean	4.0667	4.7333	5.0667	4.2667	3.7333	4.7333
	N	15	15	15	15	15	15
	Std. Deviation	1.70992	1.43759	.88372	1.27988	1.83095	.70373
Session 4	Mean	4.1111	4.3333	3.6667	3.8889	2.7778	4.0000
	N	9	9	9	9	9	9
	Std. Deviation	1.69148	1.50000	2.12132	1.53659	1.71594	1.50000
Total	Mean	4.1053	4.5132	4.1974	4.1711	3.3816	4.4211
	N	76	76	76	76	76	76
	Std. Deviation	1.53691	1.28056	1.59181	1.42724	1.84728	1.41669

Documentation and PAIR Professional Development

In other sessions, we asked them to comment on the nature of the data as well as the data collection systems in the PAIR project, discussing how we also are documenting and learning from documentation (See Table Eighteen).

Table Nineteen: Sharing Results from the Research with PAIR Participants: Notes from a Professional Development Session

What do we notice? What does it mean for this school year?
What questions does this raise for you?
Teacher Data

- *Why is the scheduling percentage (10.7%) so low?*
  - *Pay attention to what it takes to schedule the co-planning and co-teaching with the teaching artists.*
  - *Perhaps, teachers felt they didn't have enough time to plan and that is why people responded a certain way.*
  - *It is likely that people interpret the question differently*
    - *"Co-planning" versus "co-teaching"*
      - *What does co-teaching look like? What does it mean for the classroom teacher, for the teaching artist? What does it look like?*
      - *You can plan together and teach the same concepts, yet perhaps not teach at the same time.*
      - *Would it be helpful to define "co-teaching" as a group of PAIR teachers?*
- *Learning community → How do you define this?*
- *How can the teacher enhance the 10 hours with the teaching artist – when the teaching artist is there and when the teaching artist is not there?*
  - *Teacher can get in and participate, do the art making with the students; restate the teaching artists' questions and ideas for the students.*
  - *What is the intersection between what happens when the teaching artist is, and isn't, there?*
- *What does it mean to "reinforce and build on arts activities?"*
- *Year 2 data shows differences between 4th and 5th grade teacher responses – 4th grade teachers are much more confident.*

CAPE's methodology consistently incorporates the documentation of student learning in ways that communicate evidence of both process and outcomes. The Professional Development sessions during all three years of the project supported teachers' efforts to view documentation as part of their teaching, not separate from it. In one Professional Development session, teachers were asked to bring student PAIR work and develop documentation panels to demonstrate learning (See Documentation Panels in

Figures D, E, F). They then discussed the value of doing such panels, audience for the panels, and what could make the panels stronger as evidence of quality learning (See Table Twenty).

FIGURES D, E, F: Documentation Panels Constructed During PAIR Professional Development

Table Twenty: Collaborative Reflection on PAIR Documentation Panels:

Notes from Professional Development Discussion

Documentation Panels: What was the entry point?

- *Key terms – choreography, sculpture/shapes*
- *Use the checklist for effective documentation panels*
- *What do I want to see as a viewer?*
- *Needs an introductory piece – 'what is this thing?' A title is important*
- *Who is the audience? – Make the project explicit – 'THIS PANEL TELLS THEIR STORY'*
- *Find the themes – how are all the activities related? What was the team looking for?*
- *Ask yourself some questions in order to organize documentation*
- *Math concepts – problem solving, team work*
- *At first, the activities didn't seem to be related, but after awhile, the coherence emerged*
- *The role of student reflections/teacher/researcher reflections*
- *The role of the portfolio conference – framed the documentation*
- *Necessary narratives/added text for coherence/clarification*
- *Look for 'similar' artifacts and clustering them*
- *Research results helped to group the work… framing the work*
- *Analysis of the documentation becomes part of the professional development, part of the planning, part of the reflection/critique*

What caught your eye? What stood out?

- *The importance of imagery/photos*
- *Made me think of what I should do*

Why did we have you work with other schools' doc? And why in groups?

- *Take into account how much we bring to the pail – is it enough to tell our story*
- *Showing relevant stuff – what makes sense – what actually shows the process – if it was my own work, I would know what was going on; the artifact needs to speak for itself*
- *If you are able to understand the context, then that's a plus*
- *It's not an exhibition – it tells a story – it's not just a poster – a documentation panel is*

different – it is intentional; it requires context/narrative

- *Groups – pull more information through conversation – documentation panels are facilitated through discussion – it's a collaborative investigative process*
- *Trying to uncover the inquiry question…back mapping –*
- *The documentation reflects the group processes that many of the students used in the PAIR projects – the documentation should be reflective of the process in the classroom*
- *Documentation becomes the tool for sharing PAIR… how you think about your work as a team in a school….and the 'panels' that result from work in the PAILS.*

Documentation vs. Data: How are they related?

- *Documentation is not necessarily evidence.*
- *Doc. is a larger universe.*
- *Data – or evidence – answers a question.*
- *We need the context data regarding the teacher/artist process*
- *What did the teacher do?*
- *Where was the collaboration among the adults?*
- *How do we link the documentation to the teacher goals? Per the survey?*

On the PAIR Professional Development Session Survey, Section 4 asked teachers to comment on the session with respect to documentation (items 14-20). Mean scores for these items appear in Table Twenty below. Scores for these items across the 4 sessions of professional development ranged from 3.9041 (*Not currently happening/Some preparation received though in this PD*) to 4.6267 (*Happening now/progress toward making it routine thanks to this PD) (*italics = language of the Likert Scale on the survey).

This range across the items focused on what teachers are learning – and then doing – with respect to documentation is the highest of the four sections on the Professional Development Survey. These means suggest that the teachers recognize that the PAIR Professional Development has focused on documentation and that they are doing more of it in their schools as a result.

[continued on next page]

Table Twenty-One

Professional Development Session Survey Means Across Four Sessions:

Section 4 Documentation of Student Learning

Q14	Q15	Q16	Q17	Q18	Q19	Q.20
4.3846	3.9615	3.7308	3.3200	3.6154	4.2692	3.8846
26	26	26	25	26	26	26
1.32897	1.63660	1.56353	1.72530	1.67516	1.48479	1.79615
4.3600	3.9600	4.2400	4.2083	4.2000	4.4800	4.2800
25	25	25	24	25	25	25
1.55134	1.69509	1.36260	1.31807	1.44338	1.15902	1.13725
4.7333	5.0000	5.2000	4.3333	4.6667	5.2000	4.6000
15	15	15	15	15	15	15
1.22280	1.36277	.94112	1.54303	1.29099	.86189	1.05560
4.1111	4.5556	4.4444	4.0000	4.7778	5.1111	4.7778
9	9	9	9	9	9	9
1.26930	1.66667	1.23603	1.73205	.97183	.78174	.83333
4.4133	4.2400	4.2800	3.9041	4.1600	4.6267	4.2667
75	75	75	73	75	75	75
1.36652	1.63443	1.42904	1.59133	1.49811	1.23871	1.37873

Peer to Peer Sharing and PAIR Professional Development

The Session Survey asked teachers each time they attended a Professional development session to comment on the degree to which they were presenting and discussing PAIR work with other teachers at their school and outside of their school (Items 11 and 12). The Survey also asked them to comment on the degree to which the professional development was providing them with ideas for arts integrated activities in general (Item 13). These means were the lowest ranges of the four sections on the survey, ranging from 1.1806 to 3.9467, indicating that teachers acknowledged that they were not doing much sharing and that means for such sharing were not thoroughly addressed in the professional development sessions (*1 – not currently happening/not addressed in PD, 3 = Not currently happening/some preparation received though in PD)* (See Table Twenty-Two)

Table Twenty-Two: Professional Development Session Survey Means Across Four Sessions: Section 3 Peer to Peer Sharing

Q11	Q12	Q13
3.8846	3.7308	.1250
26	26	32
1.65715	1.31325	.70711
3.6400	2.7600	.0000
25	25	16
1.60416	1.50776	.00000
4.4667	4.1333	3.9333
15	15	15
1.24595	1.64172	1.86956
4.1111	3.0000	2.4444
9	9	9
1.36423	1.50000	1.42400
3.9467	3.4000	1.1806
75	75	72
1.53247	1.54220	1.93796

These data are consistent with the data received through the Year-End Curriculum Survey and other qualitative data sources (See Goal 1C above). It seems that teachers may not see themselves in cultures in which sharing is possible and adaptation of activities beyond the project itself are possible. If such goals are part of partnerships such as PAIR, then more district support such as the pending dissemination event scheduled for November 2010 are essential.

Extending and Deepening the Curriculum and PAIR Professional Development

Items 7 – 10 on the Professional Development Session Survey asked teachers to comment on how they were engaging PAIR with their focus non-arts content area. The mean scores ranged from 4.2133 to 4.4868. **This section of the survey reflected the strongest report of preparation in the professional development sessions as well as the strongest report of the teachers actually integrating what they were learning to extend and deepen their curriculum through PAIR experiences** (See Table Twenty-Three).

Table Twenty-Three: Professional Development Session Survey Means Across Four Sessions:

Section 2: Extending and Deepening the Curriculum

Q7	Q8	Q9	Q10
4.7308	4.7692	4.3846	3.6923
26	26	26	26
1.25085	1.17670	1.09825	1.73826
4.3077	4.2800	3.7200	4.1600
26	25	25	25
1.54322	1.24231	1.33915	1.02794
4.4667	5.1333	4.5333	4.8000
15	15	15	15
1.24595	.91548	1.35576	1.42428
4.3333	4.5556	4.6667	4.8889
9	9	9	9
1.11803	1.23603	1.00000	1.36423
4.4868	4.6533	4.2267	4.2133
76	75	75	75
1.33160	1.17971	1.25820	1.47312

Attendance at PAIR Professional Development Sessions

A final parameter regarding impact of the PAIR project on teachers was the attendance at Professional Development sessions offered by PAIR/CAPE staff. Table Twenty-Four below indicates teacher attendance at PAIR sessions during Year Three (2009-2010). Table Twenty-Five indicates teacher attendance at PAIR sessions across the three years of the project (2007-2010).

While it is clear that the 6th grade teachers attended fewer sessions than their peers in 4th and 5th grade, due to the structure of the program, it is still interesting that most 4th and 5th grade teachers continued to attend beyond the first year of the program. Only 3 sixth grade teachers attended 3 or 4 of the sessions in Year Three, even though they were new to the program and ostensibly would have found the Professional Development helpful.

Table Twenty-Four: Attendance at Professional Development Sessions: Year Three of PAIR Project (grade levels in parentheses)

Attended All Four General Sessions	1.40 (5) 1.20 (4) 3.20 (4) 5.60 (6)
Attended Three of Four General Sessions	1.60 (6) 3.60 (6) 3.30 (5) 5.30 (5) 7.40 (5) 7.10 (4) 9.60 (6) 11.40 (5)
Attended Two of Four General Sessions	1.50 (6) 1.10 (4) 3.40 (5) 5.40 (5) 5.20 (4) 7.50 (6) 7.60 (6) 9.50 (6) 9.30 (5) 9.40 (5) 11.60 (6) 11.30 (5) 11.10 (4) 11.20 (4)
Attended One or None of Four General Sessions	3.10(4) 3.50 (6) 7.30 (5) 11.50 (6)

*5.10 – attendance not recorded

(continued on next page)

Table Twenty-Five: Attendance at Professional Development Sessions

Across Three Years of PAIR Project

Teacher Names	school	grade	PD Attend Other	PD Attend 08	PD Attend 09	PD Attend 10	TOTAL
1.20	1	4	3	3	4	4	14
3.20	3	4	2	2	3	4	11
1.10	1	4	3	3	3	2	11
5.30	5	5	2	1	4	3	10
3.30	3	5	1	2	3	3	9
7.30	7	5	2	1	4	1	8
3.10	3	4	2	2	3	1	8
7.60	7	6	2	1	3	2	8
3.40	3	5	1	2	3	2	8
11.10	11	4	1	2	2	2	7
7.40	7	5	0	1	3	3	7
1.50	1	6	1	1	3	2	7
9.30	9	5	1	1	3	2	7
1.40	1	5	3	0	0	4	7
7.10	7	4	1	0	2	3	6
5.20	5	4	1	1	2	2	6
5.60	5	6	0	1	1	4	6
9.40	9	5	1	1	2	2	6
11.30	11	5	0	1	2	2	5
5.40	5	5	0	1	2	2	5
5.10	5	4	1	1	2		4
11.40	11	5	0	0	1	3	4
11.60	11	6	0	1	1	2	4
11.20	11	4	0	0	1	2	3
9.60	9	6	0	0	0	3	3
3.60	3	6	0	0	0	3	3
1.60	1	6	0	0	0	3	3
9.50	9	6	0	0	0	2	2
7.50	7	6	0	0	0	2	2
3.50	3	6	0	0	0	1	1
11.50	11	6	0	0	0	1	1

The project was designed to build a professional community throughout the three years, with sufficient time and collective participation in professional development both within and across partner schools. This too is a characteristic of effective professional development programs (Ingvarson, Meiers, Beavis,

2005). Teachers were repeatedly invited to consider means of collaborating within their grade levels, across grade levels, across schools and across the PAIR project as well as quite obviously with the teaching artists working at their schools. Joyce and Showers note the importance of "the proactive and productive use of peers" when teachers are learning new skills and changing their practice (2002, p. 3). The three-year PAIR project in which 4th grade teachers took the lead in planning an integrative arts project with support from CAPE and teaching artists, and then engaged 5th grade teachers the second year of the project and 6th grade teachers in the third year represents an interesting approach to problem solving, shared planning, and the transfer of skills and creative application of teacher learning.

Franke, Carpenter, Levi and Fennema conducted a study of math teachers' generative change as a result of focused professional development on student thinking and understanding (2001). One of their indicators of professional development effectiveness was the degree to which the teachers continued to implement the principles of the mathematics program four years after the intervention ended. Though the PAIR program has just completed its three-year cycle, the portfolio conference comments from teachers regarding their expectation to integrate what they had learned in the project in future teaching did emerge (See Table Twenty-Six). Some teachers made no mention of how PAIR might affect their teaching or did not respond to the interviewer's questions in that regard. Other teachers were clearly able to describe how the project will make a difference in their classrooms. Of course, ideally, the research could continue in order to follow teachers for several years after a project ends in order to more accurately determine whether there are residual effects from involvement in an innovation such as PAIR.

The Franke et al study also provides some compelling insights relative to the portfolio conference methodology. In the Franke et al study, the professional development focused on student thinking and the researchers felt that it was this focus that really contributed to substantial change in the teachers' thinking. The portfolio conferences in the PAIR project provided the teachers with a means to observe students' thinking as expressed in their explanations, descriptions, and evaluations of the work products in the conferences. The teachers could not interact with the students during the conferences, but rather observed the children interacting with each other. The students in turn were able to observe their teacher's analysis as she/he answered questions offered by the interviewer. This approach, if deepened and repeated, may be a valuable mode of professional development in and of itself for future

projects. Learning about student thinking is one important way to plan for teaching to address gaps, misunderstandings and students' inabilities to articulate what they know and how they know it.

Table Twenty-Six: Teacher Descriptions of Practice After PAIR Project Ends

Portfolio Conference Comments
And if I'm here next year, I will do the same things with the novel that I did with Charlie and Reggie even if they weren't...you know, even though they're not going to be involved in the classroom. I would still go through the same lessons that they went through. And the stuff that I got from Reggie is very useful in helping teaching students how to write using their own voice. So for my own personal, I see a major benefit in how it's going to help me teach things in the future. And we do refer to it throughout the year, so I don't feel that it was something that was just done in isolation (11.30). Well, actually, I was talking to (the teaching artist), I would actually do this same building. Last year we did something similar, but we used a different scale. We used centimeter cubes to six inches. I think you remember we made those huge buildings. And this really goes along well with our math series right now (5.20). Go beyond paper and pencils. Involve more art. Because one way it kind of like got the creativity out there of their head and also allowed them to do things that is more fun, not all like paper and pencil boring stuff, you know. And I think when they are having fun doing things that they like to do, they were able to give me a little bit more of what they were asked. So just like the biography, like this, we can use it next year, because I know how to do it. I know step by step how to do it. I think the kids feel more comfortable to express themselves if they are allowed to think through it, just like the portrait they did before – the kind of food they eat, their facial feature, the color of their face represent who they are. So instead of how we usually do, oh, a biography, you need an introduction, what day they were born, how they're...it's more...like going through this art project let them kind of like a brainstorm, let them think through it before they put it in writing. So it's kind of like help them in their thinking process and their final product. Not only just they do it because the teacher tells them to do. I learned things from it, too. So I can get what we learned here to go do it (3.20).

Portfolio Categories of Reflective Practice and Professional Development Session Attendance

If we return to those who demonstrated the reflective practice categories from the portfolio conferences and now correlate it to other data points, qualitative and quantitative, some further distinctions now occur (See Table Twenty-Seven). First, of the 12 teachers who appear in the High and High/Middle reflective occurrence categories, 7 are 4th grade teachers (1.10, 3.20, 5.10, 5.20, 7.10, 11.10, and 11.20), 5 are 5th grade teachers (1.50, 7.30, 7.40, 11.30, and 11.40). All attended at least 25 % of the

Professional Development Sessions offered and 9 of the 12 attended 50% or more of the Sessions as part of the PAIR project.

Table Twenty–Seven: Portfolio Conference Coded Categories of Reflective Practice: High and High/Middle and Number of Professional Development Sessions Attended

Yellow (High)	Green (High/Mid)
3.20 (grade 4) (attended 11 PDs)	1.10 (grade 4) (attended 11 PDs)
5.10 (grade 4) (attended 7 PDs)	1.50 (grade 5) (attended 6 PDs)
7.10 (grade 4) (attended 6 PDs)	5.20 (grade 4) (attended 6 PDs)
7.30 (grade 5) (attended 8 PDs)	7.40 (grade 5) (attended 7 PDs)
11.10 (grade 4) (attended 7 PDs)	11.20 (grade 4) (attended 3 PDs)
11.40 (grade 5) (attended 4 PDs)	11.30 (grade 5) (attended 5 PDs)

Four Teacher Profile Effects and PAIR Program Impact

The data analysis regarding teachers in the PAIR project raises some interesting hypotheses that merit further investigation. The data indicate specific teacher profiles that relate to categories of teacher impact, based on whether these teachers *demonstrated significant differences from their peers on specific variables*, as reported on the surveys, the *coded comments from portfolio conferences*, the *coded open-ended responses on the surveys*, and the *reported pedagogy identified on student work labels*. Based on a mixed methods approach that incorporates these qualitative and quantitative data, the following effects show promise with respect to impact of the program and possible relationship to student achievement:

- **The Content Expertise Effect: Pedagogical Content Knowledge Matters**
- **The Documenting to Learn Effect: Collecting Student Work Inspires Reflection**
- **The Fourth Grade Effect: Designers Have Ownership**
- **The Healy Effect: Initiatives Build on Other Initiatives in a School**

1) The data suggest that teachers in arts partnerships *who stay focused on the learning in the non-arts content* that they are responsible for are more likely to see student achievement. In the PAIR project, we might call this **"The Content Expertise Effect".** Shulman explains this phenomenon in a way that is appropriate for the discussion here. Shulman claims that teachers need more than straightforward

content knowledge. They need pedagogical content knowledge, or PCK, defined as a blend of content and pedagogy "into an understanding of how particular topics, problems, or issues are organized, represented, and adapted to the diverse interests and abilities of learners and presented for instruction" (1987, p. 8).

An arts integration innovation has the potential to contribute to teachers' pedagogical content knowledge (PCK), if the teachers are alert and focused on their non-arts content throughout the project. Some teachers in the PAIR project demonstrated more pedagogical content knowledge (PCK) than others. Teachers appearing to have strong PCK knew what students would be able to transfer in an arts integration context. They had multiple ways of describing or explaining to their students in ways that are appropriate for the grade and age they were teaching and were able to translate and make connections to their content from the art forms, even when they were not familiar with that art form. Teachers with strong PCK had a sense of the whole of the content field and the particular components of the content that they were teaching at their grade level.

According to Shulman, PCK is what distinguishes a content specialist from a true teacher. Even if these teachers did not fully invest in the professional development offered by the project or feel fully capable of experiencing the art forms offered in their classrooms, they remained focused and confident of their non arts content goals; they knew the gaps in students' understandings in those content areas; they knew what they needed to expect from the teaching artist partnerships to achieve specific content goals for their students. Having such knowledge and expertise in their subject areas enabled them to be clear on the goals for their students and their classrooms. Teacher knowledge of their content is even more critical in an innovative, inquiry-oriented classroom (Shulman, 1987). Table Twenty-Eight illustrates comments by five teachers who demonstrated focused attention on their non-arts content and were able to articulate elements of that curriculum consistent with pedagogical content knowledge. The project research did not intend to focus on this aspect of the professional development engendered through the PAIR project. But the PAIR design, which explicitly focused on specific non-arts content (mathematics, writing or culture/world language), seemed to encourage participating teachers to demonstrate their expertise in their content and focus the project specifically toward student learning in that discipline, once they were comfortable with the teaching artists and the concept of arts integration in their classrooms. There were, as noted above, some teachers who did not fully embrace arts integration as participants in the arts experiences (5.20 and 11.10); but those teachers also

demonstrated their attentiveness to the benefits of the approach for their students and attested to the value added for their classrooms. In other words, the teachers with PCK does not have to dance or even appreciate dance as an art form to know how dance can contribute to content learning for students.

Table Twenty-Eight: The Content Expertise Effect: Pedagogical Content Knowledge Matters

Portfolio Conference Comments: Focus on Non-Arts Content
I think definitely they get...this really went along with and supported the math curriculum that we do. We work with hexagons; we work with trapezoids and rhombuses and triangles and fractions, and teaching them parts of a whole, so that really helped with this. They also, which they didn't say, with making these hexagons, they did a ton of measurement to make sure the sides were equal. Some of them did better jobs than others, but they did a lot of measurement, which was something that they really, really needed for their work. We do go over the concepts, but this is like a whole – it makes them generalize the information that they've learned, which I think is kind of what you're trying to do (5.10). *Well, for myself, the value of collaborating with an artist is that I understand what their technique is and how it can help my students advance in their development of writing. And I keep referring to the writing because that's what our focus is, is developing their writing. So collaborating with them helps me gain a deeper understanding of their arts and how I can integrate it with what I do in my classroom (11.10).* *I put it in my research that I did on my master's program. We were trying to expand on what the students had learned in 4th grade and what they had learned with PAIR, and we were focusing on having the students develop their voice and recognizing point of view with main characters in a fiction story. The main purpose that we were, what we were looking to do was on all writings, regardless of whether it's ISAT or what we're using it for, we always want the student's own voice to come through and not just mimic what it is that they've read, but we want their input, their take on things (11.30).* *For me, as a teacher, it has really helped me to instruct my students how to write better, I mean, how to really look at not only when we read a story, about characterization, motivation, plot. I will tell you this class, we have read some really difficult books this year, higher level, higher thinking, and I really do think having the PAIR program has helped my students look at characters, understand plot, conflict, resolution a lot better than my other classes* (11.20). *Scale is a very difficult concept. It's a difficult concept, especially application of scale. You know, you can do it on pencil and paper, but to actually apply it is a whole other ball game. Not only can they do it on pencil and paper, but they could apply it, you know, apply it to build something or construct something now, whereas...you know, that is really difficult to get a student to comprehend that. I think it's beneficial on the test because it will help them retain it. Once you have applied it and used it, you have a tendency to retain it more than, okay, this is how you do it and it's gone. I think it's of great value because the kids were able to talk about math in a reasonable way. They truly have a much better understanding than they would have if I said open up to Page 200 and let's do scale (*7.30).

* * *

2) *Teachers who document (*including collecting student work, reflecting on teaching, outlining the curriculum plan) *and assess* regularly as part of their teaching in an arts partnership are more likely to see student achievement. In the PAIR project, some teachers more than others clearly saw the value of documenting as they were asked to do for themselves and their students, though this realization did not come easily or immediately in the three-year project. We might call this **"The Documenting to Learn Effect".**

The teachers who were more clearly invested in the PAIR Pails were also more able to articulate what they and their students were learning as evidenced in the Portfolio Conferences. In other words, for some teachers, the act of documenting and the engagement in discussion of documentation at professional development sessions contributed to the impact that the PAIR program had on their learning and their teaching. The teachers who offered reflections on their teaching in the form of student work labels were also the ones who were able to comment on the benefits of students participating in the documentation and the benefits for teachers who can learn from what they collect through a documentation process (See Tables Twenty-Nine, Thirty and Thirty-One).

Table Twenty-Nine: The Documenting to Learn Effect: Collecting Student Work Inspires Reflection

I just think that sometimes we learn things, we take a test, and then maybe the test hangs on the refrigerator for a little while and then the test goes away in the garbage or the recycling. This, you know, you have a hard time throwing something like this out. You do get to kind of peek inside their psyche and kind of get an idea of what's going on in their heads. And then we have evidence at the end, and it's evidence that they can be proud of. It's beautiful and it's something that we don't want to go away. We want to hang it and we want to be proud of it. That's what I feel is nice about this whole process (1.50*).* *I'm using things that, you know, we've done the pail, like we've finished with it, and now I'm just seeing that this should be incorporated because, again, this is so much better than what we were doing in February...no, in December, and I would really like to hand in my pail almost at the end of the year, you know, because really, we keep reiterating what we've learned (11.20).*

Table Thirty: The Documentation Effect: Teachers' Responses on Year-End Survey: Impact of PAIR Documentation of Learning by Students

Students can learn from each other (1.10).

They would say "Wow". This is what I have at first and now the final product. They were amazed to see from the beginning with just the basic and at the end they created something very amazing (1.20).

It shows the students the big picture (3.10).

They can actually see where they began and where they ended. It helps them to understand the process as well as the finished product. I think it allows them to see how far they have come from an idea to an actual finished product that represents them (3.40).

Perhaps they see how much their work is valued (3.60).

They're able to be more realistic and honest about their effort and their journal entries reflected how they were evaluating what they were doing (5.30).

When students see their work in documentation, they are able to identify the missing part (5.60).

When they reflect back on themselves and their work, they are usually very hard on themselves. They are able to see what they did well and what they would do differently in the future (7.50).

They can see their growth and assess their own learning. They can see what they might need to revisit and what they accomplished as a whole group (7.60).

Students can assess themselves and decide if or where improvement is needed. It also helps their confidence to grow (9.50).

They learn the mistakes, they correct them, and they learn not to do them again (9.60).

A student does not learn if they never see their graded/evaluated work. Providing students with feedback on their work allows students to focus on areas of need (11.10).

Students have a feeling of ownership and enjoy doing more work (11.20).

Students see themselves collaborating with their peers, working toward a common goal. Students observe the given and take needed to work successfully together (11.60).

Table Thirty-One

The Documentation Effect: Teachers' Responses on Year-End Survey:

PAIR Documentation of Student Learning by Teachers

Seeing that my students who didn't know what was going on are able to do, especially the special needs students. I was able to see the progress of the students' work. They were amazing (1.20).

Teachers have to remember a lot day to day. I feel without good documentation things get lost in our heads very easily and good ideas are forgotten (7.50).

It helps me evolve as a teacher because I see the process (9.60).

I witnessed a growth in my teaching and writing (11.20).

Similar to students, if teachers receive no feedback in their teaching methods they have no way of knowing if their teaching is effective or where they may need to make adjustments (11.30).

Having the unit document provides evidence of success and failure. Know what works and what needs improvement is key to developing an engaging, informative activity (11.30).

When viewing pictures, documenting forced me to evaluate the effectiveness of an assignment. When you document your practice it preserves the pathways you traveled to achieve your goal. It allows you the ability to review and critique what you did. Then, when you repeat the same or similar lesson, or show it to colleagues, you can delete or add on to the plan as needed (11.60).

* * *

3) Teachers *who actively contribute to the design of an arts partnership project* are more likely to fully participate throughout the life of the project and see resulting student achievement. In the PAIR project, this might be known as **"the 4th grade effect",** in that the 4th grade teachers show greater evidence of investment, ownership, and consistent documentation of achievement than do most 5th and 6th grade teachers who joined the project later.

Results from Year Two of the project revealed that the 4th grade teachers (who have been in the project for two years) reported significantly more frequent *participation in driving effective and real*

collaboration than the 5th grade teachers, new to the project this year, did *(t = -2.185, df=14, significance < .05=.046).*

This year, in Year Three, with respect to the coded reflective practice totals from teachers in the portfolio conferences, 7 of the 12 highest scoring teachers were 4th grade teachers.

Means from the Year-End Curriculum and Teaching Survey collectively were compared across the three grade levels groups through an ANOVA. Because the PAIR project was structured as a graduated model, in which a grade level was added each year, this comparison provides some worthwhile information.

There were significant differences between the 4th, 5th, and 6th grade teachers on items 5 and 6 regarding the degree to which teachers reported meeting with the artists/arts specialists and with other teachers in the PAIR project with the 4th grade teachers reported significantly more planning:

Item 5: (ANOVA, significance <.05 = .049)
4th grade teacher mean: 3.6667
5th grade teacher mean: 2.7000
6th grade teacher mean: 2.1000

Item 6: (ANOVA, significance <.05 =.027)
4th grade teacher mean: 4.4444
5th grade teacher mean: 2.8000
6th grade teacher mean: 3.1000

These results indicate that the **4th grade teachers, even though they were receiving fewer services, including artist visits, in year three of PAIR, report more collaboration than their peers in grades 5 and 6.** One could infer that the collaboration within the PAIR project remained stronger for the 4th grade teachers, perhaps due to the fact that they were key participants in the design of the project in Year One and continued to be the most invested in the program and the collaborations with the artists.

There were also significant differences across the teacher reported activities regarding arts integration with respect to movement/dance and mathematics (Item 35) and movement/dance and language arts (Item 36). These data suggest once again that **4th grade teachers, regardless of the art forms they experienced with teaching artists, report significantly more movement and dance associated with teaching mathematics and language arts than their colleagues in grades 5 and 6:**

Item 35: (ANOVA, significance <.05 = .007)
4th grade teacher mean: 3.4444
5th grade teacher mean: 2.8000
6th grade teacher mean: 1.7000

Item 35: (ANOVA, significance <.05 = .005)
4th grade teacher mean: 3.6667
5th grade teacher mean: 3.4000
6th grade teacher mean: 1.8000

Although these were the only significant differences between teachers' responses by grade level on the 40-item Year End Curriculum and Teaching Survey, they may indicate trends that other data sources can further clarify with respect to impact of the project on teachers in a project that was structured explicitly for length of involvement according to the grade level taught. There were no significance differences in the control group schools across grade levels. These data would suggest that at least certain teachers at the 4th grade level felt more engagement and participated in arts integration experiences more fully than their peers at grades 5 and 6.

There were some interesting trends with respect to individual items on the year-End Curriculum Survey regarding grade level results. There were six individual teachers from the High and High/Middle Portfolio Conference category groups (total of 12 teachers in those groups) who demonstrated significant differences or approaching significance with respect to their peers in grades 4,5,6 (See Table Thirty-Two). Five of the six teachers are 4th grade teachers who have been with the PAIR project all 3 years. However, in the case of two of those 4th grade teachers (5.20 and 11.10), the differences represented a below mean response. That is, these two 4th grade teachers, who demonstrated High or High/Middle range reflection in the portfolio conference categories were also clear about what did not happen in the project in their experience. Teacher 5.20, reported significantly less *participation as a student, practicing and learning the art form while the artist led the session*. This teacher expressed some discomfort with the performance art form during her portfolio conference, yet was clear about the advantages she perceived for her students in integrating the arts and the non-arts content area. She chose to step back from actively participating in the art making while the teaching artist was there.

Similarly, teacher 11.10, reported much less *discussion and feedback*, i.e., co-teaching, with the teaching artist(s) *for students*. She also reported fewer contributions on her part regarding *innovative arts approaches*. She seemed to experience the project, less as an active participant during the arts integration, but was capable of extending the experience for her students when the artist was not there.

These trends in 2 fourth grade teachers indicate different ways of implementing and perceiving an innovation with an external partner – ways that may in fact be no less effective than those who fully participated specifically in the arts learning with their students and the artists.

Table Thirty-Two: Teachers and Significant Differences from all PAIR Teachers Grades 4, 5, 6 on Specific Indicators as Self-Reported on Year-End Curriculum and Teaching Survey

Teacher and Grade Level	Indicator/Item	Mean Scores: Individual vs. All PAIR Teachers (higher mean = higher incidence)	Significance P = <.05
1.20 (gr. 4)	#15 – Teacher reported explaining, presenting, and getting feedback on PAIR units from other teachers at her school more often than her peers in grades 4,5,6.	6.0000 3.2857	.069 (approaching significance)
1.10 (gr.4)	# 11 – Teacher reported she develops inquiry questions for the curriculum more often than her peers in grades 4,5,6.	6.0000 3.6667	.056 (approaching significance)
	# 16 – Teacher reported explaining, discussing, and getting feedback on PAIR units from teachers who are not from her school more often than her peers in grades 4,5,6.	5.0000 3.3214	.012
	# 36 – Teacher reported using movement/dance in language arts activities in her classroom more often than her peers in grades 4,5,6.	6.0000 2.8214	.030
	#37 – Teacher reported using movement/dance to understand cultures and identity in her classroom more often than her peers in grades 4,5,6.	6.0000 2.9286	.056 (approaching significance)

(continued on next page)

7.10 (gr. 4)	# 35 – Teacher reported using movement/dance to understand math concepts/math problems in her classroom more often than her peers in grades 4,5,6.	5.0000 2.8571	.060 (approaching significance)
5.20 (gr. 4)	# 40 – Teacher reported less participation as a student when artist was teaching than her peers in grades 4,5,6.	1.0000 4.5714	.016
11.10 (gr. 4)	# 3 – Teacher reported less modeling discussion and feedback with artist for student than her peers in grades 4,5,6. #10 – Teacher reported less brainstorming and contributing her ideas to innovative arts integration approaches than her peers in grades 4,5,6.	3.0000 5.0357 2.0000 4.4286	.039 .057 (approaching significance)
7.30 (gr. 5)	# 32 – Teacher reported using music to understand math concepts or problems more often than her peers in grades 4,5,6. #35 – Teacher reported using movement/dance to understand math concepts or problems more often than her peers in grades 4,5,6.. #37 – Teacher reported using movement/dance to understand different cultures and identity in her classroom more often than her peers in grades 4,5,6.	5.0000 2.4286 5.0000 2.5357 6.0000 2.9286	.030 .060 (approaching significance) .056 (approaching significance)

Fourth grade teachers repeatedly demonstrated more ownership and independence with PAIR work as they described their involvement in the project than 5th and 6th grade teachers did in the qualitative data as well (See Table Thirty-Three).

Table Thirty-Three: The Fourth Grade Effect: Designers Have Ownership

I'm getting a lot of, like, input from them. I'm getting a lot of ideas. Without (the teaching artist) here, I'm able to go ahead, like I'm able to see what I can be able to plan for next year, or without them in here. So it's really like they give me fully great ideas and all this. With (teaching artist), I believe I still need him in here with the music. But with art, I'm able to...the only thing is, like, the materials I need, but then besides that, I'm able to do this on my own. (1.20).

For PAIR, there was little collaboration in development for me this year because ideas were developed previously. It seems that when it is a repeated program, collaboration means discussing goals, not setting goals (1.40, 5th grade).

...how much it has changed from one year to two years, especially with me being in three years. Just the comfort level and knowing what to do, and I was able to help the kids even better when the artist was outside the room, because then I would reinforce the concepts and have them write for them (7.10).

The second year I was able to implement it in some of the stories that we read, and bringing out the motivations of the characters, and how the characters changed throughout the story. And this year the students used the techniques that (the teaching artist), our playwright, taught them, to develop their own stories, and on their final pieces, the Young Authors piece that we write every year. This year it seemed to go a lot easier. We didn't collaborate as much as we have in the past, or emailed or spoke as frequently as we did in the past, and I think that comes from the fact that we've been together for three years. We worked on the same things for three years (11.10).

* * *

4) Teachers who have experience with teaching artists and arts partnerships are more likely to be independent consumers of projects like PAIR, needing less support while consistently reporting satisfaction with the collaborations, even when they are less than perfect. In the PAIR project, this might be known as **"The Healy Effect**", given that the data show Healy School teachers demonstrating significantly higher levels of satisfaction with the partnership, the professional development, and the outcomes for their students. Healy is a long-time CAPE partner school and has as a school had intensive arts integration professional development over many years.

One teacher at Healy came to the project late due to a maternity leave (1.60) and one teacher left the school. The remaining 4 teachers in the PAIR project, 1.20 (4th grade), 1.10 (4th grade), 1.40 (5th grade), 1.50 (5th grade) attended professional development sessions at a higher rate than most of their peers as follows:

1.20 – attended 14 PD sessions across 3 years
1.10 – attended 11 sessions across 3 years
1.40 – attended 7 sessions, including all 4 sessions in year 3
1.50 - attended 7 PD sessions,
1.60 - (new to project in year three) – attended 3 of the 4 PD sessions

Just these attendance figures indicate that Healy as a school community appeared to participate more fully in the project by attending Professional Development sessions offered by CAPE. Healy is also a school in which the project did disseminate evenly from 4th to 5th to 6th grade, perhaps due to the fact that at least 1 sixth grade teacher was involved, if on the periphery, from the first year.

The attendance data triangulates with the Year-End Curriculum and Teaching Survey in which there was a significant difference between Healy and the paired control group school, Holden, with respect to the *impact of professional development sessions on the arts integrated activities teachers reported using in their classrooms (#17)* with respect to one matched pair of schools. Healy teachers reported a significantly greater impact of professional development on their practice than did the control group school, and was the only one of the PAIR schools to do so.

Also, as noted in the Goal 1B section of this report (addressing the development of partnerships with teaching artists and external resources), there were significant differences between Healy and the control group school, Holden, with more occurrences reported in the treatment school teachers in the frequency of *teacher brainstorming to contribute to teacher/artist planning*, the frequency of *coming to consensus during planning meetings between artists and teachers*, and the frequency of *modeling discussion and feedback in front of students in the classroom.*

Regarding Goal 2 A (in which teachers were asked to report on their use of inquiry questions and the degree to which they reflected on their teaching), there were no significant differences in the treatment and control group teachers as groups with respect to *reporting their success at predicting the quality of student work over time (#20),* except in the case of the matched pair of Healy and Holden, in which *Healy teachers did report increasing success at predicting quality as the project continued.*

There was also a significant difference between Healy and Holden with respect to the degree to which students documenting and assessing their own learning, with Healy teachers reporting their *students documenting and assessing* more often than the control group teachers did (See Goal 2C).

In all, Healy showed positive significant differences in 8 indicators of implementation based on the Year-End Curriculum and Teaching Survey, more than any other PAIR school in comparison with their paired control group school (item # 1,2,3,17, 20, 25, 30, 31).

One possible explanation for this engagement on the part of one school that the research seems to suggest is that Healy teachers knew what to expect and knew how to engage with a community partner focused on arts integration with non-arts content in ways that the other five participating treatment schools did not. Even if all of the specific teachers in PAIR had never partnered with a teaching artist before, the culture of the school is such that these partnerships are welcomed, familiar and expected. The 2 fourth grade teachers, 1.20 and 1.10, had worked with teaching artists many times before, thus setting a norm for practice more easily perhaps than other 4th grade teachers in the other 5 schools.

It is not clear whether this familiarity with the process is actually a positive element in how much teachers were able to learn from the experience, nor is it clear whether this element contributes to students' learning to a greater degree than in other schools.

* * *

Considerations for Making Connections: Teacher Impact Effects and Student Learning Outcomes

While it is not possible to have a singular ranking of teachers with respect to their learning during PAIR or the impact of PAIR on them and their classrooms, it is possible to identify the teachers who demonstrate these 4 "effects" more clearly than their peers (See Table Thirty-Four).

Table Thirty-Four: The Four Teacher Impact Effects in PAIR Project

Content Expertise Effect	Documenting to Learn Effect	4th Grade Effect	"Healy" Effect
5.10	1.50	1.10	1.10
7.30	5.20	1.20	1.20
11.10	7.10	3.20	1.40
11.20	9.50	5.10	1.50
11.30	11.20	5.20	1.60
	11.40	7.10	
		11.10	
		11.20	

Teachers who appear in more than one of these 'effect columns are: 1.10, 1.20, 1.50, 5.10, 5.20, 7.10, 11.10, 11.20.

It may be worth investigating whether these 8 teachers' engagement in these particular ways and under these parameters affects the achievement of their students. Note that two of these effects are by virtue of grade level (4th grade) and school affiliation (Healy) and two of these effects are due to what teachers themselves have offered and reported about what they did in the project and what their perceptions of learning are (Content Expertise and Documenting to Learn).

The method of investigating individual teachers and the impact that PAIR has had on them and their practice has compelling implications for how to study the relationship between teacher learning and student achievement. It is important to offer the necessary caveats to such investigation regarding impact on teachers. Do teachers respond to such partnership initiatives differently as correlated with their years of teaching experience? There are some teachers who have already demonstrated high quality teaching; in what ways are their experiences with such partnership programs qualitatively different from less proficient teachers? Is it possible that arts partnership programs need to make accommodations for different profiles of teachers, providing a menu of professional development options that more clearly accommodate teachers' needs, interests and capacities? Is it also possible that such choice may in fact encourage continuous improvement even among the high quality teachers who are willing to take new risks to engage their students in different ways?

An investigation of the teachers demonstrating the four 'effects' noted above juxtaposed with the achievement of their students is worth exploring. Although this project design included tracking High, Average and Low Achieving students (HAL) through their participation in PAIR in grades 4,5,6, it is also worth considering the achievement of students who rotated through teachers' classrooms involved in the program more than one year. In other words, it may not be how many years the students were in the project, but rather how many years their teachers were PAIR teachers that contributes to student achievement. Anyone who has ever taught a class can attest to the value of additive expertise as one becomes more skilled and more comfortable with teaching a unit, concept or topic. It is possible that, when one or more of these project "effects" are in evidence in the teacher indicators, then their students will benefit, whether they have participated in PAIR for 1, 2, or all 3 years.

Conclusions

A. Hall and Hord's 'levels of use' (2006) is a valuable framework to adopt in considering arts integration partnerships. Different teachers embrace innovations to varying degrees, based on a variety of variables, including their years in the profession, experience with external partnerships and exposure/access to the arts in their own lives and in their schools. Examining the levels of use of teachers engaged in new ventures over time and developing more fine-tuned measures to observe teachers' movement from Nonuse to Use as Renewal and assist them to move more effectively toward that goal is a promising approach to Professional Development.

Table Fifteen (repeated from above): Level of Use of An Innovation (Hall and Hord, 2006) (*How do we know when an innovation has taken hold?)*

Users • Renewal • Integration • Refinement • Routine • Mechanical Use
Nonusers • Preparation • Orientation • Nonuse

B. The PAIR project, by Year Three, was equipped to examine qualitative and quantitative measures in order to ascertain which teachers fully developed the project in their classrooms and which engaged in the project less completely**. The four "effects" posited in this report suggest certain elements that assist teachers in implementing an innovation**:

1. *The degree to which the teacher possesses pedagogical content knowledge (PCK) of the non-arts content field addressed by the innovation*. If teachers have limited PCK, then that must also be addressed in order to have full benefit from the innovation;
2. *The degree to which the teacher learns how to document the project,* with particular attention to evidence of student learning through student products;
3. *The degree to which the teacher engages in the project over the long term* and has ownership of choices with respect to goals, experiences, and assessments;

4. *The degree to which the teacher works in a school culture that embraces external resources* and partnerships and rewards teachers for participation in them.

We often ignore the effects of professional development on individual teachers. The results from PAIR suggest the error in doing so.

C. *The PAIR design depended on teachers' ability to assess the non-arts content learning that was targeted by the arts integration projects*. Professional development specifically in meaningful classroom or grade level based assessment, would have not only served to enhance the project, but also would have been useful to teachers as they observed their students' integration of various art forms in their learning. District partnerships might focus specifically on support for assessment during an innovation provided by external partnerships in order to fully appreciate the value added of programs such as PAIR.

D. It seems clear that the design of PAIR, in which a grade was added each year, was not as effective for teachers as it may have been for the cohorts of PAIR students. The 4th grade teachers, invested in the project for all three years, were the most affected by the PAIR project. Collaborative curriculum development that is replicable would be best served by engaging multiple grades at the same time, as teachers learn to work with arts integration as *vertical as well as horizontal learning* approaches for students.

E. Ball and Cohen's reference to a "pedagogy of professional development" (1999) took on a new meaning in the PAIR project. Teachers began to reconceive of where they would learn and by what means they would learn about the benefits of arts integration in their classrooms. *Professional development, for PAIR teachers, came in the form of co-teaching with artists, observing artists working with their students, and revising the targeted curriculum to accommodate different kinds of learning that students were demonstrating*. "Workshops as professional development" are limited, but in-class co-teaching has great potential to influence practice, as this project demonstrated.

F. Teachers change their practice only when they see that change makes a difference in what their students learn and how their students are engaged. *The evidence in PAIR shows clearly that the teachers saw their students learn the non-arts content differently*. They reported that the students were able to elaborate, *give me more of what I was looking for*, as one teacher phrased

it, and expand their ideas more fully due to the tools that the teaching artists provided in the classroom.

G. Dissemination of changed practices and the impact of innovations such as this arts integration project require considerable planning and support. Schools are not designed to enable teachers to share what they know; the culture of schools are often not conducive to peer acceptance of such dissemination; teachers who hear about and view documentation from a project like PAIR are usually not compelled to do anything with it. What teachers *do* respond to is evidence of students' learning through students' products, narrated by those with strong pedagogical content knowledge from within the professional community. *Dissemination then must focus on actual products – and must be interpreted so that non-project teachers can see for themselves the value of taking the risk to engage in innovation outside of their expertise.* Dissemination did not happen as often as thoroughly in PAIR as it could. But the Magnet Cluster Schools have established a network that is conducive to such practices and could build upon the BCCLA and now the PAIR project among others to send the message that dissemination is crucial and is a necessary form of professional development for all teachers.

H. The portfolio conferences, with three students and their teacher, are a promising initiative for research and for professional development purposes. While the "three-way conference" involving a parent, student, and teacher, is not new to educators, *the concept of having a teacher watch as her/his students discuss their learning with another adult is an interesting method of assessment* that could be used by teachers and researchers in various ways.

I. The practice of collecting student work and then attaching Labels to those samples that document the teaching approach that resulted in that work is a difficult concept for teachers and takes some time to embed in their routines. The principle, however, is soundly research-based. The notion that we should be able to discuss what the teacher (or teaching artist) *did* with students that enabled students to produce a given work sample, and then learning from that insight to improve teaching the next time is crucial to continuous improvement. The categories considered as effective teaching practices within an integrative curriculum (See Table Twelve excerpted below) are representative of approaches that affect student achievement. Teachers not only reported enacting these practices more often as the project continued; they also documented those practices more often as well. The labels in the PAIR project began to be

most useful as documentation of teaching in the third year of the project. CAPE has seen the labels work to varying degrees in different projects, but it is clear that *the longer teachers have to experiment with the labels and perhaps have more ownership in what prompts should be on the labels, the more we are truly able to see into the classroom as a learning laboratory and understand what is behind the student work on the table.*

Table Twelve (Excerpted): Effective Teaching Practices Documented on PAIR Student Work Labels

Evidence of PAIR Teacher Impact based on Student Work Labels	Documented assessment	Documented students' co-creation of curriculum	Documented students teaching other students	Documented students teaching the teacher	Students participated in documentation and/or assessment	Students wrote inquiry questions	Students used new arts vocabulary

J. There was a relationship in the PAIR project between professional development attendance and other variables regarding impact of the program. Part of the rationale for that relationship, however, was in the design of the program, in which the 4th grade teachers began attending in Year One and many continued throughout the three years.

K. There were clear relationships between the PAIR professional development sessions, the documentation of the project, and the expectations for meaningful arts integration within the schools. For teachers new to professional development, the CAPE projector directors have acknowledged the need early in the project for more intentional professional development on arts integration, using the Design Seminar model.

L. Several of the PAIR professional development sessions across the four years engaged teachers and artists in conversation about preliminary research findings. This practice resulted in teachers' increased interest in results and in participating in the research with the research team. This is a common methodology for CAPE projects and continues to enhance the quality of the projects as participants see that the team is doing research *with,* not research *on* students and teachers.

M. A meta-analysis by Robinson and Timperley (2007) examined seventeen studies of professional development initiatives that have made a demonstrable impact on the students of

the teachers involved to determine what kinds of leadership practices were involved in each initiative. The analysis revealed five leadership dimensions that were critical in fostering teacher and student learning: *providing educational direction*; *ensuring strategic alignment*; *creating a community that learns how to improve student success*; *engaging in constructive problem talk*; and *selecting and developing smart tools*. The PAIR project addressed all five of these dimensions through the professional development sessions and the project support team from CAPE and Chicago Public Schools. Further strategic alignment is possible across grades 4,5,6 with future initiatives if the design works toward vertical alignment – an important outcome of this project that the district can build upon. The "smart tools" are those embedded in all high quality arts integration initiatives; some were also specifically targeted as strategies in the Effective Teaching practices. Gathering, improving, and disseminating the PAIR "smart tools" (portfolio conference process, professional development assessment measuring levels of use of material addressed, student work labels to connect product with pedagogy) would be beneficial.

In the current educational climate, much has been made of the necessity of linking teacher evaluation with student achievement as measured by standardized test scores. The recent federal competitive initiative from the Department of Education termed "Race to the Top" is encouraging the developing of "data-driven systems" in states that are competing for federal funds. New York, which has been awarded nearly $700 million, adopted a new teacher evaluation system that takes student test performance into account (*New York Times Editorial*, August 29, 2010). The challenge is of course how to determine the variables that contribute to student achievement and juxtapose those variables to the characteristics of teachers.

In a 2007 Data Quality Campaign document from the National Center for Educational Accountability, there were "10 essential elements of state longitudinal data systems" that were proposed (Bergner, Stein and Armstrong, 2007, p. 2) (See Table Thirty-Five).

Table Thirty-Five: Data Quality Campaign: Advances in State Student Data Systems

10 Essential Elements of State Longitudinal Data Systems
1. A unique statewide student identifier that connects student data across key databases across years.
2. Student-level enrollment, demographic and program participation information.
3. The ability to match individual student's test records from year to year to measure academic growth.
4. Information on untested students and the reasons they were not tested.
5. A teacher identifier system with the ability to match teachers to students.
6. Student-level transcript information, including information on courses completed and grades earned.
7. Student-level college readiness test scores.
8. Student-level graduation and drop-out data.
9. The ability to match student records between the p-12 and post-secondary systems.
10. A state audit system assessing data quality, validity and reliability.

The only mention of teachers is in element #5, noting the need for a "teacher identifier system" with the ability to match teachers to students. The proposed data points track students at all levels and through different dimensions. The formula, however, misses essential elements regarding **what teachers are actually doing in the classroom** and **whether there are specific practices that teachers across grade levels are doing to contribute to the achievement information** called for in this Campaign and rewarded by the Race to the Top program.

Although the Data Quality Campaign's system suggest the need to "identify which teacher preparation programs produce the teachers whose students have the most academic growth" (p.2), there is no mention in the proposed schema to identify the practices, strategies and approaches to learning that could be reinforced, replicated and restored in schools and classrooms where they are not as visible.

The PAIR project has had some success in helping teachers identify the effective teaching practices that emerge in arts integrated contexts and utilize them in their teaching. We have proposed a tentative and very preliminary "teacher identifier system" based on the goals of the PAIR project in general. The assumption is that involvement in PAIR has served as professional development for teachers who will use what they have learned to further contribute to their students' understanding in non-arts content areas to varying degrees, based on a number of factors that contribute to teacher quality as PAIR teachers. The research indicates that it is possible to identify teachers who have learned from the innovation and increased their level of use of the practices inherent in integrative learning and teaching. Future research is needed on how best to support and effectively provide professional development for teachers at different levels and with different needs from external partnerships and innovations and then track teachers over multiple years to determine how, when, and how often effective practices are used. Only then can we truly link student achievement to teacher evaluation.

References

Ball, D. L. and Cohen, D.K. (1999). "Developing practice, developing practitioners: Toward a practice-based theory of professional development," in *Teaching asThe Learning Profession: Handbook of Policy and Practice,* ed. L. Darling-Hammond and G. Sikes. San Francisco: Jossey-Bass, 3-32.

Bergner, T., Steiny, J. and Armstrong, J. (2007). "Benefits of and lessons learned from linking teacher and student data." *Data Quality Campaign/National Center for Educational Accountability*. Bill and Melinda Gates Foundation. www.DataQualityCampaign.org

Burnaford, G. (2009). A study of professional development for arts teachers: Building curriculum, community, and leadership in elementary schools. *Journal for Learning through the Arts*, 5(1), Retrieved from: http://escholarship.org/uc/item/18h4q9fg

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.

Darling-Hammond, L. (1999). *Teacher quality and student achievement:A review of state policy evidence*. Center for the Study of Teaching and Policy Seattle, WA: University of Washington.

Elmore, R. F. (2002). *Bridging the gap between standards and achievement: The imperative for professional development in education.* Washington, DC: Albert Shanker Institute.

Fickel, L.H. (2002). "Quality professional development: Suggestions about process and content." *The Educational Forum* 67, 47-54.

Franke, M.L., Carter, T.P., Levi, L. and Fennema, E. (2001). "Capturing teachers' generative change: A follow-up study of professional development in mathematics." *American Educational Research Journal* 38(3), 653-689.

Frechtling, J. & Killeen, D.J. (2003). Balancing summative and formative evaluation: New partnerships, new evaluation models. In Annenberg Institute for School Reform, Research perspective on school reform: Lessons from the Annenberg Challenge. Providence, RI: Brown University, 69-80.

Gordon, S. P. (2004). Professional development for school improvement: Empowering learning communities. Boston: Pearson Education.

Grossman, P., Schoenfeld, A., with Lee, C. (2005). "Teaching subject matter." In Darling-Hammond, L. and Bransford, J. Eds. Preparing teachers for a changing world: What teachers should learn and be able to do. San Francisco: Jossey- Bass, 201-231.

Guskey. T.R. (1985). "The effects of staff development on teachers' perceptions about effective teaching." *Journal of Educational Research* 78(6), 378-381.

Hall, G. E., & Hord, S. M. (2006). Implementing change: Patterns, principles and potholes. Boston: Pearson.

Ingvarson, L., Meiers, M., Beavis, A. (2005). "Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes and efficacy." *Education Policy Analysis Archives 13(10,* 1-28.

Jencks, C. (1972). *A reassessment of the effect of family and schooling in America.*New York: Basic Books, Inc.

Joyce, B. and Showers, B. (2002). *Student achievement through staff development*. (3rd ed.) Alexandria, VA: Association for Supervision and Curriculum Development.

King, K. P. and Lawler, P. A. (2003). New perspectives on designing and implementing professional development of teachers of adults. San Francisco: Jossey-Bass.

Lieberman, J.M. and Wilkens, E.A. (2006). "The professional development pathways model: From policy to practice." Kappa Delta Pi Record 42(3) 124-128.

New York Times Editorial. (2010). Continue the Race. Sunday, August 29. http://www.nytimes.com/2010/08/29/opinion/29sun2.html Downloaded 8.29.10.

Newmann, F.M., King, M.B. and Youngs, P. (2000). "Professional development that addresses school capacity: Lessons from urban elementary schools." *American Journal of Education* 108(4), 259-299.

Phillips, J. (2003). Powerful learning: Creating learning communities in urban school reform. *Journal of Curriculum and Supervision,* 18(3), 240-258.

Robinson, V. and Timperley, H. (2007). "The leadership of the improvement of teaching and learning: Lessons from initiatives with positive outcomes for students." *Australian Journal of Education* 51(3), 247-262.

Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville: University of Tennessee Value- Added Research and Assessment Center.

Shulman, L. (1987). "Knowledge and teaching: Foundations of the new reform." *Harvard Educational Review* 57(1), 1 – 22.

Vescio, V., Ross, D. & Adams, A. (2008). "A review of the research on the impact of professional learning communities on teaching practice and student learning." *Teaching and Teacher Education 24,* 80-91.

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). "Teacher and classroom context effects on student achievement: Implications for teacher evaluation." *Journal of Personnel Evaluation in Education*, 57-67.

PARTNERSHIPS IN ARTS INTEGRATION RESEARCH (PAIR) FINAL COMPREHENSIVE REPORT

Part 2: Impact of PAIR on Student Academic Performance

A Comparative Analysis of Control-Treatment School Student Standardized Academic Test Learning Outcomes by All School Types and Student Academic Classifications (Years 2007-2010)

June 27, 2011
(Revised April 2012; Update December 14, 2012)

Analyzed and Written By
Dr. Lawrence Scripp, Principal Investigator
With contributions from
Fred Sienkiewicz, Research Associate
Research Center, Center for Music and Arts in Education, Inc.
New York 10023
Professor, Graduate and Undergraduate Theoretical Studies
Chair, Music-in-Education, New England Conservatory
larry.scripp@necmusic.edu
617 585 1364

With Special Thanks for the contributions from Laura Tan Paradis, CAPE Research Associate and to the research staff at the Center for Music-in-Education: Randy Wong, Justin Stanley, Helen Liu Wong, Kristen Dirmeier Sienkiewicz.

The PAIR project was funded through a grant to Chicago Public Schools from the US Department of Education's AEMDD Program. Chicago Arts Partnerships in Education (CAPE) contracted with Chicago Public Schools for this project. CAPE then subcontracted with Dr. Larry Scripp, Center for Music in Education, Inc., to serve as a co-principal investigator for the project in collaboration with Dr. Gail Burnaford, Principal Investigator, Florida Atlantic University.

Table of Contents

PAIR Final Report Part 2 (A-G): Impact of PAIR on Student Academic Performance

Partnerships in Arts Integration Research Project Final Report Part 2: A Comparative Analysis of Control-Treatment School Student Standardized Academic Test Learning Outcomes by All School Types and Student Academic Classifications (2007-2010) 1

PAIR Final Report Part 2: Table of Contents 2

PAIR Final Report Part 2: List of Figures and Tables 3

PART 2. Student Academic Performance Impact:

Analysis of Control-Treatment School Student Standardized Academic Test Learning Outcomes by All School Types and Student Academic Classifications (2007-2010) 5

Introduction 5

2A: Review of the Background, Purpose, and Scope of the PAIR Project 5

Research Design Structural Elements 7

PAIR Student Learning and Teacher Professional Development Data Collection Sequence 8

PAIR Teacher Professional Development Data Collection 10

CAPE Teacher Artist Role in PAIR 11

Analytic Framework and Processes 12

General Presentation of Results 13

2B: External Student Learning Assessments: Control-Treatment (C-T) Analysis of Illinois Standards Academic Achievement Test (ISAT) Results 14

2C: Control-Treatment School Comparisons of PAIR Grade Level ISAT Mean Scores 16

2D: Control-Treatment Focus School (Arts vs. Academic) Comparisons of ISAT Mean Scores 19

2E: Control-Treatment PAIR School Achievement Gap Comparisons of ISAT Mean Scores According to Pre-Designated High, Average, and Low (HAL) Student Achievement Cohorts, Grades 3-6 21

2F: Control-Treatment PAIR Focus School (Arts vs. Academic) Comparisons of HAL ISAT Mean Scores and Achievement Gap Profiles, Grades 3-6 28

2G: Control-Treatment PAIR MAGNET CLUSTER SCHOOL (ELA, Math, World Language) Comparisons of HAL ISAT Mean Scores and Achievement Gap Profiles, Grades 3-6 33

PAIR Report PART 2(A-G): List of Tables and Figures

2A Table 1:	PAIR School Profiles	7
2A Table 2:	PAIR Student Learning & Teacher Professional Development Data Collection Sequence	9
2B Figure 1:	Control-Treatment Comparisons of Percent of PAIR Grade 6 Initial Cohort Students who Meet or Exceed Chicago Public Schools ISAT Benchmarks in One (Reading or Math), Both (Reading and Math) or None (Neither Read or Math)	14
2B Figure 2:	Control-Treatment ISAT Grade Level Benchmark Comparisons for Reading and Math in PAIR Follow-Up Longitudinal Cohorts	15
2B Figure 3:	Control-Treatment ISAT 2010 Grade Level Benchmark Comparisons for Reading and Math in Initial (Grade 6) and Follow-up (Grades 4-5) PAIR Longitudinal Cohorts	15
2C Figure 1:	Control-Treatment PAIR Schools ISAT Grade Level Comparisons of ISAT Combined (Averaged Reading and Math) Mean Scores, PAIR Initial and Follow-up Longitudinal Cohorts	17
2C Figure 2:	Control-Treatment Focus School (Arts vs. Academic) Comparisons of ISAT Grade Level Reading and Math Mean Scores, PAIR Initial and Follow-up Longitudinal Cohorts	18
2D Figure 1:	Control-Treatment PAIR Focus School (Arts vs. Academic) Comparisons of ISAT Mean Scores, PAIR Initial Longitudinal Cohorts, Grades 4-6	20
2E Figure 1:	Control-Treatment PAIR Schools Achievement Gap Bar Graph Display of HAL ISAT Combined, Reading and Math Mean Scores, Grades 3-6 (Initial Longitudinal Cohort)	23
2E Figure 2:	Control-Treatment PAIR Schools Achievement Gap Box Plot Analysis of HAL ISAT Combined Academic Scores, Grades 3-6	24
2E Figure 3:	Control-Treatment PAIR Schools Achievement Gap Box Plot Analysis of HAL ISAT Reading Scores, Grades 3-6	25
2E Figure 4:	Control-Treatment PAIR Schools Achievement Gap Box Plot Analysis of HAL ISAT Math Scores, Grades 3-6	26
2F Figure 1:	PAIR Focus School Control-Treatment Bar Graph Display of ISAT Combined, Reading, and Math Scores by HAL Cohorts, Grades 3-6	29
2F Table 1:	Rubric for Categorizing Statistically Significant Levels of Achievement Gap Separation of HAL Student Cohorts	30
2F Table 2:	Levels of Statistically Significant Achievement Gap Levels in HAL ISAT Combined Test Scores by Control-Treatment Focus School Designation, Grades 3-6	31

2F Table 3:	Levels of Statistically Significant Achievement Gap Levels in HAL ISAT Reading Test Scores by Control-Treatment Focus School Designation, Grades 3-6	31
2F Table 4:	Levels of Statistically Significant Achievement Gap Levels in HAL ISAT Math Test Scores by Control-Treatment Focus School Designation, Grades 3-6	32
2G Figure 1:	C-T PAIR Cluster School Achievement Gap Bar Graph Display of ISAT Combined, Reading and Math Scores by HAL Cohorts, Grades 4-6	34
2G Table 1:	Control-Treatment PAIR Cluster School Levels of Statistically Significant Achievement Gap Differentiation in HAL ISAT Combined Test Scores, Grades 3-6	35
2G Table 2:	Control-Treatment PAIR Cluster School Levels of Statistically Significant Achievement Gap Differentiation in HAL ISAT Reading Test Scores, Grades 3-6	35
2G Table 3:	Levels of Achievement Gap Differentiation in HAL ISAT Math Test Scores by Control-Treatment: School Cluster Designation, Grades 3-6	36

(continued on next page)

PAIR Final Report PART 2: Student Academic Performance Impact:

The Analysis of Control-Treatment School Student Standardized Academic Test Learning Outcomes by All School Types and Student Academic Classifications (2007-2010)

Introduction

This report is the second part of a three-part comprehensive report filed by both Dr. Burnaford and Dr. Scripp, as Co-Principal Investigators of the PAIR project.

The first report, written by Dr. Gail Burnaford [2010], focused primarily on three years of collecting evidence of progress meeting PAIR teacher professional development goals, the evolution of teacher professional development outcomes in comparison with control group teachers, and speculation on the general impact of high quality PAIR teacher practices on student learning.

The second and third parts of this report, written by Dr. Lawrence Scripp and his research team from the Center for Music-in-Education and CAPE (2011-2012], focus on the impact of PAIR on student learning. This second report (2011) reports primarily on differences among control-treatment statistical comparisons of PAIR student academic test outcomes. The following third report (2012) features an extensive review of qualitative and quantitative aspects all PAIR student arts integration outcome data. In addition, it provides a comprehensive analysis of the possible statistical links between seven teacher professional development factors and four student learning outcomes.

2A. Review of the Background, Purpose, and Scope of the PAIR Project

In this paper we report on a research project in arts integration education, conducted in the Chicago Public Schools in partnership with Chicago Arts Partnerships in Education (CAPE), a research-based organization focused on optimizing the impact of artists and arts learning in schools for the benefit of whole-school improvement in arts learning, teacher professional development, and school culture.

The PAIR project follows a previous AEMDD *Developing Early Literacies Through the Arts* (DELTA) grant that resulted in developing models of teaching artist residencies focused on fundamental concepts and skills in the musical, visual, movement, and dramatic arts as conduits for reinforcing and improving early literacy goals for the Chicago Public Schools (CPS) [Scripp and

Sienkiewicz 2007; CAPE 2008][1]. As reported in Dr. Burnaford's report, PAIR builds also upon the outcomes of a previous three-year Department of Education Professional Development Grant project titled Building Curriculum, Community, and Leadership through the Arts (BCCLA), in which 59 Fine and Performing Arts Magnet Cluster School arts specialists experienced professional development that supported the development of arts integrated curriculum, leadership, and community.

Building on DELTA professional development, curricular and assessment practices, the specific aim of the Partnerships in Arts Integration Research was to develop models of arts integration for upper elementary schools based on the intersections between arts and non-arts content learning that not only supports the development of arts learning skills and concepts, but also influences academic content learning in students. These models were developed over a three-year period and assessed through the documentation and reflection on the learning experiences of the participating grade level teachers in collaboration with CAPE artists (as reported extensively in Dr. Gail Burnaford's PAIR report), and by developing and employing multiple methods of documenting and assessing student learning (as reported in both PAIR final reports).

The assumption of the PAIR research is that arts integration differs significantly from both conventional arts and academic instruction in school settings, primarily because of its focus on a) the development and modeling collaboration between the CAPE teaching artist and the grade level teachers on the contribution of arts learning processes to academic learning in the classroom, b) the primacy of designing curriculum units based on shared arts and academic content knowledge and skills demonstrated by both the collaborating arts and grade level teachers, and c) the commitment of the classroom teachers to engage in the process of ongoing documentation, reflection, and evaluation of research-based practices as a fundamental strategy for establishing innovative practices necessary to distinguish PAIR schools from the comparison schools.

This view of arts integration also suggests arts learning should be more than just a subject to be taught by specialists alone or taught to children only for its own sake, but as a resource and approach to learning across the curriculum, for the benefit of both teachers and students. In the PAIR project the teacher professional development aspects of the PAIR project were achieved as necessary condition for "arts plus arts integration" teaching and learning that affects the whole school culture. Hence arts

[1] See capeweb.org for DELTA report.

integration is not conceived only as a strategy for teaching certain facts or concepts in the academics, but as a deep resource for melding powerful learning processes and understanding of content that is shared across arts and academic disciplines and discovered by students, grade level teachers and teaching artists alike.

Research Design Structural Elements

The research design of this project specified that six elementary schools in Chicago – two Mathematics and Science, two World Languages (foreign language) and two Literature and Writing (English language arts) Magnet Cluster Schools – were to be selected as PAIR treatment schools. A matched set of randomly selected Magnet Cluster Schools served as control schools in this study. Of these six CPS Magnet Cluster schools, three were highly arts and arts integrated learning focused and three were primarily academic (see 2A Table 1 below).

2A Table 1: PAIR School Profiles

PAIR Schools	Matched Control (Comparison) Magnet Cluster Schools	Academic/ Arts Primary Learning Approach FOCUS	CPS Magnet CLUSTER School Designation	Additional Resource Faculty Specialists	CAPE PAIR Teaching Artist 10 week Residence/ Curriculum units
Lee	CPS Control School a	Academic	ELA (Writing)	Writing, visual art	Music, drama
Eberhart	CPS Control School b	Arts	ELA	Music, visual art, arts integration	Music, drama
Thorp	CPS Control School c	Academic	MATH	Math, music, visual art	Visual art, dance
Swift	CPS Control School d	Arts	MATH	Music, visual art, arts integration	Visual art, dance
Ward	CPS Control School e	Academic	WL (World Language)	Social & cultural studies, Spanish, visual art	Music, visual art
Healy	CPS Control School f	Arts	WL (World Language)	Social & cultural studies, music, Spanish, visual arts, arts integration	Music, visual art

The aim of the research was to document and analyze the development of a general model of arts integration professional development, curriculum and assessment practices adaptable across various types of CPS Magnet Cluster Schools and their impact on teachers and student learning in comparison with a set of control schools matched according to Magnet Cluster School designation, school performance and demographic considerations.

PAIR Student Learning and Teacher Professional Development Data Collection Sequence

Student Learning Data Collection: Clearly the success of this project depends on results from both external and internal student learning assessments. Without positive indications of student learning gains compared to district averages and control school comparisons, the PAIR project would not represent a responsible or viable alternative to conventional arts and academic methods of teaching and assessing learning.

Indications of high quality teacher professional development outcomes in comparison with control group teachers described in part 1 of this report make it possible to assume that if any positive student learning outcomes do obtain, it can be reasonably be assumed that the PAIR program may be a critical link to this learning success.

Thus, in order to test the hypothesis that CAPE's arts integration methods brought to the implemented in the PAIR project we will need to see a rigorous evaluation of student academic performance data. Thus the external standard academic test data from the grade level Illinois Standards Academic Tests (ISAT) were collected each year of the project to determine to how students in the PAIR longitudinal cohorts in treatment schools compare with control group cohorts.

As indicated in the chart below the data collection was organized into three cohorts:

A) The Initial PAIR Longitudinal Cohort randomly selected from pools of pre-designated High (H), Average (A), and Low (L) academic performance rated students in both treatment and control CPS Magnet Cluster schools. These 'HAL' students were also assessed for arts learning through the internally developed Snapshots of Arts Integrated Learning (SAIL) interview protocols, PAIR portfolio conference performance assessment procedures (administered only in treatment schools) and PAIL student work samples (in treatment schools). In addition PAIR student self-report surveys inform us as to classroom culture differences between PAIR and control school classrooms.

B) The PAIR Follow-up Longitudinal Cohort 1 where data collection is limited to student ISAT test scores from entire treatment school classrooms in comparison with all other non-PAIR classrooms. This follow-up cohort will provide a window into academic performance

differences between PAIR classrooms with teachers who have 2 years of experience with the project.

C) The PAIR Follow-up Longitudinal Cohort 2 where data collection is limited to student ISAT scores from entire treatment school classrooms in comparison with all other non-PAIR classrooms. This follow-up cohort will provide a window into academic performance differences between PAIR classrooms with teachers who have 3 years of experience with the project.

The table [2A Table 2] below summarizes the data collection design elements from the viewpoint of student learning related assessments

2A Table 2: PAIR Student Learning & Teacher Professional Development Data Collection Sequence

Longitudinal Cohorts	PAIR Planning Year	PAIR Year 1	PAIR Year 2	PAIR Year 3
A. INITIAL HAL Longitudinal Student Cohort (Control and Treatment HAL samples)	Grade 3 ISAT data collected from HAL student sample	Grade 4 ISAT data from year 1 HAL student C-T samples Grade 4 Teacher PD Year 1 PD session surveys and attendance figures beginning with Grade 4 teachers	Grade 5 ISAT data collected from Year 2 HAL student C-T samples SAIL Interview 1 Student Survey 1 Grade 5 Teacher PD Year 1 PD session surveys and attendance figures now include Grade 4 -5 teachers	Grade 6 ISAT data collected from Year 3 HAL student C-T samples SAIL Interview 2 Student Survey 2 Grades 4-6 (Student Years 1-3) PAIL Portfolio Conference 1 Grade 6 Teacher PD Year 1 Year-End Curriculum and Teacher Survey PD session surveys and attendance figures; PAIL portfolio conference transcribed comments; teacher practice labels of student work; PD session documentation panels & curriculum maps

(2A Table 2 continued on the following page)

B. FOLLOW-UP PAIR Longitudinal Student Cohort 1 (Control and Treatment Classrooms)			Grade 4 ISAT data collected from Year 1 HAL students C-T samples Grade 4 Teacher PD year 2 PD session surveys and attendance figures continuing with Grade 4 teachers	Grade 5 ISAT data collected from Year 2 HAL student C-T samples Grade 5 Teacher PD year 2 Year-End Curriculum and Teacher Survey PD session surveys and attendance figures; PAIL Portfolio conference transcribed comments; teacher practice labels of student work; PD session documentation panels & curriculum maps
C. FOLLOW-UP PAIR Longitudinal Student Cohort 2 (Control and Treatment Classrooms)				Grade 4 ISAT data collected from Year 1 HAL students C-T samples Grade 4 Teacher PD year 3 Year-End Curriculum and Teacher Survey PD session surveys and attendance figures; PAIL Portfolio conference transcribed comments; teacher practice labels of student work; PD session documentation panels & curriculum maps

PAIR Teacher Professional Development Data Collection

The success of this project depended on the quality of the multi-leveled partnership between CAPE staff and six CPS Grade 4-6 schools. The professional development, curriculum design, planning and implementation processes involved intensive collaboration between artists and teachers throughout the three years of the project and the results of this work needed to clearly differentiate the treatment from the control schools in order to establish links between program development and student learning outcome variables.

As described by Dr. Burnaford in Part 1 of this report, A Year-End Curriculum and Teaching Survey was administered to 4th, 5th, and 6th grade teachers in all twelve schools (treatment and control) during Year Three of the project. Other data were also collected from the teachers in the 6 PAIR schools, including professional development session surveys and attendance figures, portfolio conference transcribed comments, student work and teacher practice labels and documentation from work completed at professional development sessions (documentation panels and curriculum maps). These data were analyzed in Dr. Burnaford's PAIR report to establish clear criteria for high quality arts integration teaching practices through survey responses and student teacher conferences. It is highly

significant that Dr. Burnaford also was able to articulate differences between treatment and control group teacher through these same data collection instruments.

In all sections of part 3 of this report, these same teacher professional development data sources are used also to explore the relationship between teacher professional development outcomes and student learning tests and performance assessments.

CAPE Teaching Artist Role in PAIR

All CAPE teaching artists were highly qualified artists, artist-teachers, and veterans of many years of formative participants in CAPE research-based projects. PAIR teachers had access to two CAPE teaching artists each year in two art forms [see Table 1: PAIR School Profiles for distribution of art forms]. These artists participated in PAIR PD sessions, met with classroom teachers to develop PAIR units every year of the project, and collaborated often with the classroom teachers in the implementation of these units. In addition the teaching artists provided continuity and focus across grade levels in ways that proved highly significant to the research design.

Essentially, the teaching artists modeled high-quality arts integration teaching practices while helping teachers to adapt CAPE methods for arts integration into the their classroom practice. This collaborative process rigorously adhered to principles of consensus-building between the teacher and teaching artist during the unit development process thus providing a significant measure of process validity and reliability that provided confidence in the distinctions between the treatment school arts integration practices in comparison with control school classroom practices. In addition, the CAPE teaching artists and PAIR teachers, without exception, remained assiduously faithful to the primary mission and focus of the particular Magnet Cluster School throughout the course of the project that validated PAIR Cluster school comparisons.

As a result of teacher and teaching artist contributions to the student work products and documentation processes, researchers were able to see evidence in the PAIL student learning artifacts of both high quality arts learning and arts integrated learning, often in the same student work products, reflective thinking documentation, and in the SAIL interviews and Portfolio Conference Performance Assessments [see chapters on SAIL and Portfolio Conference analyses in part 3 of this report for more details).

In sum, it appears to the student evaluation team that the CAPE teaching artists provided essential quality control indicators for the project particularly as the Initial Longitudinal Cohort students completed the final year of the project. In every portfolio conference, students were completely aware of detailed experiences with their teaching artist over the three years of the project, and the teachers were amazed at the positive shifts in classroom dynamics and learning culture when teaching artists visited the schools. Maintaining the same teaching artist teams in every school in the initial cohort over three years meant that the longitudinal treatment school students received a distinct and ongoing arts integration intervention in comparison with the control school students.

Analytic Framework and Processes

In the following sections of this report, the analytic framework and processes will remain largely constant, making it relatively easy to organize multivariate analysis and compare statistical results.

The analyses will proceed sequentially from the relatively simple general comparisons to the more specific and complex, mirroring the order of presentation in PAIR School Profiles [2A-Table 1 above]. Results from control-treatment comparisons will identify factors most clearly identified with the development and distinctive quality of the PAIR program in various contexts (school profiles) and looking at various types of student samples following the data collection sequence over time [see 2A Table 2 above].

Starting with the Illinois Standards Achievement Test (ISAT) scores, we will begin by studying averaged academic scores that combine reading and math, and then separate the disciplines. We will be looking for statistically significant differences by control/treatment comparisons, by magnet school focus control-treatment comparisons (Arts vs. Academic Focus), and finally by School Cluster control-treatment comparisons: (Writing [ELA], Math, or World Languages s [WL]).

Next, we will examine intensively the *initial longitudinal sample* of students as they progress through grades 4-6 as classroom teachers are first entering the project in comparison with *follow-up longitudinal student cohorts* to look for impact of classroom teacher experience with the teaching artists and its impact on student learning.

Whenever possible this pattern of analysis will be applied to *student internal assessment and survey data* that were administered to both control and treatment groups such as the Snapshots of Arts

Integration Learning (SAIL) and PAIR student surveys, and to those internal assessments only administered to the treatment school students such as PAIL portfolio conference performance assessments or PAIL student work assessments.

In the final sections of the comprehensive report, we will explore the intersections between *internal measures of teacher professional development outcomes* and *both external and internal measures of student learning*. That is, we will investigate relationships between data collected from teachers (including the Year-End Curriculum and Teacher Survey [YECS], PD session surveys and attendance figures, PAIL Portfolio conference transcribed comments, teacher practice labels of student work; PD session documentation panels & curriculum maps) and all forms student learning assessment data described above.

Presentation of Results

The presentation of results in the following sections of this paper will be in the form of narrative report informed by statistical analysis. Thus each section will be framed by a sequence of inquiry questions, themes, and vignettes designed to illuminate the stories that have emerged from the PAIR project. In the last section the implication of PAIR outcomes will be explored in ways that will help teachers, administrators and parents understand better the essential role and potential impact of research-based arts integration teaching practices and learning outcomes in upper elementary public school education.

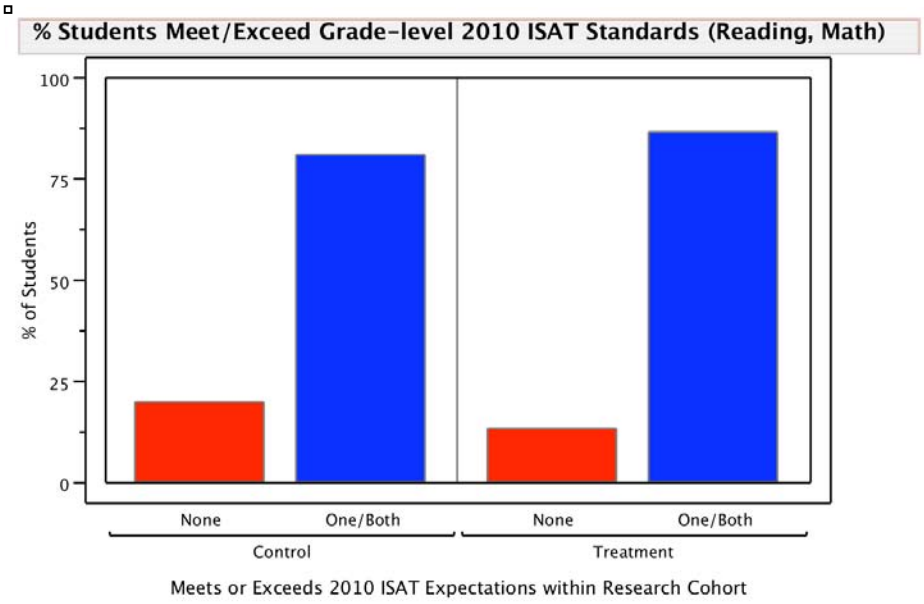
* * *

2B. External Student Learning Assessments: *Control-Treatment (C-T) Analysis of Illinois Standards Academic Achievement Test (ISAT) Results*

***Section 2B Inquiry Questions*:** To what extent are there significant differences in the percent of students 'meeting or exceeding or exceeding grade-level CPS ISAT benchmarks' among District, PAIR control and PAIR Treatment Schools in Reading and Math? What are differences between the 2010 INITIAL Longitudinal Cohort (Grade 6) and the two FOLLOW-UP Longitudinal Cohorts (Grades 4-5)

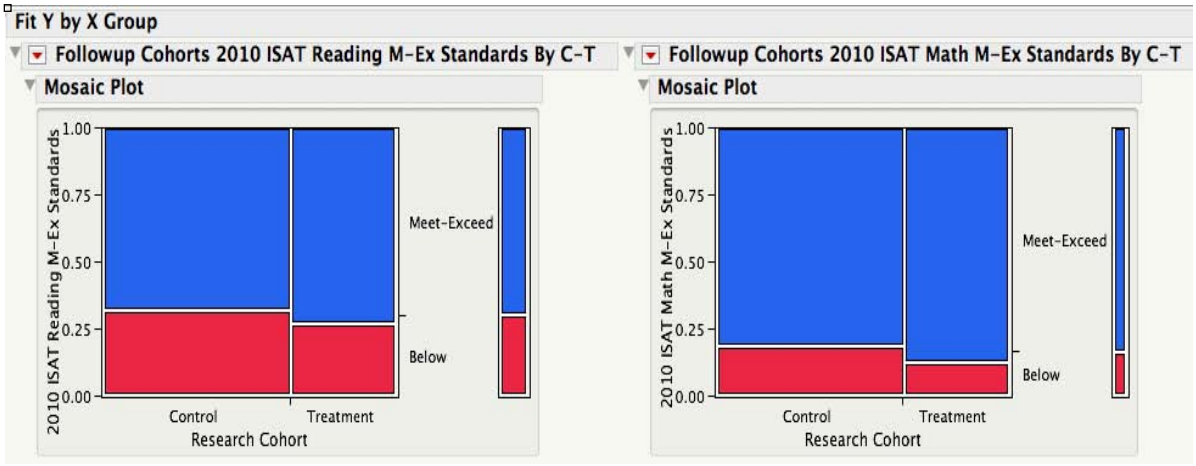
2B Figure 1 *Data Display Indications*: The bar chart below indicate clearly that, by the end of the PAIR Project, *the Treatment School initial longitudinal student cohorts contain a higher percentage (87%) of students who meet or exceed CPS grade-level benchmarks in Averaged Reading and Math ISAT test scores than students in the Control School initial longitudinal cohorts (81%)* [2B Figure 1]. This positive indication of comparatively strong general academic progress in the treatment schools is the bottom line statistic required of all successful AEMDD projects.

2B Figure 1: **Control-Treatment School Comparisons of Percent of PAIR Grade 6 Initial Longitudinal Cohort Students who Meet or Exceed Chicago Public School ISAT Benchmarks in One (Reading or Math), Both (Reading and Math) or None (Neither Read or Math)**



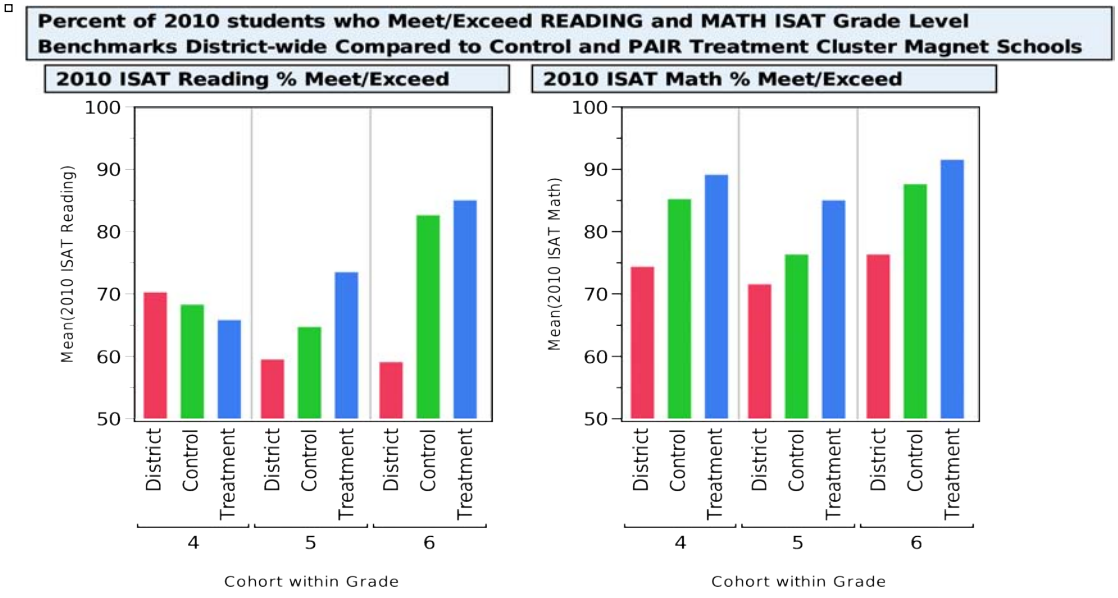
2B Figure 2 *Data Display Indications*: The mosaic chart below [2B Figure 2] expands the comparison to include the Reading and Math benchmark data to include averaged statistics from both of the two PAIR follow-up longitudinal cohorts control-treatment school comparisons. Follow-up Longitudinal Cohort 1 (grades 4-5, 2009-10) and Follow-up Longitudinal Cohort 2 (Grade 4, 2010). Comparison of benchmark data indicate again the positive impact of the PAIR program over time and across three different student samples in both Reading and Math.

2B Figure 2: **Control-Treatment ISAT Grade Level Benchmark Comparisons for Reading and Math in PAIR Follow-up Longitudinal Cohorts**



2B Figure 3 ***Data Display Indications***: The bar charts below provide a great deal more detail by distinguishing Reading and Math ISAT scores in relation to three statistics: CPS District average, PAIR Control and PAIR Treatment School cohorts. *With the exception of grade 4 reading percentiles, the initial longitudinal Treatment School cohort students are outperforming the Control School students and averaged district scores at every level in both reading and math.* The significance of these findings are highest in the second year of the sample (Reading p<.005; Math grade 5 p<.001).

2B Figure 3: **Control-Treatment ISAT 2010 Grade Level Benchmark Comparisons for Reading and Math in PAIR INITIAL (Grade 6) and Follow-up (Grade 4 and 5) Longitudinal Cohorts**



The data comparisons above [2B Figure 3] indicate consistent performance differences between the treatment schools and especially between the treatment and the overall CPS district averages. Note that the single counter-indication in control-treatment comparisons that occurred in grade 4 of the first longitudinal group in Reading occurred only at the very first year of the project when the program was least developed. Note also that both the PAIR control and treatment school cohorts compare favorably to CPS averages in Reading and Math, clearly indicating that CPS both Arts and non-Arts Magnet school cluster programs are consistently outperforming non-cluster CPS schools significantly, especially in Reading test scores.

Section 2B *Summary, Emerging Themes*: All Magnet Cluster Schools in the PAIR project significantly exceeded the CPS district percentages of 'students who meet or exceed grade level benchmarks' for ISAT Reading and Math scores. *In addition, the control-treatment school comparisons indicate that PAIR treatment students outpaced the control students most significantly by the third year of PAIR project suggesting that considerable elements of teacher professional development that support CAPE arts integrated practices are required before substantial positive student learning effects can occur. Results from PAIR Follow-up treatment classrooms indicate positive comparisons with non-PAIR classrooms after only two years, thus providing evidence that the PAIR methods are replicable and its effects are accelerating the second time around.*

The next set of data displays will represent academic performance data more precisely and with increased statistical power by using ISAT scaled scores to detect differences in student academic performance.

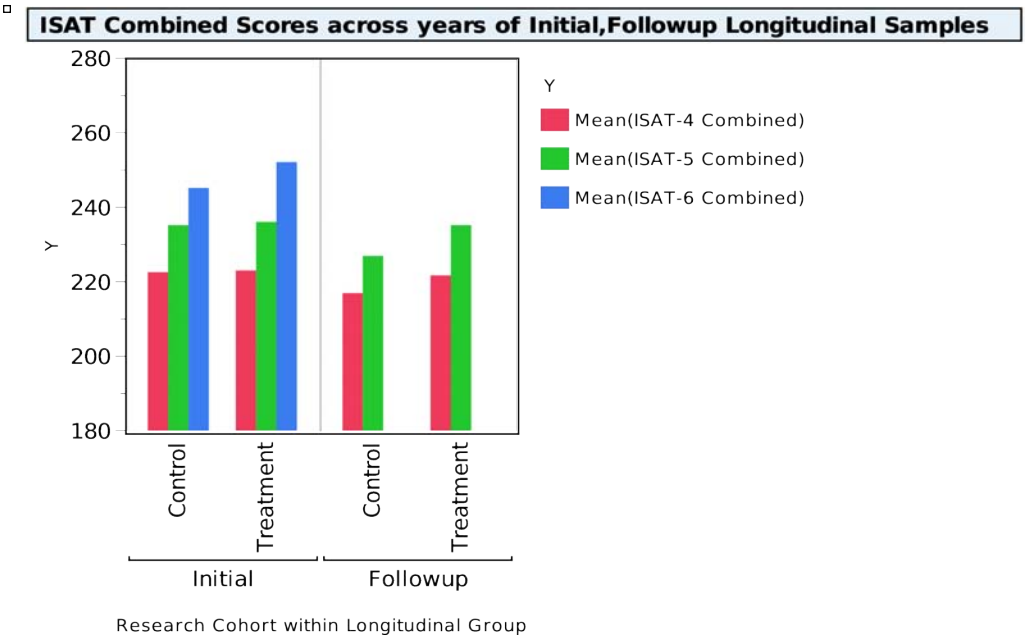
* * *

2C. Control-Treatment School Comparisons of PAIR Grade Level ISAT Mean Scores

2C Inquiry Question: To what extent did ISAT standardized academic test score results distinguish PAIR Treatment schools from PAIR Control schools?

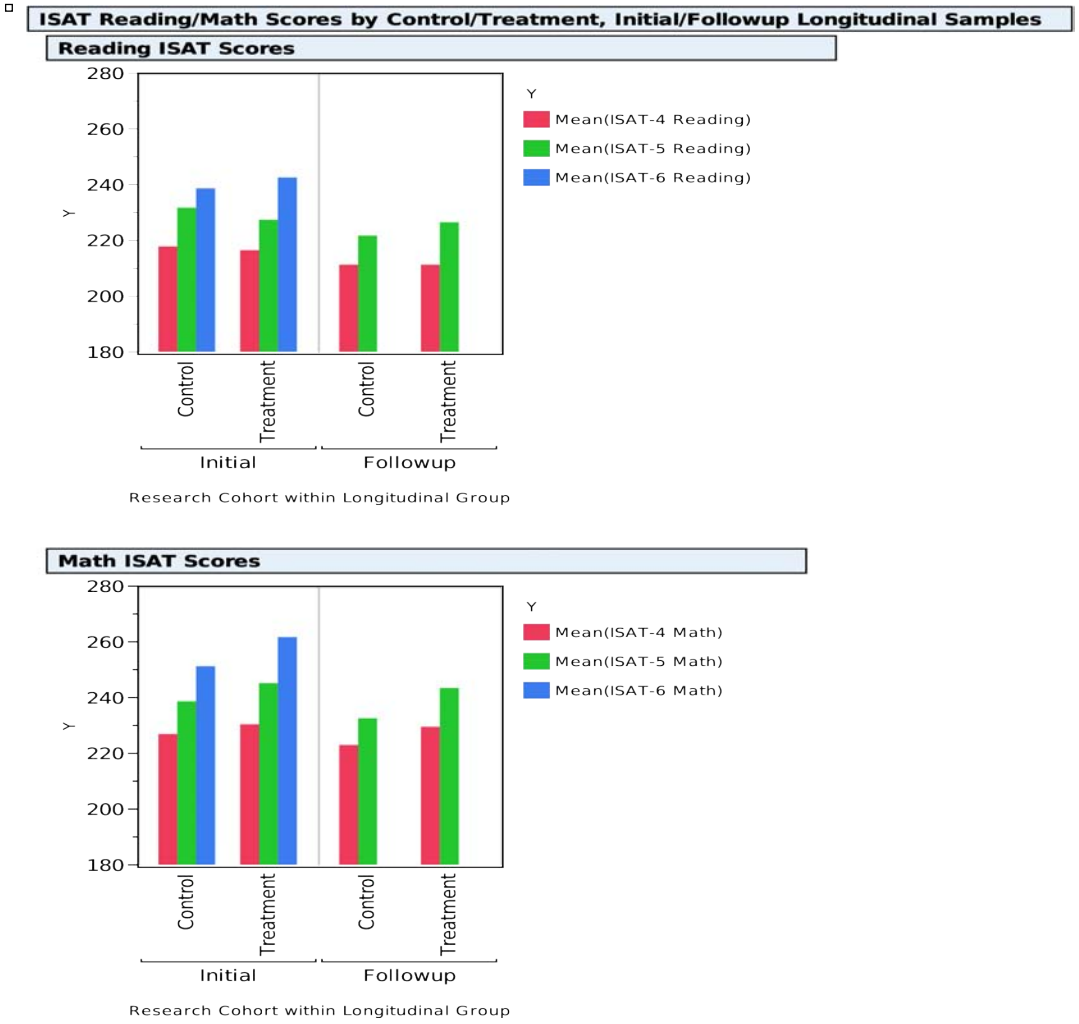
2C Figure 1 ***Data Display Indications***: The figure below [2C Figure 1] shows that student ISAT combined academic scores in the Initial Longitudinal Treatment School cohort increasingly outpaced its Control School cohort at every grade level (2008-10), yet only during the final year of the project (Grade 6) did ISAT Control-Treatment school differences finally become statistically significant (p< .01). However, both Follow-up Longitudinal Cohorts comparisons reveal immediate, highly significant differences favoring the Treatment school cohorts (grade 4 (p<.0004;) and grade 5 (p<.0001), suggesting strongly that the impact of the PAIR program increased during the second and third year implementation phases of the project.

2C Figure 1: Control-Treatment ISAT Grade Level Comparisons of ISAT Combined (Averaged Reading and Math) Mean Scores, PAIR Initial and Follow-up Longitudinal Cohorts



2C Figure 2 ***Data Display Indications***: The bar graph chart below indicates that ISAT Reading scores are not significantly outpacing the control group in the Initial Longitudinal Cohort. Nonetheless, ISAT Treatment School Reading scores are significantly outpacing the follow-up grade levels in both grade 4 (p< .04) and 5 (p< .000). In Math, the pattern of control-treatment school differences is similar to Reading, yet more pronounced. In the Initial Treatment cohort significant differences are obtained in Grade 5 (p< .04) and Grade 6 (p<.0002) and in Treatment school follow-up student cohorts the statistical significance immediately obtains in grade 4 (p< .0001) and sustained in grade 5 (p< .0001).

2C Figure 2: Control-Treatment Focus School (Arts vs. Academic) Comparisons of ISAT Grade Level Reading and Math Mean Scores, PAIR Initial and Follow-up Longitudinal Cohorts



2C *Summary, Emerging Themes*: Control-treatment cohort comparisons of combined and disaggregated Reading and Math ISAT standardized test scores provided *statistically significant* evidence of PAIR's impact on student learning outcomes. *The PAIR program impact, as measured by academic achievement test performance outcomes, became evident only after positive control-treatment comparisons were established in the Initial Longitudinal Student cohort by the third year of the project. The follow-up cohort comparisons demonstrated more immediate and more highly significant PAIR Treatment School student learning outcomes during the second year of the program.*

* * *

2D: Control-Treatment Focus School (Arts vs. Academic) Comparisons of ISAT Mean Scores

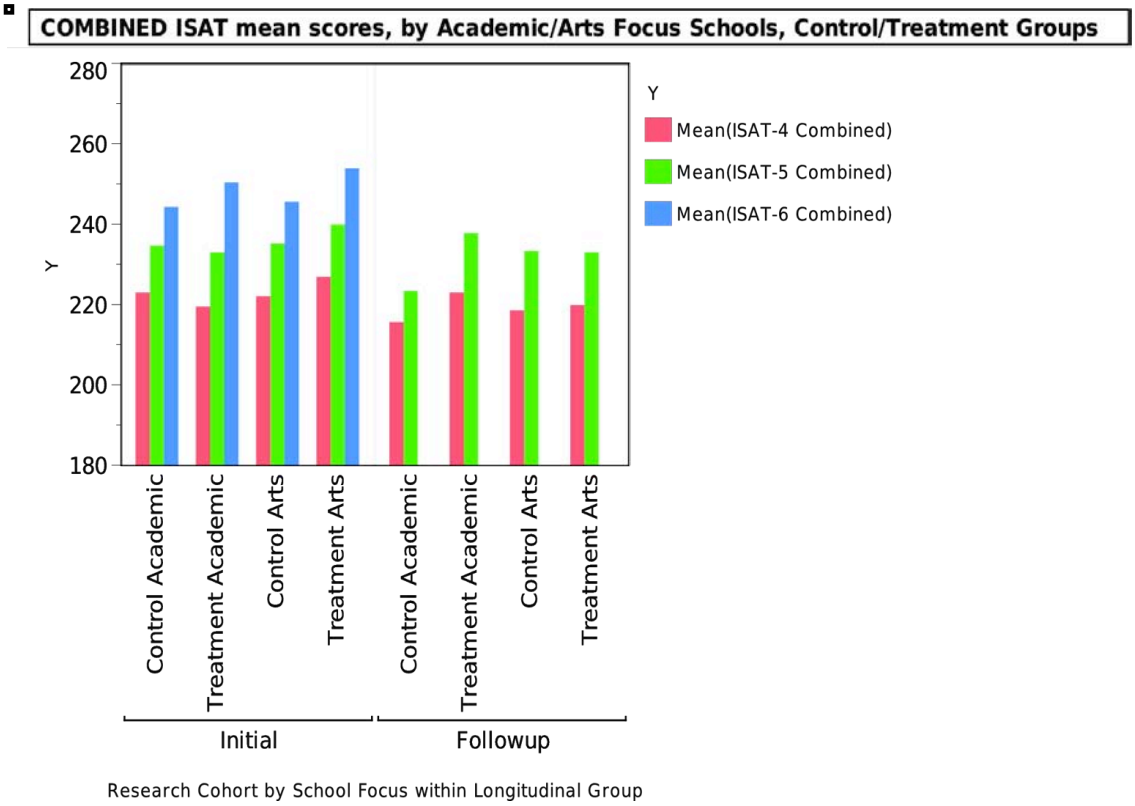
In the PAIR project, all schools were matched according to their primary focus on arts in contrast to non-arts learning. Table 2A-1 above describes how the PAIR Treatment and Control schools were evenly divided according to their designation and resources as Arts Learning Magnet Cluster Schools and various categories of academic learning such as Writing (English Language Arts), Math, or World Language (foreign language and cultural studies) Magnet Cluster Schools.

2D Inquiry Question: To what extent did the primary focus on arts versus non-arts learning predict differences in ISAT student learning outcomes?

2D Figure 1 ***Data Display Indications***: The bar graph below [2D Figure 1] displays differences in grade level ISAT Combined academic mean scores according to School Focus and differentiated by Initial and Longitudinal student cohorts. The order of comparison progresses from Academic Focus control and treatment school comparisons to Arts control and treatment school comparisons (left to right). Results displayed here suggest that, *by the third year of the project, both Arts and Academic Treatment Focus Schools have higher levels of academic scores than do the Control Focus school, suggesting that PAIR program succeeds equally well with either type of school focus. This same pattern generally obtains in the Follow-up longitudinal cohorts, with far more dramatic results developing in the control-treatment Academic Focus schools comparisons than in the control-treatment comparisons among the Arts Focus Schools.*

(continued on next page)

2D Figure 1: Control-Treatment PAIR Focus School (Arts vs. Academic) Comparisons of ISAT Mean Scores, PAIR Initial Longitudinal Cohorts, Grades 4-6.



2ID *Summary, Emerging Themes*: In the Initial Longitudinal Student Cohort, the bar graph patterns show that the *Treatment Arts Focus Schools* maintain pre-eminence as the highest scoring PAIR Focus School throughout the three years of project development, while the *Treatment Academic Focus Schools* appear to be catching up by the final year of the project (from the lowest test scores in year 1 to the second highest by year 3). *These results indicate strongly that the PAIR Treatment Arts Focus School "arts plus arts integration programs" have a more profound effect on academic achievement than do the conventional arts learning programs in the Control Arts magnet schools.*

The data here also show that it took three years for the PAIR Treatment Academic Focus School students in the Initial data sample to outperform the more conventional Control Academic Focus Schools. However, the Treatment Academic Focus School kids gained the advantage over the control schools immediately in the Follow-up student cohorts, most likely as a result of benefitting from the second and third year of classroom teacher experience with the program.

Most striking was the pattern of increasingly deflated ISAT scores in the Control Academic Focus Schools in comparison with all other schools. After three years in the Initial cohort, the Control Academic schools went from being rated second highest to lowest scoring schools. *By the time PAIR was being replicated by classroom teachers in the Follow-up Cohorts, the Control Academic Focus Schools perpetually scored lower than all other school cohorts in the study.* This study clearly shows that CPS magnet schools defined by the adherence to academic focus are less likely to rely on arts learning resources and staff training devoted to arts integrated learning and, as a result, *Academic Focus Schools increasingly lagged behind the academic performance of schools that prioritize arts and arts integration teaching and learning programs that offer a more wide-ranging approach to inter-disciplinary teaching and learning practices.*

* * *

2E: Control-Treatment PAIR School Achievement Gap Comparisons of ISAT Mean Scores According to Pre-Designated High, Average, and Low (HAL) Student Achievement Cohorts, Grades 3-6

In the PAIR project, the students were randomly selected in equal numbers from pools of individuals categorized as High, Average, and Low (HAL) according to measures of academic performance available to each grade level in each school. Typically ISAT grade 3 test scores were used to designate who the H, A, or L students were before the PAIR program began. By having these HAL designations in equal numbers from each classroom in the longitudinal samples, we are able to trace the impact of PAIR on three different learner populations throughout each phase of PAIR project implementation.

2E Inquiry Questions: To what extent did the PAIR program benefit students previously designated as high, average and low academic achieving students prior to PAIR project implementation? Were there any control-treatment differences with regard to the learning achievement gaps between high, average, and low performing learners during the three-year development and implementation of the PAIR program? Is there an overarching difference in patterns of equity and excellence that were obtained in control-treatment school comparisons?

2E Figure 1 ***Data Display Indications***: The bar charts below display ISAT Combined, Reading and Math mean scores according to HAL categories only in the Initial PAIR Longitudinal Cohort Grades 3 (baseline) and grades 4-6 implementation years of the project. The red bars represent the profile of mean scores that defined the HAL ratings; that is, the initial achievement gap that was used to categorize the students. The statistical trend in the control school data suggests that the achievement gap closes a bit between the high (H) and average (A) learners over time.

In contrast, the PAIR Treatment school profile reveals that (a) the achievement gap between PAIR Treatment School average (A) and low (L) student closes significantly in the early years of the project, and later on that (b) the overall achievement scores in Treatment scores become significantly higher than Control schools in the final year of the project and (c) the Treatment School Low students are significantly outperforming the Control group students in both Reading and Math by the final year of PAIR program development. Thus, *patterns that emerged consistently from data reports throughout various strands of the PAIR data analyses* [2E Figure 1 below] *suggests that closing the learning gap for students at risk for academic achievement leads later on to significantly higher levels of achievement for both previously High and Low ranked learners.*

(continued on next page)

2E Figure 1: Control-Treatment PAIR Schools Achievement Gap Bar Graph Display of HAL ISAT Combined, Reading, and Math Mean Scores, Grades 3-6 (Initial Longitudinal Cohort)

The following charts [2E Figures 2-4] represent the control-treatment HAL learner gap analysis in more detail. The box plots in this chart profile mean score and range of student achievement changes with each year of the project. The circles to the right of each boxplot chart indicate statistically significant differences among the HAL student cohort learning profiles. Completely separated mean score parameters indicates a statistical determination that there are statistically significant achievement gaps between the learner categories. When the circles overlap, there is now longer a significant difference among the learning categories indicating that learning gap has been closed.

2E Figure 2 ***Data Display Indications***: The circles at the right of each boxplot express the pattern of statistical significance of the achievement gaps among the HAL learners over time. The Control School pattern clearly indicates that the Low (L) learners remain significantly behind the High-Average (H-A) Learners throughout the four years of the project. In clear contract, the Treatment Schools pattern (a) begins with all three learning groups clearly differentiated academically, yet (b) the three groups are not statistically different for the remainder of the project as (c) the overall academic scores begin to favor students in the PAIR program.

2E Figure 2: Control-Treatment PAIR Schools Achievement Gap Box Plot Analysis of HAL ISAT Combined Academic Scores, Grades 3-6

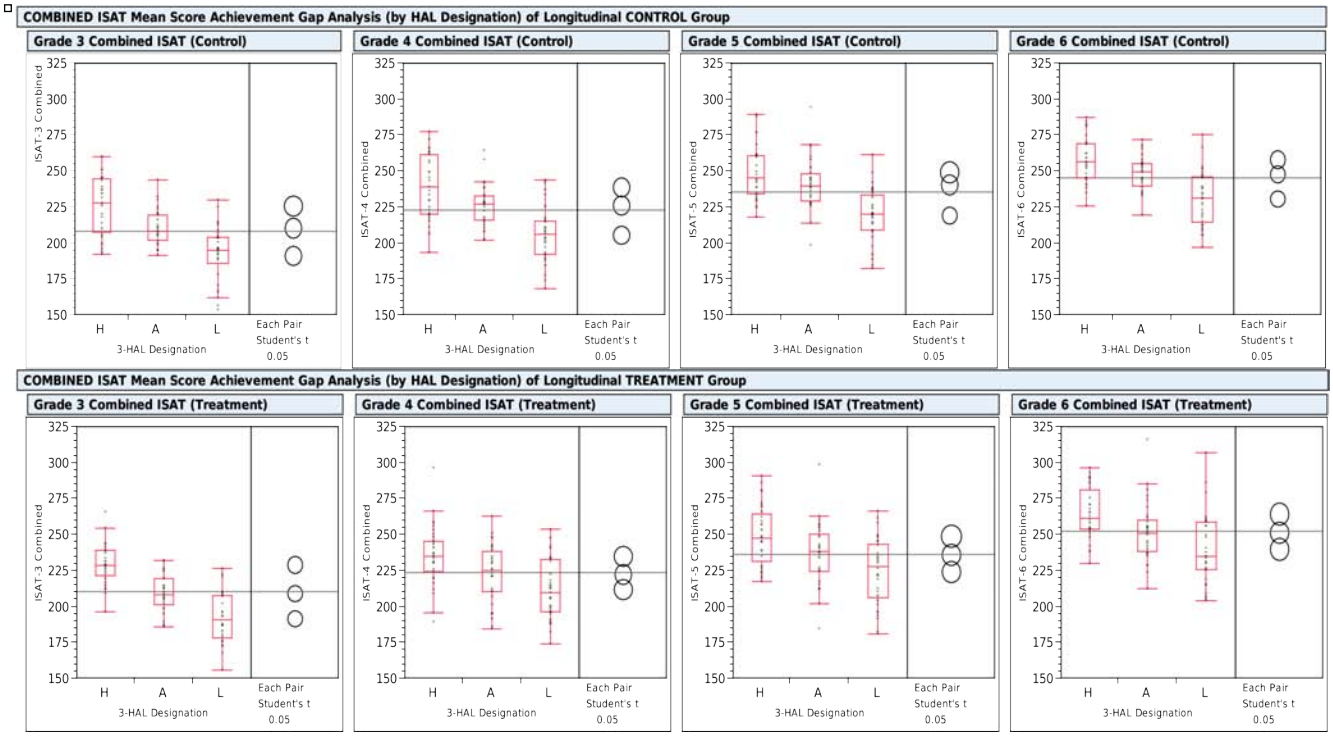
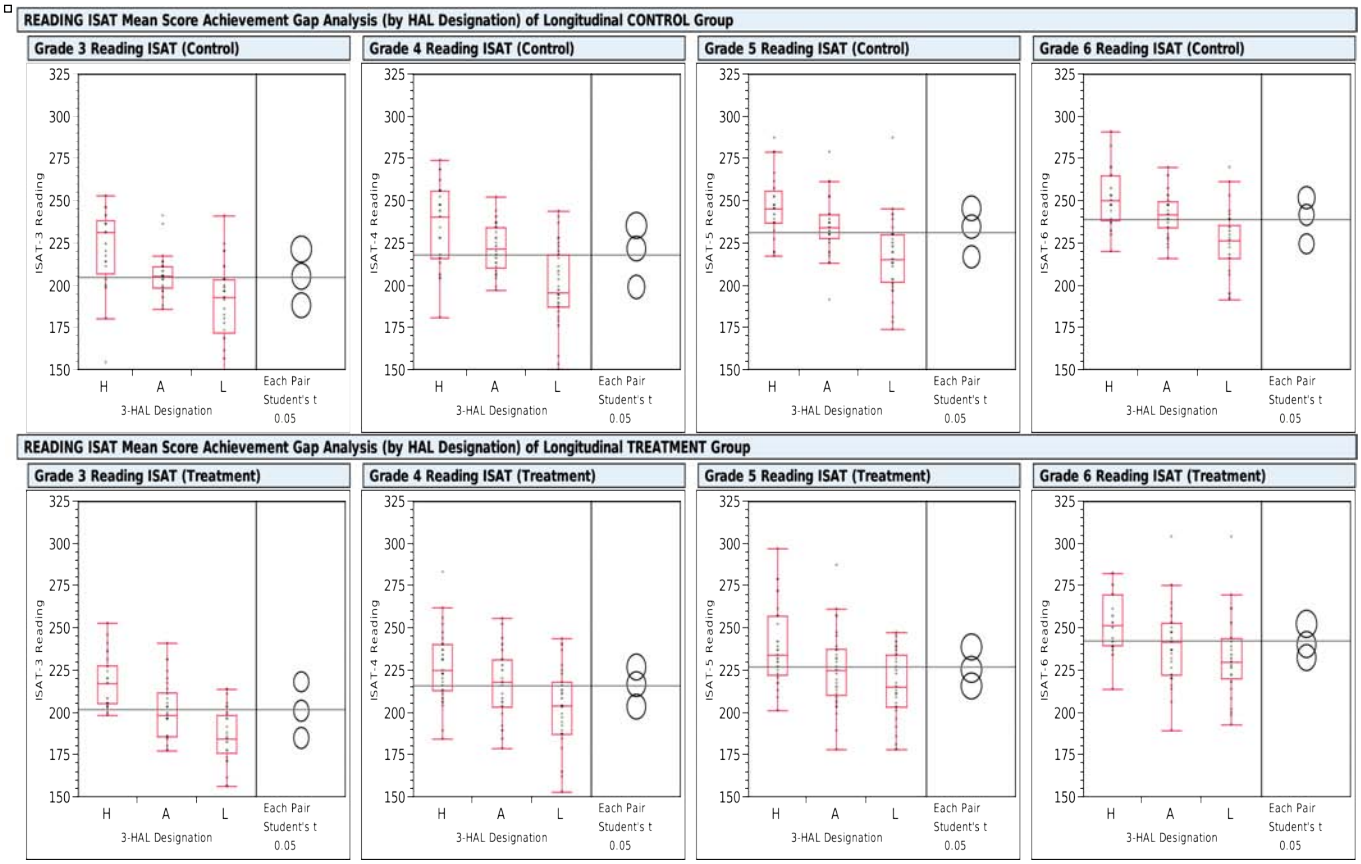


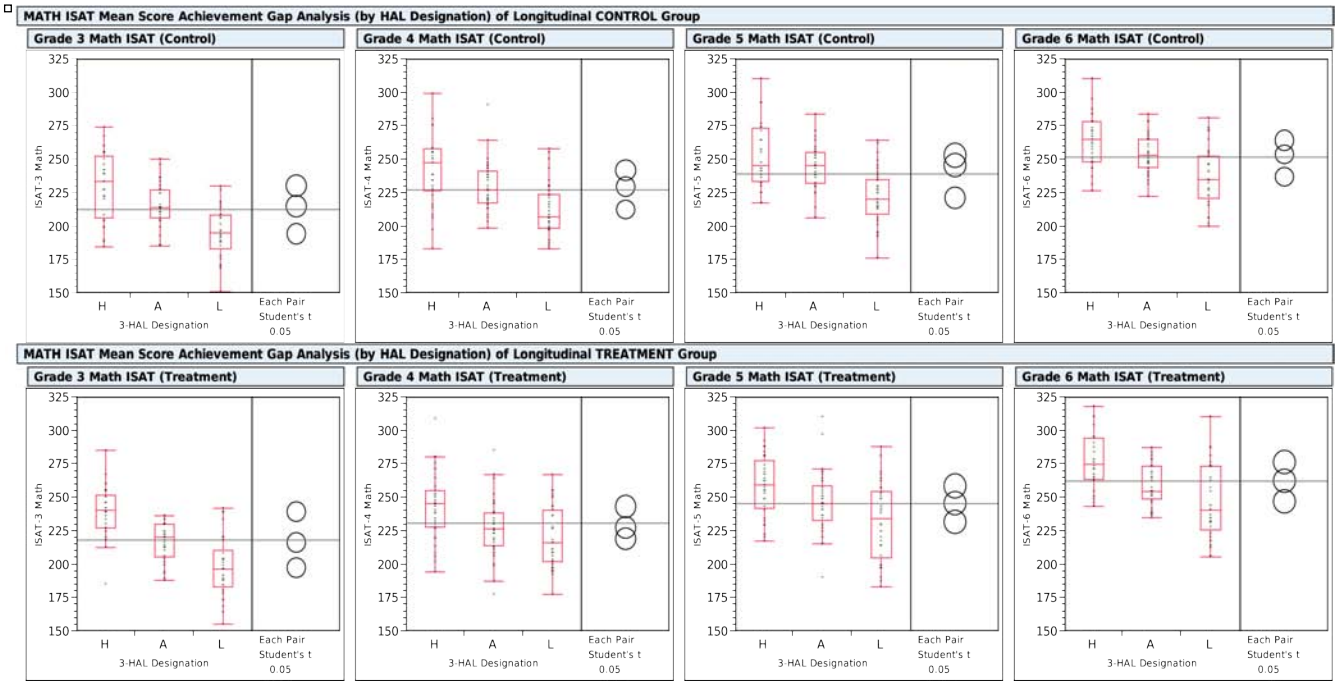
Figure 2E Figure 3: *Data Display Indications*: The following two charts display the HAL achievement gap analysis broken down by Reading and Math ISAT scores. The PAIR treatment school phenomenon of 'closing the achievement gap for Low learners precedes overall superior academic achievement' occurs in Math and especially in Reading. Note also that the achievement gap between Low learners and High-Average student cohorts appear never to reach the level of non-significance, i.e., the statistical evidence for closing the achievement gap among the three HAL student cohorts.

2E Figure 3: Control-Treatment PAIR Schools Achievement Gap Box Plot Analysis of HAL ISAT Reading Scores, Grades 3-6



(2E Figure 4 on next page)

2E Figure 4: Analysis of Control-Treatment PAIR Schools Achievement Gap Box Plot According to HAL ISAT Math Scores, Grades 3-6



2E *Summary, Emerging Themes*: Achievement gap analysis is a term used here to describe to the extent to which students previously classified as high [H], average [A] or low [L] academic achievement students in the third grade remain fixed in these categories as they progress through grades 4-6.

If there is little change in the statistically significant differences in mean score averages among Treatment HAL student cohorts, then the achievement gaps remain in place during the course of the PAIR program development. If the mean scores become more highly differentiated, the gap PAIR program will be associated with widening the learning gap between the previously designated low, average and higher learners. However, if the PAIR program results in weakening the distinctions among the HAL cohorts in Control-Treatment cohort comparisons, the PAIR program then will appear to have contributed to closing the learning gap among learners. Ideally, if PAIR Treatment schools were to close this achievement gap while simultaneously outperforming the Control Schools, then PAIR program results would suggest that changing the dynamics of widely separated, fixed student learning groups is a contributing factor to broad-based positive change in academic performance and a school culture of equitable teaching and learning practices.

The data presented so far in this report suggest that the pattern of HAL student cohort performance on ISAT tests vary considerably in control-treatment comparisons. Strikingly similar to the ideal scenario described above, the bar chart displays indicate that Treatment Schools a) consistently narrowed the learning gap for the designated Low achieving students in all areas of academic performance throughout all three years of the PAIR project and, b) by the final year of the PAIR project, the Treatment Schools were outperforming Control Schools in Combined and Math Academic Scores, while keeping pace with Reading Scores. *Most importantly, to whatever extent the ISAT Treatment School scores outpaced the Control Schools, the Treatment School PAIR Treatment Low achievement student cohorts always significantly outperformed their counterpart Control School Low cohort students in every area of academic achievement.*

The boxplot analyses provided statistically significant evidence that the boundaries among the HAL classifications are less rigid in PAIR Treatment schools than in the PAIR control schools. The pattern and shape of these data represent unmistakable evidence of a pattern of "equity leading to excellence" in the PAIR treatment schools.

The analysis of HAL cohort data provides an important window onto the nature and problems of achievement gaps among various sorts of students in public education. In this study it appears that to some degree schools that did not employ arts integration teaching and learning practices in their schools for the most part essentially reinforce predisposed assessments of students as high, medium or low achievers. *In PAIR Treatment schools statistical results support the hypothesis that arts integration programs can reduce the achievement gap for low-achieving students at risk for falling further behind the pre-classified high and/or average students – and that leveling the playing field for all learners can lead to yet higher levels of whole school academic performance.*

Further analyses below will look into this hypothesis in more detail through further external student learning comparisons based on School Focus (primarily Arts or Non-Arts) and School Cluster designations (Writing (ELA), Math, and World Language (& Cultures).

* * *

2F: Control-Treatment PAIR Focus School (Arts vs. Academic) Comparisons of HAL ISAT Mean Scores and Achievement Gap Profiles, Grades 3-6

As previously discussed, PAIR schools were selected in equal numbers from pools of CPS Magnet Cluster Schools categorized as Arts Learning Focus Schools that are also assigned an Academic Focus in Writing (ELA), Math, or World Language (WL) Cluster Schools. Thus student learning outcomes in the PAIR project can be grouped into the following four categories of schools: Control Academic, Treatment Academic, Control Arts, and Treatment Arts schools [see Table IA above].

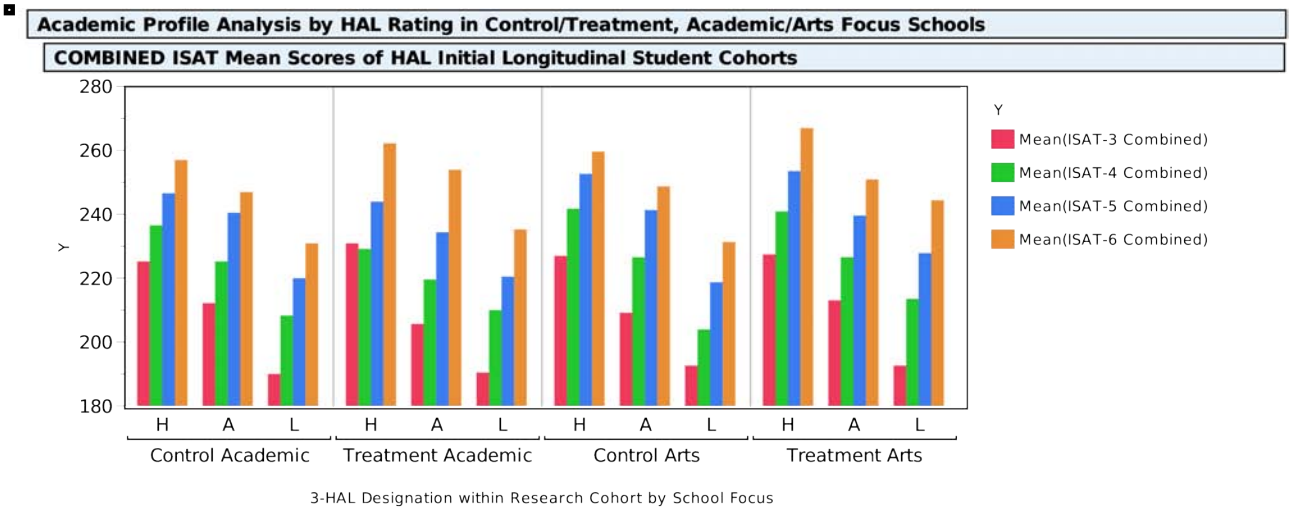
2F Inquiry Questions: To what extent does the category of School Focus (primarily arts v. academic) affect standardized test outcomes? To what extent does the category of School Focus affect any control-treatment differences with regard to the learning gaps among high, average, and low performing learners during the three-year development and implementation of the PAIR program?

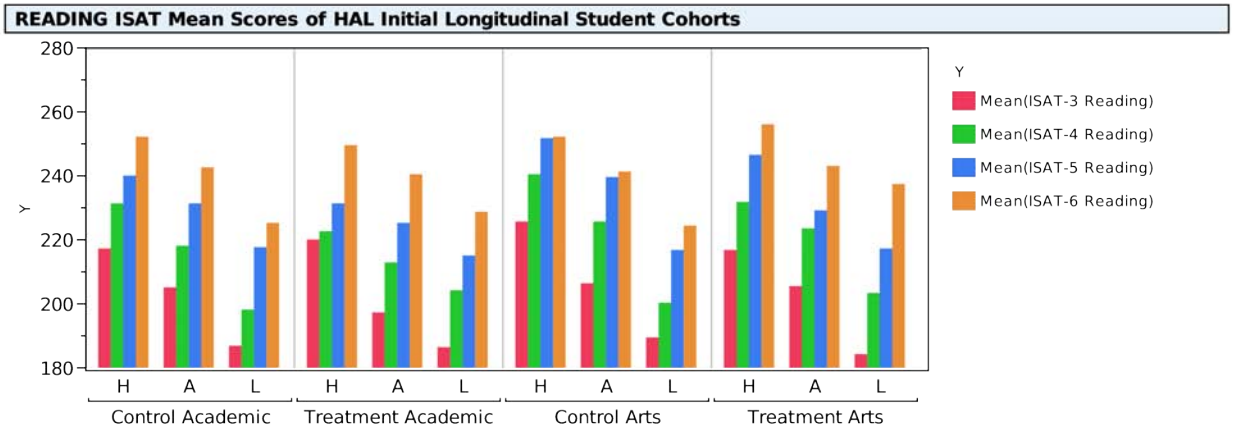
2F Figure I *Data Display Indications*: The bar chart display below demonstrates differences in ISAT mean scores among the four types of PAIR Focus Schools. Reading from left to right, the red bars (project planning year data) are extremely similar in scope in all school cohorts by definition, because the HAL student cohorts were selected and matched according to relative High, Average, and Low levels of achievement according to ISAT test results in the 3rd grade.

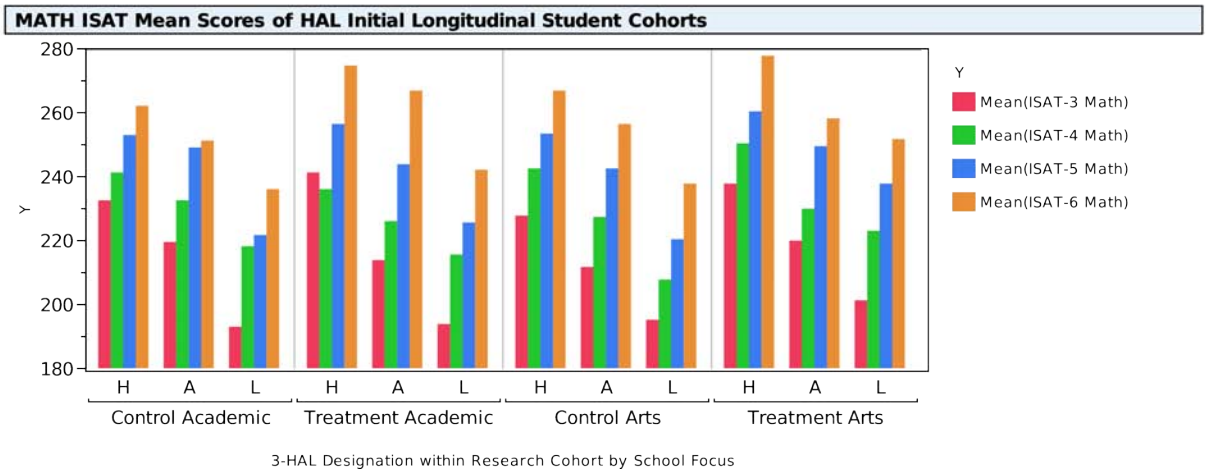
In the first two years of project implementation, the data show that both control and treatment ARTS FOCUS schools scored higher in combined academic test scores than control and treatment Academic Schools, suggesting at that time that Arts Focus, and not the PAIR program was the more powerful predictor of academic achievement results.

However, by the final year of the project we see that both treatment Arts and treatment Academic PAIR schools posted the highest Combined ISAT scores, with the Treatment Arts Schools showing the largest gains from year 2 to year 3. In the final analysis it is the Treatment Arts FOCUS schools that not only have the highest test scores, but also demonstrate the clearest evidence for closing the achievement gap with previously designated below average learners, especially in Math [2F Figure 1].

2F Figure 1: PAIR Focus School Control-Treatment Achievement Gap Bar Graph Display of ISAT Combined, Reading, and Math Scores by HAL Cohorts, Grades 3-6







2F Table 1 ***Data Display Indications***: The patterns of significant differences associated with the data results displayed in Figure 2E-1 also obtain in the following three tables.

In the table below, an interpretive rubric is provided to provide a lens for measuring change in the grade level HAL ISAT data results over time. The placement of the slash marks (/) indicate highly distinct, statistically significant boundaries exist at various divisions among the HAL learner cohorts at all levels of the project. The degree of separation in the HAL cohort ISAT outcomes ranges from 'most differentiated' (i.e., separated according to significant differences between the HAL cohorts) to 'least differentiated' performance' (in some cases, no significant differences whatsoever among the three HAL levels).

2F Table 1: Rubric for Categorizing Statistically Significant Levels of Achievement Gap Separation among the HAL Student Cohorts.

Most Differentiated ----------Least Differentiated			
Level 1: H/A/L	Level 2a: HA/L or Level 2b. H/AL	Level 3: HA/AL	Level 4: HAL

The table below [2F Table 2] represents the various pattern of HAL 'achievement gap analysis' of HAL ISAT Combined scores in both the Control and Treatment Schools. This table reveals that the statistically significant achievement gaps in Control Schools never change among the HAL group cohorts and that the achievement gap is always most severe between low (L) learner cohorts and the rest of the students.

In contrast, the pattern of decreasing differentiation among the HAL cohorts obtains much more strongly in the Treatment Schools. In both Arts and Academic Focus Schools, the highly differentiated academically rated populations (High, Average, Low Learners) obtained at the beginning of the project become far less separated as the project implementation proceeded. Most significantly, the initially Low learning students are no longer performing at a statistically significantly different level from the previously designated Average Learners.

2F Table 2: **Levels of Statistically Significant Achievement Gap in HAL ISAT Combined Test Scores by Control-Treatment School Focus Designation, Grades 3-6**

	CONTROL *ACADEMIC FOCUS* CPS Magnet Schools	TREATMENT *ACADEMIC FOCUS* (Plus Arts Integration) CPS Magnet Schools	CONTROL *ARTS FOCUS* CPS Magnet Schools	TREATMENT ARTS FOCUS (Plus Arts Integration) CPS Magnet Schools
Grade 3 (baseline)	HA/L	H/A/L	H/A/L	H/A/L
Grade 4 (PAIR yr1)	HA/L	HA/AL	H/A/L	H/AL
Grade 5 (PAIR yr2)	HA/L	HA/AL	HA/L	HA/AL
Grade 6 (PAIR yr3)	HA/L	HA/L	HA/L	H/AL

The separated Reading and Math ISAT results in the next two tables [2F Tables 3-4] indicate that this same pattern of reduction of the learning gap far more significantly in PAIR Treatment School cohorts compared to the Control Schools in both disciplines.

2F: Table 3: **Levels of Statistically Significant Achievement Gap Differentiation in HAL ISAT Reading Test Scores by Control-Treatment Focus School Designation, Grades 3-6**

	CONTROL *ACADEMIC FOCUS* CPS Magnet Schools	TREATMENT *ACADEMIC FOCUS* (Plus Arts Integration) CPS Magnet Schools	CONTROL *ARTS FOCUS* CPS Magnet Schools	TREATMENT ARTS FOCUS (Plus Arts Integration) CPS Magnet Schools
Grade 3 (baseline)	HA/L	H/AL	H/A/L	H/A/L
Grade 4 (PAIR yr1)	HA/L	HA/AL	H/A/L	HA/L
Grade 5 (PAIR yr2)	HA/L	HA/AL	HA/L	H/AL
Grade 6 (PAIR yr3)	HA/L	HA/L	HA/L	HA/AL

2F: Table 4: **Levels of Statistically Significant Achievement Gap Differentiation in HAL ISAT Math Test Scores by Control-Treatment School FOCUS Designation, Grades 3-6**

	CONTROL *ACADEMIC FOCUS* CPS Magnet Schools	TREATMENT *ACADEMIC FOCUS* (Plus Arts Integration) CPS Magnet Schools	CONTROL *ARTS FOCUS* CPS Magnet Schools	TREATMENT ARTS FOCUS (Plus Arts Integration) CPS Magnet Schools
Grade 3 (baseline)	HA/L	H/A/L	H/A/L	H/A/L
Grade 4 (PAIR yr1)	HA/AL	H/AL	H/A/L	HA/AL
Grade 5 (PAIR yr2)	HA/L	HA/AL	HA/L	HA/AL
Grade 6 (PAIR yr3)	HA/L	HA/AL	HA/L	H/AL

2F *Summary, Emerging Themes*: The bar chart displays and tables that trace patterns of significance distinctions among the HAL cohorts presented here show that, over time, *the PAIR Treatment Schools generally outperform the Control Schools regardless of their designation as an Arts or Academic (Non-Arts) FOCUS School. This finding suggests that PAIR arts integration treatment improves school performance for both arts and academic focused schools with equal effect.*

In addition, the achievement gap analysis establishes that the Arts Treatment Schools – those schools that focus both on arts and arts integration teaching and learning – not only outperform all other schools as previously reported, but also demonstrate the most compelling school profile for reducing the learning gap for low achieving students. These findings suggest that PAIR arts integration methods and practices best demonstrate how raising test scores and closing the gap for the lowest level achievers is also a potent strategy for improving school culture and academic improvement simultaneously.

The next section presents case study differences in ISAT and HAL cohort data according to the shared academic focus of the PAIR school collaborations.

* * *

2G: Control-Treatment PAIR MAGNET CLUSTER SCHOOL (ELA, Math, World Language) Comparisons of HAL ISAT Mean Scores and Achievement Gap Profiles, Grades 3-6

As previously discussed, PAIR schools were selected in equal numbers from pools of CPS Magnet Cluster Schools assigned an Academic Focus in Writing (ELA), Math, or World Language & Cultures (WL) Schools [Table IA above]. Throughout the PAIR project Arts and non-Arts teachers were assigned to collaborate with two CAPE teaching artists to create and implement arts integration units designed to reinforce on essential concepts that would optimize learning across two arts disciplines. Treatment Writing (ELA) Cluster schools worked with CAPE music and drama teaching artists; Treatment Math Cluster school teachers worked with CAPE dance and visual arts teaching artists; Treatment World Language s (WL) Schools worked with CAPE music and visual arts teaching artists [Table 1A above]. This essential structural feature of the PAIR treatment schools challenged the

CAPE teaching artists to co-create and implement units throughout the three years of implementation that addressed the essential focus and needs of each Cluster School type. Thus the 'pairing' of arts integration specialists with classroom teachers constituted, along with the supporting professional development sessions and requirements, represents the element of the PAIR program that distinguished the Treatment from the Control Cluster Schools.

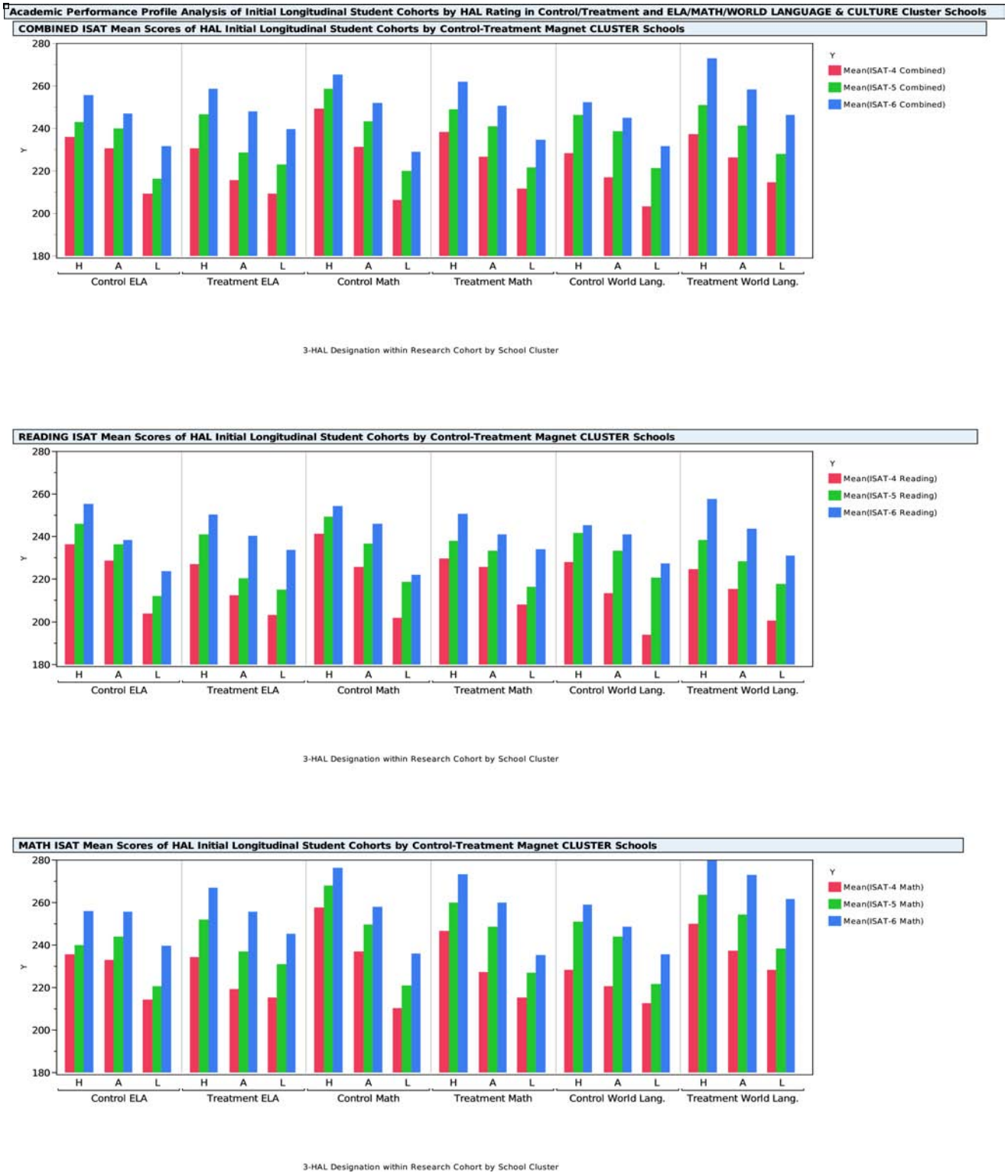
2G Inquiry Questions: To what extent does the category of Cluster School Academic focus (ELA vs. Math vs. World Languages) affect ISAT standardized test scores and the HAL student cohort learning outcomes? To what extent does the category of School Cluster affect any control-treatment differences with regard to the learning gaps among high, average, and low performing learners during the three-year development and implementation of the PAIR program?

2G Figure 1 ***Data Display Indications***: *Statistical trends in the ISAT Combined Scores indicate that all he PAIR Treatment Cluster Schools, regardless of any Academic focus classification, are more likely than Control Schools to generally outperform or to close the achievement gap for the most at risk learners by the third year of the project.* It is also clear that the one particular set of schools, the World Language Cluster Schools, consistently significantly outperform all other Cluster School types and manage to narrow the gap between the high, average and low performing students.

Although comparisons in Reading ISAT scores reveal that there may be a few cases where reading achievement in the highest HAL cohorts may be superior in the Control ELA and Math Cluster Schools compared to the Treatment schools, the Low Achiever cohort gap is still significantly less present in all the Treatment schools. *And in no cases do the Control School Low performing students outperform the low performing students in Treatment schools.* Once again there is no doubt that the Treatment World Language schools are outperforming their control group counterparts in all three HAL cohorts, all three years of project implementation.

Although the Math ISAT scores for the Math Cluster schools are virtually the same in both the Control and Treatment Schools, the ELA gained an edge over their comparison schools in the last two years of the project. WL Treatment schools outpaced their Control Schools completely at every phase of the project in terms of ISAT mean scores while, in this case, both Control and Treatment schools managed to close the achievement gap significantly.

2G Figure 1: Control-Treatment PAIR Cluster School Achievement Gap Analysis of ISAT Combined, Reading, and Math Scores by HAL Cohorts, Grades 4-6



Section 2G Tables 1-3 below display the various pattern of HAL 'achievement gap analysis' in HAL ISAT Combined, Reading, and Math scores in both the Control and Treatment Schools. These tables reveal that, *regardless of Cluster School Type, Control Schools are never able to demonstrate change among the HAL group cohorts and that the gap is always most severe between low (L) learner cohorts and the rest of the students.*

In contrast, the Treatment Schools demonstrate that the initially designated Low learning students are no longer performing at a statistically significantly different level from the previously designated Average Learners in the first year of the PAIR project. By year 3 of the project most Treatment Cluster schools have virtually removed the gap between previously designated High, Average, and Low Achievers.

2G Table 1: Control-Treatment PAIR Cluster School Levels of Statistically Significant Achievement Gap Differentiation in HAL ISAT Combined Test Scores, Grades 3-6

	CONTROL *Writing-ELA* Cluster Schools	TREATMENT *Writing-ELA* Cluster Schools	CONTROL *Math* Cluster Schools	TREATMENT *Math* Cluster Schools	CONTROL *WL* Cluster Schools	TREATMENT *WL* Cluster Schools
Grade 3 (baseline)	H/A/L	H/AL	H/A/L	H/A/L	HA/AL	H/A/L
Grade 4 (PAIR yr1)	HA/L	HA/AL	H/A/L	H/AL	HA/L	HA/AL
Grade 5 (PAIR yr2)	HA/L	HA/AL	HA/L	HA/L	HA/L	HA/AL
Grade 6 (PAIR yr3)	HA/L	HAL	H/A/L	HA/AL	HA/AL	HA/AL

2G Table 2: Levels of Achievement Gap Differentiation in HAL ISAT Reading Test Scores by Control-Treatment School Cluster Designation and by Grade

	CONTROL *Writing-ELA* Cluster Schools	TREATMENT *Writing-ELA* Cluster Schools	CONTROL *Math* Cluster Schools	TREATMENT *Math* Cluster Schools	CONTROL *WL* Cluster Schools	TREATMENT *WL* Cluster Schools
Grade 3 (baseline)	H/A/L	H/A/L	HA/L	HA/L	HA/AL	H/A/L
Grade 4 (PAIR yr1)	HA/L	HA/AL	HA/L	HA/AL	HA/L	HA/AL
Grade 5 (PAIR yr2)	HA/L	HA/AL	HA/AL	HA/AL	HA/AL	HA/AL
Grade 6 (PAIR yr3)	H/A/L	HAL	HA/L	HAL	HA/AL	HA/AL

2G Table 3: Levels of Achievement Gap Differentiation in HAL ISAT Math Test Scores by Control-Treatment School Cluster Designation and by Grade

	CONTROL *Writing-ELA* Cluster Schools	TREATMENT *Writing-ELA* Cluster Schools	CONTROL *Math* Cluster Schools	TREATMENT *Math* Cluster Schools	CONTROL *WL* Cluster Schools	TREATMENT *WL* Cluster Schools
Grade 3 (baseline)	H/A/L	H/AL	H/A/L	H/A/L	HA/AL	H/Ad/L
Grade 4 (PAIR yr1)	HA/L	HA/AL	H/A/L	H/A/L	HA/L	HA/AL
Grade 5 (PAIR yr2)	HA/L	HA/AL	HA/L	HA/L	HA/AL	HA/AL
Grade 6 (PAIR yr3)	HA/L	HAL	H/A/L	HA/AL	HAL	HA/AL

2G *Summary, Emerging Themes*: The PAIR Cluster School data analysis reaffirm that the PAIR Treatment Schools academically outperform and more consistently narrow the academic achievement gap for previous Low rated students in comparison with Control Schools regardless of their designation as an ELA-Writing, Math, or World Language Magnet School. In addition, it appears that the PAIR World Language Treatment Schools – those schools that focus on both arts and arts integration teaching strategies to address their particular academic focus on multiple languages and social studies - outperform all other cluster school types in standardized test scores thereby demonstrating the most compelling school profile for boosting academic achievement while reducing the learning gap for low achieving students.

A research question stated in Part 1 of this report – Do teachers who demonstrate high levels of committed participation in the PAIR project positively influence student achievement? – will be explored vigorously through further analyses of student arts integration learning assessments and the patterns of statistical connections that will offer evidence of causal links between teacher professional development variable and student outcomes in the third and final sections of this comprehensive report.

* * *

PARTNERSHIPS IN ARTS INTEGRATION RESEARCH

(PAIR) FINAL COMPREHSENSIVE REPORT

Part 3: Analysis of PAIR Student Arts Integration Assessments and their Intersections with Teacher and Student Performance Outcomes

PAIR Student Learning and Survey Results and a Comprehensive Statistical Analysis of their Links with Teacher Professional Development and Student Standardized Test Outcomes by all Demographic Factors (2007-2010)

May 17, 2012
(December 14, 2012 Update)

Analyzed and Written By
Dr. Lawrence Scripp, Principal Investigator
With contributions from
Fred Sienkiewicz, Research Associate
Research Center, Center for Music and Arts in Education, Inc.
New York 10023
Professor, Graduate and Undergraduate Theoretical Studies
Chair, Music-in-Education, New England Conservatory
larry.scripp@necmusic.edu
617 585 1364

with Special Thanks for the contributions from Laura Tan Paradis, CAPE Research Associate and to the research staff at the Center for Music-in-Education:
Randy Wong, Justin Stanley, Helen Liu Wong, Kristen Dirmeier Sienkiewicz.

The PAIR project was funded through a grant to Chicago Public Schools from the US Department of Education's AEMDD Program. Chicago Arts Partnerships in Education (CAPE) contracted with Chicago Public Schools for this project. CAPE then subcontracted with Dr. Larry Scripp, Center for Music in Education, Inc., to serve as a co-principal investigator for the project in collaboration with Dr. Gail Burnaford, Principal Investigator, Florida Atlantic University.

PAIR Final Report PART 3: An Analysis of PAIR Student Arts Integration Assessments and their Intersections With Teacher and Student Performance Outcomes

Introduction to the Seven Sections of Part 3

While the previous two parts of PAIR report focused entirely on the impact of PAIR on teacher professional development and on student standardized academic test results, Part 3 of the report is organized into seven sections that present the analysis of multiple student arts integration learning assessment results and the intersection among teacher-student outcome variables by the final year of the project. The results are reported in seven different sections, each featuring its own table of contents, list of figures and tables, and an appendix:

A. Snapshots of Arts Integration (SAIL) Interview Response Ratings analyzed for control-treatment and within-treatment school differences in students' understanding of arts integration processes and connections;
B. PAIR Student Survey Responses analyzed for control-treatment school differences in the perception of arts integration practices in their classrooms and control-treatment school differences in the presence of classroom culture practices most highly associated with PAIR professional development goals and outcomes;
C. PAIR Partnership Arts Integration Learning (PAIL) Student Work Samples analyzed for qualitative differences among within PAIR treatment school classroom practices and in relation to the documentation and assessment goals for the PAIR project;
D. PAIR Portfolio Conference Performance Assessments of teacher verbal reflections and student individual and group performance assessments analyzed for qualitative differences in PAIR treatment school PAIR student work and portfolio conference performance assessments.
E. PAIR Portfolio Conference Performance Assessments of student individual and group performance assessment data analyzed statistically for their relationship to SAIL assessments, PAIL classroom ratings, and teacher portfolio conference performance data.
F. PAIR Treatment School Teacher-Student Outcome Intersections analyzed for statistically significant degrees of association between teacher professional development variables analyzed and student learning outcome data.

G. PAIR Treatment School Summary Charts, Conclusions, Implications for Future Research. This chapter reviews the sequence of PAIR data collection and analyses as a six step process for determining possible causal links among the eleven factors of the PAIR program outcomes in this report:

Step 1: Determine quantified measures of individual teacher engagement in PAIR professional development processes and events over time and measure their possible influence on teacher and student performance factors;

Step 2: Rate the relative level of teacher attitudes about, and reflections on, their professional development experiences in PAIR as revealed by teacher surveys and exit slips after professional development sessions and determine its possible influence on teacher and student performance factors;

Step 3: Rate the relative quality and sophistication of teacher performance outcomes as revealed throughout their response to the PAIR portfolio conference protocol and determine its possible influence on teacher and student performance factors;

Step 4: Determine individual student awareness of change in their learning environment resulting from PAIR as revealed by survey response ratings and the possible influence on student performance factors;

Step 5: Student arts and arts integration learning outcomes as revealed by PAIR interviews and portfolio conference performance assessments and their possible influence on academic student performance factors;

Step 6: Student academic performance as measured by state standardized test results.

* * *

PARTNERSHIPS IN ARTS INTEGRATION RESEARCH (PAIR) FINAL COMPREHSENSIVE REPORT

Part 3: Analysis of PAIR Student Arts Integration Assessments and their Intersections With Teacher and Student Performance Outcomes

May 17, 2012

(Updated December, 2012)

Analyzed and Written By

Dr. Lawrence Scripp, Principal Investigator and Associates
Research Center, Center for Music and Arts in Education, Inc.
New York 10023
Professor, Graduate and Undergraduate Theoretical Studies
Chair, Music-in-Education, New England Conservatory
larry.scripp@necmusic.edu
617 585 1364

3A: Analysis of Snapshots of Arts Integration (SAIL) Interview Response Ratings

(December 16, 2012 Update)

3A: Table of Contents

Part 3A: Analysis of PAIR Student Arts Integration Assessments and their Intersections With Teacher and Student Performance Outcomes	1
3A: Analysis of Snapshots of Arts Integration (SAIL) Interview Response Ratings	2
3A: Table of Contents	2
3A: List of Tables, Figures and Appendices	3
3A: Introduction	4
Purpose of the SAIL Interview Assessments	4
Description of the SAIL Interview Protocols	5
Ensuring the validity and reliability of SAIL data collection and analysis	6
Rating SAIL student responses for levels of articulation and perspective on their understanding of arts and arts integration learning	7
3A-1: Control-Treatment and Within Treatment Comparisons of PAIR Final Year SAIL Student Response Ratings	9
3A-2: Control-Treatment Comparisons of PAIR Final Year SAIL Student Response Ratings According to Pre-Classified HAL Academic Achievement Ratings	13
3A-3: PAIR Final Year SAIL Student Response Ratings Analyzed According to Pre-Designated HAL Academic Achievement Classifications and by Cluster School Types	16
3A: Summary, Emerging Themes	18

3A: List of Tables, Figures and Appendices

3A Table 1:	SAIL Interview and Performance Task Protocol (Writing Magnet School Example)	5
3A Table 2:	PAIR Snapshots of Arts/Arts Integration Learning Scoring Rubric Categories and Sample Student Response Indicators	8
3A-1 Figure 1:	Control-Treatment Comparisons of Grade 6 SAIL Interview Average Response Ratings by School Focus and School Cluster School Type	10
3A-1 Figure 2:	Statistically Significant Comparisons of SAIL Interview Averaged Response by School Focus Type (Arts vs. Academic)	11
3A-1 Figure 3:	Control-Treatment Schools Comparisons of Grade 6 Snapshot of Arts Integration Learning (SAIL) Interview Average Response by School Focus (Arts vs. Academic) breakdown by Question Categories	12
3A-2 Figure 1:	Fit of 2010 ISAT Averaged Math and Reading Scores by SAIL Interview Task Ratings from all students in PAIR Control and Treatment Schools.	14
3A-2 Figure 2-3:	Comparative Fit of 2010 ISAT Averaged Math and Reading Scores by SAIL Interview Task Ratings between students in PAIR Control and Treatment Schools	14
3A-2 Figure 4:	Scatterplot displays depicting the relationship between ISAT and SAIL Responses by Control-Treatment Arts vs. Non Arts Focus Schools	15
3A-2 Figure 5:	Degree of Correlation Between Student SAIL and ISAT Academic Ratings	16
3A-3 Figure 1:	Control-Treatment Comparison of SAIL Ratings by HAL Academic Designations	17
3A-3 Figure 2:	SAIL Mean Score Bar Graph Comparison of pre-designated HAL Student Cohorts by World Language Control and Treatment Student Cohorts	18
3A Appendix 1:	ELA Writing-Arts Cluster School SAIL Protocols Grade 5-6, 2009-2010	21
3A Appendix 2:	ELA Math-Arts Cluster School SAIL Protocols Grade 5-6, 2009-2010	23
3A Appendix 3:	World Languages & Cultures-Arts Cluster School SAIL Protocols Grade 5-6, 2009-2010	25
3A Appendix 4:	PAIR SAIL Interviewer Guidelines	27
3A Appendix 5:	PAIR SAIL Interview Scoring Rubric	28

3A: Introduction

Purpose of the SAIL Interview Assessments

The primary purpose of the individual PAIR SAIL interviews is to provide an authentic assessment vehicle for rating young children's level of understanding of arts learning and arts integration learning processes in both control and treatment schools. The validity of the analysis of SAIL protocol responses was ensured by asking all students (a) to describe and reflect on their personal experience and understanding of learning processes in several art forms that occurred in their classroom and then (b) to speculate on learning in the arts that had been or could be connected with learning in various academic subjects. The reliability of the analysis was ensured by a defined protocol conducted by an outside interviewer, video documentation and written transcription of each interview session, and the development of an outside scoring team trained to rate each child's level of response according to a common scoring rubric.

The ratings of the SAIL responses are intended to provide control-treatment school comparisons of student level of sophistication that could then validate the impact of the PAIR program in the treatment school classrooms. The SAIL assessments are intended also to provide evidence of differences in levels of meta-cognitive understanding of arts and arts integration learning process depending on the focus and mission of the various participating magnet school cohorts and depending on the high, average, or low academic level rating assigned to each student prior to the beginning of the PAIR project.

Description of the SAIL Interview Protocols

The SAIL interview and performance assessment protocols, developed by Dr. Larry Scripp in collaboration with his research colleagues at the CMIE Research Center and CAPE research staff led by Laura Tan Paradis, is designed to determine levels of student understanding of arts and arts integration learning in both PAIR control and treatment schools.

The SAIL Interviews and Performance Assessment Tasks were customized to address the particular art forms featured at each PAIR Cluster School site [see 3A: APPENDICES 1-3] and administered by research assistants trained to follow the questions and tasks on the protocol and provide each student an equal opportunity to elaborate his/her response. These protocols provided opportunities for students

in the project to demonstrate their understanding of the following seven topics (in this case customized to the Writing Magnet schools). The question topics originated from a previous CAPE research project, in which artists contributed to discussions on what was vital to practicing their artistic discipline. These topics formed the basis for the SAIL interview and Performance Task protocol. Questions were purposely designed to be appropriate for both control and treatment schools:

3A Table 1: SAIL Interview and Performance Task Protocol (Writing Magnet School Example[1])

Topic 1: Describe and compare the philosophy of two art forms and one primary academic discipline focus at your school.
a. What is music? Can you give me an example?
b. What is theater? Example?
c. What is writing? Example?
d. How are music, theater, and writing similar to or different from one another?

Topic 2 Describe and compare the process of two art forms and one primary academic discipline focus at your school.
a. How do you make music? Can you give me an example?
b. How do you perform or act out something? Example?
c. How do you write a story? Example?
d. How are making music, acting, and writing a story similar to or different to each other?

Topic 3 Describe and compare the concept of a mistake of two art forms and one primary academic discipline focus at your school.
a. What is a mistake in music? What do you do when you make a mistake while playing music?
b. What is a mistake in drama performance? What do you do if you make a mistake while performing?
c. What is a mistake in writing? What do you do when you make a mistake?
d. How is it similar or different when you make a mistake in music, when performing, or when writing?

Topic 4 Describe and compare the skills employed in two art forms and one primary academic discipline focus at your school):
a. What skills do you use to make music? Example? How?
b. What skills do you use to act? Example? How?
c. What skills to you use to write? Example? How?
d. How are the skills you use to make music, act and write similar to or different from one another?

Topic 5 Describe and compare the ability to create meaning in two art forms and one primary academic discipline focus at your school.
a. How do you tell a story through music? Can you give me an example?
b. How do you tell a story through drama performance? Example?
c. How do you tell a story using writing? Example?
d. How are telling a story through music, performance, or writing similar to or different from each other?

(Table continued on next page)

[1] The SAIL protocol differs according to the PAIR magnet school designation in questions (a) and (b); according to the CAPE teaching artist disciplin3es in questions (c); in terms of particular academic focus (math, writing, or world languages) in question (d); and according to the arts integration among the disciplines represented in questions (a), (b), and (c).

Topic 6 Describe and compare the ability to express yourself in two art forms and one primary academic discipline focus at your school s (show) through music?

a. What can you express through a drama performance? How can you express things in drama?
b. What can you express through music? How can you express things in music?
c. What can you express through writing? How can you express things in writing?
d. How are the things that you can express in music, theater, and writing the same or different?

Topic 7 Describe and compare the ability to use your imagination in two art forms and one primary academic discipline focus at your school.

a. Where do your ideas come from when you are making music? Example?
b. Where do your ideas come from when you are performing drama? Example?
c. Where do your ideas come from when you are writing a story? Example?
d. How do your ideas come from the same or different places when making music, performing, and writing a story?

Performance Assessment Tasks:

- Please give them the story-paragraph to read silently and then ask: What is in this story? Now, come up with a different story, one that you make up on your own, that uses the same elements as this story? [i.e. themes, things, characters, concepts, places, etc.] whatever the student answered for the first part.]

- Play music, and when finished ask: What do you hear in this music? Invent a story that explains what happens or what you hear in this music and tell me about it.

The performance assessment tasks administered during the final portion of the SAIL interview sessions proved to be particularly useful opportunities for students to *demonstrate* their ability to express, and not just talk about, their understanding of the seven aspects art forms, arts skills and artistic processes introduced in the interview portion of the SAIL protocol.

Ensuring the validity and reliability of SAIL data collection and analysis

The detailed SAIL Interviewer Guidelines [see 3A APPENDIX 4] provided a complete script meant to ensure uniformity and equal opportunity for every treatment and control school student in the central longitudinal cohort to respond to the interview questions and tasks. The guidelines include a student-friendly description of the PAIR research questions and scripted prompts for interviewer to elicit multiple appropriate responses and a tremendous emphasis on providing ample opportunity for each student to elaborate extensively on their interview responses. Even if it meant expanding the time or scope of the interview a bit, the most important underlying principle for the interview facilitators was to elicit the most thoughtful responses possible from every student without giving the impression that the interviewer was looking for a particular 'correct' response. Although there was the expected amount of attrition in the entire student longitudinal sample, there were not enough differences in the treatment and control school sample size by the end of the project to warrant corrective statistical procedures.

All SAIL interviews were recorded with digital video cameras, checked for uniformity and fealty to the 'spirit and letter' of the administrative aspects of the protocol, and then later rated by teams of researchers not having direct experience with the PAIR project.

Rating SAIL student responses for levels of articulation and perspective on their understanding of arts and arts integration learning

The rating of student responses was achieved through the creation and application of the SAIL Interview Scoring Rubric [see 3A Table 2 below]. The table below shows a broad outline of five categorically different levels of interview/task response. This scoring rubric provides descriptors of levels of complexity of student responses and provides examples of student response that served as anchors for establishing inter-rater reliability throughout the data coding process.

At first, videos were scored by pairs of raters until near 100% percent reliability was achieved with the five categorical scores (0, 1, 2, 3, 4) and over 90% inter-rater reliability on ratings within each categorical level (0.0-0.5, 1.0-1.5, 2.0-2.5, 3.0-3.5, 4.0-4.5). All team ratings that did not achieve reliability within .5 of each other were averaged in the dataset. Later on, when a scorer expressed uncertainty of his or her rating, another scorer was brought into the process and the two scores were averaged. Eventually this process resulted in statistically reliable 10-point scaled student response ratings that provided enough statistical power to comparing these data with other factors in the PAIR study.

(3A Table 2 on next page)

3A Table 2: PAIR Snapshots of Arts/Arts Integration Learning Scoring Rubric Categories and Sample Student Response Indicators

Categories of Response	Sample Anchor Statements
Level 0: **No Relevant Response** Irrelevant or indiscernible response; silence;	Answers, "Nothing", "I don't remember", "I don't know," skips question; doesn't understand or respond to the question because of language problems; unintelligible mumble; OR Answer is not relevant/does not address the question.
Level 1: **Single Dimensional Responses** Concrete, un-detailed response. Generic statements, singular perspective. Unspecific, unfocused, diffused. No elaboration, no detail, no personal specifics or procedural relationships. Lists undifferentiated elements. "One dimensional thinking"	"We made up stories." "I liked making up songs" "We danced together." "I drew pictures of my family." "We played drums with Charlie. It was fun."
Level 2: **i.Multiple Single Dimensions** Concrete connections, some occasional detail, some elaboration, or emerging specificity; Some coordination of elements, like a clearly ordered procedure. Specific personal insight.	"We acted out stories from books, then we acted out our own stories." "We made up songs for our own country and sang them with bells." "I drew picture of food I like. I liked different fruits than other kids." We danced in different ways and we had to keep the beats." "I made drawings of buildings and then we had to make the buildings with paper." "We made up words and then we made up beats to go with them."
Level 3: **Coordination of Dimensions of Understanding** Detailed descriptive relationships. Often provides elaborative detailed statements. Evidence of higher-order relational thinking, including elements of *inter*-personal insight and purpose, artistic aesthetic, and/or historical references.	"We told stories that had beginning, middle, and end. We were all different characters and we had a problem to fix. My story had a surprise at the end." "I make a drawing of all my family and me. Then we all had to draw a new nice place for our families to live with other families. We had cards with our pictures on them so others would know who we are." 'We had to make draw buildings with the numbers on them so we could build a building with the same shape, but much bigger." "Our song expressed the feeling of our new country, so we all had to like the notes and the beats of the song and sing it together." "We made up our own beats for the characters in the story, and then the beats would change if some got mad or sad." "We would all dance different motions together but it had to be a fraction too. We counted the beats so the different motion had the right fraction."
Level 4: **Systemic Understanding** Substantial detail and specificity. Causal statements. Compare and contrast relationships. Critical perspective, highly complex, multiple relationships.	"Story telling is better when we acted them out cause you can see how all the characters move and talk and argue with each other. My story got better because we had fun making the story funnier when we did it for the class.' "When we did the dance it was really math, too. We had to count. We had to get the fractions right. It had to be right so everyone could do it together." "Our drawings are art, but they are math too, because all the numbers add together and it has to look good, too. Sometimes we didn't do the numbers right and it didn't look right and we had to fix it." "Our song expressed the feelings of the words and told about what our country is. The most important words got the highest notes so everyone would know what our country stands for. And it wasn't done until we could all sing together and that was hard to do."

* * *

3A-1: Control-Treatment and Within Treatment Comparisons of PAIR Final Year SAIL Student Response Ratings

3A-1 *Inquiry Questions*: 1) To what extent are there significant differences in Snapshots of Arts Integration Learning (SAIL) interview response ratings between PAIR Control and Treatment Schools? 2) How do SAIL scores vary according to particular academic focus (Writing, Math, or World Languages & Culture)? 3) How do SAIL ratings vary among PAIR Treatment and Control Schools according to their primary focus on academics versus arts learning?

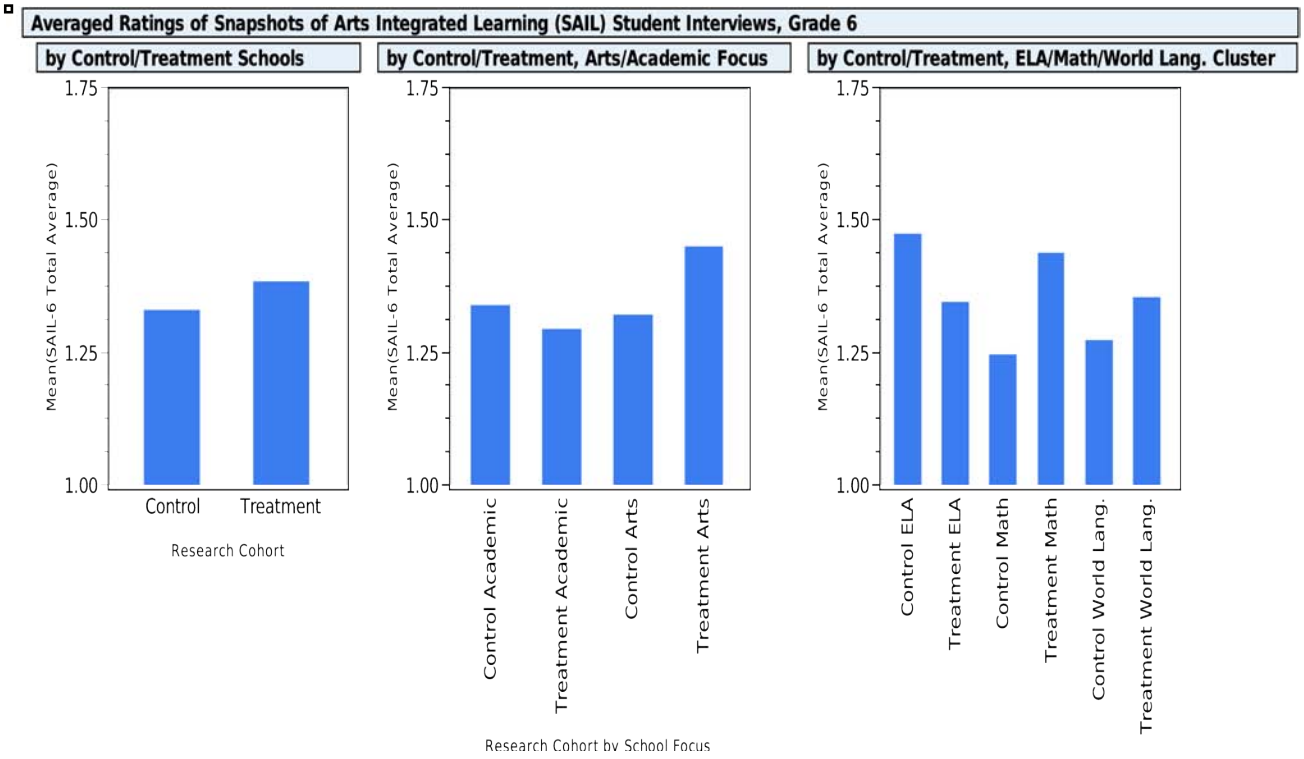
3A-1 Figures 1-3 *Data Display Indications*: First, the bar chart below on the left indicates that, in the final two years of the project, pre-post comparisons of levels of sophistication in SAIL student interview responses reveal that the *Treatment schools overall SAIL scores were higher than Treatment schools in grade five and increasingly so in grade 6* (though only attaining a p value of .09).

In the far right bar chart below it appears that the control-treatment comparisons of SAIL ratings by academic focus school PAIR Magnet CLUSTER School type suggest that only the Treatment World Language Treatment schools are outperforming their 'paired' control school cohort. However, this finding may be significant because the World Language School cohort proved to be, as demonstrated in previous analysis of ISAT test scores, also the school cohort with the highest increase in academic achievement.

From a third perspective, *there is strong evidence for overall highly significant differences between the Arts Focus Treatment school cohort compared to the Academic Focus Treatment and to both types of Control School cohorts* (center bar chart). Thus, we can conclude that *there is much stronger evidence for the consistent link between the SAIL assessments of students' understanding of arts and arts integration concepts in the Arts Focused Schools that explicitly developed PAIR arts integration practices without a particular attention given to one particular strand of academic achievement.*

(3A-1 Figure 1 on next page)

3A-1 Figure 1: **Control-Treatment Comparisons of Grade 6 SAIL Interview Average Response Ratings by School Focus and School Cluster School Type**

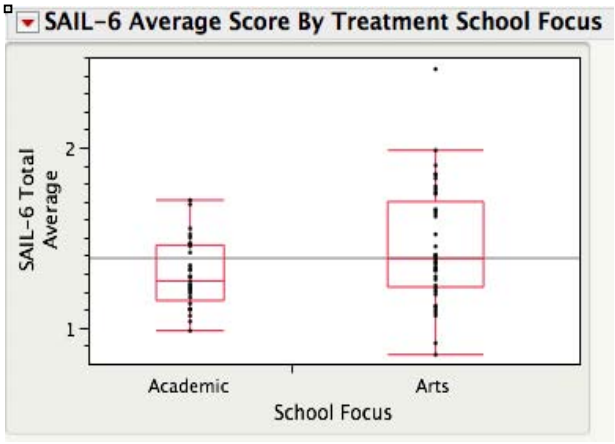


Taken together, the three data displays examined above suggest that, although PAIR Treatment student cohorts may be slightly outperforming the Control schools in SAIL interview rating overall*, it is principally the PAIR Treatment Arts Focus Schools* – and not the PAIR Treatment Academic Focus Schools – *that have demonstrated the highest levels of understanding of arts and arts integration in their classroom*. And – as seen previously in Part 2 of this report – *it is the Treatment Arts Focus school cohort that demonstrated the largest differential between treatment-control group academic scores by the final year of the project*.

3A-1 Figure 2 below displays statistical evidence for the supremacy of SAIL ratings in Treatment Arts Schools through the direct *within-treatment* school comparison between Arts and Academic schools. The box plot representation of SAIL student responses averaged across all categories of types of questions demonstrates statistically significant differences [ANOVA F Ratio 5.34, $p < .02$] between PAIR Treatment Academic and Arts Focus Schools not only in mean scores, but also in the range of student responses. The combined effect of these statistically different scores in the Treatment Schools

suggests that there is a sub-cohort of students who are highly articulate with respect to the SAIL tasks, a discrete cohort that does not exist in the Treatment Academic Focus School cohort. *In contrast to the PAIR Treatment schools, the Control school comparisons revealed no statistically significant differences between the Arts vs. Academic Focus Schools suggesting that the increased SAIL ratings in the PAIR Arts Focused Treatment Schools a principal outcome of the project.*

3A-1 Figure 2: **Statistically Significant Comparisons of SAIL Interview Averaged Responses by School Focus Type (Arts vs. Academic)**



Further analyses of student response to various categories of SAIL Interview question types [see 3A-1 Figure 3 below] indicate clear distinctions between the Arts and Academic Focused Treatment School responses to three groups of questions distributed throughout the SAIL protocol outlined earlier:

- Questions (a, b) focused on arts learning disciplines provided by the CAPE teaching artists,
- Questions (c) focused on particular academic disciplines featured in the PAIR project treatment according to the magnet school classification, and
- Questions (d) focused on arts integration learning processes relevant to the disciplines addressed previously in questions (a, b, c).

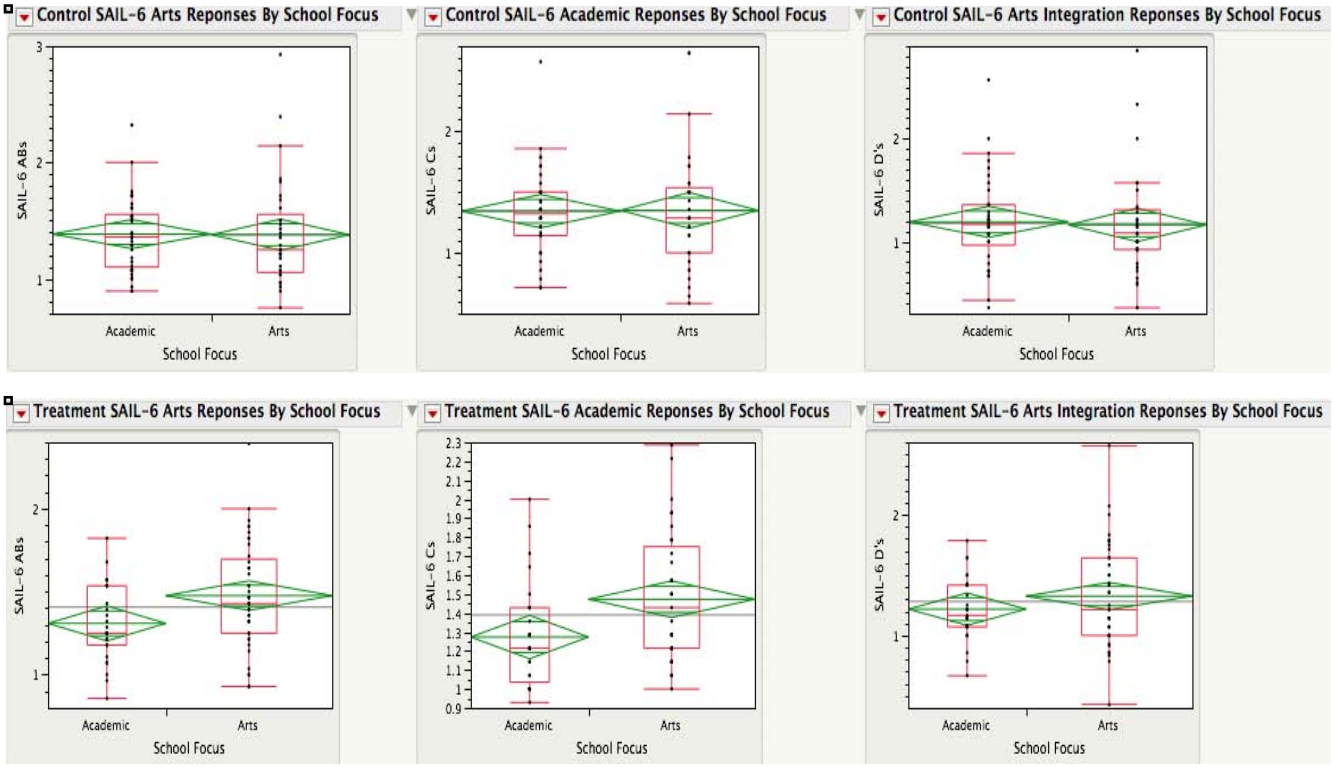
Two findings were obtained from the SAIL interview rating charts below. First, although there are no statistical differences between the Arts Focus and the Academic Focus Schools in the control schools student samples (see top row of charts below), results from the treatment schools indicate that the *Arts Focus School students score higher than do the Treatment Academic Focus School students on all three question types* (bottom row of charts). *Thus in Arts Focus PAIR treatment schools, arts*

integration units appear to heighten the ability to articulate learning processes in arts, academic and arts integration contexts.

Second, statistical significance is attained mainly with regard to questions focused on learning in the arts disciplines featured in each school's particular arts discipline focused PAIR curriculum [bottom left SAIL 'arts response' responses, ANOVA F Ratio 5.82 p <.02] and on learning in the PAIR prioritized academic discipline [bottom center SAIL 'academic discipline' responses, ANOVA F Ratio 7.06 p <.01). In contrast, results from the SAIL questions about the integration of arts and academic learning (right chart SAIL "arts integration" questions) are not significantly different within the Treatment Arts and Academic Focused school cohorts. This second finding suggests that *only in the Treatment PAIR Schools do the Arts Focused school outperform the Academic Focused students in the SAIL interviews in the arts discipline focused questions.*

3A-1 Figure 3: **Control-Treatment Schools Comparisons of Grade 6 Snapshot of Arts Integration Learning (SAIL) Interview Averaged Response by School Focus (Arts vs. Academic) breakdown by Question Categories**

(left = focus on PAIR arts disciplines, center = focus on prioritized academic learning according to Cluster School groups, right = focus on arts integrated learning)



3A-1 ***Summary, Emerging Themes:*** It appears that the PAIR program generally has impacted the Treatment Schools differently from the Control Schools in terms of their average level of response to SAIL interview tasks. Within-Treatment School profile analysis reveals that the Treatment Arts School student responses differ from the Treatment Academic Schools with regard to their understanding of arts or academic disciplines, but not in their grasp of arts integration teaching and learning practices. Looking across all three question areas of the SAIL protocol, *it appears that the PAIR Treatment Arts School students are able to articulate arts, academic, and arts integration questions equally well, whereas the Treatment Academic Focus School students are relatively more articulate about the process of arts integration across disciplines than they are discussing the processes and content of the discrete subject areas.*

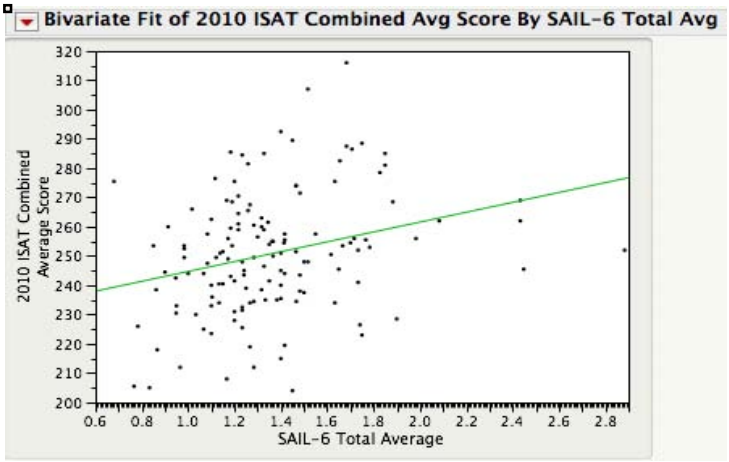
* * *

3A-2. Control-Treatment Comparisons of PAIR Final Year SAIL Student Response Ratings According to Pre-Classified HAL Academic Achievement Ratings

3A-2 ***Inquiry Questions*****:** To what extent do the SAIL scores predict academic achievement? To what extent are there significant differences in SAIL Interview Tasks between different levels of pre-designated HAL academic achievers in PAIR Control and Treatment Schools?

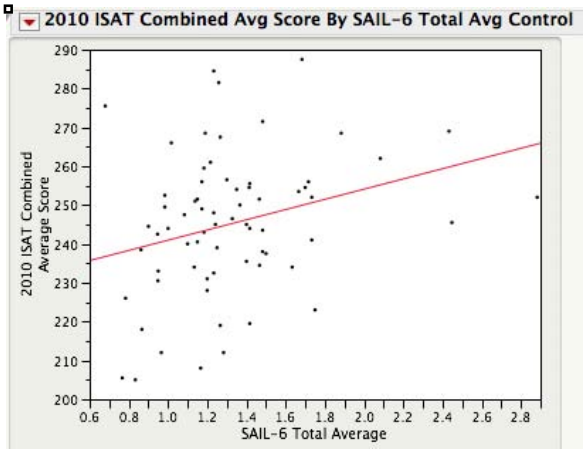
3A-2 Figures 1-3 Data Display Indications: The data displays below indicate that the SAIL Interview ratings do predict academic achievement in both Control and Treatment schools. *Thus individual student SAIL ratings constitute a central indicator of all students' understanding of arts and arts integration teaching and learning and a significant predictor of student standardized test regardless of receiving the PAIR arts integration program or not.* The figure below first displays the overall correlation between ISAT and SAIL scores including both control and treatment schools [ANOVA F Ratio 11.47 $p < .001$].

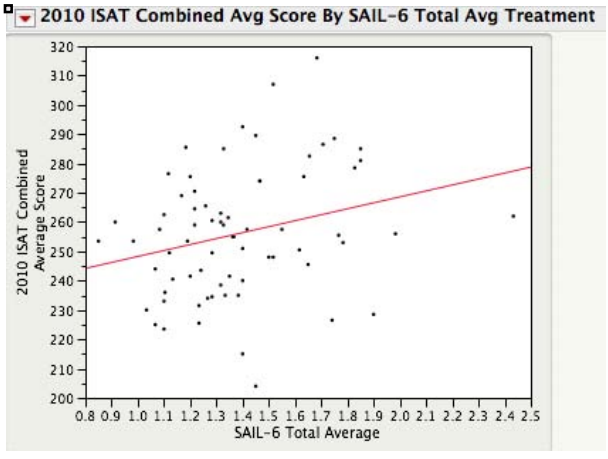
3A-2 Figure 1: Fit of 2010 ISAT Averaged Math and Reading Scores by SAIL Interview Task Ratings from all students in PAIR Control and Treatment Schools



3A-2 Figures 2-3 directly below establish that the positive correlation index between ISAT scores and SAIL ratings is generally the same in both control [ANOVA F-Ratio 5.85, $p < .02$] and treatment schools [ANOVA F Ration 5.12 $p < .03$]. This finding validates the SAIL interview ratings as an important tool for measuring student understanding of arts and arts integration teaching and learning practices as a predictor of academic achievement in any public school.

3A-2 Figure 2-3: Comparative Fit of 2010 ISAT Averaged Math and Reading Scores by SAIL Interview Task Ratings between students in PAIR Control and Treatment Schools

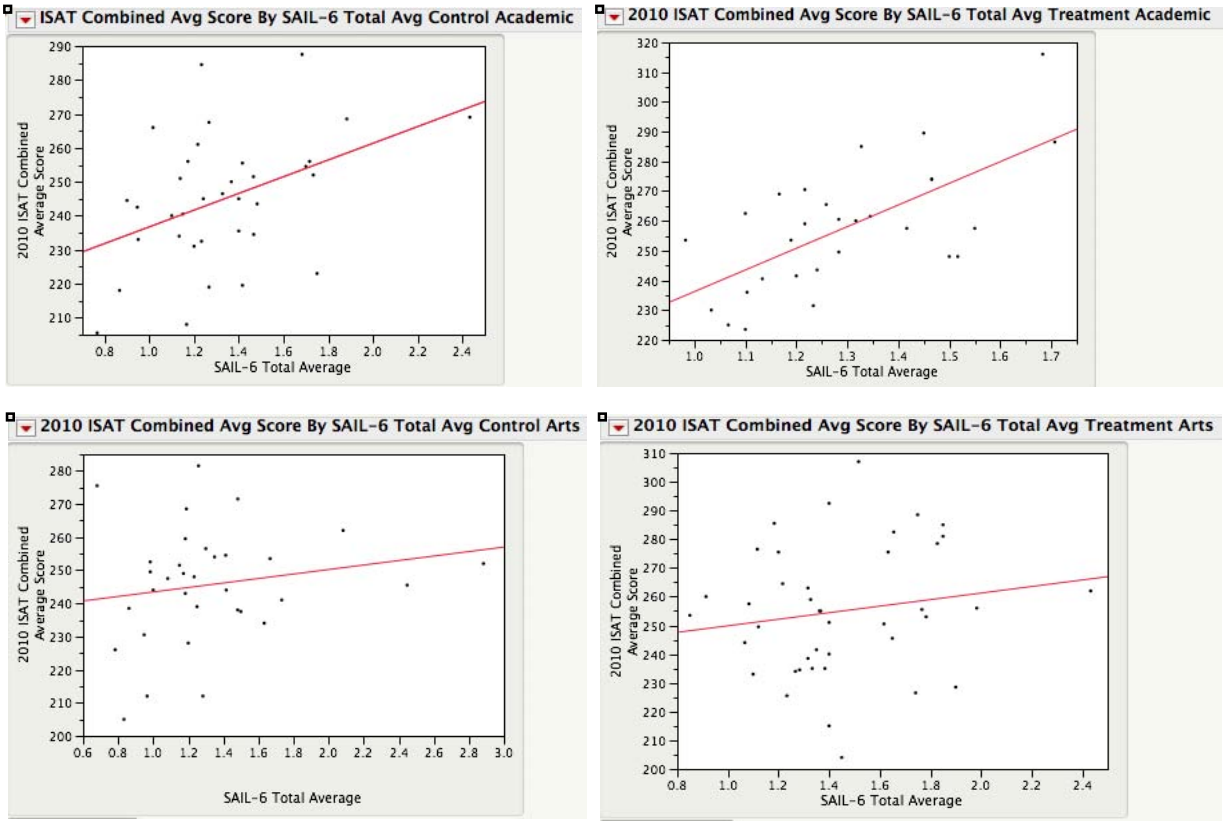




3A-2 Figure 4 *Data Display Indications*: The statistical relationship between SAIL protocol response and ISAT Standardized Test scores *within* the Focus School types is revealed by the scatterplots in figure below. The two upper row scatterplots show that the correlation between SAIL protocol ratings and ISAT test performance scores in both Treatment and Control *Academic Focus* School student cohorts are highly significant (p value .0001 and p value .01 respectively) compared to *Arts Focused* Schools, suggesting that performance on SAIL instruments has a higher predictive value for academic achievement in schools that first and foremost prioritize academic learning.

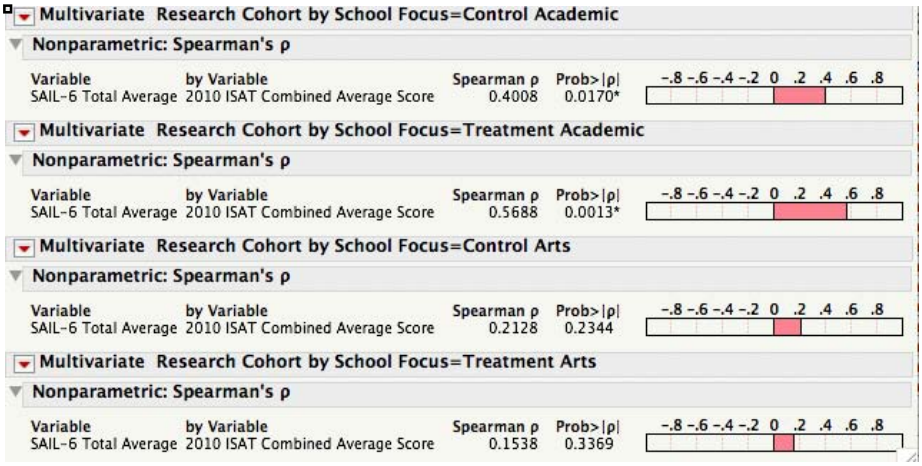
Conversely, the two lower row scatterplots [also in 3A2 Figure 4 below] indicate that there is no statistically significant association between test scores and SAIL scores in either of the control or treatment Arts Focus Schools, suggesting that the SAIL data may be generally less predictive of academic performance in schools that concentrate more on high standards of arts learning for all students.

3A-2 Figure 4: Scatterplot displays depicting the relationship between ISAT and SAIL Responses by Control-Treatment Arts vs. Non Arts Focus Schools



The correlation chart below confirms these findings through the calculation of highly significant correlation between arts and academic learning in the two Academic Priority Schools at the top of the table compared to the non-significant ratings in the magnet Arts Schools below.

3A-2 Figure 5: Degree of Correlation Between Student SAIL and ISAT Academic Ratings



* * *

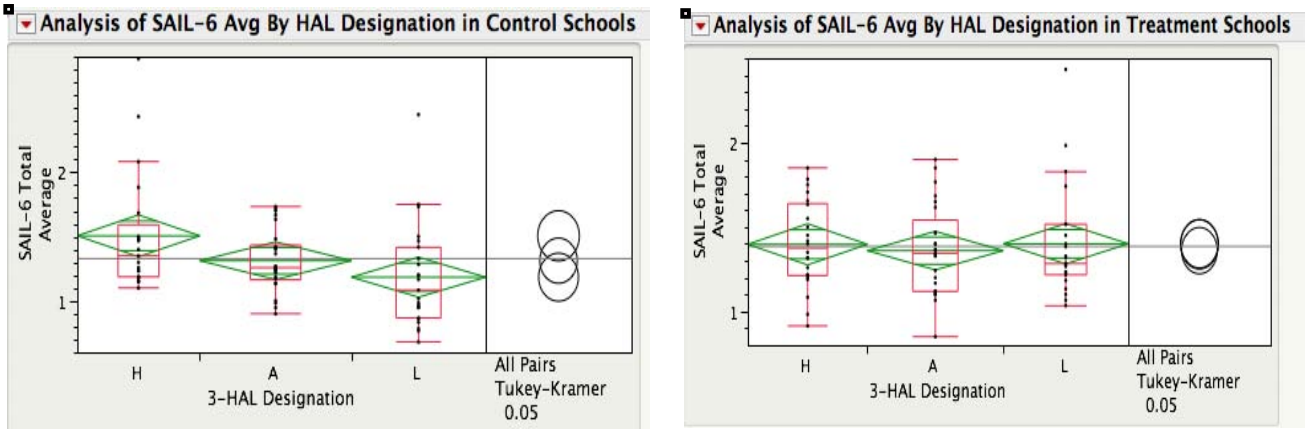
3A-3: PAIR Final Year SAIL Student Response Ratings Analyzed According to Pre-Designated HAL Academic Achievement Classifications and by Cluster School Types.

3A-3 *Inquiry Questions*: To what extent do the student HAL ratings predict averaged SAIL arts integration ratings? To what extent are there significant differences in SAIL Interview Tasks according to various types of Cluster Schools?

3A-3 Figures 1-2 *Data Display Indications*: The analyses of SAIL data indicate that the pre-designated classification of the HAL learners generally does not predict performance on the SAIL questions and tasks thus suggesting that the *SAIL ratings provide an alternative avenue for demonstrating understanding of arts and its connection to academic studies independent of prior academic achievement levels.* Box plot charts below in 3A4 Figure 1 provide evidence that *closing of the achievement gap also is detected principally in SAIL data in PAIR Treatment Schools*. Note that the SAIL means comparisons are significantly different in the Control Schools according to HAL

designation whereas in the Treatment Schools there is no significant difference among the HAL student cohorts by the end of the PAIR project. *The analyses of SAIL student learning outcomes supports previously reported findings that the PAIR Treatment Schools as a whole, and especially the Treatment Arts Focused Schools, provide an alternative and more equitable path to academic excellence through arts plus arts integration practices that the Control Arts Focus schools – and especially the Control Academic Focus Schools – do not*.

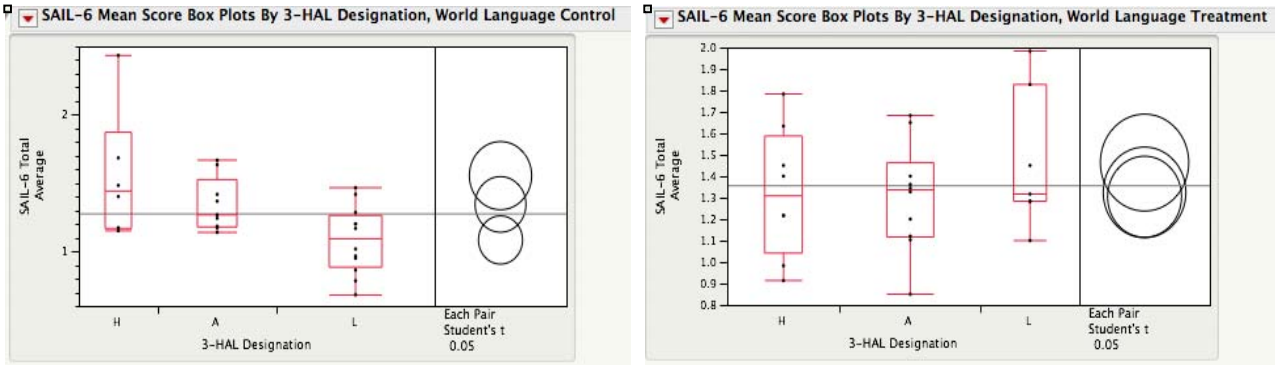
3A-3 Figure 1: Control-Treatment Comparisons of SAIL Ratings by HAL Academic Designations



In one small case study of PAIR Cluster Schools, for example, PAIR World Language Control-Treatment cohort comparisons reveal that, without the benefit of the PAIR program, measures of arts and arts integration understanding derived from the SAIL interview and performance assessment tasks in World Language CONTROL groups (see Figure 3A-3 left box chart) conform closely to the HAL profile gaps between high and low learners (High-Low comparisons $p < .004$).

In contrast to the CONTROL School SAIL HAL profile, *results from SAIL Treatment World Language school data* (Figure 3A-3 Figure 2 below, right box chart) *demonstrate the possibility of 'reversing' the expected achievement gap: that is, the pre-designated LOW ISAT academic achievement students in the treatment school outperformed the formerly designated HIGH and AVERAGE academic achievers in the control schools by the final year of the PAIR project*.

3A-3 Figure 2: SAIL Mean Score Bar Graph Comparison of pre-designated HAL Student Cohorts by World Language Control and Treatment Student Cohorts



3A *Summary, Emerging Themes*: Designed in collaboration with the CAPE staff and teaching artists, the SAIL interview and performance assessment tasks demonstrate valid and reliable measures of student understanding of creative and problem solving processes in multiple art forms and in their connection to other academic subject areas.

Results from SAIL assessments reviewed here suggest strongly that the *PAIR program substantially heightens understanding of arts and arts integration learning processes in Treatment Arts Focus Schools in comparison with matched control schools.* Reflecting on the arts and academic control-treatment comparisons as a whole, the analyses of the SAIL interviews data reveal that the treatment student cohorts, led by the Treatment Arts Focus students, emerge as the primary sample of students best exemplifying the effects of the PAIR program as measured both by the ISAT scores (as reported in Part 2) and by the SAIL Interview and Performance Assessment Task responses reported here. That the *Treatment Arts Focus School SAIL overall scores in particular outpace the other Treatment Academic and all other Control group schools suggests that PAIR program advances understanding of arts learning in ways that primarily academic or conventional arts focused school programs do not*.

Analysis of the SAIL ratings in relation to state standardized test scores suggests that the measures of student understanding of arts and arts integration may reveal unanticipated differences in the impact of PAIR on between the two types of PAIR Focus Schools. Whereas the Treatment Arts Focus school groups show that higher levels of understanding of arts and arts integration in SAIL assessments were

obtained across the broad spectrum of student academic classifications, the impact of the PAIR program in *Academic* Focus School SAIL results turn out to be more precisely associated with levels of academic achievement in both Control ($p < .01$) and other Treatment ($p < .0001$) schools. The data also show that the PAIR Treatment *Academic* Focus schools score relatively better on SAIL assessment items that are more focused on arts integration learning processes, a finding that suggests that PAIR arts integration practices are equally effective for both Treatment Arts and Treatment Academic Focus School when the emphasis in PAIR is weighted more toward arts integration processes than on arts learning in isolation from academic studies.

And furthermore, since previous chapters determined that PAIR treatment groups eventually did outpace the control groups in ISAT achievement scores by the final year of the project, we can now hypothesize that *the SAIL assessments function as a relatively independent and important intermediary[2] measure of student learning that has a direct and highly significant association with state standardized test scores.* Conversely, the Control Arts Focus groups, deprived of PAIR treatment program benefits, demonstrate (a) neither SAIL or ISAT scores equivalent to the Treatment Schools, (b) no indicators of success for pre-classified low-level learners compared to the Treatment Schools, and (c) no indication of significant correlation among the internal (SAIL) and external (ISAT) student learning outcomes that are strongly characteristic of the Treatment School statistical profiles.

In sum, students who demonstrate a more comprehensive and articulate view of arts learning and its integration across subject areas also are the students in schools most closely associated with the positive academic excellence and a culture of equity (as reported earlier) by the final year of the project. This result suggests that the SAIL assessment processes have captured critical evidence for explicit learning transfer effects in these schools through arts integration practices. These findings altogether suggest that 'arts plus arts integration' teaching and learning practices featured most clearly in the Treatment Schools may now be considered the optimal condition for students to benefit from both arts and academic based teaching and learning.

---

[2] The word "intermediary" is used to place the SAIL assessments in the sequence of associated factors of multivariate analysis that begins with A) PAIR school teachers receiving arts plus arts integration program services and professional development guidance (internal independent variable) that result in B) student ratings of understanding of arts plus arts integration teaching and learning SAIL assessment (internal dependent variable), that in turn can be linked with C) lISAT standardized student academic test scores (external dependent variable).

Additional HAL student achievement gap analysis shows that the SAIL arts learning assessments are correlated significantly overall with standardized ISAT test scores in Treatment schools despite the fact that there pre-designated levels of academic achievers no longer predict academic test scores by the end of the project. In one case study control-treatment group comparison between the SAIL ratings in World Language Arts and Academic Treatment Schools, the SAIL assessments appear to demonstrate a 'reversal' of expectations for formerly classified Low academic achievers. In PAIR World Language Schools Low-rated academic achievers sometime demonstrate a deeper understanding of both arts and arts integration learning than do the pre-designated Average or High achievers in the control World Language Schools. One possible explanation for PAIR's optimal impact on this particular cluster of Arts Focused and Academic Focused World Language case study schools is that arts integration may take on a deeper impact in schools committed to bi-lingual learning environment that may be more supportive of 'multiple literacy' integrated learning in both language and fine arts.

* * *

[3A Appendices 1-5 begin on the next page]

3A APPENDIX 1:
ELA Writing-Arts Cluster School SAIL Protocols
Grade 5-6, 2009-2010

Please ask the student for their name, teacher, and school before the questions begin. With each question guide the student to understanding the question and always ask the follow-up questions.

- Question 1 (Philosophy):

a. What is music? Can you give me an example?
b. What is performance or theater? Example?
c. What is writing? Example?
d. How are music, theater, and writing similar to or different from one another? Why?

- Question 2 (Process):

a. How do you make music? Can you give me an example?
b. How do you perform or act out something? Example?
c. How do you write a story? Example?
d. How are making music, acting, and writing a story similar or different to each other? Why?

- Question 3 (Judgment):

a. What is a mistake in music? What do you do when you make a mistake while playing music?
b. What is a mistake in drama performance? What do you do if you make a mistake while performing?
c. What is a mistake in writing? What do you do when you make a mistake?
d. How is it similar or different when you make a mistake in music, when performing, or when writing? Why?

- Question 4 (Concrete Experience):

a. What skills does a musician use? Example? How are these skills useful?
b. What skills does an actor use? Example? How?
c. What skills does a writer use? Example? How?
d. How are the skills used by a writer, a musician, and an actor similar or different from one another? Why?

- Question 5 (Concrete Experience):

a. How do you tell a story through music? Can you give me an example?
b. How do you tell a story through drama performance? Example?
c. How do you tell a story using writing? Example?
d. How are telling a story through music, performance, or writing similar or different from each other? Why?

- Question 6 (Imagination):

a. What can you express (show) through music? How?
b. What can you express through drama performance? How?
c. What can you express through writing? How?
d. How are the things that you can express in music, theater, and writing the same or different? Why?

- Question 7 (Imagination):

a. Where do your ideas come from when you are making music? Where else ... ?
b. Where do your ideas come from when you are performing drama? Example?
c. Where do your ideas come from when you are writing a story? Example?
d. Do your ideas come from the same or different places when making music, performing, and writing a story? Why?

- Please give them the story-paragraph to read silently and then ask: What is in this story? Now, come up with a different story, one that you make up on your own, that uses the same elements as this story? [i.e. themes, things, characters, concepts, places, etc... whatever the student answered for the first part.]

- Play music, and when finished ask: What do you hear in this music? Invent a story that explains what happens or what you hear in this music and tell me about it.

* * *

3A APPENDIX 2:
ELA Math-Arts Cluster School SAIL Protocols
Grade 5-6, 2009-2010

Please ask the student for their name, teacher, and school before the questions begin. With each question guide the student to understanding the question and always ask the follow-up questions.

- Question 1 (Knowledge):

a. What is a dance performance? Example? (if dance has no meaning to student, use just movement "performance")
b. What is visual art? Can you give an example?
c. What is math? Example?
d. How are dance, art, and math similar to or different from one another? Why?

- Question 2 (Process):

a. How do you make a dance performance? Can you give me an example?
b. How do you make visual art? Example?
c. How do you solve a math problem? How do you use math to solve problems outside of school? Example?
d. How are making a dance performance, making art, and solving a math problem similar or different? Why?

- Question 3 (Judgment):

a. What is a mistake in a dance performance? What would you do if you made a mistake in a dance performance?
b. What is a mistake in visual art? What to you do when you make a mistake in art?
c. What is a mistake in math? What do you do when you make a mistake?
d. How is it similar or different when you make a mistake in a dance performance, in art, or in math? Why?

- Question 4 (Personal experience):

a. What skills does a performer or dancer use? Can you give an example? How are these skills useful?
b. What skills does an visual artist use to make art? Example? How?
c. What skills does a mathematician use? Example? How?
d. How are the skills that a performer, artist, and mathematician use similar or different from one another? Why?

- Question 5 (Personal experience):

a. How do you tell a story through a dance performance? Example?
b. How do you tell a story through visual art? Example?
c. How do you use math to communicate ideas? Example?
d. How are telling a story / communicating ideas through dance, art, and math similar to or different from one another? Why?

- Question 6 (Imagination):

a. What can you express through a dance performance? Can you give me an example of how?
b. What can you express through visual art? Can you give me an example of how?
c. What can you express through math? Can you give me an example of how?
d. How are the things that you can express through math, performance and art similar or different from one another?

- Question 7 (Imagination):

a. Where do your ideas come from when you are creating a dance performance? Where else...?
b. Where do your ideas come from when you are making visual art?
c. Where do your ideas come from when you are working on a math problem?
d. Do you get your ideas for dance, art, and math from similar places or different places? Why?

- Show them the art image, and ask: Describe what you see in this image and what it means to you. Can you invent a story based on what you see and tell me about it?

- Show them the performance video, and ask: Please describe what you see in this video and what it means to you. Can you invent a story based on what you saw in this video and tell me about it?

3A APPENDIX 3:
World Languages & Cultures-Arts Cluster School SAIL Protocols
Grade 5-6, 2009-2010

Please ask the student for their name, teacher, and school before the questions begin. With each question guide the student to understanding the question and always ask the follow-up questions.

- Question 1 (Knowledge):

a. What is music? Can you give me an example?
b. What is visual art? Example?
c. What is language and what is culture? What would be an example of a person's language/culture?
d. How are music, art, and culture similar or different from one another? Why?

- Question 2 (Process):

a. How do you make music? Can you give me an example?
b. How do you make visual art? Example?
c. What are different ways that you (or someone else) show or express your culture? How does that show your culture?
d. How can music or art express a person's culture? Why?

- Question 3 (Judgment):

a. What is a mistake in music? What do you do when you make a mistake while playing music?
b. What is a mistake in visual art? What do you do when you make a mistake while making art?
c. What is a mistake when you are using a new language? What do you do when you make a mistake in a language?
d. How is it similar or different when you make a mistake in music, a mistake in art, or a mistake in language? Why?

- Question 4 (Concrete Experience):

a. What skills does a musician use to make music? Example? How are these skills useful?
b. What skills does a visual artist use to make art? Example? How?
c. What skills do you use to learn a language? Example? How?
d. How are the skills you use for making music, art, and learning language similar to or different from one another? Why?

- Question 5 (Concrete Experience):

a. How do you tell a story through music? Can you give an example?
b. How do you tell a story through visual art? Example?
c. How does a culture tell its stories? What other ways does a culture tell its stories?
d. How is it similar or different when you tell stories through music, art, and culture? Why?

- Question 6 (Imagination):

a. What can you express or show through music? Example?
b. What can you express or show through visual art? Example?
c. What can you express through language? Example?
d. How are the things you can express through music, art, and language the same or different?

- Question 7 (Imagination):

a. Where do your ideas come from when you are making music? Where else ?
b. Where do your ideas come from when you are making visual art? Where else?
c. Where do your ideas come from when you are telling a story? Where else?
d. Do you get your ideas for music, art, and storytelling from similar or different places? Why?

- Show them the art image, and ask: Describe what you see in this image and what it means to you. Can you invent a story based on what you see and tell me about it?

- Play music, and when finished ask: Please describe what you hear in this music and what it means to you. Can you invent a story based on what you hear in this music and tell me about it?

3A APPENDIX 4:
PAIR SAIL Interviewer Guidelines

INTRO Script:
Hi I'm _____ and I work with an arts education organization in Chicago. We support the arts in public schools. For these three years, we're basically interested in what 4th -6th graders think about the arts – including visual art, theater, performance, dance, and music – and some academic areas like math, language and literature.

So the kinds of questions I'll ask you will be about these topics. They're like "What is art? What is _____ (content area that school is working on)? How do you make art? How do know when you've finished learning (math, language and culture, story)?" These questions **have no right or wrong answers**. They **are also not trick questions**. So if you don't know the answer to a question, you can just tell me or ask me to clarify, or ask me to repeat it.

These questions are ways for us to understand how you think about these things. We're going to ask these for three years, and the idea is to see if your thinking about them changes over time.

Does this make sense? (answer any of their questions)

NOTE: If they ask why we're asking them these questions – just tell them we're trying to understand what they think about these topics, because most people don't ask 4th graders these questions, and they don't really know what 4th graders think. Don't mention the integration part – that we're interested in learning how one topic relates to another. It may lead them to answer the questions differently.

END:
Do you have any questions for me?

Two important principles in conducing the interviews:

Freedom to Interview
Our goal is to get each student to demonstrate as much as possible of their complete understanding of a topic or question. Because we are scoring on complexity of understanding and not "correct answers", you cannot give the student the "right answer" (false positive) by prompting or helping them understand. However, they can very easily fail to demonstrate what they are capable of (false negative) because the question is asked without follow-up or because they do not understand the meaning of the question. Please support the student in their responses and tailor your prompts to their individual needs.

It is important that if a student has a fundamental misunderstanding (thinks "making music" means playing a CD; doesn't understand culture; language issues etc.), you should offer support and correction to assist them in understanding the point of the question.

Encouraging Elaboration in Every Student
In each bulleted question (a, b, c, d) there are two essential parts. The question generates a first response. Every student should then be asked a follow-up question (How? Example? Why?) that prompts them for elaboration. If they answer the "stock" follow-up in their initial answer, please ask for a more complex follow-up. You can ask them for a second (contrasting) example, or ask them to compare their two examples, or improvise another follow-up based on their response to get a more detailed, personal, and complex answer.

3A APPENDIX 5:
PAIR SAIL Interview Scoring Rubric

	Example
Level 0: **No Relevant Response** Irrelevant or indiscernible response; silence;	Answers, "Nothing", "I don't remember", "I don't know," skips question; doesn't understand or respond to the question because of language problems; unintelligible mumble; OR Answer is not relevant/does not address the question.
Level 1: **Single Dimensional** Concrete, un-detailed response. Very generic statements, singular perspective. Unspecific, unfocused, diffused. No elaboration, no detail, no personal specifics or procedural relationships. Lists undifferentiated elements. "One dimensional thinking"	"We made up stories." I liked making up songs" "We danced together." "I drew pictures of my family." "We played drums with Charlie. It was fun."
i.Level 2: **ii.Multiple Single Dimensions** Concrete connections, some occasional detail, some elaboration, or emerging specificity; Some coordination of elements, like a clearly ordered procedure. Specific personal insight.	"We acted out stories from books, then we acted out our own stories." "We made up songs for our own country and sang them with bells." "I drew picture of food I like. I liked different fruits than other kids." We danced in different ways and we had to keep the beats." "I made drawings of buildings and then we had to make the buildings with paper." "We made up words and then we made up beats to go with them."
Level 3: **Coordination of Dimensions** Detailed descriptive relationships. Often provides elaborative detailed statements. Evidence of higher-order relational thinking, including elements of *inter*-personal insight and purpose, artistic aesthetic, and/or historical references.	"We told stories that had beginning, middle, and end. We were all different characters and we had a problem to fix. My story had a surprise at the end." "I make a drawing of all my family and me. Then we all had to draw a new nice place for our families to live with other families. We had cards with our pictures on them so others would know who we are." 'We had to make draw buildings with the numbers on them so we could build a building with the same shape, but much bigger." "Our song expressed the feeling of our new country, so we all had to like the notes and the beats of the song and sing it together." "We made up our own beats for the characters in the story, and then the beats would change if some got mad or sad." "We would all dance different motions together but it had to be a fraction too. We counted the beats so the different motion had the right fraction."
Level 4: **Systemic Understanding** Substantial detail and specificity. Causal statements. Compare and contrast relationships. Critical perspective, highly complex, multiple relationships.	"Story telling is better when we acted them out cause you can see how all the characters move and talk and argue with each other. My story got better because we had fun making the story funnier when we did it for the class.' "When we did the dance it was really math, too. We had to count. We had to get the fractions right. It had to be right so everyone could do it together." "Our drawings are art, but they are math too, because all the numbers add together and it has to look good, too. Sometimes we didn't do the numbers right and it didn't look right and we had to fix it." "Our song expressed the feelings of the words and told about what our country is. The most important words got the highest notes so everyone would know what our country stands for. And it wasn't done until we could all sing together and that was hard to do."

• * *

3B: PAIR Analysis of Student Survey Response Ratings.

3B: *Table of Contents*

3B: PAIR Student Survey Response Ratings analyzed for Control-Treatment and With-Treatment Schools differences in Students' Perception of Changes in Classroom Teaching Practices Relevant to High-Quality Arts Integration Instruction and Learning	1
3B: Table of Contents	1
3B: List of Tables, Figures and Appendices	1
3B: Introduction	2
Purpose of the PAIR Student Survey	2
Construction of the PAIR Student Survey	2
3B: Control-Treatment Comparisons of Grade 6 PAIR Student Survey Average Responses by School Focus (Arts vs. Non-Arts) and School Cluster (Writing, Math, or World Languages)	5
3B: Preliminary Summary, Emerging Themes	7
3B: Final Summary, Emerging Themes	10

3B: *List of Tables, Figures, and Appendices*

3B Table 1:	Key Categories of the PAIR Student Survey	4
3B Figure 1:	Control-Treatment Comparisons of Grade 6 PAIR Student Survey Average Responses by School Focus (Arts vs. Non-Arts) and School Cluster (Writing, Math, or World Languages)	7
3B Figure 2-3:	Control-Treatment Box Plot Comparisons of Grade 6 PAIR Student Survey Average Responses by School FOCUS	9
3B Figure 4:	Control-Treatment Box Plot Comparisons of Grade 6 PAIR Student Survey Average Responses by School Focus	9
3B Appendix 1:	ELA Writing-Arts Cluster School SAIL Protocols Grade 5-6, 2009-2010	11

3B: Introduction

Purpose of the PAIR Student Surveys

The purpose of the Student Survey data analysis was to test the assumption that the PAIR program results in a distinctly different classroom practices compared to control school classrooms. In order to substantiate this assumption, the analysis of PAIR survey responses needed to indicate that Treatment School students perceived increasingly higher levels of *arts integration* emphasis and classroom practices that included higher levels of PAIR program development factors, specifically *collaboration*, *reflection*, *arts learning presentation*s and *arts learning participation.* Hypothetically, it was assumed that the presence of these factors were necessary, though not sufficient, to support, optimize and sustain high-quality program development in their classroom.

Positive control-treatment comparisons of the Student Survey responses would also validate the role of the teaching artist and the effect of ongoing teacher professional development on classroom practices described in the PAIR Teacher Report (Part I of the PAIR report), and provide evidence of program outcomes in the Treatment School that could explain why the PAIR project may have been linked to positive external standardized test and internal SAIL student learning outcomes reported here.

Construction of the PAIR Student Survey

The PAIR Pre-Post Student Survey [3A APPENDIX 6] items were included to mirror to a certain extent the construction and content of the PAIR Year End Teacher Survey. In both cases the responders were asked to rate how often certain key elements of exemplary teaching practices were present in the classroom on an eight point scale ranging from Never (1) to Rarely (2-3) to Sometimes (4-5) to Often (6-7) to always (8). Thus control-treatment differences in student response could be used to validate the extent to which teachers succeeded in having students experience the benefits of the PAIR program development factors.

Statements rated by the students were coded according to key elements of the PAIR program from the teacher professional development perspective. The responses were then categorized as primarily measures of the students' perception of the following characteristics of their classroom culture: Collaboration (teacher and student), Reflective Practices, Focus on Arts Integration, Frustration with writing assignments, Opportunities to present arts learning work, and Engagement in arts learning outside of the classroom.

3B Table 1: Key Categories of the PAIR Student Survey

Key PAIR Program Development Factors	Sample Statements Linked to Each Program Development Factor to be Rated from 1-8
Collaboration [C]	• I have opportunities to teach and be taught by other students in class. [C] • There are assignments where students plan together in class. [C] • My classmates and I think of new ideas before and when we work together. [C, R] • My classroom teacher and my in-school art, music, performance or dance teacher work together in my class. [C] • My classroom teacher works with outside teaching artists, musicians, performers, or dancers who come into our classrooms to teach. [C, A
Reflective Practices [R]	• My teacher meets just with me to review my work when it is completed and/or when it is in progress. [R] • My teacher meets with my working group to review my work when it is completed and/or when it is in progress. [R] • I have the opportunity to write in class to reflect on or think about what I learned. [R]
Focus on Arts Integration [AI]	• Students in my school are learning about the arts, and also explore its connection to math, science, or world languages and cultures. {AI] • I use images, music, movement or dance to understand books, stories or vocabulary in class. [AI} • I use images, music, movement or dance to understand math concepts or math problems in class [AI} • I use images, music, movement or dance to understand different cultures and my own or my family identity in class. [AI} • My teacher explains and/or we discuss why we use images, music, movement or dance to understand books, stories, writing or vocabulary in class. [AI} • My teacher explains and/or we discuss why we use images, music, movement or dance to understand math concepts or math problems in class [AI} • My teacher explains and/or we discuss why we use images, music, movement or dance to understand different cultures and my own or my family identity in class. [AI}
Frustration with writing assignments [F]	• I get bored or frustrated writing in class, especially when I am asked to write long paragraphs. [F]
Opportunities to present student arts learning work [PA]	• Student work is publicly displayed and/or performed in the classroom. [PA] • Students in my class have opportunities to show their work in an exhibition, performance, or presentation for students and/or parents from the school. [PA] • Students in my class have opportunities to show their work in an exhibition, performance, or presentation to students from another school. [PA]
Engagement in arts learning outside the classroom [OA]	• I take part in art, music, theater or dance classes outside of school. [OA]

Ensuring the validity and reliability of PAIR Student Surveys data collection and analysis

The data were collected early and late in every year of PAIR Project implementation. Rated responses were coded and averaged according to their category so that the data could be used with sufficient statistical power to determine statistically significant control-treatment comparisons and explore the relationships between student survey responses and other student learning and teacher professional development outcomes in this project.

* * *

3B PAIR Student Survey Responses: Control-Treatment Analysis of Student Perceptions of Classroom Culture and Elements of Arts Integration Teaching and Learning Practices

3B *Inquiry Questions*: 1) To what extent are there significant differences in PAIR student survey response ratings between Control and Treatment Schools? 2) How do PAIR Student Survey scores vary according to particular academic focus (Writing, Math, or World Languages & Culture)? 3) How do PAIR Student Survey ratings vary among PAIR Treatment and Comparison Schools according to their primary focus on academics versus arts learning?

The Student Survey was developed piloted in the first two years of the project, and administered as pre-post tests in the last two years of the project. Over time, the statements became focused on determining the students' willingness to recognize increasingly specific factors of a high-quality PAIR classroom culture as it unfolded during the course of the academic year. Thus survey responses were analyzed for differences in pre-post ratings of these key factors. From the CAPE staff and teaching artist perspective, the strongest evidence that high quality factors of PAIR classroom practices were established from the student perspective would be if students become increasingly willing to rate the presence of Arts Integration factors more positively by the end of the school year.

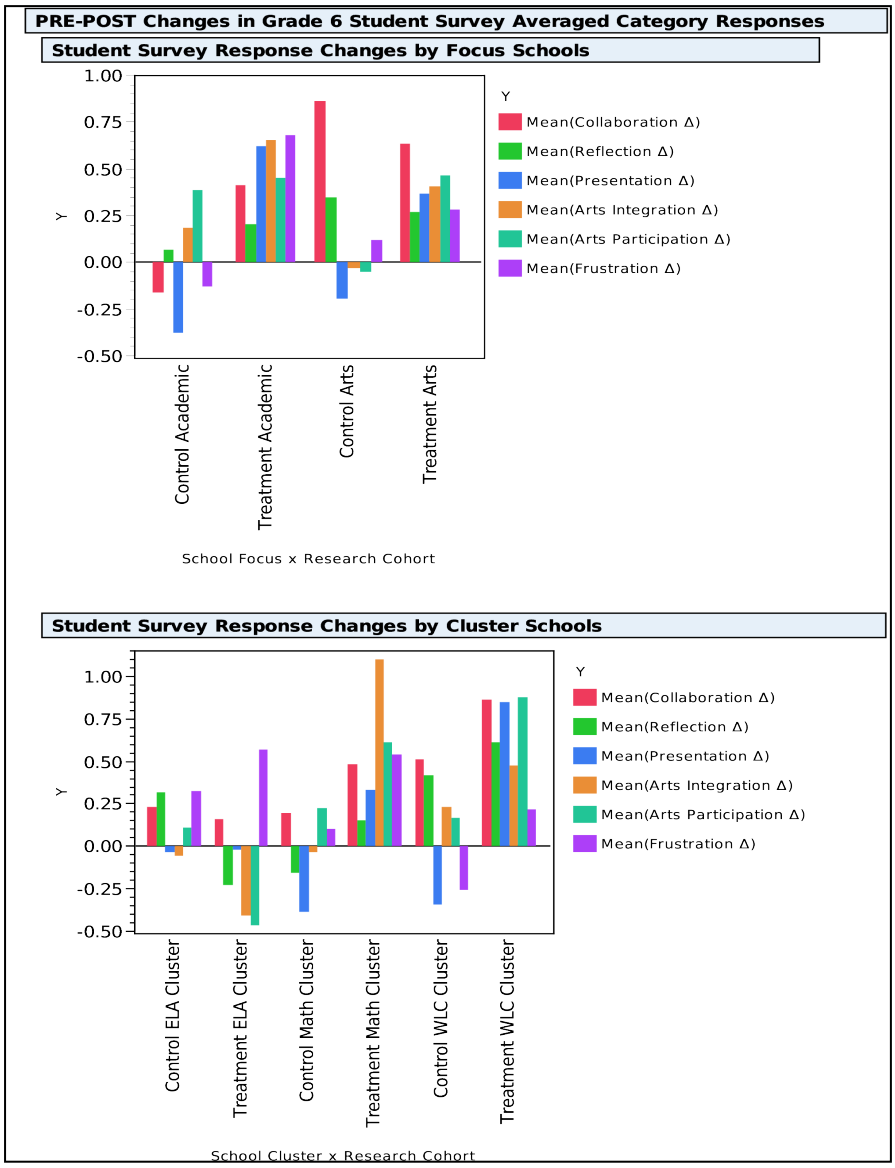
3B Figure 1 *Data Display Indications*: Students in treatment schools, with one exception, register far higher student perception of change in classroom culture than do students in the control schools.

In 3B Figure 1 below the top row bar chart display indicates clearly the difference between Control-Treatment comparisons between PAIR Arts and Academic Focus schools. Analyzed according to key features of the PAIR classroom culture that were explicitly supported by ongoing teacher professional development, students in Treatment schools show patterns of gains over time that contrast significantly with the control school student responses. With the exception of the highly positive response to the presence of 'collaboration' in the Control Arts Focus Schools and the starkly enthusiastic response to questions about 'participation in the arts' in the Control Academic Focus Schools [see top charts in 3B Figure 1]. *Treatment Arts and Academic Focus Schools register gain scores in every dimension of the program*, including student reflection, arts presentation, arts integration, and arts participation.

The bottom row chart indicate that the overall pattern of positive gain scores in Treatment Schools according to school cluster type, as a whole was particularly strong in Math and World Language Schools but did not happen in the Treatment ELA Schools.

(3B Figure 1 on next page)

3B Figure 1: **Control-Treatment Comparisons of Grade 6 PAIR Student Survey Average Responses by School Focus (Arts vs. Non-Arts) and School Cluster (Writing, Math, or World Languages)**



3B *Preliminary Summary, Emerging Themes*: The data display here supports strong evidence for the increasing presence of the PAIR program in Treatment Schools thus validating Control-Treatment differences between classroom cultures from the student perspective. The relatively high student ratings of collaboration, reflective practice and arts integration in the Treatment schools over time distinguishes the PAIR Treatment versus Control Schools in terms of classroom practices and culture.

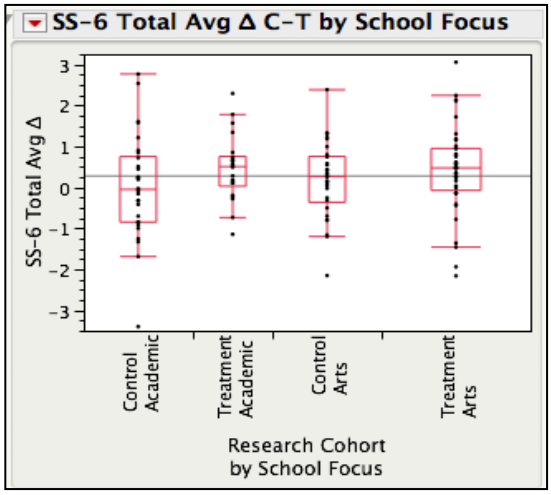
Differences Control Arts and Academic focused schools also provides evidence that could explain why the Academic Control Schools generally underperformed in most factors compared to the Arts Control Schools.

The case study school cluster analysis showed that student in one type of PAIR school were not inclined to notice increasing levels of arts integration classroom practices. The lone exception of the under-rated Treatment ELA Schools may be explained in part by the fact that expectations and the presence of arts integration at these schools initially may have been much higher from the beginning of the project compared to the Math and World Language Schools. The presence of the PAIR program may have only reinforced a pre-existing culture of arts integration embedded in the language arts (through illustration and dramatization) which would explain the strong presence of support in the post survey results that did not exceed initial expectations as seen in the pre-post comparisons. Thus it is reasonable to assume that student attitudes toward PAIR in schools that had already established as a part of the classroom culture would not as likely become more positive over time compared to the Math and World Culture schools where the perception of PAIR classroom culture factors increased significantly.

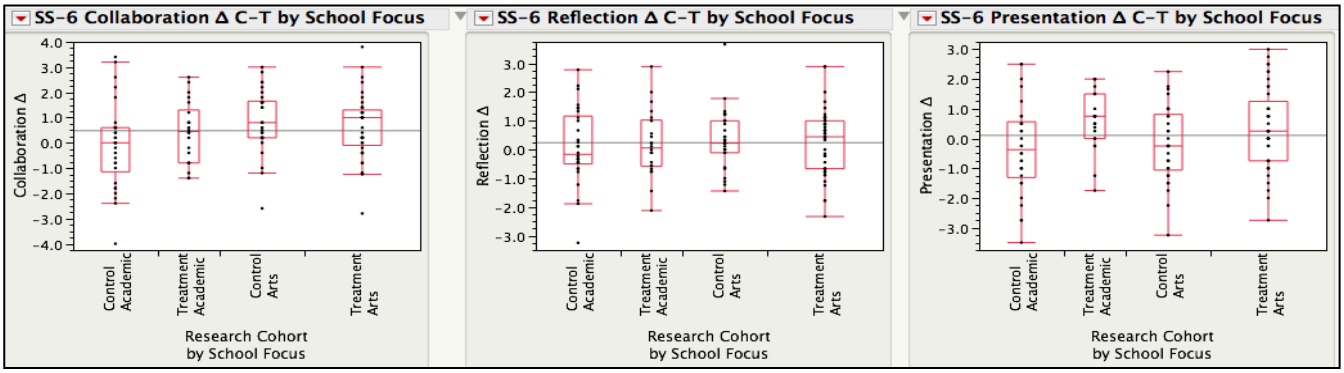
Another explanation, however, can be inferred from a parallel anomaly in the SAIL Interview and Performance Assessment results presented in the previous section. As previously reported, The SAIL student assessment averages in the ELA Treatment Schools did not exceed the levels of their counterpart Control ELA schools. In this case, students surveyed in the ELA Treatment Schools likewise did not express increased enthusiasm for features of arts integrated learning compared to *any* of the other school cohorts. These results suggest the possibility of a link between the PAIR students' perception of their Arts Integrated learning environment factors and the SAIL assessment ratings of arts and arts integrated learning.

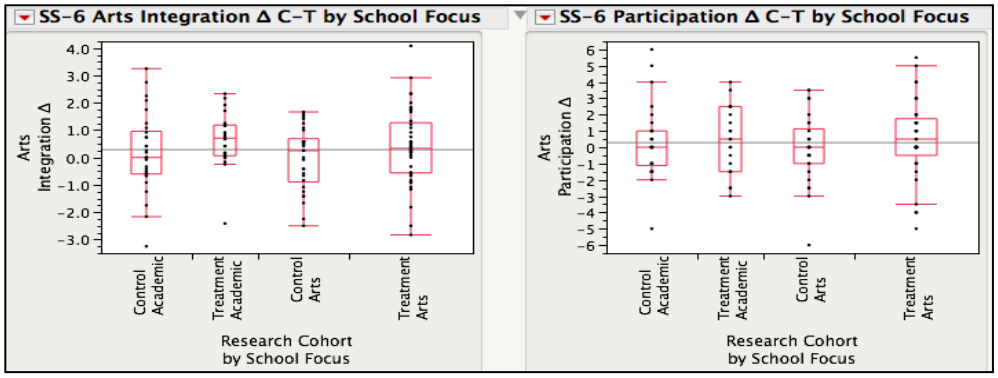
3B Figures 2-3 *Data Display Indications*: 3B Figures 2-3 box charts provide another look at the differences in pre-post survey scores according to the PAIR Arts vs. Academic Focus Schools. Note here that the *Treatment students rated three key factors higher than both types of Control FOCUS school students* in Total Average, Average of Key Factors, and every single PAIR Student Survey statement. Note also that the Control Academic Schools, in particular, rated these factors lowest on virtually all categories of key aspects of the intended PAIR classroom culture.

3B Figures 2-3: **Control-Treatment Box Plot Comparisons of Grade 6 PAIR Student Survey Average Responses by School FOCUS (left: overall averages, right: average change of key indicators Arts Integration, Reflection, and Presentation of Arts Learning Work)**



3B Figure 4: **Control-Treatment Box Plot Comparisons of Grade 6 PAIR Student Survey Average Responses by School Focus (three primary question categories)**





Section 3B *Preliminary Summary, Emerging Themes*: Positive changes in students' perception of classroom culture drawn from pre-post Student Survey data indicate that key elements of PAIR arts integration practices emphasized in the CAPE teacher professional development program outcomes were perceptible to the Treatment School students by the end of the project. Results obtained by the Initial Cohort indicate that these changes are associated to some extent with the emergence of significant trend comparisons between the treatment and control school student learning assessments. Unfortunately, follow-up cohort data was not collected, making it impossible to determine if these effects obtained more quickly or with more statistical significance when the PAIR teachers had more experience with the project.

Furthermore, control-treatment school comparisons reveal that the Treatment Arts Focus Schools were the cohort most likely to show positive change on the Student Survey responses in matters of teacher collaboration, participation in art activities, and reflection on arts learning. In contrast the Treatment Academic Schools were more likely to indicate higher levels of arts integration processes and presentation of arts integration work. Taken together, it is evident that student ratings of classroom culture provide a way to understand student judgments of key PAIR program factors that depended on the type of PAIR School they attended. Overall the rank order effect of student ratings of key program features of PAIR proceeds from strong to weak according to the level of school focus and treatment:

from Treatment Arts → Treatment Academic → Control Arts Schools →Control Academic Schools.

As similar rank ordered ratings of the School type trends are determined consistently in other areas of data analysis as has been demonstrated repeatedly in this report, this rank-order effect suggest that both arts and academic teaching and learning are optimized by the PAIR Treatment School emphasis on arts plus arts integration practices and that schools with neither an arts nor an arts integration focus are consistently underperforming academically and artistically compared to all other schools in this study.

APPENDIX 6:

PAIR Student Learning Survey

Student Name: _____ School: _____
Grade: _____ Classroom Teacher: _____
Date: _____

Instructions: How often do these things happen in your classroom or school?
Score the sentences on a scale of 1 to 8, with 1 being "this never happens" and 8 being "this happens all the time."

never		rarely		sometimes		often		always
1	2	3	4	5	6	7	8	

Example:
1 2 3 4 5 6 7 8 In class, I get to write on the chalkboard
Circling "2" means that students write on the chalkboard very rarely. More than never, but not every week. What if you wrote at the chalkboard but no more than once or twice a week? Which number would you circle? Maybe or 3 or 4?

1 2 3 4 5 6 7 8 In class, I get to write on the chalkboard
Circling "8" means that your write on the chalkboard always, maybe even several times a day. What if it were almost every day? Maybe a 7? And what if you marked a 5 or 6, What would that mean? Maybe three or four times a week?

For the following sentences, circle ONLY ONE number

Never ---------> always	Questions coded for evidence of Collaboration [C}, Reflective Practices {R], focus on Arts Integration [AI], Frustration with writing assignments [F], Presentation of Arts Learning [PA], Engagement in Arts outside the classroom [OA] from the student perspective
1 2 3 4 5 6 7 8	1. I have opportunities to teach and be taught by other students in class. [C]
1 2 3 4 5 6 7 8	2. There are assignments where students plan together in class. [C]
1 2 3 4 5 6 7 8	3. My teacher meets just with me to review my work when it is completed and/or when it is in progress. [R] 4. 4B. My teacher meets with my working group to review my work when it is completed and/or when it is in progress. [R]
1 2 3 4 5 6 7 8	5. I have opportunities to give feedback to the teacher about how we feel about our class activities. [R]
1 2 3 4 5 6 7 8	6. [omitted in final version]
1 2 3 4 5 6 7 8	7. My classmates and I think of new ideas before and when we work together. [C, R]
1 2 3 4 5 6 7 8	8. My classroom teacher and my in-school art, music, performance or dance teacher work together in my class. [C]
1 2 3 4 5 6 7 8	9. I have the opportunity to write in class to reflect on or think about what I learned. [R]
1 2 3 4 5 6 7 8	10. I use images, music, movement or dance to understand books, stories or vocabulary in class. [AI} 11. 10B I use images, music, movement or dance to understand math concepts or math problems in class [AI} 12. 10C I use images, music, movement or dance to understand different cultures and my own or my family identity in class. [AI} 13. 10 D My teacher explains and/or we discuss why we use images, music, movement or dance to understand books, stories, writing or vocabulary in class. [AI} 14. 10EMy teacher explains and/or we discuss why we use images, music, movement or dance to understand math concepts or math problems in class [AI}

	15. 10F My teacher explains and/or we discuss why we use images, music, movement or dance to understand different cultures and my own or my family identity in class. [AI}
Never ---------> always	16. [omitted in final version}
1 2 3 4 5 6 7 8	17. [omitted in final version]
1 2 3 4 5 6 7 8	18. My teacher uses ideas about dance, theater, art, music, or image-making to teach other things about reading or writing in class. [AI} 19. 12B My teacher uses ideas about dance, theater, art, music, or image-making to teach other things about math concepts or problem solving in class [AI} 20. 12C My teacher uses ideas about dance, theater, art, music, or image-making to teach other things about family or individual cultural identities or values in class. [AI}
1 2 3 4 5 6 7 8	21. As we start a new unit, I have the opportunity to write about what I know, and towards the end of the unit, I write about what I learned. [R]
1 2 3 4 5 6 7 8	22. I get bored or frustrated writing in class, especially when I am asked to write long paragraphs. [F]
1 2 3 4 5 6 7 8	23. During class students get the opportunity to look at each others' work and talk about it together. [R]
1 2 3 4 5 6 7 8	24. Student work is publicly displayed and/or performed in the classroom. [PA]
1 2 3 4 5 6 7 8	25. [not included in data analysis]
1 2 3 4 5 6 7 8	26. In our class, the students have the opportunity to decide how their work gets graded. [C. R]
1 2 3 4 5 6 7 8	27. Students in my class have opportunities to show their work in an exhibition, performance, or presentation for students and/or parents *from the school*. [PA]
1 2 3 4 5 6 7 8	28. Students in my class have opportunities to show their work in an exhibition, performance, or presentation to students from *another school*. [PA]
1 2 3 4 5 6 7 8	29. I and/or my group take the time to brainstorm for ideas as part of my work in class. [R]
1 2 3 4 5 6 7 8	30. When a project or problem is or becomes really hard, I don't want to do it anymore. [F]
1 2 3 4 5 6 7 8	31. In class I have the opportunity to make choices about my own work. [R]
1 2 3 4 5 6 7 8	32. In class, I have the opportunity to explore different ideas and tools to learn more than what I know. [R]
1 2 3 4 5 6 7 8	33. [omitted in final version]
1 2 3 4 5 6 7 8	34. I take the time to look back at the work I've done and tell others what I learned about it in class. [R]
1 2 3 4 5 6 7 8	35. Students in my school are learning about the arts, and also explore its connection to math, science, or world languages and cultures. {AI]
1 2 3 4 5 6 7 8	36. I am aware of and take part in the after school art activities that take place at my school. [OA]
1 2 3 4 5 6 7 8	37. I take part in art, music, theater or dance classes outside of school. [OA]
1 2 3 4 5 6 7 8	38. My classroom teacher works with outside teaching artists, musicians, performers, or dancers who come into our classrooms to teach. [C, AI]

3C: Analysis of PAIR Student Partnership in Arts Integration Learning (PAIL) Work Samples

3C: *Table of Contents*

3C: Analysis of PAIR Student Partnership in Arts Integration Learning (PAIL) Work Samples	1
3C: Table of Contents	1
3C: List of Tables, Figures and Appendices	2
3C: Introduction	3
The importance of student work documentation in the PAIR project	3
Evaluating student work documentation in the context of arts integration	3
Systematic review of the PAIL student work samples	4
Four PAIL classroom exemplars	5
3C-1: Analysis of Four PAIL Classroom Student Work Case Studies	6
3C-1 Case Study #1: Grade 4 PAIR Arts-Focused School World Language Classroom	6
3C-1 Case Study #2: Grade 5 PAIR Academic-Focused School Writing Classroom	7
3C-1 Case Study #3: Grade 4 PAIR Academic-Focused School Math Classroom	9
3C-1 Case Study #4: Grade 6 Arts-Focused PAIR School World Language Classroom	11
3C-1: Summary of PAIL Case Study Findings	13
3C-2: Statistical Analysis of Within-Treatment School Comparisons of PAIR Final Year PAIL Classroom Student Work Sample Ratings.	14
3C: Summary, Emerging Themes	19

3C: *List of Tables and Figures*

3C-1 Table 1:	Grade 4 PAIL Classroom Student Work Ratings Profile	6
3C-1 Table 2:	Grade 5 PAIL Classroom Student Work Ratings Profile	7
3C-1 Table 3:	Grade 5 PAIL Classroom Student Work Ratings Profile	9
3C-1 Table 4:	Grade 6 PAIL Classroom Student Work Ratings Profile	11
3C-2 Figure 1:	PAIR Within-Treatment School Distribution of Final Year (Grades 4-6) PAIL Quantity and Quality Classroom Student Work Profile Ratings.	15
3C-2 Figure 2:	PAIR Within-Treatment School Mean Score Comparisons of Final Year (Grades 4-6) PAIL Student Work Sample Quantity and Quality Ratings According to School FOCUS (Arts vs. Academic)	16
3C2 Figures 3:	Correlations between Quantity and Quality of Classroom PAIL work sample ratings According to combined and separate factors of School Focus (Arts vs. Academic)	16
3C-2 Figure 4:	PAIR Within-Treatment School Mean Score Comparisons of Final Year (Grades 4-6) PAIL Classroom Student Work Sample Quantity and Quality Ratings According to School Cluster Types	17
3C-2 Figure 5:	Analysis of Variance of PAIL Quantity and Quality Ratings by Number of Teacher Years in the PAIR Project	18
3C-2 Figure 6:	Non Parametric 'Degree of Correlation' Analysis of PAIL Student Work Data and Years of Teacher Participation in the PAIR project	19

3C: Introduction

IIIC *Overall Inquiry Questions*: *1) What do the analyses of PAIL student work samples reveal about the nature of arts integration teaching and learning? 2) To what extent do the PAIL student work samples validate the teacher and student learning goals of the PAIR program?*

The importance of student work documentation in the PAIR project

In the first section of this report, Dr. Burnaford reports how her research team isolated specific indicators of Effective Teaching Practices through examining how teachers a) annotated documentation of student work in PAILS (Partnerships in Arts Integrated Learning Samples), b) created and discussed documentation panels that capture the context and coherency of this student work [see page 56 Table Twenty: Collaborative Reflection on PAIR Documentation Panels: Notes from Professional Development Discussion] , and c) reflected on the quality of student work and its implication for effective arts integration teaching and learning while participating in PAIR Student Portfolio Conference sessions.

Analyses of teacher reflections in Section I revealed that "Teachers who document (including collecting student work, reflecting on teaching, outlining the curriculum plan) and assess regularly as part of their teaching in an arts partnership are more likely to see student achievement" and that this "Documentation to Learn Effect" was a major outcome of the PAIR professional learning program. In addition, Dr. Burnaford claims that "the teachers who were more clearly invested in the PAIR PAILS were also more able to articulate what they and their students were learning as evidenced by their student work labels and their ability to discuss student work in the Portfolio Conferences. According to teachers, rigorous student documentation practices in the PAIR program "... allows [students] to see how far they have come from an idea to an actual finished product", to "see their growth and assess their own learning" and to see themselves "collaborating with their peers, working toward a common goal" [and] to observe "the give and take needed to work successfully together."

Evaluating student work documentation in the context of arts integration

For many teachers, PAIL student work sample documentation provided an important medium for evaluating the impact of arts and arts integrated teaching and learning in their PAIR classroom. In comparison with standardized tests for math, for example, documentation of arts integration provides an aesthetic dimension to student learning assessment that is more likely to include the ongoing

discovery of exciting and concrete evidence of student learning over time in the art work itself – a discovery that benefits both teachers and students. As one teacher put it during her PAIR portfolio conference:

> "*I just think that sometimes [students] take a test, and then maybe the test hangs on the refrigerator for a little while and then the test goes [into] the garbage or the recycling. ... you have a hard time throwing something like this (PAIL artifact] out. You do get to kind of peek inside their psyche and kind of get an idea of what's going on in their heads. And then we have evidence at the end, and it's evidence that they can be proud of. It's beautiful and it's something that we don't want to go away. We want to hang it [up] and we want to be proud of it. That's what I feel is nice about this whole process."*

Systematic review of the PAIL student work samples

Because all PAIR treatment school teachers were asked to create PAIL student work samples, researchers were obliged to review the contents of the PAILS for evidence of the extent and quality of arts and arts integration teaching and learning in each of the PAIR grade level classrooms by the final year of the project. Accordingly, researchers reviewed representative artifacts from each Treatment School grade-level (4-6) classroom and provided the following categories of analysis:

(1) Number of Artifacts (ranging from 4-67);

(2) General summary comments (e.g., "Lots of reflective writing, yet very little artwork," "Quality of reflections increases with time," "Description of artwork and processes is detailed, complete and connected with reflective comments at various stages of completeness," etc.);

(3) Specific commentary on pieces (for example, "Demonstrates connections between art, calligraphy, self-esteem/self-awareness," or "Only one student took good notes about the sound of congas and how the pitch changes based on the player's touch," etc.); and

(4) Ratings of PAIL samples using 5-point (1.0-1.5-2.0-2.5-3.0) scaled rubrics for evaluating:

a. *Overall Quantity/Breadth/Completeness* of PAIL Student Work Samples Relevant to Arts and Arts integrated Teaching and Learning
Level 1 - Little or no work represented in PAIL Work Sample.
Level 2 – Incomplete/unbalanced sample of some or moderate work represented in PAIL Work Sample.

Level 3 – Balanced, extensive work represented in PAIL Work Sample.

b. *Overall Qualitative Differences in PAIL Student Work Samples* Relevant to Arts-Arts Integrated Teaching and Learning

Level 1 - Novice or Beginning (generic, diffuse, single perspective, weakly relevant, undetailed)

Level 2 - Developing or Emergent (moderately specific, detailed, relevant, elaborative, coherent, relational)

Level 3 – Proficient or Advanced (highly specific, detailed, relevant, elaborative, coherent, relational, systemic)

The PAIL rubrics listed above enabled the researchers to rate and average the overall breadth and quality of each classroom's work samples, which provided an opportunity to both assess the impact of arts integration practices in the classroom and eventually to compare these averaged classroom levels by treatment school differences in CPS Magnet Cluster school types (Math, ELA, World Language) focus, and finally by three different levels (years) of teacher engagement in the PAIR project.

Four PAIL classroom exemplars

Data profiles were produced for each treatment school classroom based on analyzing, commenting and rating the entire set of PAIL student work samples and the detailed descriptions of their PAIR units. The averaged ratings of student work in each classroom were determined through analysis of all the samples and rating the amount, breadth, consistency, relevance and cognitive complexity of the work as a whole.

The four PAIL profiles examined next are based on PAIR exemplary units (across grades 4-6) provide robust examples of a strong emphasis on visual art, music and dance concepts and skills and their relationship to academic subject topics and skills typically found in arts and academic focused PAIR Math, Language Arts, World Language and Arts magnet schools.

* * *

3C-1 Analyses of Four PAIL Case Studies

3C-1 PAIL Case Study #1: Grade 4 PAIR Arts-Focused School World Language Classroom

3C-1 Table 1: Grade 4 PAIL Classroom Student Work Ratings Profile

Exemplary Visual Art –World Language Unit (not including Music-World Language Project)

Teacher Name	Grade	# of Artifacts	General Comments	Exemplary Artifact Comments (ID by Filename)	Averaged Quantity Rating	Averaged Quality Rating
Classroom A	4	31	Art-arts integration work (detailed and complete) and written reflection are focused on the themes of identity, family, and culture	CIMG8821.jpg, CIMG8822.jpg, CIMG8972.jpg: shows connections among world cultures and student's individuality through art and poetry	2.0	3.0

The first PAIR Visual Art – World Language Unit focuses on the concepts of identity, family and community in the context of parallel activities in self-portraiture ('calligram') and the creation of a biographic poem ('biopoem').

After writing very personal poems (this sample copied from the photographed original in the left figure below) and including a personal photo and other identifying factors (that had to be redacted in this report due to concerns for the protection of student identities), students were encouraged next to capture elements of the poem in their follow-up self portrait. The final self portrait combined words of the poem with images of self, that included personal artistic choices in color, design, and iconographic images of favorite foods, clothes, etc., included in the final work (right figure below).

Melody's (pseudonym) Bio-poem

Funny, nice, friendly, smart,

Daughter of Richard,

Lover of Haunted Trails, ice cream and movies,

Who feels excited for holidays, and when people die, and bored of speeches,

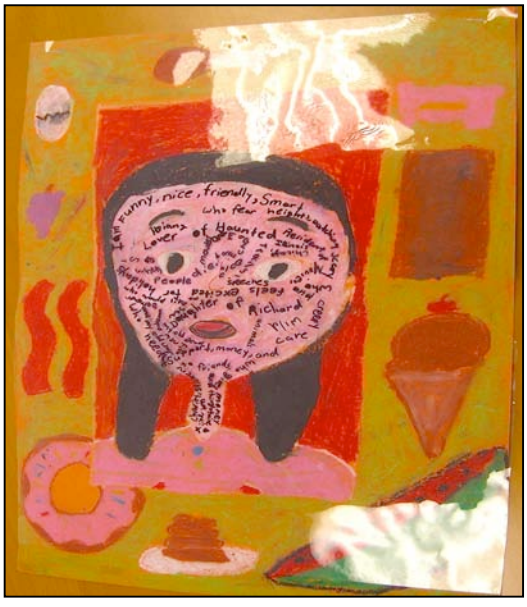
Who needs support, money, and care,

Who fears heights, watching scary movies, and creepy wild animals,

Who gives money to charity, stickers to friends, and things of mine like toys.

Who would like to see my dead pet the fish, new Pokèmon games, and my old classmates.

- Resident of Chicago, Illinois



The resulting visual work (similar to the example above) stems from the challenge given to all students to explain their compositional choices in terms of their own identity, family and community culture and history. In this exemplary PAIR unit, the biopoem-self portrait (calligram) functions as a medium for reflective understanding of arts integrated learning processes that students can readily use to express learning connections across the World Language Arts curriculum. In this school's "create an imaginary nation" project (discussed below in Example 4 in this chapter and in the portfolio conference chapter), for example, 6th grade students describe how they leveraged their self-identified traits and cultural contexts to describe their imaginary nation's form of government, community action and their role in their newly created society. Furthermore, when the students composed a new national anthem, the lyrics were drawn from these same bio-poems.

* * *

3C-1 CASE STUDY #2: Grade 5 PAIR Academic-Focused School Writing Classroom

3C-1 Table 2: Grade 5 PAIL Classroom Student Work Ratings Profile

Exemplary Dance – Mathematics Unit (not including the Visual Arts – Mathematics Unit)

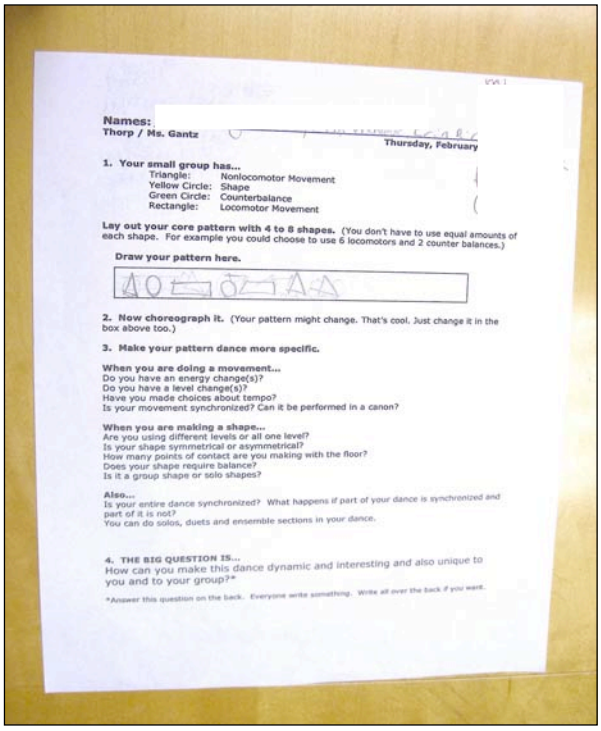
Teacher Name	Grade	Number of Artifacts	General Comments	Exemplary Artifact Comments (ID by Filename)	Averaged Quantity Rating	Averaged Quality Rating
Classroom B	5	38	Worksheets provide descriptions of explicit, detailed connections among math, science, arts, aspects of the projects.	CIMG0289: Worksheet showing integration of math and dance class; CIMG0299 & CIMG0300: connections between patterns, math, and choreography	2.0	2.5

As indicated by the review sheet for our *second* exemplary PAIL profile (above), this classroom rates relatively well in breadth and completeness of students work samples and receives a high rating in terms of level of quality. The general comments and specific content of the exemplars make clear that the connections among fundamental concepts shared across disciplines in this classroom is particularly relevant to the PAIR project goals both in terms of teacher professional development and student learning outcomes.

As demonstrated later in the portfolio conference analysis section (IIID), the more the PAIL work contains thorough evidence of curriculum planning, student process documentation, student final work, and reflective writing, the more opportunities there are for further reflection by teachers and students on arts and arts integrated learning in the presence of these artifacts.

The following classroom artifacts illustrate exemplary levels of articulation of PAIL student work and reflections that illustrate the students' experience of connecting principles shared between dance and mathematics.

The first PAIL *artifact* to the right stems from a PAIR Dance Unit Planner (Teacher and Student names redacted) that specifies unit inquiry questions, activities, and essential vocabulary for dance (choreography, locomotor and non-locomotor movement) that focus on varied fundamental concepts shared between dance and other arts/academic disciplines (shape, points of contact, balance-counterbalance, symmetrical-nonsymmetrical, energy, synchronization, tempo, canon).



Names:
Thorp / Ms. Gantz
Thursday, February

1. Your small group has...
Triangle: Nonlocomotor Movement
Yellow Circle: Shape
Green Circle: Counterbalance
Rectangle: Locomotor Movement

Lay out your core pattern with 4 to 8 shapes. (You don't have to use equal amounts of each shape. For example you could choose to use 6 locomotors and 2 counter balances.)

Draw your pattern here.

2. Now choreograph it. (Your pattern might change. That's cool. Just change it in the box above too.)

3. Make your pattern dance more specific.

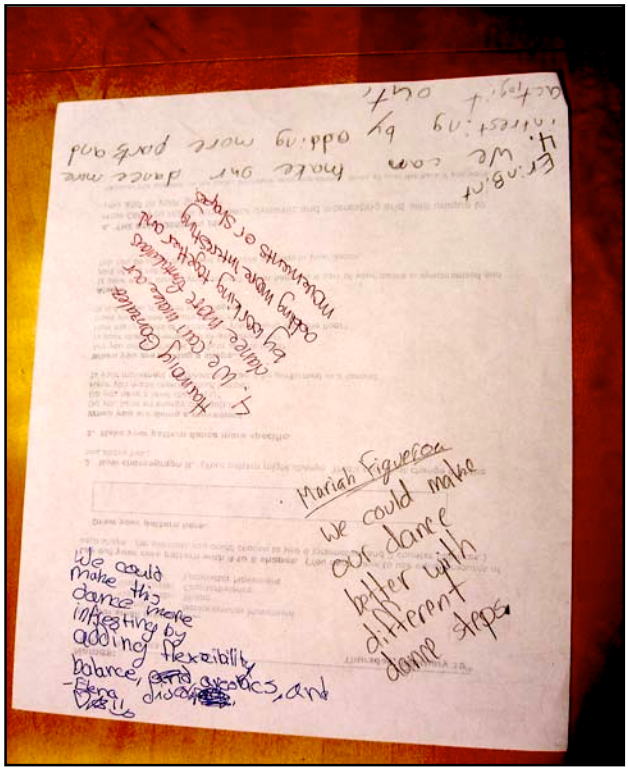
When you are doing a movement...
Do you have an energy change(s)?
Do you have a level change(s)?
Have you made choices about tempo?
Is your movement synchronized? Can it be performed in a canon?

When you are making a shape...
Are you using different levels or all one level?
Is your shape symmetrical or asymmetrical?
How many points of contact are you making with the floor?
Does your shape require balance?
Is it a group shape or solo shapes?

Also...
Is your entire dance synchronized? What happens if part of your dance is synchronized and part of it is not?
You can do solos, duets and ensemble sections in your dance.

4. THE BIG QUESTION IS...
How can you make this dance dynamic and interesting and also unique to you and to your group?*

*Answer this question on the back. Everyone write something. Write all over the back if you want.



Later on in the sequence of math-dance activities, PAIR students are challenged to demonstrate a high level of understanding of their choreography as defined by their understanding of geometric shapes and strict proportion of types of motion both in their culminating dance event and during their portfolio conferences.

Post activity comments by the four student collaborators (left figure) demonstrate a fluent reflective awareness and understanding of their arts integration unit concepts (e.g., balance, flexibility, shape) introduced 10 weeks earlier.

* * *

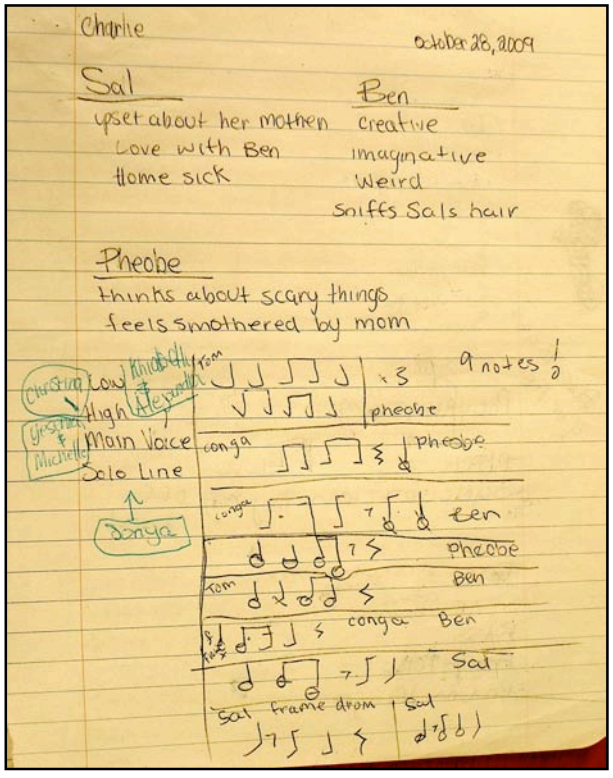
3C-1 CASE STUDY #3: Grade 4 PAIR Academic-Focused School Math Classroom

3C-1 Table 3: Grade 5 PAIL Classroom Student Work Ratings Profile

Exemplary Music – Language Arts Reading Unit (not including Drama – Language Arts Project)

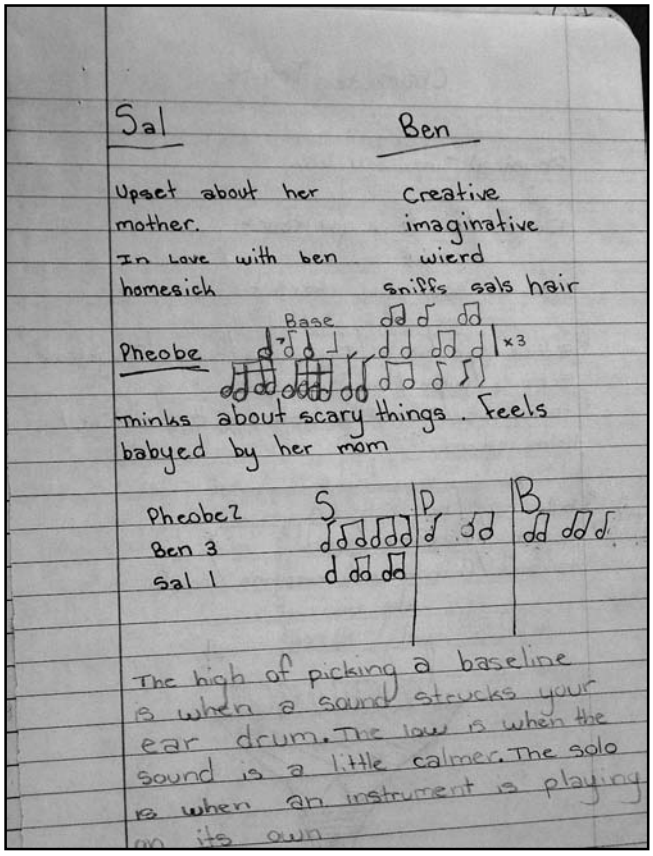
Teacher Name	Grade	Number of Artifacts	General Comments	Exemplary Artifact Comments (ID by Filename)	Quantity Rating	Quality Rating
Classroom K	5	14	Limited amount of finished artifacts but the quality of creative rewriting of stories, music composition and character unit reflection work quality is high	CIMG8788.jpg, CIMG8786.jpg, CIMG8784.jpg process notes for character analysis and drum pattern compositions,	1.5	2.5

The *third* classroom profile suggests that the understanding of character and actions in the grade 5 Language Arts classroom benefits from use of music and receives a relatively high rating for quality despite having relatively fewer artifacts. Informed by their work with character analysis sketches with their CAPE music teaching artist and acting exercises with the drama teaching artist, each student is encouraged to identify critical traits of each character in their reading assignment and how these traits are informed by the actions of the characters throughout the book.

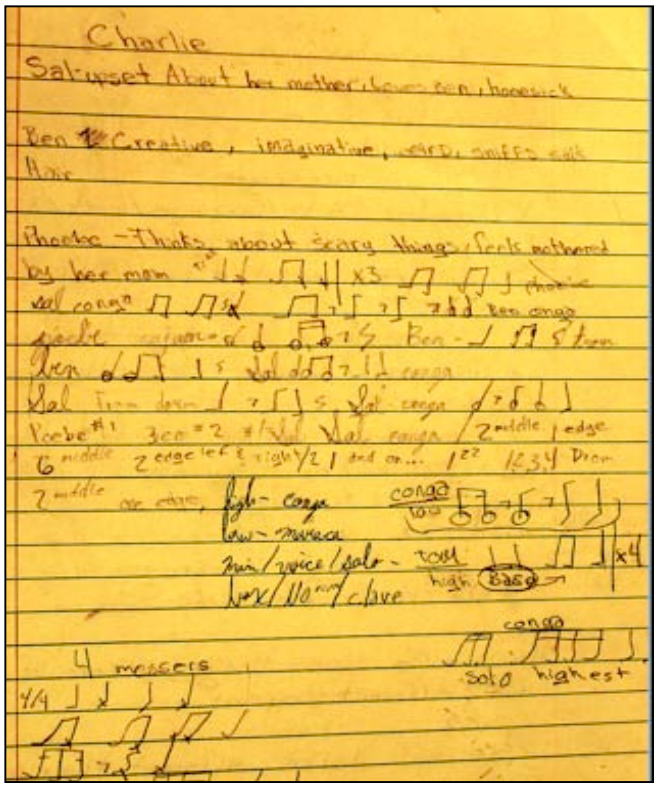


Three artifacts reveal how the student worked with the music teaching artist to compose the drum patterns that capture differences among the characters in terms of their dispositions, actions and the kind of music that would represent their character during a dramatic reading of the book.

In the first journal artifact (figure to the left), the student list certain details in each character portraits and then goes on to compose percussion instrument patterns intended to represent each characters' actions, mood, and disposition in the story. Drawing on contrasts in the musical rhythms, silences, and choice of instruments, the musical character sketches could be performed and critiqued by the entire class.



The second journal entry artifact (figure to the left) contains multiple elements of music that were considered as ways to capture the essence of each character in relation to the other. Although still a work in progress, the student provides several renditions of the drum character patterns and an illustration of how a character (Phoebe) can go through different actions and mood changes. At the bottom of the page, the student provides a description of musical terms and how to interpret the music symbols in order to perform the patterns as intended by the student composer.



The final artifact of this unit (to the left) provides yet another example of how a student can work with elaborate combinations of musical structures to depict character. By considering a yet wider array of musical patterns and their association with character traits and actions (e.g., "Phoebe thinking about scary things" or "feeling mothered by her mom"), the final sequence of interleaved drum patterns can convey an understanding the similar and contrasting personal and emotional dispositions of the principal characters as the conflicts and resolutions described in the story unfold over time.

* * *

3C-1 CASE STUDY #4: Grade 6 Arts-Focused PAIR School World Language Classroom

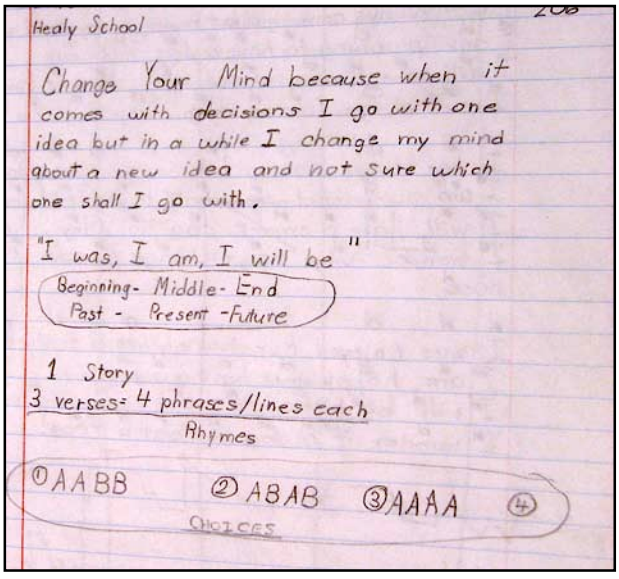
3C-1 Table 4: Grade 6 PAIL Classroom Student Work Ratings Profile

Exemplary Music – World Languages Unit (not including Visual Arts – World Language Project)

Teacher Name	Grade	Number of Artifacts	General Comments	Exemplary Artifact Comments (ID by Filename)	Averaged Quantity Rating	Averaged Quality Rating
Class-room C	6	41	Limited number of finished projects yet the reflection on process and the quality of music composition skills and reflection work is high	CIMG9479.jpg, CIMG9485.jpg, CIMG9488.jpg process notes for lyrics, melody, and structure	2.5	2.5

As this *fourth* classroom profile suggests, the focus on music reading, composition and performance provides a medium for arts integration in World Languages through the creation of lyrics that expound on imagined critical junctures for each child's maturity. Informed by their work with bio-poems/self portraits with their previous CAPE visual teaching artist, each student group is encourage to identify imagined critical junctures in their future path toward personal maturity.

The following four artifacts capture the range of integrative design and content variables in song composition ranging from elements of language prosody and meaning to musical structure and melodic design.


Healy School

Change Your Mind because when it comes with decisions I go with one idea but in a while I change my mind about a new idea and not sure which one shall I go with.

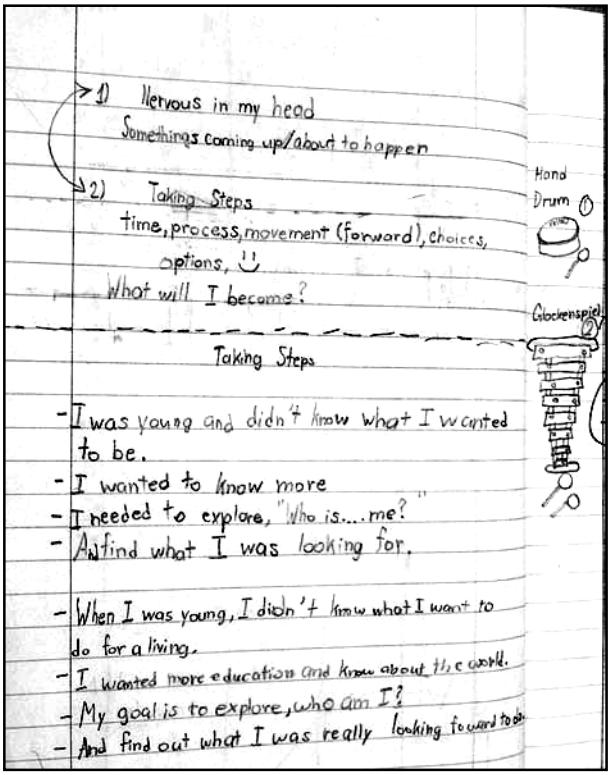
"I was, I am, I will be"
Beginning- Middle- End
Past - Present -Future

1 Story
3 verses= 4 phrases/lines each
Rhymes

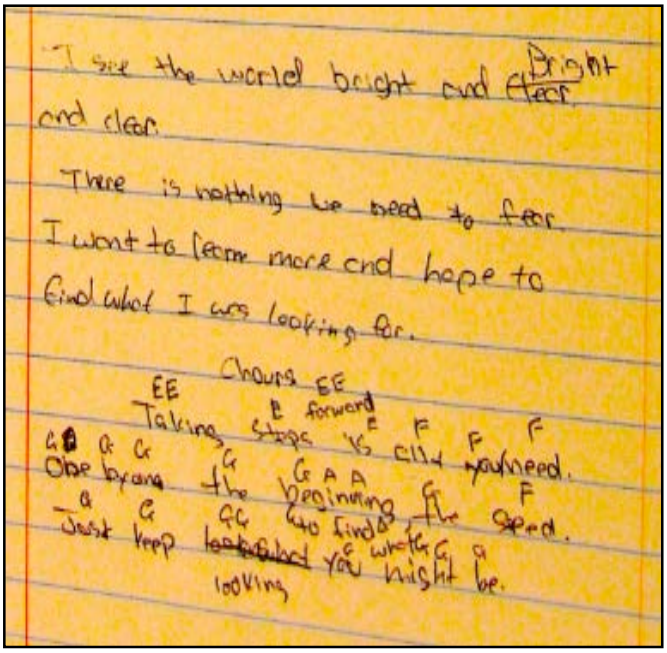
①AABB ②ABAB ③AAAA ④
Choices

In the first journal artifact (figure to the left), one student chooses the title "Change Your Mind" to reflect on the nature of choices and decisions that are made in the context of new ideas. Drawing on the parallels among the concepts of "I was, I am, I will be", Beginning-Middle-End, and Past-Present-Future, this document shows how this student is considering alternative song structure choices to frame ideas about their future.

The second journal entry artifact (figure below) contains multiple elements of music and music integration understanding that will be needed to compose a song that captures the elements of the maturation process: drawings that show how musical instruments work, notes on "time, process, movement, moving (forward), choices, options", and asking "What will I become?" The page also traces the evolution from the scansion of the lyrical lines into stanzas of the song "Taking Steps" where the quest for self-identify (previously explored in the PAIR visual arts unit) now is transformed into the following rhyme scheme

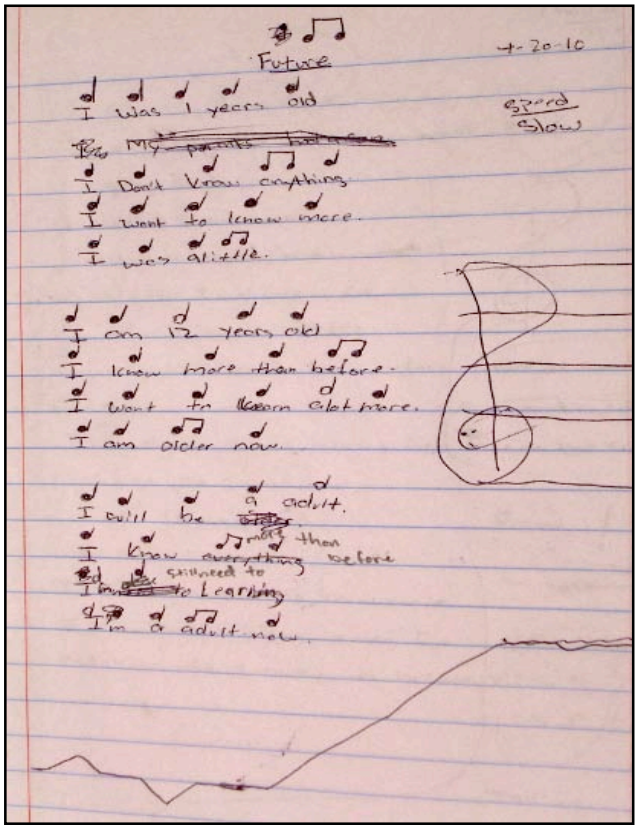


– I was young and didn't know what I want[ed] to be
– I wanted to know more,
– I needed to explore, 'Who is …. Me?'
– And find what I was looking for."



The third artifact to the left provides an example of how students can connect lyrical structure with song structure and concepts of self-identity. By considering all these factors, the final product can convey a rich perspective of the interconnectivity of past and future events in the context of one's own life decisions. This understanding is also informed by the lyrical content, which serves also to express the social-emotional factors that accompany the spectrum of life changing circumstances and events.

The final artifact to the right traces in increasing detail the evolution of the song composition, showing how contour, rhythmic grouping, note choices and form are integrated in both the subject matter and prosodic (text setting) challenges of the PAIL unit. Note how the overall shape of the energy and pitch of the music rises as the path to adulthood is achieved by the last line of the lyrics, "I'm an adult now."



* * *

3C-1 Summary of PAIL Case Study Findings: The PAIL classroom student work case study samples collected by teachers and teaching artists in the final year of the project were rated by the research team according to amount and quality of the artifacts per unit. The relatively higher rated exemplars described above provide a clear glimpse of the nature of PAIR arts integrated teaching and learning practices in multiple art forms that connected with core concepts within and integrated across English language arts, math, and social studies in grades 4-6.

This documentation provides an important window on how students work with core concepts across disciplines and provides a direct insight into the cognitive (within discipline) and meta-cognitive (across discipline) complexity of all the PAIR unit tasks. In PAIR World Language units, mapping poems onto self-portraits and developing musical lyrics that represent a deep understanding of a student's self-identity in the context of family and community is a cognitively rich task that was fulfilled through teaching artist guidance across complementary art forms. The growth in art and music skills serve to highlight the level of interest sixth grade students have in their social and personal

development. In PAIR Language Arts units, the artful representation of story characters through percussion instruments allowed students to go deeper into concepts of character that were authentic to both language arts and music. Stories have characters, and musical patterns and instruments can bring new dimension to a deeper understanding of each individual character's disposition, response to others, and the motivation to act. Transformation in characters through a narrative can be informed also by transformations in musical patterns with themes. In the PAIR Math units, precise and meaningful calculations, measurements and concepts of distance and proportion are well demonstrated in choreography and with visual blueprints for sculpture and buildings. By the end of the project, students in these classrooms often expressed the sentiment that the math work in the context of creating dance is 'real math', just as the choreography that came from mathematical concepts and calculations was 'real dance.'

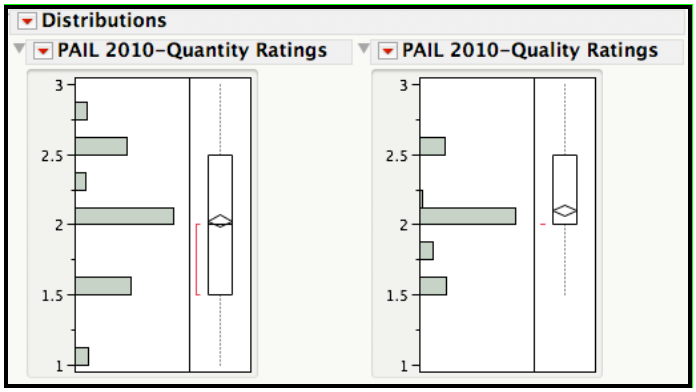
* * *

3C-2: Statistical Analysis of Within-Treatment School Comparisons of PAIR Final Year PAIL Classroom Student Work Sample Ratings

3C-2: *Inquiry Questions*: *1) To what extent do PAIR classroom student work sample ratings vary according to School Focus (Arts vs. Non-Arts)? 2) To what extent do PAIR student work sample ratings vary according to a particular arts integration and academic focus of the Cluster School (Writing, Math, or World Languages & Culture)? 3) What is the relationship between PAIL classroom quantity ratings, quality ratings, and the experience level of PAIR Treatment School teachers?*

3**C-2 Figure 1** ***Data Display Indications*:** The data display below [demonstrates that a normal statistical distribution ('bell curve') profile classroom PAIL ratings resulted from aggregating all PAIR classroom quantity and quality ratings. This finding is important because it demonstrates a valid statistical basis for comparing the quantity and quality of student portfolio work across different school types (Arts vs. Non Arts Focused) and schools with different academic priorities (math vs. writing vs. World Languages).

3C-2 Figure 1: **PAIR Within-Treatment School Distribution of Final Year (Grades 4-6) PAIL Quantity and Quality Classroom Student Work Profile Ratings.**



The PAIL ratings displayed in the distribution chart also demonstrate that the range and levels of sophistication of the artifacts varied significantly among the classrooms. In contrast to the four exemplary profiles presented in the previous qualitative analysis of student work samples, there were often cases of PAIL classroom work profiles that showed an uneven amount or distribution of types of documentation. Some PAIL folders contained little student work compared to reflection on the learning process. In other cases, there was an imbalance in focus between arts and academic work. In a few cases there was no documented evidence that students were asked to focus on the interconnections between concepts, processes, or academic content designed into the unit that were shared between the arts and academic subject areas.

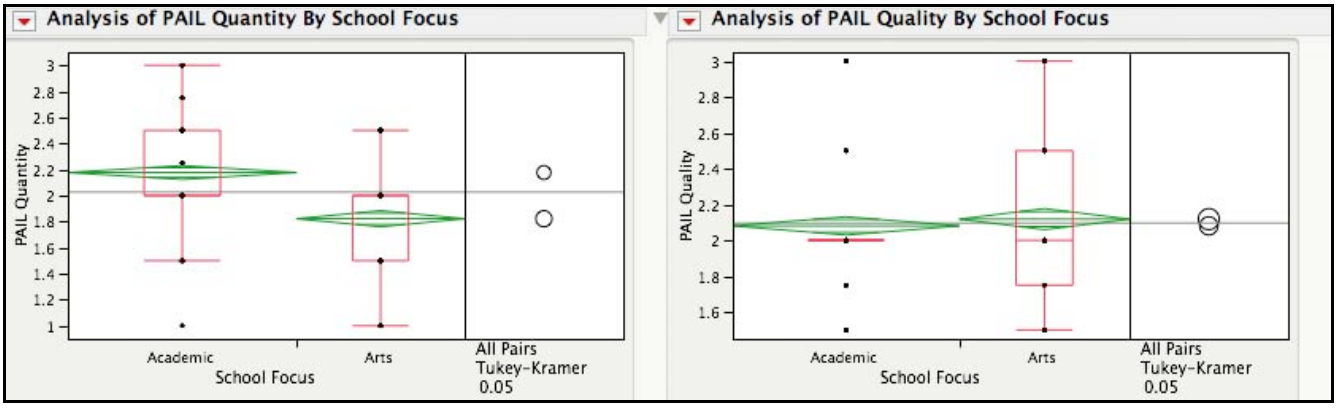
Thus, the distribution of PAIL classroom work profiles analyzed below indicate that the averaged level of teacher arts integration documentation practices and its impact on consequent opportunities for reflection on work samples by both teachers and students differed significantly across the treatment school classrooms.

* * *

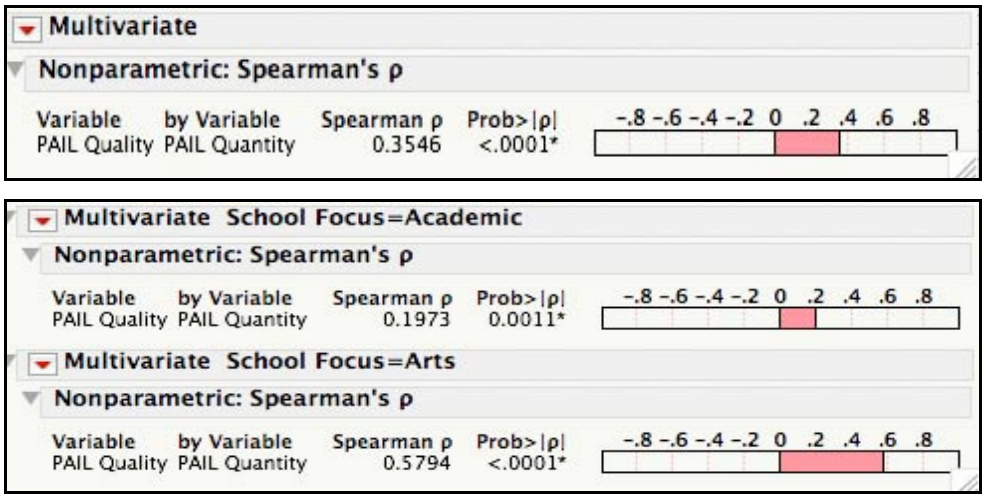
3C-2 Figures 2-3 *Data Display Indications*: The PAIL mean scores in IIIC figure 1 below demonstrate that the PAIR Academic treatment classrooms are far more likely to produce and organize student work documentation samples than the Arts Focus Treatment classrooms (ANOVA p value <.0001; Wilcoxon/Kruskal-Wallis Tests Chi Square Approximation p value <.0001). Conversely, the averaged quality ratings of the student work samples between Academic and Arts Focus classroom mean scores are virtually identical, suggesting at first glance that the quality ratings may have little relevance to the quality ratings of the student work. However it appears from correlation chart below

the (3C-2 Figure 2) that the *quantity and quality* of classroom portfolio work sample ratings are significantly correlated in both types of schools, yet much more strongly linked in the Arts Focus Schools (Spearman r=.58 p value <.0001). Thus we can conclude that the Academic Schools decidedly produce more documentation than the Arts focus schools, but with much less predictive value as to its quality.

3C-2 Figure 2: **PAIR Within-Treatment School Mean Score Comparisons of Final Year (Grades 4-6) PAIL Student Work Sample Quantity and Quality Ratings According to School Focus (Arts vs. Academic)**



3C2 Figures 3: **Correlations between Quantity and Quality of Classroom PAIL work sample ratings According to combined and separate factors of School Focus (Arts vs. Academic)**

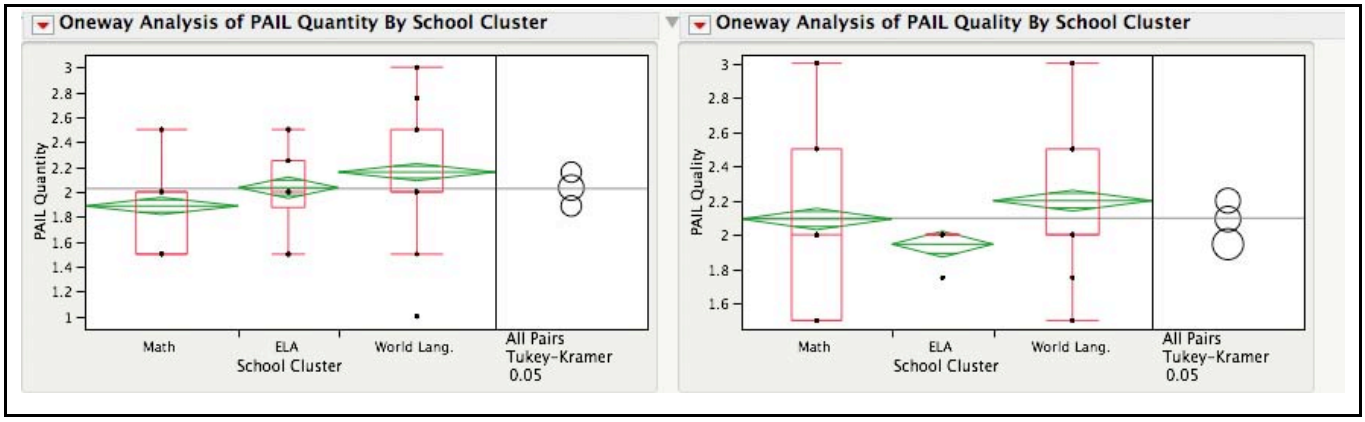


* * *

3C-2 Figure 4 ***Data Display Indications***: IIIC2 Figure 4 below shows how distinctions in school response to the PAIR project varied according to the different types of arts integration teaching artists different types of academic emphasis among the PAIR Cluster Schools (Math-Dance-Visual Art, ELA-Drama-Music, and World Language-Visual Art-Music). Significant differences occurred among the three Cluster School types, both in terms of quantity and quality of PAIL student work sample organized by PAIR teachers in grades 4-6 (ANOVA p value <.0001; Wilcoxon/Kruskal-Wallis Tests Chi Square Approximation for categorical data P value <.0001).

The mean scores displayed directly below indicate that the World Language classrooms decisively outperformed the ELA schools in breadth of documentation processes and that World Language, Math, and ELA classroom are rank-ordered by their categorical differences in the averaged *quality* level of classroom student work documentation. One possible factor in this *quantity* findings is that visual art, music composition, and creative writing tasks (in the Math and World Language classrooms) tend to result in broader array of artifacts compared to dance, acting, and musical improvisation tasks. If quantity does predict the quality ratings of the artifacts we may conclude that commitment to systematic and comprehensive student work documentation will more likely support higher levels of meta-cognitive understanding of the processes of arts integration learning processes and products.

3C-2 Figure 4: PAIR Within-Treatment School Mean Score Comparisons of Final Year (Grades 4-6) PAIL Classroom Student Work Sample Quantity and Quality Ratings According to School Cluster Types

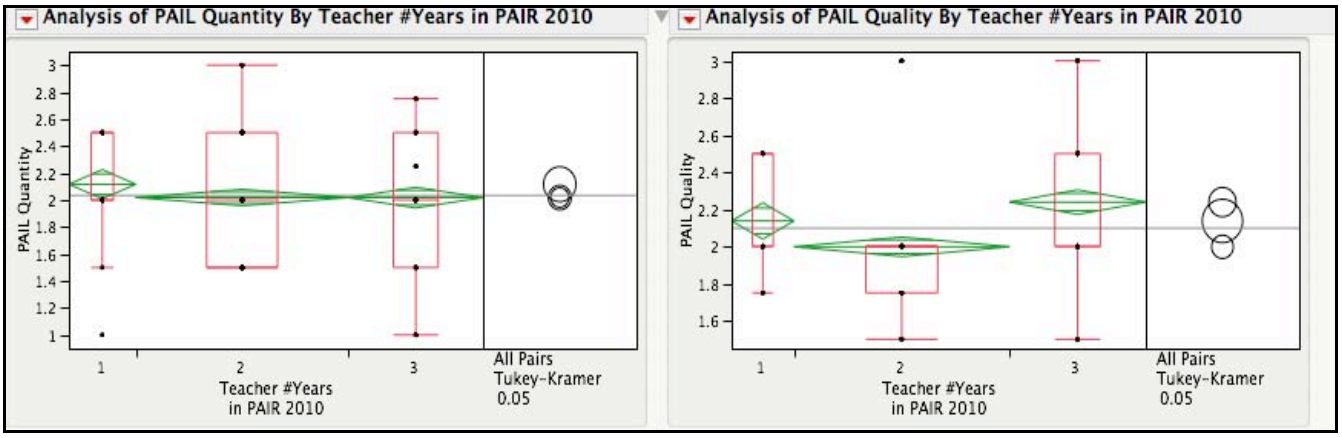


* * *

3C-2 Figure 5 *Data Display Indications*: The next perspective on the PAIL classroom data directly focuses on teacher performance by exploring the question: Does the number of years in the PAIR program predict the quantity and/or quality of PAIL student work documentation ratings?

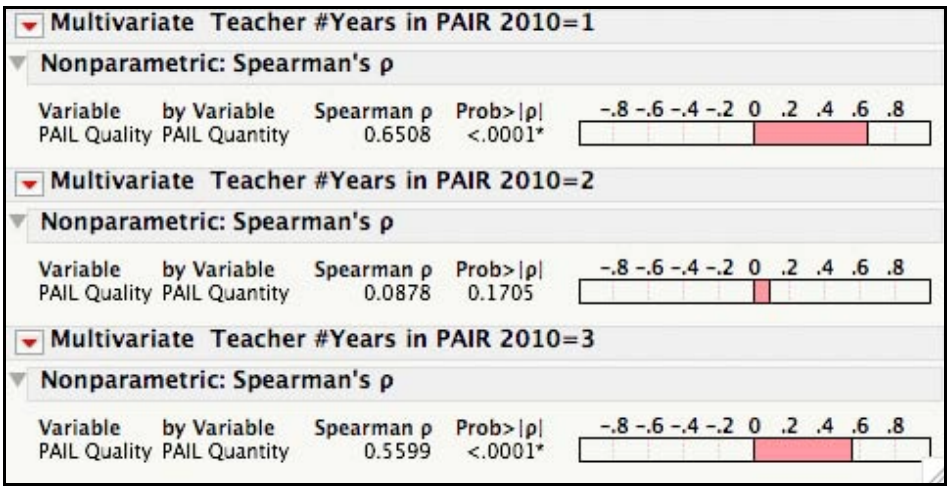
The relationship among statistical averages of the *quantity* of classroom work ratings and teacher years of experience in the PAIR project were statistically non-significant. In terms of breadth and quantity of documentation it appears that the teachers with one, two, and three years experience exhibit very similar profiles; that is, the three teacher cohorts share the capacity to produce similar amount of student [chart below to the left}. The chart to the right however, reveals statistically significant differences in *quality* ratings among the three teacher cohorts. Those teachers with 1 or 3 three years of experience with the project, when compared with teachers with two years of experience, achieve significantly higher ratings of quality with comparatively similar level of documentation produced.

3C-2 Figure 5: Analysis of Variance of PAIL Quantity and Quality Ratings by Number of Teacher Years in the PAIR Project



The correlation tables below provide evidence that the amount of documentation may also predict the quality of the documentation, depending on the depth of teacher and student experience with PAIR. These data suggest it is only the PAIR treatment school classrooms with either with the most experienced teachers (grade 4 cohort) or the most experienced students (grade 6 cohort) that demonstrate statistically significant relationship between quantity and quality of the documentation. The grade 5 classroom cohorts, whose teachers *and* students had only two years experience with the project, had both the lowest rating of quality documentation (3C-2 Figure 5) and the least significant degree of association between the quantity and quality ratings of the documentation (3C-2 Figure 6).

3C-2 Figure 6: **Non Parametric 'Degree of Correlation' Analysis of PAIL Student Work Data and Years of Teacher Participation in the PAIR project**



Due to a design flaw in the PAIR experiment that prevents, these findings cannot be fully explained. Since all teachers did not start the PAIR project at the same time, the researchers were unable to determine the relative impact of teacher professional development on the overall quantity or quality of the classroom work documentation independent of the degree of their experience with the project.

Nonetheless, when averaging data from all the PAIR treatment schools, the degree of documentation accomplished by the teachers, regardless of grade level, does predict the quality ratings of student work significantly. The research team further speculated that, when a higher quantity of work was demanded of the students, there were more samples to choose from, and therefore students and teachers were more likely to select higher quality work for their PAILS and PAIR Portfolio Conferences (see next sections 3D and 3E for qualitative and statistical analysis of the individual student portfolio conference ratings and their relationship with PAIL work samples).

* * *

3C *Overall Summary, Emerging Themes*: The importance of evaluating student work documentation became increasingly crucial to the validation of arts and arts integration practices as the PAIR project progressed over time. The teachers and teaching artists were increasingly challenged to produce documentation that captured the essence of arts and arts integrated learning in the context of their teaching practices. By the end of the project, the collection of sample student work by the classroom

teachers and teaching artists at all grade levels (Grades 4-6) made it possible to explore the breadth and qualitative aspects of PAIR student learning in each classroom, to make some classroom rating comparisons according to school focus and mission, and to determine to some extent whether the years of classroom practice in the project may have influenced these measures of student work.

However, overall ratings of each PAIR classroom samples indicate significant imbalances in the amount and quality work products collected. The overall process of student work documentation was not sufficiently uniform in quantity or breadth or differentiated within the classrooms for the researchers to create rank-ordered *individual* assessments of student work quality statistically meaningful. Thus the statistical links established previously between PAIR practices and the measures of student arts integration understanding in SAIL assessments (Section 3A) and perceptions of classroom culture in student surveys (Section 3B) could not be connected to the quality of the work product on an individual student basis.

There were findings from the PAIL data that did meet the level of statistical significance at the level of treatment school cohort analysis and therefore warrant further investigation.

It appears, for example, that the same type of Treatment Cluster School [i.e., the World Language Schools] that outperformed the control schools in SAIL assessments also have the highest-rated evidence of PAIL Quality student work [C-2 Figure 4], a finding that suggests there may yet be a strong link between positive measures of student reflective understanding of arts plus arts integrating learning (from survey and interview data) and the work they can produce in PAIR classrooms. It is particularly interesting to the researchers that the World Language (social studies-bilingual, visual art-music focused) classrooms produced a much higher amount and breadth of arts integration documentation and higher rated quality of student work than the other PAIR schools. And since the World Language Treatment schools outperformed all other schools in standardized test scores as reported previously in this study, it appears possible that the *commitment* to rich and comprehensive PAIR arts integration classroom documentation practices – whether or not the schools arts or academic focused schools – may constitute an optimal condition for creating a stronger connection with ISAT test scores.

While the "Content Knowledge Matters Effect" was well substantiated in the qualitative analysis of exemplary student work in the Part 1 of the entire report, the finding that the number of years in PAIR can predict the quality of student work documentation supports further the "Fourth Grade Effect" (long

term engagement in the project design and implementation impact matters) and the "Documenting to Learn Effect" (collecting student work inspires higher levels and quality of teacher and student reflection) findings reported in the Part I of this report.

In addition, PAIL classroom student work quality profiles also support the positive trends in follow-up Grade 4 and 5 Treatment School cohorts ISAT standardized test data reported in Part 2 of the whole report. The special case of the World Language Cluster School PAIL results reported in this section re-affirms the "The Healy Effect" (initiatives build on other initiatives in a school as reported in Part I of this report) and its link to accelerated ISAT performance (as reported in Part 2 of the overall report.

The World Language PAIR schools case study continues to be important throughout this report because it best synthesizes the PAIL findings in this chapter by suggesting that optimal arts integration practices results when (a) exemplary levels of arts integration teaching and learning develop over a long period of time that promote (b) the ability of classroom teachers and teaching artists to produce student work products increases, which then predicts (c) a higher quality of that student work that demonstrates the positive impact of arts plus arts integrated teaching and learning on school performance. With the presence of higher quality of student work, there is d) a greater likelihood that student understanding of PAIR practices (reported in the SAIL assessments) will increase in ways that can be linked to the external measures of academic achievement which are now known to be associated with both equity and excellence factors related to arts integration practices as reported in Part II.

The limitations of data cited above prevented making a direct analysis of the association between *individual* PAIL student work and ISAT results. While these trends are significant, the ability to rate student individually on independent work in arts and arts integration learning would have provided much stronger evidence of its impact on other measures of student learning.

Nonetheless, two other possible causal links of PAIL classroom data to student performance measures – the evaluation of student reflection on their work in portfolio conference sessions (3D and 3E) and a multivariate analysis that includes an evaluation of the connections between teacher survey data and reflections on academic results – will be explored and summarized in the final two sections (3F and 3G) of the overall report.

3D: PAIR Student-Teacher Portfolio Conference Qualitative Analysis

Teacher Reflective Understanding of PAIR, Student Arts Integration Performance Assessments, and the Qualitative Analysis of Student Cognitive, Meta-Cognitive and Social-Emotional Development Outcomes

3D: *Table of Contents*

3D: PAIR Student-Portfolio Conference Qualitative Analysis	1
3D: Table of Contents	1
3D: List of Tables, Figures and Appendices	2
3D: Introduction	3
The Purpose and Structure of the PAIR Portfolio Conference	3
The Portfolio Conference Interview and Performance Assessment Protocol	5
Synopsis Of The PAIR Portfolio Conference Protocol	6
3D-1: Qualitative Analysis of the PAIR Portfolio Conference Teacher Reflections on Student Learning	8
PAIR Portfolio Conference Qualitative Data as a window onto meta-cognitive and social-emotional developmental outcomes of arts integration learning	13
3D-2: Student Reflective Understanding of Their Personal and Collaborative Art Integration Learning: Four PAIR Portfolio Conference Vignettes	14
3D-2 Portfolio Conference Vignette I: *Math-Arts Integration Case Study School*	14
3D-2 Portfolio Conference Vignette II: *World Language-Arts Integration Case Study School*	16
3D-2 Portfolio Conference Vignette III: *Language Arts-Arts Integration Case Study School*	18
3D-2 Portfolio Conference Vignette IV: *Math-Arts Integration Case Study School*	19
3D-3: General Discussion and Summary of PAIR Portfolio Conference Vignettes	20
3D-4: Differences in PAIR Portfolio Conference performance among previously designated HAL academic achieving students	22
3D-5: Evidence of meta-cognitive aspects of student arts integration learning outcomes in PAIR Portfolio Conference data	26
3D-6: Rating Student and Teacher Responses in PAIR Portfolio Conferences	29
Rating Scales for Teacher and Student Portfolio Conference Responses	29
Summary of PAIR Portfolio Conference Response Scoring Process	29

Inter-rater scoring reliability	31
Individual differences in PAIR units, possible limitations of the rating system	31
3D–7: Tracking Qualitative Differences in Teacher and Student Reflective Understanding: PAIR Portfolio Conference Transcript Analysis Methods	32
Introduction	32
Methodological Notes	34
Organization and Purpose of the Teacher and Student Portfolio Conference Scoring Exhibits	34
3D-8 Four Portfolio Conference Scoring Exhibits: PAIR Classroom TEACHER Portfolio Conference Reflections (Excerpted Transcripts, Ratings, and Rater Summary Comments)	35
3D-8 Portfolio Conference Scoring Exhibit 1: Classroom TEACHER Reflections and Rater Comments: *Grade 4 World Language (Visual Art and Music)*	35
3D-8 Portfolio Conference Scoring Exhibit 2: *Classroom TEACHER Transcript Reflections and Rater Comments: Grade 5 Math (Visual Art and Dance)*	36
3D-8 Portfolio Conference Scoring Exhibit 3: *Classroom TEACHER Reflections and Rater Comments: Grade 5 Writing (ELA) (Drama and Music)*	38
3D-8 Portfolio Conference Scoring Exhibit 4: Classroom TEACHER Reflections and Rater Comments: Grade 6 World Language (Visual Art and Music)	40
3D-8: Summary of the Four Classroom TEACHER Portfolio Conference Scoring Exhibits	42
3D-9 Four Portfolio Scoring Exhibits: Excerpts from PAIR STUDENT Portfolio Conference Performance Responses and Reflection Ratings	42
3D-9 Portfolio Conference Scoring Exhibit 1: STUDENT Portfolio Conference Responses and Rater Comments: *Grade 4 World Language (Visual Art and Music)*	43
3D-9 PORTFOLIO CONFERENCE SCORING EXHIBIT 2: STUDENT Portfolio Conference Responses and Rater Comments: Grade 5 Math (Visual Art and Dance)	46
3D-9 PORTFOLIO CONFERENCE SCORING EXHIBIT 3: STUDENT Portfolio Conference Responses and Rater Comments: Grade 5 Writing (ELA) (Drama and Music)	51
3D-9 Portfolio Conference Scoring Exhibit 4: STUDENT Portfolio Conference Reflections and Ratings: Grade 6 World Language (Visual Arts and Music)	57
3D-10 Summary of Student and Teacher Portfolio Conference Interview and Performance Assessments	63

3D: *List of Tables*

3D-6 Table 1:	Scoring Rubric for Levels of Response for Teacher and Student Response Data	30

* * *

3D. Introduction

3D *Overall Inquiry Questions*: 1) What is the structure and purpose of the PAIR Portfolio Conference Interview and Performance Assessment Protocols? What do the portfolio transcript samples reveal about the nature of teacher and student reflective understanding of arts integration teaching and learning processes? What constitutes qualitative evidence of student cognitive, meta-cognitive and social-emotional development in the conferences? How can teaching and student participation in the portfolio conference be rated?

The Purpose and Structure of the PAIR Portfolio Conference

The PAIR Portfolio Conference interview and performance assessment protocols adapted for the PAIR project were designed to engage students and teachers to reflect separately on their teaching and learning experiences related to the PAIL work samples (see section 3C of this report) gathered from two PAIR arts integration units administered within one academic year.

The primary purpose of the PAIR Portfolio Conference assessments was to provide an authentic assessment vehicle for rating young children's level of understanding of arts learning and arts integration learning in the context of describing, discussing and demonstrating aspects of their own and their peer's work. In conjunction with the SAIL assessments described earlier in section 3A of this report, the portfolio conference provided a second tool for measuring student meta-cognitive awareness and understanding of arts and arts integration learning goals and processes.

A secondary purpose of the portfolio conference was to give teachers an opportunity to articulate their views on the mission and goals of the PAIR project and then to reflect on their observations of student performance in the portfolio conference in relation to their previous statements. Transcription and rating of teachers responses was intended to provide qualitative evidence and quantitative ratings of teacher's meta-cognitive understanding of the PAIR project as a result of their professional development (Section I of this report] and its possible impact on student learning in this section of the report.

The validity of the analysis of PAIR portfolio conference responses was ensured by the presence of student work chosen by the teacher and students for the conference to represent their best

examples of student learning process and products. The reliability of the analysis was ensured by a defined protocol conducted by an outside facilitator, video documentation and written transcription of each entire session, and a outside scoring team trained to rate each child's and teacher's level of response according to a common scoring rubric.

In *preparation* for their PAIR portfolio conferences, teachers were required to collect and categorize student work stemming from the PAIR Units according type of documentation such as:

- reflective writing – ongoing class journal writing, reflection sheets, etc. focused on goals, vocabulary and activities that the teaching artist and classroom teacher presented;
- planning and process notes (pre-post examples of blueprints, character studies, lyrics, melodic or rhythmic design sheets, etc.);
- culminating products (finished paintings, buildings, sculptures, bio-poems, stories, songs and instrumental pieces with text, choreography, video recorded performances, etc.);
- and final reflections on the quality of work and its value for teacher learning and assessment practice.

The teachers also were required to *randomly select* three students for each portfolio conference from pools of students pre-designated at three categorically different levels of academic achievement. This selection process ensured that teachers had the first-hand opportunity to observe "HAL" students (that is, the High Achieving, Average Achieving and Low Achieving students) as they reflected on their work and interacted with other students while the facilitator conducted the portfolio conference protocol.

Once the teachers categorized the documented student work according to PAIR "student work labels" (as described in the Part I of this report) for the HAL students selected for the conference, the CAPE program staff members were responsible for the final selection of student work items to be discussed in the portfolio conferences. The samples selected from each of the two PAIR Units included individual student work sample, a group work, a video of a performance and/or work in progress and a culminating even. This process ensured that the students and teachers would have strong arts-integrated work to discuss, and that what is discussed would be as consistent as possible across classroom and school comparisons.

Finally, the teacher participants in the portfolio conferences were able to prepare themselves for answering questions about their understanding of the PAIR project and its impact on their professional learning and their student learning outcomes.

* * *

The Portfolio Conference Interview and Performance Assessment Protocol

[designed and facilitated by Dr. Larry Scripp, Principal of the Center for Music and Arts in Education and co-principal investigator of the PAIR project with assistance of Laura Tan Paradis, Research Associate of CAPE]

The object of the video recorded conference was

- *in the first section*, to create an occasion for teachers to briefly express their views of the PAIR project goals and projected outcomes by answering questions posed by the facilitator while students are listening and thinking about what work they will be presenting in the conference. These interview questions coincided with several questions in the teacher surveys thus providing a) an opportunity for the classroom teacher to elaborate on their views of the project and b) a chance for analysts to explore the connections between the teacher descriptions of their class culture and practices and the results from the teacher survey ratings.
- *in the second section*, to create a forum for the students - with no interruption from the teacher – to answer general questions about the program again through expert facilitation. As with the teacher questions in the previous section, the students were challenged to elaborate on many questions posed in the student surveys. In addition, each student had at least six opportunities to explain and interpret their own work products and the work of their peers chosen throughout this 30-minute section of the conference. The facilitator's job was to create a 'flow' of reflection among the students by always challenging all students equally to elaborate on their remarks, to point out details in the work sample, and to provide multiple 'performances of understanding' using two artistic media to demonstrate connections to academic content whenever possible.
- *in the final section,* to allow the teacher – with the students no longer present - to reflect on the performance of her or his students throughout the previous section of the conference. Here the facilitator posed new questions unrelated to the teacher survey that

challenge the teacher to compare their preciously stated goals and anticipated outcomes of the project with what was observed on the part of the students in section two of the conference. The questions stimulated the teachers to reconcile as objectively as possible the student performance in the session with their own expectations for the project and their own and their teaching artist's objectives as co-teachers of the PAIR Units.

The following synopsis of the PAIR portfolio conference protocol provides a way to understand the analytic frameworks and principal questions that will be explored in order to document to rate the teacher and student responses.

* * *

Synopsis Of The PAIR Portfolio Conference Protocol

Portfolio Conference SECTION I. Excerpted Principal Questions from the Pre-Session Question Protocol for the Classroom Teacher (students listening quietly)	
1. Please summarize your view of the PAIR project and its goals. 2. How have the goals of the PAIR project and PAIR project teaching evolved from your work in your classroom? 3. To what extent is collaborative teaching with the teaching artist important to you? 4. To what extent do the PAIR PAIL samples of student work brought to this conference represent fairly the quality and goals of your arts integrated learning projects? 5. To what extent do you think the work that was chosen for this session represents good examples of 'high quality arts integrated curriculum teaching and learning practices?	Teacher's overview of PAIR project, goals, and teaching practices

PPC SECTION II. Focus on Student Participation and Work in Teaching Artist-Classroom Teacher Units (no comments from Classroom Teacher are allowed throughout this section) This Section is administered TWICE, once for each Teaching Artist Unit featured at each PAIR Treatment School	
1. How does the work you brought today reflect your work and your peers' work with [Teaching Artist #1]? (Each student answers individually). 2. You studied [name of book, unit, or culture] with a teaching artist 1. 3. Please briefly tell me about what you did with this artist. What's it essentially all about? (e.g., How did you work with characters? Math Principles? Three-dimensional art? Musical composition, performance, and listening? Dance and choreography? How does it work together? Music and character, math and choreography, etc.) 4. Please show and describe examples of work with Teaching Artist 1 that connections to math, science, language arts or world cultures.	Each student has 6 chances to reflect on their own work, their opinion of the two other student work in two different arts integration projects.

PPC **SECTION II. Focus on Student Participation and Work in Teaching Artist-Classroom Teacher Units** (no comments from Classroom Teacher are allowed throughout this section) **This Section is administered TWICE, once for each Teaching Artist Unit featured at each**

PAIR Treatment School	
1. How does the work you brought today reflect your work and your peers' work with [Teaching Artist #1]? (Each student answers individually). 2. You studied [name of book, unit, or culture] with a teaching artist 1. 3. Please briefly tell me about what you did with this artist. What's it essentially all about? (e.g., How did you work with characters? Math Principles? Three-dimensional art? Musical composition, performance, and listening? Dance and choreography? How does it work together? Music and character, math and choreography, etc.) 4. Please show and describe examples of work with Teaching Artist 1 that connections to math, science, language arts or world cultures.	Each student has sex chances to reflect on their own work, their opinion of the two other student work in two different arts integration projects.
1. Why did your teacher do PAIR arts integration units in your class this year? 2. Should these kinds of units be done in other schools? Why or why not?	Final Reflection Questions

The portfolio conference protocol above is repeated for Teaching Artist Unit 2

SECTION III: Post-Session Teacher-Teaching Artists Question Protocol **(Students are not present):**	
1. What did you notice about the work that the students shared in the portfolio conference session? 2. What did you notice about students' comments on their own and their peer's work? 3. What did you learn about your students from their remarks in the portfolio conference that you might not have known from previous observation during the class activities? 4. Is that the sort of assessment that happens in school and we were just sort of reenacting it, or is this a different way to look at what they're learning in school? 5. To what extent do students understand that parts of the writing/culture/math curriculum also relates to art integration practices?	Reflections on Student work, their ability to explain their work and the work of their peers, their ability to reflect on their learning in arts integration units.
1. From your observation of the student participation and responses, what were differences among the HAL students? Were there significant differences? Were you surprised about the quality of response with regard to the student designation as H, A, or L students. 2. If other teachers were to observe this conference and were not aware of the HAL designations, would they be able to be able to tell which designation applied each of the HAL students? Why or why not? Is this important?	Teacher reflection on differences in response to PC protocol according to pre-designated H (high), A (average), and L (low) rated students.
1. Looking forward in your teaching career, are there ways that you can see PAIR project work impacting your students, your teaching practices, and your own professional development in the future?	Final teacher reflection on future implementation of PAIR influenced teaching practices.

* * *

3D-1: Qualitative Analysis of the PAIR Portfolio Conference Teacher Reflections on Student Learning

Although the PAIR Portfolio Conference protocol was designed initially by Dr. Scripp and his associates to focus primarily on student learning assessment, the teacher evaluation teams noted that the teachers benefitted also from the opportunity to observe, corroborate and reflect on aspects of student learning as they emerged in their observation of portfolio interactions. Dr. Burnaford remarks that

> "The portfolio conferences, with three students and their teacher, are a promising initiative for research and for professional development purposes. While the "three-way conference" involving a parent, student, and teacher, is not new to educators, *the concept of having a teacher watch as her/his students discuss their learning with another adult [facilitator] is an interesting method of assessment* that could be used by teachers and researchers in various ways." [Burnaford et al., 2011]

From Burnaford's point of view, the portfolio transcript analysis provides two kinds of evidence for assessing teacher professional development outcomes: 1) the conferences resulted in valid anecdotal evidence for the presence of effective teaching practices in the PAIR project that are known to be linked to positive student learning outcomes and 2) the teachers benefitted also from the professional learning experience that resulted from observing first-hand the impact of the PAIR project based on students 'performance of understanding' and reflection on their learning in relation to their work samples the teachers had documented and labeled. Based on the portfolio conference interview transcripts and some corroborating evidence from teacher surveys, Dr. Burnaford's team discovered that not only did the PAIR Portfolio Conferences provide an opportunity for teachers to gauge the impact of their teaching practices in the PAIR units, but it provided also an occasion for teachers to reflect on the nature and evolving quality of arts integration *student learning* in the PAIR project.

For the teacher evaluation team, the portfolio conference transcripts became a major data source for discovering four teacher evaluation hypotheses that emerged in the course of studying a mixture of teacher reflective thinking data sources. These hypotheses suggest that the success of the project could be understood best in terms of four characteristic phenomena:

1) The Content Expertise Effect: PAIR arts integrations practices succeed comprehensively when pedagogical content knowledge matters to teachers as they design and implement their arts integration units, *and when teachers assess the quality of discipline-specific content knowledge in relation to arts integration student learning experiences*.

2) The Documenting to Learn Effect: PAIR impact is made visible and most effective when teachers documented student work inspires reflection on teaching and student learning *for both teachers and students.*

3) The Fourth Grade Effect: PAIR Teachers who had repeated opportunities to implement and revise arts integration units over each year of the three-year longitudinal study gained expertise in the ownership and sustainability of the arts integration units which, in turn, should lead to improved levels of student learning in the fourth grade PAIR classrooms over time.

4) The Healy Effect: The PAIR project works best in schools where its new arts integration program initiatives can build on other initiatives that are conducive to integrated teaching and learning practices. The Healy effect is the impact of PAIR on a school culture that already committed to arts learning.

Thus, in this section of the PAIR report, the PAIR Portfolio Conference anecdotal data takes on added significance with regard to hypotheses about student learning.

For example, the 'Content Expertise Effect' as distilled from the portfolio conference transcript data in the teacher evaluation report, also can be interpreted from the learning assessment perspective as evidence that positive student learning outcomes most likely resulted from exemplary PAIR units where content knowledge was most emphasized. The quotations below, used previously to indicate positive teacher evaluation outcomes, now demonstrate the extent to which the teachers witnessed persuasive evidence of student learning based on the fundamental academic concepts and skills featured in the Arts Integration Unit. Thus the reflections of PAIR teachers in the portfolio conference clearly suggest that those arts integration units - which reinforced academic content - also enabled students to develop a deeper and more extensive understanding of concepts fundamental to academic disciplines through its integration with alternative modes of artistic expression.

Math Magnet School PAIR classroom teachers assert in these first few quotations during their Portfolio Conference that they can detect authentic mathematical learning outcomes that stem directly from their students' work with visual arts integration with math projects (designing building or sculptures):

> **Classroom Teacher**: "I think definitely they get...this [the PAIR project] really went along with and supported the math curriculum that we do. We work with hexagons; we work with trapezoids and rhombuses and triangles and fractions, and teaching them parts of a whole, so that really helped with this. They [students] also, which they didn't say [in the PPC], with making these hexagons, they did a ton of measurement to make sure the sides were equal. Some of them did better jobs than others, but they did a lot of measurement, which was something that they really, really needed for their work. We do go over the concepts, but this [art integration] is like a whole – it makes them generalize the information that they've learned, which I think is kind of what you're trying to do."

> **Classroom Teacher** "Scale is a very difficult concept. It's a difficult concept, especially application of scale. You know, you can do it on pencil and paper, but to actually apply it [in the PAIR unit] is a whole other ball game. Not only can they do it on pencil and paper, but they could apply it, you know, apply it to build something or construct something new,...you know, that is really difficult to get a student to comprehend that. I think it's beneficial on the tests because it will help them retain it. Once you have applied it and used it, you have a tendency to retain it more than, okay, this is how you do it and it's gone. I think it's of great value because the kids were able to talk about math in a reasonable way. They truly have a much better understanding than they would have if I said open up to Page 200 and let's do scale."

Similarly, teachers in the PAIR English Language Arts Magnet Schools remarked in their portfolio conferences that deep understanding of linguistic concepts had been enhanced by student work in their PAIR acting and music units:

> **Classroom Teacher:** "The main purpose that we were, what we were looking to do was on all writings, regardless of whether it's ISAT or what we're using it for, we always want the student's own voice to come through and not just mimic what it is that they've read, but we want their input, their take on things. So this [PAIR Unit work in drama and music] was helping them develop confidence in their writing so that their own true personality and their voice would come through. When they have to make connections, it's harder for them to make a connection if it's not their own voice and their own real experiences."

> **Classroom Teacher:** "For me, as a teacher, it [PAIR] has really helped me to instruct my students how to write better, I mean, how to really look at not only when we read a story, about characterization, motivation, plot. I will tell you this class, we have read some really difficult books this year, higher level, higher thinking, and I really do think having the PAIR program has helped my students look at characters, understand plot, conflict, resolution a lot better than my other classes."

Other teachers commented on the need for teachers to look more deeply into what really is evidence of conceptual understanding that cannot be derived from relatively superficial assessment of writing or reading skills.

> **Classroom Teacher:** "Rather than look for their simple grammatical errors and sentence structures and development of the story, I kind of put that to the side now—and I don't look at their mistakes in spelling. I look at the ideas and things developed and the process that they took in writing their stories. That's a big difference."

The teachers also report that learning more about the arts disciplines themselves helped students advance their language literacy skills:

> **Classroom Teacher:** "This form of instruction is very beneficial for students. It provides them with two different interpretations or styles of teaching…"
>
> "Well, for myself, the value of collaborating with an artist is that I understand what their technique is and how it can help my students advance in their development of writing. And I keep referring to the writing because that's what our focus is, is developing their writing. So collaborating with them helps me gain a deeper understanding of their arts and how I can integrate it with what I do in my classroom ... "

> **Classroom Teacher:** "We were trying to expand on what the students had learned in fourth grade and what they had learned with PAIR, and we were focusing on having the students develop their voice and recognizing point of view with main characters in a fiction story. And Reggie [CAPE drama teaching artist] specifically worked on how do you identify a person's point of view and how do you recognize someone else's point of view … "

> **Classroom Teacher:** "And then what Charlie [CAPE music teaching artist] did was really expand on the voice area of it. And we worked on developing a theme song for the main characters, so if the main character's personality were put to a musical beat, what would it sound like. And then they had to incorporate what they were doing with Reggie in character description to come up with what they would think the beat of someone's theme music would be. So they had to look at the physical characteristics, emotional characteristics, how people interact with each other."

The portfolio conference process and the role of systematic documentation in providing qualitative evidence of student learning is a by-product of the "Documenting to Learn Effect" hypothesized above. The conference depends on the quality of documented work samples and the degree to which teachers are able to reflect on student learning outcomes in conference protocols where students are challenged to interpret their own work and speculate on their learning that has resulted from this work.

Knowing that teacher reflection on student learning is bourn from observing student performance in portfolio conferences, the teacher testimony becomes significantly more compelling in its conclusions.

> **Classroom Teacher:** "… you know, you have a hard time throwing something like this [PAIR Portfolio work] out. You do get to kind of peek inside their psyche and kind of get an idea of what's going on in their heads. And then we have evidence at the end, and it's evidence that they can be proud of. It's beautiful and it's something that we don't want to go away. We want to hang it and we want to be proud of it. That's what I feel is nice about this whole process."

> **Classroom Teacher:** "I would...I see that now, when we're writing and we do writer's workshop almost every day, and their characters are better defined, there is more detail about their characters. We're learning...I really think I've seen a progression, and even for the time that, you know, what I had in their portfolios and now, months later, it's a better awareness of how you write a story, that there has to be a story. You just can't write then…, then…, then…, then.. , then…, [but] that there's characters in their stories – and this goes for my whole class, not just these three students – that they're really getting the idea of how to put a story together. But what I see is each piece gets better and better."

Furthermore, in conferences where the "Fourth Grade Effect" and/or the "Healy Effect" of long-term experience with arts integration collaborative PAIR curricular and teaching practices is present, testimonials from teachers about student learning garner even greater credibility.

> **Classroom Teacher:** "Well, the original goal was to see how the arts itself affected the students' abilities as far as their writing, and how the different arts, for example, music, because we had a percussionist and a playwright, would enhance the child's development of characters and stories. The first year was pretty new. We still had to develop some skills as far as what the artists were teaching the children and my own understanding of what it was. This year I think it's been...I've seen...to get the kids to actually think more critically. Last year it was okay, okay, it's going through the machinations, for the most part. This is what I'm supposed to do, this is what we did."

> **Classroom Teacher:** " … I think these students are much better at relating it to math than last year. I think they were much more...they expressed themselves much better mathematically, which is because when I was in the room, we really stressed this is not only art. I mean, we talked about it before, but we really stressed it this time because I didn't want it to happen again where they didn't see the value of the math in the project."

> **Classroom Teacher:** " … how much it [PAIR] has changed from one year to two years, especially with me being in three years. Just the comfort level and knowing what to do, and I was able to help the kids even better when the artist was outside the room, because then I would reinforce the concepts and have them write for them."

> **Teacher:** "The second year I was able to implement it [PAIR] in some of the stories that we read, and bringing out the motivations of the characters, and how the characters changed throughout the story. And this year the students used the techniques that (the teaching artist), our playwright, taught them, to develop their own stories, and on their final pieces, the Young Authors piece that we write every year. This year it seemed to go a lot easier. We didn't collaborate as much as we have in the past, or emailed or spoke as frequently as we did in the past, and I think that comes from the fact that we've been together for three years. We worked on the same things for three years."

* * *

PAIR Portfolio Conference Qualitative Data as a window onto meta-cognitive and social-emotional developmental outcomes of arts integration learning

From the viewpoint of qualitative data analysis of student learning, the purpose of the facilitated PAIR portfolio conferences was to provide evidence for arts integration learning outcomes in the context of observing student reflection on their own work and the work of their classmates in the project.

The PAIR Portfolio Conference protocol required that student choose their most meaningful work, explain their work to their peers, and demonstrate their understanding with small performance tasks such as demonstrating their choreography, drumming, storytelling or improvisatory acting, or commenting on their written work, drawings, sculpture, blueprints for buildings, or songs.

The goal of this format was to elicit aspects of cognitive dimensions of arts and academic learning, meta-cognitive aspects of learning transfer and integration across subject areas, and to distill aspects of social-emotional learning that occurred in the PAIR program units.

The following five vignettes illustrate the emergence of these themes that emanate from both the perception of the PAIR project from the teacher and the reflective understanding of the students who participated fully in the PAIR project.

* * *

3D-2: Student Reflective Understanding of Their Personal and Collaborative Art Integration Learning: Four PAIR Portfolio Conference Vignettes

* * *

3D-2 Portfolio Conference Vignette I: *Math-Arts Integration Case Study School*

The cognitive and social contexts of arts integrated math teaching and learning

During the math-dance-visual art portfolio session at Thorp, the three HAL students A, C, and K all talk in the portfolio conference about their collective work from the point of view of the integration of math, sculpture, and social studies. It was completely evident to the conference facilitator and those analyzing the video that these students were as much excited to talk about the process or creating sculpture and its relationship to their mathematically precise blueprints of the various phases of their design work as they were anxious to explore social issues that extended beyond the boundaries of their school studies into their families and community.

In the first section of the portfolio conference, the teacher extols the value of her PAIR project emphasis on students' reflection on their arts integration work as an essential strategy for assessing alternative aspects of progress in their elementary school education. From her perspective, upper elementary grade students – and perhaps *especially* in math magnet schools – are evaluated only by test scores and district defined benchmarks in ways that do not capture process and diverse social context to the extent of their PAIR arts integration learning units. Consequently, teaching practices tend to reflect a certain impersonal and socially de-contextualized aspects of learning. For many PAIR teachers, arts integration thus becomes a way for teachers to structure and evaluate students learning experiences through creative process and reflection of the personal, collaborative and socially relevant aspects of their work.

The following statements culled from the conference data are typical of the many discussions that emerged after the PAIR units were completed and teachers and students faced the challenge of describing the most important features and results of the project.

In this case, the teacher begins with her passionate plea for more individual and more varied forms of assessment:

> **Classroom Teacher:** "Well, those are the people that they want to test score … yet these are individual children with individual needs, and this program allows us to address individual parts of the child."
> "This [program] makes it more enjoyable for them to learn it. They do learn it because it comes in different forms. It's not just reading it in a book. It's using your brain. It's being imaginative. These kids don't get to use their imagination anymore."

After hearing the teacher respond to questions, the student must speak for themselves with no further comment from the teacher for the next 30 minutes of discussion about their work. After much questioning about their portfolio work, and insisting that each type of HAL child have equal time to reflect on their arts integration work both individually and collaboratively, a common theme emerges: *arts integration teaching and learning is personal, highly engaging and relevant to their understanding of themselves in the context of important social issues.*

Student discourse about their own work and the work of their peers reveals the intensive interest students have in investing in social, personal and emotional aspects of their arts and arts integration learning processes.

> **Student C**: "I'm very happy that we get to have this program because it teaches us how to work on a team and it also teaches us math with sculpting."
>
> "And I was [thinking] that they colored, they painted the gun green because money is green. And I see a little bit of blood, red paint, on top of the money and on the gun."
>
> **Facilitator**: Look what's coming out of the front of it. Why is there money coming out of the gun?
>
> **Student K**: Because there's a lot of money that's going up and a lot of taxes is going up, and a lot of people don't have that much money.
>
> **Facilitator**: Wow. And. so what do you think this all means, altogether [pointing to the sculpture work]?
>
> **Student K**: ...it means altogether it's a war."
>
> **Facilitator**: Mm-hmm. Now, some people might think that kids shouldn't think too much about that, and I'm sure you don't just think about death. But why do you think you gravitate to that in this particular task?
>
> **Student A:** I think about death and war right now because my uncle was in Iraq and Afghanistan, so...and then he didn't die, though."

* * *

3D-2 Portfolio Conference Vignette II: *World Language-Arts Integration Case Study School*

When students connect positive disposition, good work, and leadership with the process of imaginative learning in and through the arts

In this vignette, a different teacher expresses the desire that her students 'go deeper' to understand themselves in social contexts in ways that are aligned with the stated mission of world language schools.

> **Classroom Teacher: … "**the artist, the artwork of it and just how you're reaching out to the kids with the different foods and different countries, things of that nature, … I can tie that in so much with my curriculum, with writing and with history, so that the kids can take this and grow with it."
>
> "You know, a lot of kids, when you ask them 'who are you?', at first they're going to tell you their name, my birthday, I live here. But it [this project] made the kids think a little deeper and dig really to find out who are they exactly, where do they come from. So it helped the students open up with their thinking and take it to the next level. You know, and it helped them, too, to go about doing research and finding out more."

At first, what the students say about their imaginative work is surprisingly specific. Details about how they created a new nation through their visual artwork (maps and passport cards) and musical composition (a national anthem) with the PAIR teaching artists are important topics that the portfolio conference facilitator can use to draw out the meaning of the project from each student's perspective.

All Students recite:	"Oh, say Symbolia, the great Symbolia, how dear you are to me. Your soda flowing [figs and popping] beneath your starry tree. We promise to take care of you and never, ever roam. Oh, great Symbolia, our home sweet home." …
Student S:	It's our anthem.
Facilitator:	It's your anthem? What's an anthem?
Student N:	Like a pledge.
Student S:	It's like a song that we promise to our country.
Facilitator:	Oh. Is that what a pledge is? A promise, right?
Student S:	It's like the pledge of allegiance. We sing that to make a promise to America.
Facilitator:	So what have you promised to your own Symbolia?
Student A:	We promise to never, ever roam and...
Facilitator:	What does it mean never, ever roam?

Student A: We would never go somewhere else. …

Facilitator: Ah. Okay, what do you think, Nicholas? What's an anthem? What's a pledge? What is this all about?

But when asked the importance of their work, the students in another classroom responded in terms of their own social-emotional values such as patience or pride that were specifically represented in the text of the anthem and their self portraits, and, in addition, observed during the process of creating their arts integration work.

Student R: "Patience is like relaxing, just relaxing, just like not getting mad, not, well, not being mean, like that. Just being patient and being nice.

Facilitator: And why is that important?

Student B: Because I don't like getting mad at people."

Student K (following up on previous remarks): Because if we mess up, then she will say, like, you should do it like that, and then, like, she'll come back later and she wouldn't be mad if you didn't get [through] the whole project that day.

Facilitator: So why is that important?

Student K: It's important because if you rush, then you probably won't do a good job, and then it'll be a waste of time."

Student D: "and guess what? We voted for a name for our nation. Then we voted for a president, and I won!

Facilitator: (laughing) So how did it feel to be president of an imaginary nation?

Student D: (in a serious tone) Well, I was proud."

Whereas in Vignette I the students insisted on social context, in Vignette II students personally self-identified with the concepts and dispositions that they used as the basis for their creative work and their understanding of social-emotional values in their work. The next vignette illustrates how student demonstrate their understanding of the social-emotional aspects of character, motivation and conflict resolution in the stories they read through multiple artistic media.

*** * ***

3D-2 Portfolio Conference Vignette III: *Language Arts-Arts Integration Case Study School*

Expressing understanding of characters and their actions through multiple artistic media

In this vignette the teacher focused on her students' ability to understand characters and what motivate characters to act. Reflecting on the quality of her student reflections during the portfolio conference, she was very pleased to see that her ELL students in particular were able to speak articulately about the personality traits of characters in the book they read and the creative writing they did. What astonished her was the students' ability to continue to come up with news ways of expressing their understanding several months after the project had ended. Not only did the students describe how the characters felt and what they did, but they also were willing and able also to demonstrate those character traits and actions *and their own feelings about the characters* with drumming patterns spontaneously demonstrated in the portfolio session.

Classroom Teacher: The main purpose that we were looking to do was to focus on all areas of writing, regardless of whether it's ISAT or what we're using it for, we always want the student's own voice to come through and not just mimic what it is that they've read, but we want their input, their take on things. So this was helping them develop confidence in their writing so that their own true personality and their voice would come through. When they have to make connections, it's harder for them to make a connection if it's not their own voice and their own real experience.

Student M: [plays on the hand drum *two uneven slaps, two even slaps, one long boom, two even slaps, one long boom* and repeats the whole pattern*)*

Facilitator: So why is that express the character's niceness? Because … ?

Student M: I think it has feeling in it and it makes me feel happy. … how to describe things, how to play instruments, how to express your feelings playing instruments.

Facilitator: And why is that important to do?

Student M: Cause you don't want your feelings to be all bottled up. It's good to let them out.

Student J*:* [announces that she will drum the feeling of her favorite character being 'paranoid']

Student V: [listens to Student J and then announces that he will "use his drum to describe the 'mysterious side' of his character].

[The teacher then announces it time for lunch although the students are very reluctant to leave the session because they say they need more time to play their drum characters!]

* * *

3D-2 Portfolio Conference Vignette IV: *Math-Arts Integration Case Study School*

Students who gain self-confidence in the context of arts integration learning projects

In this final vignette the classroom teacher describes the need to raise student confidence as learners during both the beginning and ending segments of the conference. Student M also emerges as a leading example for self-confidence in the student part of the portfolio conference, prompting the teacher to say more about how much M overcame his lack of self confidence at the beginning of the school year as he excelled in the PAIR math-visual arts activities

> **Classroom Teacher:** "… we got a little bit of a late start on really getting in and diving into creating our blueprints and buildings, but I feel like their projects are now really coming along beautifully, and the students are really working and using their math skills. You can see their measurement skills in their drawings and I could see how these skills informed their discussions during the conference, which I think were great, very successful."
>
> *[speaking about student engagement in the portfolio conference]* **…** and that's like Student M to stutter when he gets nervous, … yet watching him over the course of this project, he really took a leadership role in this, and he, like, really directed the project. The project became his basic vision and he gave everyone a job, and other people were measuring and doing, … and the group work has been really successful. Whereas before maybe he'd be quiet and just kind of sit back and do his own stuff by himself."

The following dialog among the portfolio conference illustrates this perception of student M's growing math skills and social-emotional development through the PAIR project.

> **Student A: "**The diameter, so if you're saying that the [ladder] is right here, then this is 12, where from here it's 12, and another six inches, so it might be 18.
>
> **Student M**: Yeah, but I'm saying that the dome is going to be equal to the size of the bottom, which is—
>
> Facilitator: Where is the dome now, from where to where?
>
> **Student M**: From here to here...see this drawing.
>
> Facilitator: That whole thing is going to be equal to what?
>
> **Student M**: This is inside the dome right here.
>
> Facilitator: Right.
>
> **Student M**: See these lines? They meet right here. That's—
>
> Facilitator: So how long is it from here to here?

Student M: From the—

Facilitator: Here to here.

Student M: Oh, you're thinking of the height, which is going to be six inches." *(answer is not only correct but more specific than the question required)*

* * *

3D-3: General Discussion and Summary of PAIR Portfolio Conference Vignettes

The four vignettes comprise a small sample case-by-case analysis of the portfolio conference data collected in the final year of the PAIR project. The portfolio conference vignettes capture the rhythm of the protocol through sequential excerpts of teacher interview responses, facilitated small group interview discussion and follow-up performance tasks. In each case the teacher provided initial assumptions about the broad goals and outcomes of the project, the students were challenged to interpret and reflect on their work, and finally the teachers were asked to reflect on each student's performance in the conference in light of their assumptions.

Although the portfolio protocol framework was standardized, many cognitively and socially charged discussions emerged that concepts and skills indigenous to academic subjects (such as math reasoning and measurement skills) with rigorous arts learning and skills. The discussions, work samples and problem solving cognitive skills are brimming with enthusiasm about connections among these disciplines that appear to be equally distributed among visual art, dance and math, or reading, writing, drama and music, or social-cultural studies, visual art, and music – depending on the PAIR unit and the focus of each magnet school.

Analysis of the dialogue among students, however, suggests that meta-cognitive skills combined with social-personal understanding are also primary outcomes of this work. While creating sculptures from blueprints may sound like a marriage between cognitive skills in math and visual art, first and foremost for most of the students was the importance of social context, social values and social-personal development in their PAIR work. Depicting objects accurately required cognitive skills from math and visual art, but the interpretation of the work, yet the portfolio discussion among the students became most highly engaged over the choice of colors green and red for a gun sculpture that turned out to be all about the social context of war and greed in our

society, a searing discussion that revealed how children melded their most personal and family experiences into the creation of their school work.

If arts integration teaching and learning is personal, highly engaging and relevant to their understanding of themselves in the context of important social issues, then portfolio assessment protocols appear to serve as a powerful authentic assessment of this wide range of arts integration learning outcomes.

In the world language portfolio conference vignette, arts integration didn't merely serve to illustrate what they were learning in academic subjects or arts classes; it meant employing self and group exploration activities in order to go deeper into their own identity in ways that allowed them to imagine better, more idealistic societies. The music, lyrics, maps and descriptions the students created for these new societies teemed with the valued dispositions such as patience, pride, and good work.

Based on the exemplary vignette analysis, we can determine with increasing confidence that the portfolio conference format provided PAIR project participants with opportunities for

a) PAIR teachers to test their assumptions about arts integrated teaching and learning through observation and reflection on student work and performances throughout the facilitated conference protocols;
b) The facilitator to challenge students to demonstrate a significant grasp of cognitive and meta cognitive aspects of arts integrated learning transfer across disciplines; and
c) Both teachers and student to nurture the essence of social, personal and emotional understanding throughout their joint experience of the arts integration teaching and learning process.

The next two sections provide a more detailed qualitative analysis of portfolio conference outcomes according to differences in pre-designated academic performance student cohorts and meta-cognitive aspects of student arts integration learning outcomes.

* * *

3D-4: Differences in PAIR Portfolio Conference performance among previously designated HAL academic achieving students

Blurring the boundaries among high, average and low achieving students

If the PAIR portfolio conference transcripts suggest that arts integration teaching and learning practices can be best assessed in the form of interactive interview and performance task formats, we can now ask, "To what degree are portfolio conference data independent of, or linked with, standardized academic or arts learning assessments?" To what extent do the HAL student performance in the portfolio conferences suggest that students can be categorized in the same levels as their previous designation as high, average or low academic performing students? Is there any evidence that arts integration assessment devices may bring more equity and opportunity for pre-designated low rated students?

One of the most distinctive findings distilled from analysis of portfolio conference qualitative analysis is that PAIR teachers consistently report a surprising change in their perception of their student academic standing in their classroom.

Before every portfolio conference teachers were told to bring three students they had previously designated as high, medium and low according to their academic standing as defined by standardized test scores. However, the teachers were told not to reveal the designation of academic standing of their students to the facilitator until after the conference was concluded, thus ensuring that each student had the same amount of time to answer questions and expectations for follow-up questions and demonstrations of understanding were equally applied to all three students.

Early on in the process of data collection, the facilitator reported a surprising perception: it was sometimes easy to spot the over-achieving demeanor of the highly rated academic students in relation to the attitude and speed of response of the average and low rated students, but it was more difficult to distinguish the quality of their responses. It was often more difficult still to distinguish the low and average designated academic learners.

Typically, in the course of the conference protocol final reflection questions (when students were not present) the teachers revealed that it was hard to tell what academic level the student were categorized if an observer did not previously know the students:

"**Classroom Teacher**: Student E was chosen as the high rated student, and he spoke quite eloquently. He uses a nice vocabulary choice words, while Student J is the below performance because she has trouble expressing herself, and she gets flustered, and you can tell, you know, there are certain parts where goes "I don't know, 'cause I get confused," and she does that normally in the classroom. And Student N is my average student."

Facilitator: Yet, just in the way Student J [L] actually came up with and discussed what was going on in the project, were there big differences in the sense of what she was saying and what Student J [A] was saying?

Classroom Teacher: No, no.

Facilitator: So in that sense, there wasn't a big gap between those two?

Classroom Teacher: There was not a big gap."

As indicated in the portfolio conference protocol, the standard question asked of all teachers in the last section of the conference, sets the stage for discussing the phenomenon of 'blurred boundaries' of previous academic standing with regard to the student's 'performance of understanding' during the session:

Facilitator: … for anybody watching this tape, would they know which one of these kids is the pre-designated highest performer in your classroom, or is the portfolio a completely different way to rate the academic performance of a student?

Classroom Teacher: People watching the video clip that don't know them?

Facilitator: Yes, the people don't know anything about the students they watching on video documentation for the first time.

Classroom Teacher: I think that would be hard to tell.

Facilitator: Is that important?

Once the teacher admits that it is hard to tell who the high rated academic students are in the portfolio conference, the follow-up question raise questions about the whole purpose of the PAIR project:

Facilitator: [Is that important?] Does that mean that it's [the portfolio conference responses are] less related to academic achievement?

Classroom Teacher: I think so. I think it would be very important because it shows how they each...I mean, they're independent in their own way, and yes, as far as from what I know in the classroom, there was a low, medium and high, but when it comes to the arts and music, they work together, and it's not as obvious.

Facilitator: Is that a good thing?

> **Classroom Teacher**: I think it's a great thing.
>
> **Facilitator**: Because?
>
> **Classroom Teacher**: I think it's a great thing because, I mean, it brings everybody on the same level. It seems to. I mean, in this case it brought everyone on the same level, and they weren't stereotyped or...people watching the tape wouldn't be able to know. And I think that's important because you're not going to say, oh, you're the low group or you're the high group. Who knows? You couldn't tell from that."

It is not uncommon for the teachers who participated in the PAIR Portfolio Conferences to report that the stereotypical characteristics of high academic achievers based on their paper and pencil work in class and standardized test taking acumen do not reliably predict their ability to think on their feet during the portfolio sessions. It appears evidence of arts integrated learning is based instead on the student's ability to interpret their work and the work of their classmates as they adapt to the general discussion. Students observed in small group discussion (without teacher intervention) create a wider range of group learning dynamics in action, and in some cases, an assessment of individual student learning capacities that are at odds with their previous HAL designations of academic achievement.

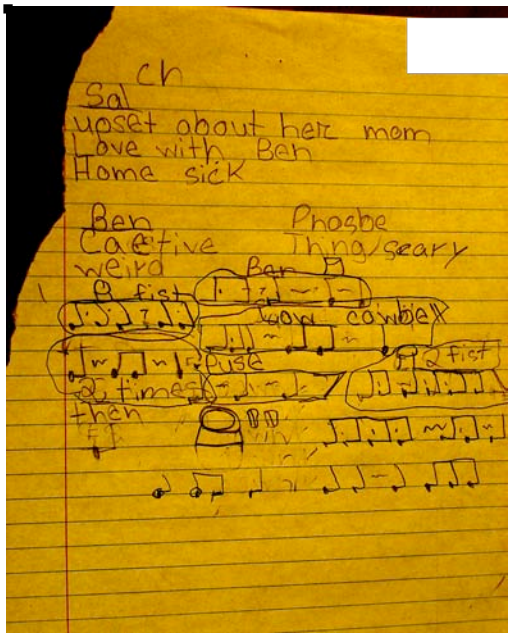
Another teacher adds that, although highly rated academic skills may be easy to see when only looking at differences in the quality of the student writing sample in the PAIR project [see writing example below], low rated students appear to be more engaged in the process of reflecting on their choices and comment on new ideas as they emerge in discussion. As one teacher notes, her designated low writing skill student seemed to enjoy the writing process more the other students and, according to her judgment, this student's writing increasingly revealed so much more of the thought process, though there are formidable technical writing problems.

> **Classroom Teacher**: "I don't know. I'm just saying that every year I look at students that are my low level or very disinterested in school. I mean, Student I and Student M are very high-level students. Student A is not. But yet it seems that for some reason, and I think it's the artists bringing...showing them that they can write, a different approach to it, that it's fun, it makes them think, oh, I can do that."
>
> **Facilitator**: [Can you say more about that?]
>
> **Classroom Teacher**: "What I've also found, writing was a chore [for my students], and I'm seeing that it's not a chore, that they're beginning to enjoy writing - especially the low-rated Student. She writes everything backwards. I mean, her papers are flipped, so I do think she has some type of dyslexia or something like that. But her writing, even

> though it's not, you know, the highest caliber, I've seen so much more thought process, and I think she really is...she almost enjoys it more than the other two students that I brought [to the conference]."

Another teacher reiterates why NOT being able to spot the stereotypical behaviors of high or low academic proficiency in the context of the portfolio conference is a desirable and positive social outcome for the PAIR project:

> **Classroom Teacher**: "I think it's a great thing because, I mean, it brings everybody on the same level. It seems to. I mean, in this case it brought everyone on the same level, and they weren't stereotyped or...people watching the tape wouldn't be able to know. And I think that's important because you're not going to say, oh, you're the low group or you're the high group. Who knows? You couldn't tell from that."



The conference artifact on the left is an example of a character study artifact chosen for discussion from the PAIR Portfolio Conference vignette described above. As many teachers report, the obvious superficial signs of technical writing problems such as handwriting and spelling errors does create a bias about the cognitive capacities of the particular student. Yet, in this portfolio conference, the student displays the ability to discuss her perception of character in their reading assignments and is able to demonstrate her ability to create percussion patterns that she imagines match of the character's state of mind in the story.

As the Language Arts Magnet School teacher of the student sample above reflects in the last section of the portfolio conference:

> **Classroom** Teacher: "Now that they're [the students] not here, I can speak freely about their growth and development. They're language learners, English Language Learners, so during this year I didn't see, or didn't get the growth in the writing as I expected to see in my previous classes, especially the first year with the project in a mono English classroom. This year they were still developing the writing skills. It's the last step that they develop as bilingual students. However, their confidence and ability to speak to others using the English language has grown significantly, and I think it's because of the fact that they can express their knowledge through other means rather than the simple

> paper and pencil. Because paper and pencil basically means they're going to be graded as far as what they write and how they spell things. But when they have the opportunity to say it to you, they know that they can express their ideas, what they know and what they've learned, a whole lot more freely than they can on paper and pencil."

Many other teachers finish the PAIR Portfolio Conference reflection section with similar statements about preconceptions about inherent differences in cognitive capacities among HAL-rated students. As in the case below, teachers voice the need to assess students in ways that more resemble the interactive arts integration portfolio conference demonstrations they were able to experience first-hand in the PAIR project:

> **Facilitator:** "[Do] You think people watching the tape would know that they're low, medium and high?"
>
> **Classroom Teacher**: "I don't think so. That's why I'm very glad, because all three were answering, all three had something to say, and it wasn't tense. They were really putting things together and everything else. And I don't think...this is part of the program where I see that they're integrated, and you can't really tell the difference."
>
> "I don't think if you just looked at them you could say like, oh, well, this one really struggles with math and this one doesn't– I mean, I don't think you can really tell across the projects."

* * *

3D-5: Evidence of meta-cognitive aspects of student arts integration learning outcomes in PAIR Portfolio Conference data

One explanation of blurring the distinction among the HAL classified learning types may rest in what may be called more global aspects of meta-cognitive learning outcomes that can be culled from teacher characterizations of learning revealed in the PAIR Portfolio Conferences.

According to teacher reflections during the portfolio conferences, the value of arts integration is that it brings a heightened awareness of multiple aspects of learning through its engagement of 'multiple Intelligences' and multiple modalities of learning, its demand for the expression of understanding through multiple media, and the presence of meta-cognitive strategies about learning that are both implicitly and explicitly supported in the PAIR unit activities. As a result

of these 'optimal conditions' for learning, teachers hope that the PAIR unit will also result in evidence of higher-level critical thinking skills for their students.

> **Classroom Teacher**: "And I find that it [PAIR] appeals to different intelligences. They actually get up and get out of their seats. They get around and move. It's not just visual and listening skills. I think a lot of students, in this class especially, are kinesthetic. They need to get up and move and do things with their hands. I think it helped them to develop a better understanding of the material. As we read, we tried to connect to a novel in reading. I think all the students were engaged. I think they were very excited about learning. It was kind of fun to be in the classroom. And I think it...I'm hoping it helped develop higher level thinking skills."

When the facilitator asked teachers to comment on what they saw with particular attention to whether PAIR student learning outcomes (as measured on ISAT test scores) were evident in the portfolio assessment session, it was clear that many teachers viewed the conference as an occasion for a unique form of assessment that makes connections with creative thinking and collaborative learning as opposed to 'paper and pencil' evaluations based solely on single answer or multiple choice responses.

> **Classroom Teacher**: "I think it brings a lot of creativity into my classroom and it really gives students a lot more vocabulary that they don't use daily. It kind of like gave them another way to express themselves and to learn in a different way other than the way that we teach traditionally, you know, through papers and through the board …"
>
> "Go beyond paper and pencils. Involve more art. Because [in] one way it kind of like got the creativity out there of their head and also allowed them to do things that is more fun, not all like paper and pencil 'boring stuff', you know. And I think when they are having fun doing things that they like to do, they were able to give me a little bit more of what they were asked. So just like the biography, like this [as presented in the portfolio conference], we can use it next year, because I know how to do it. I know step by step how to do it. I think the kids feel more comfortable to express themselves if they are allowed to think through it, just like the portrait they did before – the kind of food they eat, their facial feature, the color of their face represent who they are. So instead of how we usually do, oh, a biography, you need an introduction, what day they were born, how they're...it's more...like going through this art project let them kind of like a brainstorm, let them think through it before they put it in writing. So it's kind of like help them in their thinking process and their final product. Not only just they do it because the teacher tells them to."

The impression that many teachers garnered from the portfolio conference sessions is that reflective understanding of arts integration in collaborative assessment conferences may bring about a new way to appreciate and evaluate student academic performance. These alternative

assessments were more apt to draw out reflective thinking skills in the context of interactive discussion that supported students' articulation and understanding of academic and artistic content in arts plus arts integration practices featured in the PAIR project. Furthermore many teachers expressed the idea that these alternative assessments would provide evidence of learning that would exceed expectations teachers would normally have for students based on their prior academic work. Portfolio conference transcripts reveal this alternative form of assessment allowed teachers to view the HAL student' capacities in new ways that resulted in understanding student level of subject area skills and understanding differently.

Moreover, the teacher portfolio conference reflections reveal that most students, and especially those pre-ordained as average or low performing students, seem now less inclined to deliver short, 'correct' answers akin to responses expected on their standardized tests, but much more inclined to elaborate on their working or learning processes, more likely to imagine new possibilities and responses to challenging work they have accomplished, and more apt to focus on their social-personal awareness, meta-cognitive skills and their emotional commitment to their individual work and with others.

Overall, the qualitative findings from the relatively small sample of portfolio conference data suggest that the value of the PAIR project may be best indicated most vividly in the range of student response to portfolio conference tasks. The evidence of long-term memory of the PAIR units and the students' ability to reflect on purpose and value of their arts integration work suggest also that these attributes may also be strongly associated with other kinds of student learning outcomes previously discussed. By rating reliably the student and teacher responses according to levels of articulation, perspective, and cognitive, the portfolio conference outcomes can be statistically examined for their strength and pattern of association with the previously reported teacher professional development, student arts and student academic student outcomes in the final chapters of the comprehensive report (parts 3E-3G).

* * *

3D-6: Rating Student and Teacher Responses in PAIR Portfolio Conferences

Rating Scales for Teacher and Student Portfolio Conference Responses

While qualitative analyses of teacher and student portfolio comments were explored according to their focus on topics related to teacher professional learning and types of student learning outcomes, this section will present a scoring rubric for rating both teacher and students responses according to levels of articulation of PAIR learning goals, processes, and conceptual understanding of interdisciplinary content integration.

Once these results are quantified, these data can be used to test for the degree of association of portfolio conference responses with other student learning outcomes distilled from the student surveys, SAIL interviews, and even ISAT test scores. Later on in this report, there will be a discovery of links between the PAIR portfolio conference results and other teacher survey and professional development data.

Summary of PAIR Portfolio Conference Response Scoring Process

The process for quantifying the level of complexity of teacher and student portfolio conference responses is as follows:

Scoring:

Teachers:

- Teachers were each scored on discrete answers (1-3 paragraphs) within the transcript and observed on video replay, in six separate categories. The first three were scored from teacher remarks before students spoke: (1) perceived goal of the PAIR project, (2) relationship (including planning and lesson execution) with the teaching artists, (3) expectations of for student achievement. The 4th, 5th, and 6th categories were asked during the teacher's reflection on student performance: (4) summary of student work within the art form, (5) summary of student art-work in relation to other disciplines, and (6) ability to articulate change or an opinion in their own teaching and assessment practices
- Teachers have two sets of scores, one scored from a piloted sample in 2010, and the second from a comprehensive administration of the conferences for all grade levels in the summer of 2011.

Students:

- Each student has twelve separate scores per conference, or six per art form. These include: (1) level of understanding of project activity, (2) level of meaning making expressed about student's own work within the art form, (3) level of meaning making expressed about other students' work within the art form, (4) level of meaning making of the student's own art work in relation to its application – and understanding of content – in other disciplines, (5) level of meaning making of other students' art work in relation to its application – and understanding of content – in other disciplines, (6) ability to summarize discussion and critique of student work in terms of arts and arts integrated learning.
- Multiple categories of response could be applied to a single response, as a student's responses in a certain category may well inform another (for instance, in comparing one's own artwork to one's peers).
- Especially in the language arts schools, there was not always a clear division between the art form (drama/character study) and the core discipline (language arts/reading). Therefore, those student scores often inform each other in categories 2-5.

3D-6 Table 1: Scoring Rubric for Levels of Response for Teacher and Student Response Data

NA – not applicable *(if a question is not asked)*
NR – not relevant to the question

- 1.0-1.5 – Single Dimensional: Generic Response; *diffuse, highly, unelaborated general statements (like "I don't know" or "It was fun") that lacked detail*
- 2.0-2.5 – Multiple Single Dimensions: Several Concrete Concepts*; some detail, but lacking relational thinking, cohesion, and cause and effect statements*
- 3.0-3.5 – Coordination of Dimensions: Detailed Concepts and Explanation: *and a focus on relationships and causal links among the elements of the response*
- 4.0-4.5 – Systemic Reflective Understanding: Evidence of Comprehensive Understanding: *demonstrating systematic knowledge and understanding of concepts, processes and other examples described and their persuasive description of the links and associations among all factors described.*

Inter-rater scoring reliability

- The rating team achieved reliability in the scoring of the responses by scoring different types and levels of the response. Responses focused on all questions (as indicated above) were scored on a trial basis for inter-rater reliability of complexity responses. In addition, responses that focused on arts and academic topics, and those centered on cognitive, meta-cognitive and social-emotional development issues were all included in these scoring reliability sessions

Individual differences in PAIR units, possible limitations of the rating system

- The *comparability of scoring* is challenged by the diversity of content explored in the portfolio sessions. For the Math Arts Magnet Schools and their partnering Arts Focus Schools, the content was focus on conventional fundamental concepts of measurement and quantification in math-visual arts-dance problem-solving arts integration tasks. Similarly, the English Language Magnet Schools and their partnering Arts Focus Schools focused on traditional issues of character, setting and plot in literature. The goal of the world language schools and their Arts Focus School partners with respect to the PAIR project fundamentally differed from the other two cohorts. The mission of the World Language schools of was to get their students to focus on self-identity, family history, and community and its connection to social studies and cultural studies. In the 4th grade these schools focused students focusing on their 'past', in 5th grade on their 'present', and in 6th grade on their 'future'. Because these topics allowed for a wide range of interpretation by the teacher, it was therefore harder to understand the content learning goals with regards to success or change. Nonetheless students often expressed themselves at various levels of complexity concerning their arts integration work with the same teaching artists at every grade level - whether in their 'calligrams' (art projects with symbols like hieroglyphics), songs, or imaginary countries. Thus in general, it can be assumed that all PAIR students were challenged to articulate their view of arts integration learning on a level of complexity comparable with classes incorporate English Language Arts in the context of acting drama and music or math with dance and visual art.
- The scoring format that focused on discrete sections of text may have been at odds with the overarching span of conversation focused on integrative teaching and learning. The

somewhat segregated scoring format may have suppressed the determination of levels of complexity of the overall responses to the list of questions.

- In general, the students in all schools cohorts and at all grade levels found it challenging to articulate understanding of arts integration processes across these three disciplines into their portfolio conference discussions. It was difficult for them to consistently achieve Level 3 or 4 responses to the portfolio conference tasks. This may simply be a result of not being developmentally prepared to comment with high levels of complexity on their work across three disciplines. Nonetheless these data reveal a growing ability of students to express increasingly sophisticated levels of understanding of arts integration teaching and learning in the context of their portfolio conference reflections.

* * *

3D–7: Tracking Qualitative Differences in Teacher and Student Reflective Understanding: PAIR Portfolio Conference Transcript Analysis Methods

Introduction

This section of the report illustrates how teacher and student comments throughout the three sections of the Portfolio Conference were rated for levels of reflective awareness and understanding according to the scoring rubric described above. Sample qualitative ratings of excerpted teacher remarks on their project goals and observations of student work (Sections I and III of the portfolio conference) are followed by qualitative transcript excerpt ratings of individual student discussing the process and meaning of their PAIL (Section II of the portfolio conference).

The purpose of providing sample ratings of portfolio conference system is two-fold:

A. To illustrate how the PAIR Portfolio Conference protocol can be used as a tool for understanding

- how teachers can look for evidence of student learning in the context of their own understanding of arts integration and the role of arts integration in their classroom culture.
- how the 'rhythm and flow' of the portfolio conference proceeds through distinct

phases of support for individual and group student reflection as the facilitator provides multiple occasions for each student to (a) respond to general questions about their learning, (b) to then make distinctions in their individual differences in response to their own work and peer work, and finally (c) to draw conclusions as to the meaning of their work and to cite evidence of their own learning in the project.

- how ratings of student responses can stem from assessing short initial responses that are then followed by responses based on prompts for further elaboration of concepts, interpretation of work, relevant personal experience, additional work samples, or a summation of previous discussion -- always probing for more detail and definition from students who at first may only offer general statements or lists of activities.
- to what extent teachers can draw conclusions about the impact of the program based on evidence of student reflection on both the content knowledge and processes explored in the project.
- how teachers may be able to consider alternative assessment of student learning that reveal new perspective on differences among pre-designated high, average and low performing students by adopting arts integration portfolio conference practices.

B. To illustrate how ratings from the portfolio conference excerpts can be used:

- to determine criteria for reliable qualitative judgments of portfolio conference responses.
- to provide exemplars of various levels of rubric systems that capture the increasing sophistication of responses from (a) general-to-specific, (b) singular-to-multiple, and (c) relational-to-systematic thinking, that serves as the basis for producing reliable judgment of both teacher and student responses;
- to explain how qualitatively validated quantitative ratings for teacher and student reflection may provide for statistical basis of judgment of the kind and degree of student learning that has resulted from the project.

* * *

Methodological Notes

This analysis builds on the vignettes presented previously in this chapter and on the analysis of portfolio student sample work discussed in the previous chapter. The organization of the portfolio conference reflection exhibits draws on the same four case studies presented in the previous PAIL student work sample analysis:

Classroom A: PAIR World Language Cluster School (Visual Arts-Music) Grade 4
Classroom B: PAIR Math Cluster School (Visual Arts-Dance) Grade 5
Classroom C: PAIR Writing (ELA) Cluster School (Drama-Music) Grade 5
Classroom D: PAIR World Language Cluster School (Visual Arts-Music) Grade 6

These classroom exhibits taken together will provide a comprehensive view of differences in reflection at three different kinds of PAIR cluster schools, three different grade levels, and in the context of arts integration units informed by four different arts disciplines.

Organization and Purpose of the Teacher and Student Portfolio Conference Scoring Exhibits

- The sequence of classroom exhibits follows the same order of the PAIL analysis in the previous chapter.
- Each exhibit will present all teacher reflection data first followed by rated student reflections in the same classroom. Together this juxtaposition of data provides a look at the qualitative basis for comparing teacher and student data across grades levels, arts integration units, and schools.
- In a few cases the teacher and student quotations already explored in the PAIL chapter of the portfolio conference vignettes above will now be presented in the context of exemplification of the reflection rating system.
- In the final section of this chapter, averaged teacher reflection ratings from a comprehensive sample of classrooms in all three grade levels will provide a statistical rating for each teacher and averaged ratings from a larger sample of student reflections will provide a statistical rating for each student who participated in the portfolio conference sample in the final year of the PAIR project.

* * *

3D-8 Four Portfolio Conference Scoring Exhibits: PAIR Classroom TEACHER Portfolio Conference Reflections (Excerpted Transcripts, Ratings, and Rater Summary Comments)

Each of the four classroom teacher exhibits provide contrasting levels of reflective awareness and understanding in their articulation of the goals of the PAIR project and evidence of impact on student learning. The exhibit 'boxes' contain quotations from the transcripts that have been edited for clarity, a rating score based on the scoring rubric (see 3D-6 Table 1), and summary of reasons (in italics) provided by the scorers.

* * *

3D-8 PORTFOLIO CONFERENCE SCORING EXHIBIT 1: Classroom TEACHER Reflections and Rater Comments: *Grade 4 World Language (Visual Art and Music)*

The first exhibit illustrates how the classroom teacher provided positive testimony of project goals and student growth in the PAIR project, yet could only offer general, undetailed statements in both the introductory and post-student observations sections of the portfolio conference protocol [see excerpted scorer remarks *in italics* in each of the boxed reflection statements].

Portfolio Conference Section I TEACHER Reflections: *To what extent did student performance in the portfolio conference meet your expectations for the PAIR project?*

"… the kids really did a great job on this, because it represent who they are with the self-portrait, which is that the art would be curriculum."

"I think it's really a learning process, not just for my students, it's also for myself, too. Learning how to put this in the curriculum, and then my students learning who they are, where they came from, and all this and that, connecting to their cultures. So it's really, to me, both of us are, like, learning."

"You can see the growth from the students, the knowledge that they gain from it. Like first of all they start off with nothing. They don't know. They start out with nothing, and once they finish with this whole, entire thing, that's when you see the big surprise. Like wow, they did this, they did this."

Rater Comments: *1.5 - There is a strong expectation yet very vague description of what is meant by growth of understanding of elements of the arts integrated curriculum.*

Portfolio Conference Section III Classroom Teacher Reflections: *To what extent did the portfolio conference reveal evidence of arts integration across disciplines?*

> "They're becoming more expert than I am."
>
> "From first from not knowing anything to be able to pick up where they left off from and continue to tell me what's going on, so you're able to see that they have better knowledge of – they're gaining more knowledge from doing all of this."
>
> "It's really connected to all subjects. There's writing, there's reading, there's, as she was saying, math also, which the student connected to symmetry that they just learned, which is connected back to their portrait. It's really connected to all subjects, even social studies, about their own culture, the food that they eat, pictures that they take. Pictures that they take at home representing who they are. So it's everything connected together."
>
> Rater comments: *1.5 – The teacher provides general statements that all things are connected, but the she didn't elaborate on the connections between symmetry in math and in visual art, for example, nor provide any details about gains in content knowledge.*

In the transcript above, the grade 4 World Language Magnet school teacher gives impassioned testimony that there is more knowledge gained and connectivity in her arts integration unit without articulating evidence of student knowledge of concepts and their connectivity through concepts or processes shared between arts and math. In addition she does not cite evidence from her observation of the student learning in the conference to confirm or disconfirm her assumptions.

* * *

3D-8 PORTFOLIO CONFERENCE SCORING EXHIBIT 2: *Classroom TEACHER Transcript Reflections and Rater Comments: Grade 5 Math (Visual Art and Dance)*

Exhibit B represents a more highly articulated viewpoint on arts integrated teaching and learning by a Math Magnet School teacher who chooses to share her observations of the students' performance in the conference to substantiate her conclusions. The transcript in section III of the conference protocol includes the full dialogue with the facilitator to show how the scoring of teacher reflection levels requires taking these interactive statements into account.

Portfolio Conference Section I Classroom TEACHER: *What were your student learning goals for the PAIR project**?***

"This year I think it's been...I've seen...to get the kids to actually think more critically. I think it's to have them...we, as the instructors, the visual artist, the movement artist, in our case, and the teacher, and all of us combined working together … to have the kids learn to problem solve, or to think more critically about things around them. … It was a non-stressed academic setting kind of thing, and I love that. Any time I don't have to stand up there and yak in front of them, it's great."

"Case in point, their projects. … We had an issue with this [blueprint-sculpture project], because one sculpture is a gun. It has a trigger. They then had to write what is this, and what's the meaning of it, and they all wrote a beautiful paper. So as long as they could explain it, … but they all could support their project. And that they would never have been able to do without going through this whole process, I think."

Rater Comments: *2.5 – This teacher provides several examples that support her view that critical thinking is best achieved through collaborative teaching and the posing of problem solving tasks in the context of highly personalized projects. The idea that current events study, and ample reflection and creation processes come together to form a more holistic learning process is heading toward level 3 but lacks enough specificity to know how this process really works.*

Portfolio Conference Section III Classroom TEACHER Reflections: *What was revealed about student learning in the portfolio conference?*

Facilitator:	*...how would you frame the evidence for integrative thinking that's going on in what you saw today?*
Teacher:	Student A is a prime example. He does not write as well as he verbalizes his information.
Facilitator:	*So what does that have to do with integration?*
Teacher:	This [PAIR] is a way to integrate. He doesn't have to put [everything] on paper. He can build it. He can move it. And he's going to tell me a story. He's going to tell you the information. And understand it. And an understanding of what the concept is. …that's part of the culture that Juan Carlos (Visual Art TA) and Jessica (Dance TA) also add to the classroom. ... Arts integration helps build that part of our school's curriculum or classroom's curriculum, -- they got fractions [in dance and visual art], and that was a bonus because, you know, I'm thinking, well, at least it's something mathematical [the students] actually can do. And that was great.

Facilitator:	*How does the student participation in this portfolio conference and PAIR in general relate to academics? ... does it eventually help test scores? What would you say?*
Teacher:	Well, of course it does because you have to learn how to correctly think in order to problem solve. So I think working together in groups, as one student said. … part of it was teamwork and working together with other people. – unless you're living on an island by yourself, you need other people around you. And this also helps that. … Every one of my kids has this empathy.
Facilitator:	*That's what you're after?*
Teacher:	Absolutely.
Facilitator:	*And you're saying this program might help you do that?*
Teacher:	I know this program helps do it.

Rater Comments: *3.0 - This lively dialogue brims with ideas about the importance of arts integration learning processes, multiple modes of expressing what you know, in relation to a classroom culture of collaborative teaching and learning processes that result in a quality of empathy among the students – all of which she believes will contribute to academic excellence based on the development of problem-solving skills.*

* * *

3D-8 PORTFOLIO CONFERENCE SCORING EXHIBIT 3: *Classroom TEACHER Reflections and Rater Comments: Grade 5 Writing (ELA) (Drama and Music)*

Exhibit C showcases relatively highly rated comments than are quite detailed with respect to the depiction of classroom teaching and learning dynamics and, as is the case in the previous exhibit, a focus on the impact of arts integrated learning on other forms of assessment.

In this first case the teacher is rated much more highly in her description of goals of the project than with her ability to link these goals to evidence of students learning observed during the portfolio conference. The disparity in these two scores with this particular reflection suggest that at the beginning of the conference, this PAIR teacher did not fully realize to what extent these goals were manifestly present in her student's work until she had observed students in Section II of the protocol.

Portfolio Conference Section I TEACHER Rating: *What were your student learning goals for the PAIR project?*

"We were trying to expand on what the students had learned in fourth grade and what they had learned with PAIR, and we were focusing on having the students develop their voice and recognizing point of view with main characters in a fiction story. And Reggie (Theater TA) specifically worked on how do you identify a person's point of view and how do you recognize someone else's point of view."

"And then what Charlie (Music TA) did was really expand on the voice area of it. And we worked on developing a theme song for the main characters, so if the main character's personality were put to a musical beat, what would it sound like. And then they had to incorporate what they were doing with Reggie in character description to come up with what they would think the beat of someone's theme music would be. So they had to look at the physical characteristics, emotional characteristics, how people interact with each other."

"The main purpose that we were, what we were looking to do was on all writings, regardless of whether it's ISAT or what we're using it for, we always want the student's own voice to come through and not just mimic what it is that they've read, but we want their input, their take on things. So this was helping them develop confidence in their writing so that their own true personality and their voice would come through. When they have to make connections, it's harder for them to make a connection if it's not their own voice and their own real experiences."

Rater Comments*: 4.0 - This teacher describes acting and music as extensions on what has already happened in her curriculum, concrete areas of study across disciplines, elaboration on TA roles in the classroom culture, all connected to her main detailed purpose to have children make connections in writing with the support of their own developed personal voices. She persuasively asserts that there is a strong relationship between her own classroom goals and those of the Teaching Artists in PAIR. She remarks also suggest a systematic understanding of relationship between theater, music, and reading/writing in her project.*

Portfolio Conference Section III TEACHER Rating*: To what extent did the portfolio conference reveal evidence of student learning?*

Well, [Student V] got that conflict is in every story you're ever going to read or movie you're going to see. There's always three attempts [to solve the conflict], the characters always fail at first, and then the last one works itself out, which was one of the points we were trying to get the students to comprehend. When they're writing and they're telling a story, it can't all be solved in the first paragraph, otherwise there's no story, and nobody's going to want to finish reading it. So I think it helped them be able to extend what it is that they're writing and provide more detail and more information."

*"*There was no culminating activity [with the music], and I think that's what helped keep it fresh in their minds. And that's why some of them were struggling with remembering what the beat was that they came up with, and how, although the three of them positioned their hands differently on the drums, basically it's the same beat.*"*

"And the kids made the connection that Phoebe's [character in the book] was kind of all over the place and very paranoid, and thought everybody was out to hurt her, and Sal was more grounded and that they complemented each other."

Rater Comments: *2.0 – There were some general examples of something that happened during the student portion of the conference and in the work of the teaching artists, but she was unable to relate her anecdotes to an overarching conception of arts integrated student learning.*

* * *

3D-8 PORTFOLIO CONFERENCE SCORING EXHIBIT 4: Classroom TEACHER Reflections and Rater Comments: Grade 6 World Language (Visual Art and Music)

In this final teacher grade 6 exhibit, the teacher is rated better at articulating the goals and outcomes of the project *after* viewing the student performance in the portfolio conference. That fact that many teachers were able to rate more richly on student learning after observing their students reflect on their own work suggest the validity and practical value of building in portfolio conferences into the project design.

Portfolio Conference Section I Classroom TEACHER Reflections: *What were your student learning goals for the PAIR project?*

"We pretty much had this [PAIR Units] kind of mapped out since probably at least the start of the project. They knew that fourth grade was going to talk about the past, and fifth grade was going to talk about the present, and sixth grade was going to talk about the future…"

"And then as part of our classroom activity that we did, we created compositions where they had to research high schools, colleges, careers that they were trying to get to, and so then they wrote compositions about what they needed to do in order to achieve those goals."

"But as far as...every time they've come in, you just...you look at things in a different way, you look at it from a different side, and I love that about teaching with the arts because you think of different ways to teach things, and different ways to get things

across. And of course it also works with different types of learners. I mean, you don't have the same type of learners.

Every single kid's not the same kind of a learner, so you have kids that are definitely hands-on learning and auditory learners, and this is great for that. So differentiated learning is going on when you're using this kind of a thing, which I think is fantastic. And that's, of course, a huge push for ever teacher, is try to differentiate your instruction."

Rater Comments *2.0 – teacher provides background information, connects the projects in class to the perceived curriculum for the PAIR project and speaks about the importance of the arts integration as a powerful 'differentiated instruction' strategy for her students – yet she does not site specific examples, and detail is limited.*

Portfolio Conference Section III TEACHER Reflections: *What might you change or revise in your teaching or assessment practices based on what occurred in the conference?*

"I love the idea of the musical part and having them create songs. The only thing is, I only played one instrument in my life for, like, six weeks and that was it, so I don't have any musical background. But as far as writing lyrics goes, that was a very fun way for them to write versus what they're always having to do with their five paragraph essays and their, you know, the stuff that they're being taught to do because they have to know how to do that, too. This is fun, though, for them. And then also, too, when we did the writing for Miss Mirtas, and they did do writing on there, which you can kind of see in the calligram stuff."

"And last year we made poems at wintertime doing the calligram as an extension. And they liked doing that. And I think they pretty much had a little bit of that every single year with this project with the calligramming. And that's definitely something, a different way of writing a poem and doing it in a calligram. It definitely would be something I would use.
… And too, I mean, everything is going to...thinking about how different kids learn, and it goes, all of this goes to that, and is extremely important."

"And it gives us – I don't know if you can say it – like an "excuse" to go this way with teaching. And, you know, so much of – there's so much pressure put on homeroom teachers and classroom teachers about tests and scores and this and that, and so many times you want these ideas in your back pocket so you can do things in a different way, because it's valuable in what goes on."

Rater Comments 2.5 – *Some concrete examples of what teachers can use from the project as extensions of their own ability to teach arts integration and by posing the 'excuse' given to teachers to use arts integration as a valuable alternative to her usual teaching activities.*

3D: Summary of the Four Classroom TEACHER Portfolio Conference Scoring Exhibits

These four exhibits provide a lens for rating teacher reflection in the context of the 'before and after' sections of the PAIR Portfolio Conference protocol. These reflections were analyzed and rated according to the ability of the teachers to describe PAIR project goals and to articulate the extent to which their expectations for student learning were met. By using the portfolio conference rubric to rate examples of teacher reflection from three different grade levels and three different schools cultures, the research and evaluation team was able to use these data to measure the relationship between the PAIR teachers' ability to articulate their reflective understanding of the PAIR project with factors of student achievement and their own professional development during the course of the three-year study.

* * *

3D-9 Four Portfolio Scoring Exhibits: Excerpts from PAIR STUDENT Portfolio Conference Performance Responses and Reflection Ratings

The student reflection exhibits constitute a comparatively larger set of transcripts then did the teacher exhibits. More time is needed in order to assess individual responses in these three-student group portfolio settings because a careful process of teasing out each student's statements about their work and reflection on their experience with the PAIR arts integration learning processes is required for valid and reliable comparisons among the students. Furthermore, the substance and level of articulation of each student's responses needs to be understood in the context of the dialogue and action among their peers.

The following four student reflection exhibits are matched to the teacher reflections just reported. Informed by the analysis of their work products as described in the previous PAIL work samples and the previous portfolio conference vignettes, these reflections will provide the most comprehensive indication in this study of the level of each student's reflective understanding of the value and quality of their own arts integration work.

* * *

3D-9 PORTFOLIO CONFERENCE SCORING EXHIBIT 1: STUDENT Portfolio Conference Responses and Rater Comments: *Grade 4 World Language (Visual Art and Music)*

In this first group reflection students are exploring the meaning of their song lyrics in relation to their imaginary nation project. The first three excerpts reveal initial differences among the three students. According to the raters, Student A at first is better able to provide detail than students P or M.

Portfolio Conference Section II: *STUDENT Responses to Questions focused on the "Level of meaning-making of the one's own art work in relation to its application to and understanding of content in other disciplines."*

Facilitator**:** *What is the important word in your song lyric?*
Student P**:** 'Country' is important word in the lyrics of the national anthem.
Facilitator: Tell me why your reason is for picking country. What does country mean to you?
Student P: It's like where to...country and the country that you live in, that's why it's important.

Rater comments*: 1.5 – One general response, some attempt at elaboration.*

Student M: Country is it because that's...we talk about the country we live in, so we just sing the song words ... it is not the tower (pointing to the drawing of a tower in the country)
Facilitator: Sing a line that has the country in it.
Student M: "The country that I love."
Facilitator: Can you sing it?
Student M: I can play the notes. [Plays melody on the xylophone.]

Rater Comments: *1.5 – She explains her choice of one word over another and demonstrates her knowledge of the 'national anthem' for their imaginary country.*

Student A: The Duf tower *[featured in the map of their imaginary nation]* …it's in the lyrics. … "To see the Duf tower."
Facilitator: Why is the Duf tower so important?
Student A: It's in the city.
Facilitator: Well, lots of things are in cities. Why is that so important?

Student A: It's the biggest thing.

Facilitator: So a nation is defined by the things in it?

Student A: Uh-huh

Facilitator: And what makes a country? What's the most important thing about a country? I know it's got buildings in it. But think about a country as people. What's the most important idea about a country to the people in it besides the buildings? What do you think?

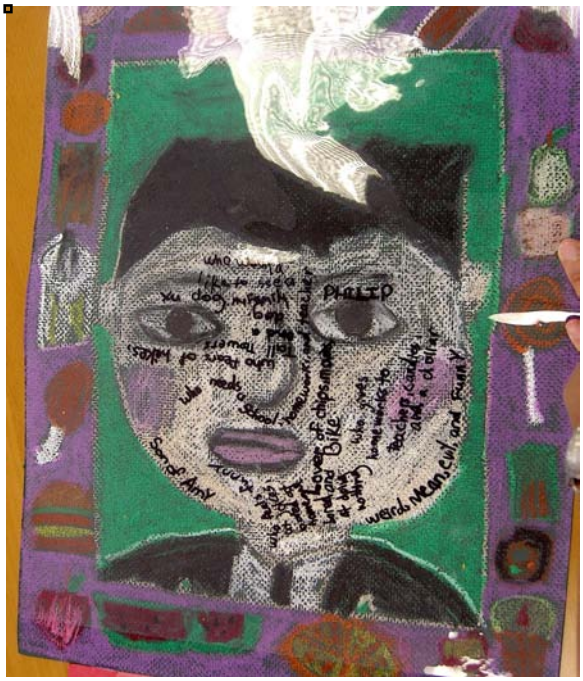
Student A: Their rights.

Facilitator: What do you have a right to do?

Student A: Right to do anything. … Like do anything, but not bad things.

Rater Comments*: 2.0 – After follow-up questions, this student is beginning to comment on the importance of civil rights, although this sentiment is not found in the lyric.*

In the next section students P and M comments rate the same as Student A, indicating that the portfolio conference format is working equally well to stimulate this second tier of ratings for all three students.



[Calligram-biopoem' self portrait of Student P, an artifact that informed the next round of reflections in the conference]

Portfolio Conference Section II: *STUDENT Responses to Questions focused on th*e "*Ability to summarize discussion and critique of student work in terms of arts and arts integrated learning."*

Student P: We learned about...we learned how to draw more better.

Facilitator: Take your time. You learned not just to color...

Student P: Yeah.

Facilitator: But instead you learned what?

Not just to color pretty, but what?

Student P: But you could do whatever you want in there. … Yes, because every time when I see some books that's math, there's...they shows the shapes of it.

Rater Comments*: 2.0 - in a very cryptic fashion, this student mentions arts skills and choices in coloring, and a general connection here to the concept of shape*

Student M: Now I know how to draw myself in a self-portrait.

Facilitator: When you draw shapes, does that have anything to do with math?

Student M: Kind of. …Because it has, like, symmetry. It's also part of math because you have to use, like, symmetry and congruent to do like that. You have to do the same shape and the same size.

Rater Comments*: 2.0 – Mathematical terms are mentioned after prodding, not only does art 'have' symmetry, but you have to use symmetry and other conceptual skills in art making.*

Student A: On a person's face, you don't have to, like, put all just one color. Some parts of the face are different colors. …You can combine writing and a picture.

Facilitator: Uh-huh. For what purpose?

Student A: So it can be prettier. Because the shapes, they also tell about math. Like if you see a square, it's part of math, too… The words in the bio poem—it can help you, like, know that person.

Rater Comments*: 2.0 — Without much prodding she is able to provide examples on a few connections among disciplines*

Facilitator: What do the numbers in your song tell you?

Student M: It's like you put them in, each one are in order, so it has to have numbers on the side of it so it shows which one is in order."

Student M: It's [the words are] better with melody. It's better with this.

Student M: It's better to do it [words] with music numbers sometimes because sometimes when you don't know the melody, then it's a little off key.

[*Melissa demonstrates the ability to play a melody backwards without rhythm*]

Facilitator: What is the number six sound like? .

[*Student* M *sings the sixth scale degree, accurately with her eyes were closed, than sings and plays the bells at same time*]
Facilitator: So when letters or numbers are changed to music, what does it create?
Student M: A melody.

Rater Comments*: 2.0 – While she didn't elaborate much on the meaning of music in relation to words, her ability to perform playing, singing, and listening exercises proved a developed understanding within the art form and its integration with mathematical concepts that validated the intentionality of her composition choices.*

Overall, none of these fourth graders were able to go much beyond their initial listing of multiple general statements. The lack of connections among their statement or elaborated comments may be due to developmental factors. However, this general lack of articulation may also be a by-a product of the classroom music integration process that did not allow time for the students to gain enough musical concept knowledge or the skill development necessary to elaborate on their arts integration aspects of their work in a more sophisticated manner.

* * *

3D-9 PORTFOLIO CONFERENCE SCORING EXHIBIT 2: STUDENT Portfolio Conference Responses and Rater Comments: Grade 5 Math (Visual Art and Dance)

In Scoring Exhibit 2 below, there is an interesting difference between the students who are reflecting on one another's work compared with reflective comments by the creator of the artwork being discussed. In this example, Student A is able to draw some insightful inferences from the sculpture and Student C affirms these speculations. Student K, who created the sculpture, elaborates at some length on the theme of war and money in ways that suggests a systemic awareness of interacting factors between war and the economy.

Portfolio Conference Section II: *STUDENT Responses to Questions focused on the "Level of meaning-making of the one's own art work in relation to its application to and understanding of content in other disciplines."*

Facilitator: I want all three of you to talk about your sculpture in terms of imagination and ideas, and let's see how deep we can go, all right. So...and you, and I want you to look at each other's sculptures. So let's start with this one right here. Student A, what do you see when you look at this sculpture?

Student A:	[looking at Student C's sculpture) I see...I see a gun, and then I also can see that how, when she said the money, I can see, and then I can see blood. I think it's blood.
Facilitator:	… So if this is blood and this is money and this is a gun, what are we imagining here? What is this all about?
Student A:	The war and economy.
Facilitator:	And what's the relationship between the two?
Student A:	Um...I don't know

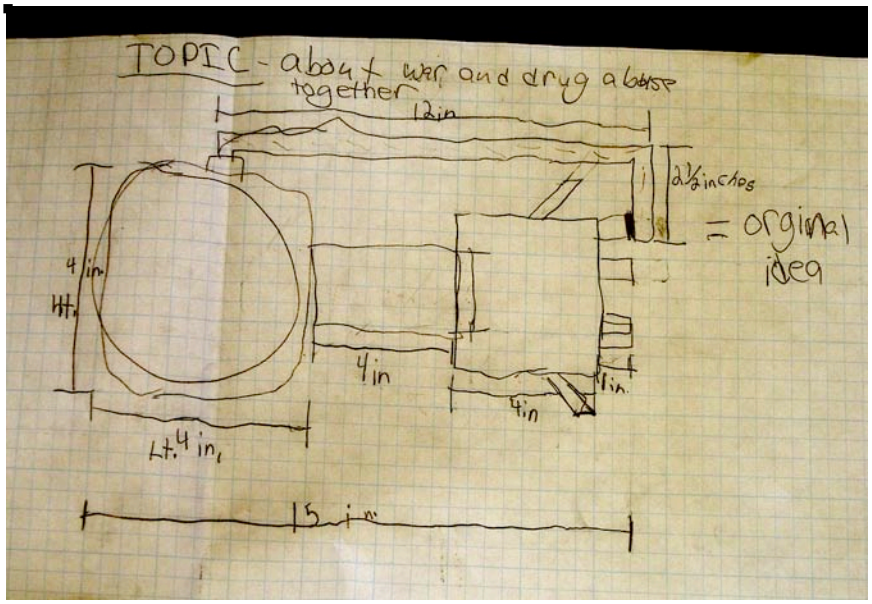
Rater Comments*: 2.0 - Student A notices possible themes with some speculation on their meaning that Student C and K affirm.*

Facilitator:	What do you see?
Student C:	I see when she said that it was about money all those things that--
Facilitator:	Do you see money here?
Student C:	Yeah, on the top.
Facilitator:	On the top, right?
Student C:	And I was [thinking] that they color, they painted the gun green because money is green. And I see a little bit of blood, red paint, on top of the money and on the gun.

Rater Comments*: 1.5 - Student C now notices specific color detail that lends credibility to Student A's statements.*

Facilitator:	So what was your imagination? What were you thinking when you made this?
Student K:	That it was about the war and the environment and the taxes.
Facilitator:	Where's the environment?
Student K:	Like you will see green, like the grass is green.
Facilitator:	Oh, so this is environment. Oh. The money is green and the environment is green?
Student K:	Yeah.
Facilitator:	Tell me more. … Look what's coming out of the front of it. Why is there money coming out of the gun?
Student K:	Because there's a lot of money that's going up and a lot of taxes is going up, and a lot of people don't have that much money.
Facilitator:	Wow. And...so what do you think this all means altogether?
Student K:	About...it means altogether it's a war.

Rater Comments*: 2.5 - Student K now begins to elaborate on connections among multiple concepts that connect the details observed in the sculpture and a general meaning of war.*



[Visual arts projects resulted in math-rich scaled blueprints such as the example to the left. These same projects also connected with connections to social issues that students discussed with considerable detail in the portfolio conferences.]

Later on in the conference, students had to face the challenge of demonstrating the math aspects of their PAIR work, such as the concept of scale and calculation with reference to their sculpture blueprints. Below is an excerpt from the conference that concerned the relationship of math concepts to choreography in the student PAIR unit.

Assessing the grasp of mathematical concepts can be tricky when looking for evidence in dance, a problem-solving domain rarely approached in conventional math classes. Careful questioning techniques and analysis of transcript analysis are needed to determine differences among student responses based on their conceptual grasp of mathematical principles that extend beyond just listening for correct calculations.

In the following facilitated dialogue, Student A's grasp of math concepts and their connection to concepts of choreography begins to outpace Student C's destabilized understanding. Student K offers little reflection in this discuss other than to list all the calculations of the fractions at once.

Portfolio Conference Section II: *STUDENT Responses to Questions focused on the "Level of meaning-making of the one's own art work in relation to its application to and understanding of content in other disciplines."*

Student A:	I got...we had...with Miss Jessica (CAPE dance teaching artist), she …had us do two things. She made us do where we had to have a fraction of ourselves….
	First I did, like, what I am, like my race … and I'm going to say it to you like how I wrote it. "I am a fraction. I am 33% Mexican, 33% Puerto Rican, 33% Honduran, and 1% American. I am whole." She made us do this. She said we can do one part too, and I did two. Want to hear the other one?
Facilitator:	Yes, please.
Student A:	"I am a fraction. I am 50% kind and 15% mean, and 35% happy. I am whole."
Facilitator:	[*Laughs.*]. I like the proportion there. Now, why was Jessica, who does dance, care about fractions in relationship to who you are? That's not even about dance. Why did she do that?
Student A:	Because for fractions, 'cause, like, how Student K was saying we had to have different fraction counts for how much we move. And then she also kind of wanted us to do a little personality in our dance routines.

Rater Comments*: 3.5 - Connects fractions/percentages with choreographic structure and self-identity, substantially detailed explanation.*

Facilitator:	… All right, Student C, what have you got?
Student C:	We had to do -- first of all we had a group, and we picked, like, a theme for our group. My group's theme was respect and love. And we had to do a dance that had fractions in it, like for our locomotor movement was one-fourth of the [dance]. …And then it would be eight counts, because we had to put in counts also. The whole dance had to be 32 counts.
Facilitator:	Oh. So if you do eight counts of locomotor motion, what is the fraction of time that you're going to be spending on your locomotor? What part of the time?
Student C:	Maybe like four minutes.
Facilitator:	Four minutes out of how much time?
Student C:	I'm not sure… Like 30 or maybe like 45 [minutes].

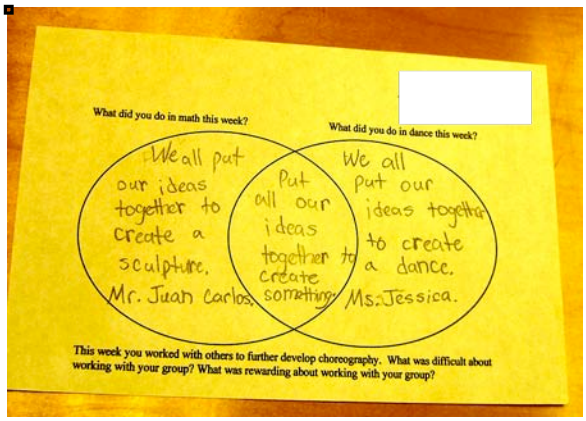
Rater Comments*: 2.0 - Student C seems to understand fractions to a certain extent yet she confuses 1/8 with 4 minutes (1/8 of 32 minutes), and 8 locomotor parts within 32 parts (1/4). She cannot seem to express time in fractions at this point without mixing up concepts of minutes, hours, etc with her calculations. Were there no confusion among these concepts, and her concept of fractions remained stable, 4 minutes could have been clearly connected with its determination of being 1/8th of the 32 minutes.*

In the context of Student A's demonstration of problem-solving skills below, Student C seems to regain her composure and gives more accurate answers about fractions. Student K provides calculations with virtually no explanation or reflection on the computational process.

Portfolio Conference Section II: *STUDENT Responses to Questions focused on the "Level of meaning-making of the one's own art work in relation to its application to and understanding of content in other disciplines."*

Facilitator**:**	Student A. What do you think? If there were 32 counts, and you spent eight counts on locomotor motion, what fraction is that?
Student A:	One-fourth of the whole, of 32.
Facilitator	Student C, [… according to the choreography analysis], if you spend a quarter of the time on locomotive dance motions, how much time are you spending on non-locomotive?
Student C:	A quarter, too.
Facilitator:	So that would be the same amount of counts? How many counts would that be?
Student C:	Eight.
Facilitator:	Which [type of movement], Student K, did you spend the least amount of time on?
Student K:	The least amount was the synchronizing movement.
Facilitator:	And how much of the time was that?
Student K:	One-eighth.
Rater Comments*: These short, accurate responses, while not scored individually, affirmed the mathematical validity of the portfolio conference reflections.*	

Although the Venn diagram below may only suggest a very broad description of the dynamics of arts integration teaching and learning, this type of artifact often served as a tool for student reflection on interdisciplinary learning during the portfolio conferences. Documentation of the underlying logical-mathematical thinking processes necessary for student to complete the sculpture project could just as easily been applied to character development, choreography, or composing with words and music.



Various type of conceptual thinking frameworks became objects for categorical and elaborative thinking as students reflected on their learning processes engaged by various form of interdisciplinary arts learning during their PAIR projects

In the Math exhibit below, the student reflection profiles evolve out of detailed conversation about their arts integration projects that depend on their ability to quantify design elements of sculpture and choreography. Within a distinct focus on collaborative process, the mathematical aspects of the PAIR projects appear to be expressed clearly by all the students, though with distinctly different level of sophistication with regard to mathematical concepts of scale, proportion and calculations. Students spoke with unanimity about math's applicability to visual arts and dance and the *validity* of those applications. Students vouched for authenticity of math in their blueprints, and for the validity of their artistic vision in the context of their math work. Clearly, these students had no more trouble articulating the importance of math in their ability to create and analyze choreography, than they did reconciling the importance of dance in the service of expressing their understanding of proportion and fractions.

* * *

3D-9 PORTFOLIO CONFERENCE SCORING EXHIBIT 3: STUDENT Portfolio Conference Responses and Rater Comments: Grade 5 Writing (ELA) (Drama and Music)

In PAIR Writing Magnet Schools, PAIR units are focused on enhancing the study of language arts through drama and music. In this classroom the PAIR unit focused primarily on the study of character and story telling by taking time to study the elements of story telling, creative writing and illustrating conception of character, mood and action in percussion.

Since the previous chapter reported that the music artifacts included in the ELA school portfolio were rated relatively highly for quality, in this chapter we can test whether or not highly-rated

artifacts portfolio and conference ratings result in higher levels of reflective understanding of the inter-relationship among the main characters [Sal, Phoebe, and Ben) in the book (*Walk Two Moons*) these children (Students V, M and J) studied.

Portfolio Conference Section II: *STUDENT Responses to Questions focused on the "Level of meaning-making of the one's own art work in relation to its application to and understanding of content in other disciplines."*

Facilitator:	So tell me about Sal's character
Student M:	I think she [Sal] was a very outgoing person. She liked to be around with her friends, do [things].
Facilitator:	… what kinds of things?
Student M:	Like spy on her friends and ...spy on a lady. And there was this guy [Ben], and they kept following her, and she was trying to figure out who he was. And then she saw a picture of the man on a police officer's desk. … So she thought it was his son. I think that she's more afraid than bad, like she doesn't want to, like, face anything because, like, she doesn't want to know that she's wrong. So she acts out…. Ben's just, in the beginning I thought Ben was kind of weird.
Facilitator:	Weird because?
Student M:	Of how he acted. He drew a picture of...what's her...oh, yeah. He drew a picture of Sal as a lizard with long hair, and the hair turned into a chair, and she was sitting in it."

Rater comments*: 2.0 – Description of mindset, personality, and motivation of character connected to some elements of the story – connection between the two not really substantiated. He did that again to a much lesser extent with the second character, but he had responded strongly to Ben's 'weirdness' and the intriguing drawings he created.*

(continued on next page)

Facilitator: [to Student J]: Tell me about Phoebe.

Student J: Phoebe, she was kind of mean to people. … She got annoyed easily. She didn't like to get babied by her mother, but I think that Phoebe kind of put on an act when she was around people, and that she was kind of like …Facilitator: Is it kind of an act that she's mean, she's not really mean, or is she mean?

Student J: No, that she acts like she's trying to...like she's...how to explain this. Like she doesn't...care or she doesn't...

Facilitator: So she acts like she knows things she doesn't know, is that what you mean?

Student J: That she doesn't, like...she tries to put on an act, but she's...um...really smart, but she's kind of not, yeah. … She's not mean to people, not necessarily mean. She...I think sometimes she is mean, but she doesn't realize she's mean… Sal is like, she's nice to people, but Sal, she doesn't like to be hugged or anything. And Sal doesn't want to face the fact that her mom's not going to come back. And she doesn't really know how she feels about Ben.

Rater comments*: 2.5 – Student J is able to articulate conflicting elements within Phoebe's character and actions, and to note subtle differences between Pheobe's and Sal's 'meanness'.*

Facilitator: Do you think Phoebe is mean?

Student V: Not necessarily. It's because she wants to know the truth, but in the end she really knows the truth. It's because she's being mean because she doesn't know the truth. She thinks that she knows everything, but she really doesn't. And she's kind of creative in a lot of different ways, but she can't really explain her point, but she thinks she knows everything, until she found the truth. That's when she started being more nice.

Facilitator: Tell me about Ben

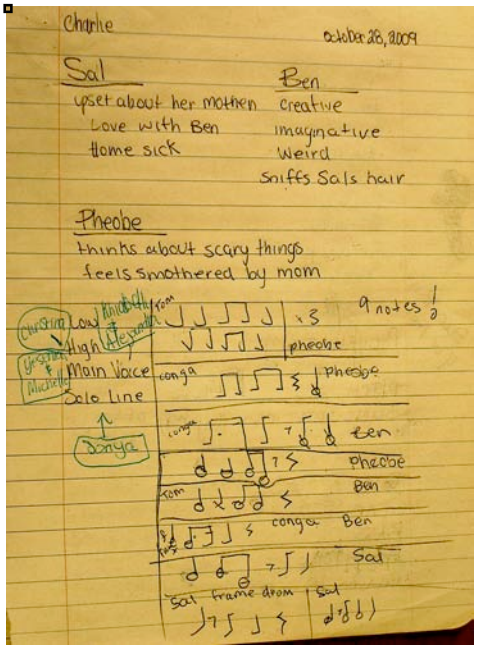
Student V: Ben, he's a creative person. He draws things.

Facilitator: So Ben and Phoebe are both creative? [Student V nods]

Facilitator**:** How are they different in their creativity?

Student V**:** Because Phoebe wants to know more of the truth, and Ben just draws pictures, like, of [memories]. …he comes up with things that you really don't [know what they are]...he made this flower. He drew a flower and then nobody would come up with...nobody knew what kind of flower it was… Phoebe, she always wants to know the truth. Since her mother disappeared, she always wants to know the truth. She thinks that somebody kidnapped her and then her dad says no, because she left this note, what do you think she'll do? Let the kidnapper write her a note? So she always wants to know the truth, but she really doesn't know it [what the truth is].

Rater Comments*: 2.0 – This response contains more than generic distinctions among the characters as written on his worksheet. Student V sensed the conflicting elements of Phoebe's character and provided a very interesting description about Ben's mysterious drawings. Although there is considerable detail about Phoebe's quest for truth, these descriptions don't rise to the level of clear comparisons among the characters beyond the characteristics previously written in his notebook.*



In the final segment of the portfolio conference, the students are challenged to use their own drum patterns to augment their descriptions of the characters. The music drum pattern artifacts (such as the PAIL student work sample displayed to the left), add a new dimension for scoring the student portfolio conference because students were able to make further distinction about character based on the generic descriptions they had written on paper worksheets. The following dialogue excerpts, laced with musical illustrations, provide additional evidence of the students' ability to demonstrate their understanding of character, mood, and action in the story.

Portfolio Conference Section II: *STUDENT Responses to Questions focused on the "Level of meaning-making of the one's own art work in relation to its application to and understanding of content in other disciplines."*

Facilitator:	Give me some words and some music about Ben.
Student V:	Ben is really creative, so watch... [*Drums.*] (dotted eighth - sixteenth, four eighths, one quarter, repeat) [ti---ti ta ta ta ta boom] sounding like a march]
Facilitator:	Good. Tell me more about Ben.
Student V:	Ben, he's more of a mystery, but in the end, Sal knows a lot about him. And Ben...but Sal started knowing a little bit more about Ben.
Facilitator:	All right, let's hear that side about Ben. The mysterious side.
Student V:	[*Drums.*] quarter, eighth-eighth, quarter, quarter (last quarter low in center of drum) [ping, tap-tap, ping, boom] – (first section without the repeat was a tiny bit out of rhythmic flow, which might make sense for the mysterious quality).

Rater Comments*: 2.0-Contrasting rhythm patterns support differences within Ben's character.*

Facilitator:	All right, [Student M], what do you say? Creative Ben.
Student M:	I think creative Ben is a little bit more like with feeling in it than mysterious Ben, when it sounds a little scary. Like he's all [sneaking around]
Facilitator:	[to Student M] All right, let's hear Sal's drum pattern. Say two things about Sal and then play them. Sal. You can quote from yourself [from your character worksheet].
Student M:	I think Sal is nice and she likes to hang out with her friends.
Facilitator:	OK, let's hear that.
Student M:	[*Drums] (at the end of the drumming, the pattern was dotted eighth, dotted eighth, dotted eight, eighth, quarter, repeat. (sort of bouncy Afro-Cuban rhythm: 1&a 1&a 1&a 1& 1e&a]*
Facilitator:	And why is that nice? Because?
Student M:	I think it has feeling in it and it makes me feel happy.
Facilitator:	What's another side of Sal?
Student M:	She was afraid that...she doesn't want to know that her mom is dead.
Facilitator:	OK, I want to hear that.
Student M:	[*Drums tapping quietly*]
Facilitator:	All right, now let's hear SAL getting along with her friends.
Student M:	[*Drums the previous pattern*]

Rater Comments*: 2.5 - More detail about the character emerges in this discussion that supports both linguistic and musical expression of the social-emotional understanding intrinsic to each character's traits, moods, and actions.*

Facilitator:	All right, Student J, let's hear about Phoebe.
Student J:	Phoebe, she's afraid of what might happen to her mom and she doesn't...
Facilitator:	Let's hear that.
Student J:	[*Drums.*] *(very similar to Student M's last pattern)*
Student J:	And Phoebe, she's also very paranoid.
Facilitator:	Paranoid, meaning what?
Student J:	That she's always on guard. She's like scared.
Facilitator:	Good. Let's hear that.
Student J:	[*Drums*] (*much different, but perhaps no pattern at all)*

Rater Comments*: 1.5 Although contrast is clear between the two drum patterns, the musical lack differentiation from the previous drumming by student and therefore appear to be generic responses.*

Facilitator:	All right. Does Sal meet up with Phoebe?
Student M:	[*Nods.*]
Facilitator:	What happens when Sal meets Phoebe?
Student M:	Phoebe starts liking her and Sal is kind of getting used to her.
Facilitator:	All right, let's hear it together. That's the scene. Let's hear it. You can speak and play the drums. *[Student M and J Drum together.*]
Facilitator:	Do they get along? Student K, Student V, come in. What kind of mood [is your character] in?
Student V:	[*Drums with the other students*] (*at the end of the session all three students were doing the 'play' drumbeat created by Student M earlier*)

Rater Comments*: Although these students were not rated individually, their composite rhythm lends validity to their consensus about how the two characters are getting along.*

By both describing their character portraits in both words and music, these Writing Magnet School students were quite willing to reflect on the specific nature of their musical experience and its implications for story comprehension. The rater comments directly above and below attest to the social learning value of bonding collaborative work in two art forms and literature together around the central concept of dialog and character.

The following unrated transcript excerpts provide evidence of their collective understanding of a three-dimensional arts integration process in their classroom that manifests in enhanced arts learning skills (music and drama), understanding of structure and content in story, and their own awareness of social-emotional aspects of their language arts curriculum.

Portfolio Conference Section II: Student Responses to Questions focused on the "Level of meaning-making of the one's own art work in relation to its application to and understanding of content in other disciplines."

Facilitator:	What did you learn, what do you think you really learned by working with Reggie and Charlie? What did you *really* learn?
Student M:	How to describe things, how to play instruments, how to express your feelings playing instruments.
Facilitator:	And why is that important to do?
Student M:	'Cause you don't want your feelings to be all bottled up. It's good to let them out.
Facilitator:	OK. What did you really learn from Charlie and Reggie?

Student J:	That everybody has a different point of view and it's like your point of view is not the same as everybody else's, so everybody else might, like, feel different.
Facilitator:	And so how did Charlie help you do that?
Student J:	By showing us different – [*drums*] – and helping us do the...making the different theme songs of characters.
Facilitator:	And what did you learn from Charlie and Reggie?
Student V:	Charlie, I learned where to come in at the right time, like if we're playing. Like dun, dun, dun and then another part comes in, you stop, and then you go in. I learned from that, from the instruments. And with Reggie, if you ever make up a story or a book or direct it, there's always three trials, they all fail. The last one always works, and it takes time to do all of those. And then with music, the emotion is way better.

Rater Comments*: Collectively, these students describe the complementary nature of arts integration is achieved through working with teaching artists in drama and music.*

* * *

3D-9 PORTFOLIO CONFERENCE SCORING EXHIBIT 4: STUDENT Portfolio Conference Reflections and Ratings: Grade 6 World Language (Visual Arts and Music)

This final portfolio exhibit provides a forum for students to reflect on their three-year experience with PAIR arts integration processes. Previously World Language (WL) students focused on their past and present identity, family history, and cultural roots. In the third year of the project the World Language units now focus on each student vision of their future life aspirations, relationships and accomplishments through visual art and music.

In year 3 the students created maps of their future. These portfolio conference artifacts begin with developing personal images and tracing a journey through imaginary milestones that will shape their success in life. Once completed, these maps and the songs that complement them become primary sources for students to explore and reflect on the meaning of their work in this PAIR unit.

The figure below displays a collection of individual 'future' maps so that students can reflect on common representations of passageways (ladders, pathways, footsteps), symbols of possible

achievement in their future (college degrees, possessions, career opportunities), and calligrams (drawings created with words) and self-portraits (sometimes bio-poems) that suggest evolving personal traits (determination to succeed), tough choices (in career paths or relationships), and states of mind (confusion, confidence) about their future.

The following transcript excerpts reveal that the level of student reflection appears to change when the attention shifts from comments on *artistic process in isolation* to interpretative remarks on *artistic products in the broad context* of their arts integration unit. The first set of portfolio conference responses are rated according to the level of reflective understanding these students express in their description of their artistic processes in the PAIR unit.

Portfolio Conference Section II: STUDENT Responses to Questions focused on the "Level of meaning-making of the one's own art work in relation to its application to and understanding of content in other disciplines."

Student J:	These are just the colors that...up here?
Facilitator:	Yeah. Tell us about the rest of your work.
Student J:	These are just the colors that represents this part of it.
Student J:	Miss Mirtas told us to twist the words around.
Facilitator:	For what reason?
Student J:	So that it ... it looks different from normal writing.

Rater Comments*: 1.0 – Extremely generic statements about a very specific artistic technique*

Facilitator:	What about these circle words and so forth?
Student K:	Mirtas taught us to write our words in different shapes and sizes.
Facilitator:	Why? What was important?
Student K:	So it doesn't all look the same and plain.
Facilitator:	And why is that important? What does it help to do when they don't look the same?
Student K:	So people just don't get bored by looking at it.

Rater Comments*: 1.5 – A concept of aesthetic purpose is expressed, but otherwise just describing a superficial aspect of artistic technique.*

In the following array of student responses to the same topic, more attention is given to *interpreting* the purpose and capacity of the art to express the uncertain hope of these students attaining their life goals. In this context the level of reflection tends to be more sophisticated than when speaking only to the artistic techniques of rendering their future maps.

Student K:	First I was sad because next year I'm going to a different school. Then I was nervous to go to that school.
Facilitator:	Uh-huh.
Student K:	But then I graduate that school, but then I'm nervous again because I'll be going to a different high school.
Facilitator:	Mmm, keep going.
Student K:	And then I also wanted to learn how to swim.
Facilitator:	And are you going to get three degrees? Is that true?
Student K:	[Shakes head no.]
Facilitator:	No, well, then what do these things mean? So what did that mean?
Student K:	I get a paper that tells me I graduated that school.
Facilitator:	OK, and this one?
Student K:	That's my degree.
Facilitator:	Oh, that's the final degree, all right. So it's a huge process here. Are you still nervous at the end?
Student K:	[Nods.]
Facilitator:	What are you nervous about at the end?
Student K:	About my job, like how to do it."

Facilitator:	Oh, so this is the story you were telling us about. This one, who did this one? All right, Student K, tell us about this one.
Student K:	It shows our paths of what we were going to do. [*referring to the future paths of two people represented in elaborate symbols and interweaving, pictures*]

Rater Comments *2.0 – strong projection on the emotional (anxiety) of anticipating events in one's future. Strong ability to explain her story via the symbols on her creation, yet not much elaboration beyond that.*

Facilitator:	What do you see?
Student E:	Footsteps.
Facilitator:	What do footsteps have to do with the future?
Student E:	It's like you're learning to become what you want as your career. Because you're taking little steps on going to your goal…
Facilitator:	How did you figure out what that path would be and what your future is going to be?
Student E:	What we're interested in is what we're thinking of having for our careers.
Facilitator:	Tell me why it is important to think about your future.
Student E:	Because it's important about your education and how your living is going to be.
Facilitator:	So it has something to do with your education?
Student E:	Yes. My second point is that you just want a job that's really good so you can get money and do whatever you wanted to do.

Rater comments: *2.5 - This statement articulates the importance of taking steps toward goal attainment in the future, of developing a career out of one's own interests and of education and the freedom to pursue one's interests and get good jobs that education affords.*

Student E later describes another factor in his map: the possible crossing of career paths among friends in the future that may influence career choices.

Facilitator:	Let's read that [reflection sheet in Student E's portfolio].
Student E:	"My team member and I drew a diagram and planned how to create the board. We decided to use oil pastels. My team member and I are similar...we have similar careers because we both wanted to become veterinarians. We drafted that we will meet at a [cloud]. Also, we want to make a difference."
Facilitator:	OK, and this one [image] tells about?
Student E:	This one tells about the...the learning from other teachers.

Facilitator:	OK, and they both you and your friend get to the same place, which is what?
Student E:	You want me to tell about the steps I take?
Facilitator:	[*pointing at the symbols in the future map*] These steps. This is a bunch of people, right? So I want you to tell me what's important about this?
Student E:	It tells about two people's careers. And then after that, they were separated and doing different careers.
Facilitator:	All right, so what is this? It's a story or is it a map or is it...what is it?
Student E:	A story. It's a story about two [people].
Facilitator:	So...and point out one detail. What's the relationship between these two people?
Student E:	They were doing their own thing and suddenly they met each other again, and then shocked. After that they went up and went with this .... [pointing to another spot on the map with animals] and went with working with animals [as a veterinarian].

Rater Comments: *3.0 - Here student E adds another dimension: the crossing of paths and the motivation for having a career that makes a difference. If these reflections are joined with the previous they should be rated together as 3.5 based on the detail and inter-related cognitive complexity of his explanation of his future career map.*

In Student K's case the ratings are not as high as Student E, although she does rate slightly higher with respect to levels of sophistication rated earlier in the conference.

Facilitator:	Why is it important, thinking about the future important to you? Put it in your own words.
Student K:	About how much you can learn.
Facilitator:	All right, Student K, tell us about this one.
Student K:	It shows our paths of what we were going to do. [points to two people. very elaborate symbols, interweaving, pictures]
Facilitator:	Detail?
Student K:	That we were going to get our diplomas and get a car after we get our license.
Facilitator:	OK. And why is it important to do this?
Student K:	So you could travel, and also you can be able to get a good job.

Rater Comments: *2.0 – Student K now offers slightly more detail and connections among her statements than before*

Facilitator:	Student J, did you do this one [future map]?
Student J:	Yes.
Facilitator:	All right, tell us.
Student J:	She's a lawyer and I'm a police and, like, we see each other when we work with each other because she's a lawyer and I could help her out with the things she might do as a lawyer.
Facilitator:	How can you help her?
Student J:	For example, like if I catch a thief, I could put him to the lawyer, I mean, to her and she can help.

Rater Comments: *2.0 – Student J now offers more detail by describing concrete connections than occur in socially aligned careers.*

In this final rapid-fire discussion students appear to be building toward consensus by offering a set of cumulative, inter-related reflections on the value of combining music with language arts components of their PAIR units.

Facilitator:	All right, you did lots of great work here. You did drawings, you wrote about your future, you told me a lot about your future, you read each other's drawings, and you also put music with this. Is this a good thing to do in school or is this something that maybe just some schools should do and some schools don't need to do? If you were going to do this again, do you think all sixth graders should do this or do you think not necessarily? Is this important or not?
Student J:	Yes.
Facilitator:	Why?
Student J:	Because it shows what we feel.
Facilitator:	And why is it important to show how you feel in school? In a place where you're learning all these things, getting your diplomas, everything, why is it important to know how you feel
Student J:	To know what we're thinking…. Know more about yourself.
Facilitator:	And why is that so important to know more about yourself when you're right here in school?
Facilitator:	What do you think, Student E?
Student E:	I think next year's sixth graders should do [what we did].
Facilitator:	Why?
Student E:	'Cause it express about their feelings, people can know them more, so that...they can express to others how they feel, to get to know each other a little bit more.

Rater Comments: Contextual qualitative data that supports the validity of arts integrated teaching and learning in this classroom.

* * *

3D-10 Summary of Student and Teacher Portfolio Conference Interview and Performance Assessments

These four student sample exhibits demonstrate how student reflection ratings were derived from sample responses in section II of the PAIR Portfolio Conference protocol. During these procedures, all three students from all four classrooms were given relatively equal opportunity to describe the quality of their participation and their 'best' work in two arts disciplines in relation to PAIR arts integration units in either math, writing (ELA), and/or world languages (social and cultural studies). The facilitator had to randomize the order of student responses so that each student had the opportunity to provide both initial and follow-up responses to questions, constantly seeking elaboration through standard 'tell me more' or 'put this in your own words' prompts. These prompts assured that students were given time to provide added detail to their comments and time to come up with their own connection across disciplines and tasks without comments or help from their teachers during the conference.

Besides providing examples of ratings, the rhythm and flow of the student discussion reveals the impact of focusing on individual learning in the context of discussion about collaborative working processes. Ratings at the first stages of the protocol resulted relatively general, unspecific responses initially and, for most students, their responses became more detailed and relational later as they began to feed off each other's comments and gain confidence in their voicing their own opinions.

Teacher comments after observing student group discussion validated the effectiveness of the portfolio conference protocol for evaluating student learning in several ways.

First, most teachers came to understand the portfolio conference as an important alternative mode of assessment for student understanding of their content knowledge that was missing from their classroom practice. In the portfolio conference, students could take the time to formulate and elaborate on their understanding of content knowledge and its application to the creative project work in ways that seemed wholly different from their usual method of classroom academic assessments.

Second, teachers discovered new appreciation for evidence of teaching effectiveness and student learning in the context of the problem solving tasks that the portfolio conference provided. Students were constantly challenged to go beyond short 'correct responses' and give additional reasons for their response by using their work products and writing samples to invent new ways of expressing their understanding of content knowledge as it developed throughout their arts integration units.

Although most teachers (especially the teachers who repeated their PAIR units during this project) were not surprised that the students expressed enthusiasm for their work with the teaching artists and their collaborative work in their peer groups, they were often surprised how well individual students expressed their understanding of content knowledge without the benefit of teacher intervention as in the class of their usual classroom discussion or written work. Teachers often commented on how student performance in the PAIR Portfolio Conferences exceeded their expectations for students that were based on their prior evaluation of high, average and low academic rated students chosen for the conference.

For their previously designated high academically performing students, teachers reported that the Portfolio Conference format demanded more than pat, correct responses and some of these student did not excel with this challenge. Average academically rated students sometimes displayed unexpected levels of sophistication in their range of description and interpretation of their creative work in ways that suggested a deeper understanding of their application of content knowledge and creative process previous unrevealed in their conventional academic assessments. Previously designated low rated students, although still revealing their challenges with formal expression in writing, speaking, or the capacity to provide 'quick responses' to questions, did come up with surprisingly substantial responses when given the time to think and respond to their work in the context of their group work. And as previously reported, teachers were sometimes astonished that the academically average and low rated students performed at an equal, if not occasionally superior level compared to the high academic rated students when judged by their performance in these portfolio conferences.

Overall, teachers reported in final section of the conference that they were able to find evidence of learning in the context of their children's own understanding of arts integration in their

classroom culture by paying close attention to how their individual students respond to general questions about their learning, make distinctions between descriptions of their own work and their peers, and draw conclusions as to the meaning of their work based on evidence of their own learning in their portfolio artifacts. Through first-hand experience with the portfolio conference protocol, several PAIR teachers felt they should implement portfolio conferences in their own classroom as a necessary tool for prompting student reflection aimed at eliciting increasingly sophisticated description of their learning and urging all students to take the time to summarize their learning experiences based on their class work examples and their personal engagement in collaborative working processes.

In sum, the examples from the conference transcripts illustrate how the student portfolio conference interactions provide validated criteria for determining qualitatively different levels of individual student response to facilitator questions about their work process and products. We now have a catalog of exemplary responses that range from vague to specific, relational to systematic thinking that provide the basis for reliably rating evidence of student learning in the context of the portfolio conference protocol. As a result of these analytic processes, qualitative ratings for teacher and student reflection now can provide the statistical basis of judgment of kind and degree of student learning explored in sections 3E-G of this report.

3E: Statistical Analysis of Student Portfolio Conference Response Ratings

An analysis of the Inter-correlations of PAIR Student Portfolio Conference Performance Assessment Ratings with Student SAIL Arts Integration Surveys, Classroom Teacher PAIL Student Work Ratings, Teacher Portfolio Conference Response Ratings, and Academic Test Scores (by Academic and School Focus Cohorts).

3E: *Table of Contents*

3E: PAIR Student-Portfolio Conference Qualitative Analysis	1
3E: Table of Contents	1
3E: List of Tables, Figures and Appendices	2
3E: Introduction	3
3E-1 Inquiry Question 1: *To what extent are there significant differences in PAIR Treatment School Student Portfolio Conference student ratings within treatment school types and across HAL classified learners?*	3
3E-1 *Summary, Emerging Themes*	4
3E-2 Inquiry Question 2: *To what extent do PAIL individual classroom averaged student work ratings predict individual student Portfolio Conference Ratings?*	5
3E-2 *Summary, Emerging Themes*	6
3E-3 Inquiry Question 3: *To what extent do PAIR Treatment School Student Portfolio Conference ratings predict individual student SAIL Interview response ratings?*	7
3E-3 *Summary, Emerging Themes*	8
3E-4 Inquiry Question 4: *Are there significant links between Treatment School Student Portfolio Conference Ratings and Classroom Teacher Ratings?*	9
3E-4 *Summary, Emerging Themes*	12
3E-5 Inquiry Question 5: *Are there significant links between student Treatment School Student Portfolio Conference Ratings and ISAT standardized Test Scores? How do results vary according to school types?*	13
3E-5 *Summary, Emerging Themes*	16

3E: *List of Tables and Figures*

3E-1 Figure I:	Total Student Portfolio Conference Averaged By PAIR Focus School (arts vs. academic), PAIR Magnet Cluster Schools Types (Math, ELA, World Language), and 3-HAL Student Designation across all PAIR schools.	4
3E-2 Table 1:	PAIR Student Portfolio Conference Interview Rating Correlations with classroom PAIL portfolio work quantity and quality ratings	5
3E-3 Figure 1:	Analysis of Treatment School Individual Student Portfolio Conference Ratings Associations with Student SAIL Arts Integration Interview Response Ratings.	7
E-3 Figures 2-3:	Bivariate Fit of Student SAIL Average Scores by Portfolio Conference Ratings by Academic Focus Schools and Arts Focus Schools	8
3E-4 Table 1:	Teacher and Student Years of Participation by the Final Year of the PAIR Project	9
3E-4 Figure 1:	Teacher and Student Portfolio Conference Performance Averages According to Years of Teacher Participation in the PAIR project	10
3E-4 Figure 2:	Teacher and Student Portfolio Conference Performance Averages according to School Focus (Arts vs. Non Arts)	11
3E-4 Figure 3:	Pearson Correlation Table measures the degree of association between Teacher and Student Portfolio Conference Ratings by Focus Schools (Arts vs. Non Arts)	11
3E-4 Figure 4:	Pearson Correlation Table measures the degree of association between Teacher and Student Portfolio Conference Ratings by Magnet School Cluster Designation [Math, Writing (ELA), and World Languages]	12
3E-5 Figure 1:	Degree of Association of Combined Student Portfolio Conference Ratings and ISAT Combined Academic Test Scores	13
3E-5 Figure 2:	Degree of Association of Student Portfolio Conference Ratings and ISAT Combined Academic Test Scores According to Pre-designated Academic Achievement Levels (High, Average, and Low)	14
3E-5 Figure 3:	Degree of Correlation among Student Portfolio Conference Ratings, Student SAIL Interview Response Ratings, and ISAT Combined Academic Test Scores	15
3E Figure 4:	Differences in the Degree of Correlation between Student Arts Integration and Academic Outcomes when comparing Arts Focus to Academic Focus PAIR Treatment Schools	15
3E-5 Figure 5:	Differences in the Degree of Correlation between Student Arts Integration and Academic Outcomes when Comparing the Particular Academic Focus of PAIR Treatment Cluster Schools (Math, ELA, World Language)	16

3E: Introduction

This chapter provides statistical profiles of student portfolio conference response ratings that are consequently examined for their relationship to academic performance cohort designations, type of schools, PAIL class student work ratings, teacher portfolio conference ratings, and student academic test results. Five inquiry questions define each segment of this report:

3E *Inquiry Questions*:

1) To what extent are there significant differences in PAIR Student Portfolio Conference student ratings within treatment schools types and across HAL-classified academic achievers?

2) To what extent do PAIL individual classroom averaged student work sample ratings predict individual Student Portfolio Conference ratings?

3) To what extent do PAIR Treatment School Student Portfolio Conference ratings predict individual student SAIL Interview response ratings?

4) Are there significant links between Treatment School Student Portfolio Conference Ratings and Teacher Portfolio Conference Ratings?

5) Are there significant links between student Treatment School Student Portfolio Conference Ratings and ISAT standardized Test Scores? How do results vary according to school types?

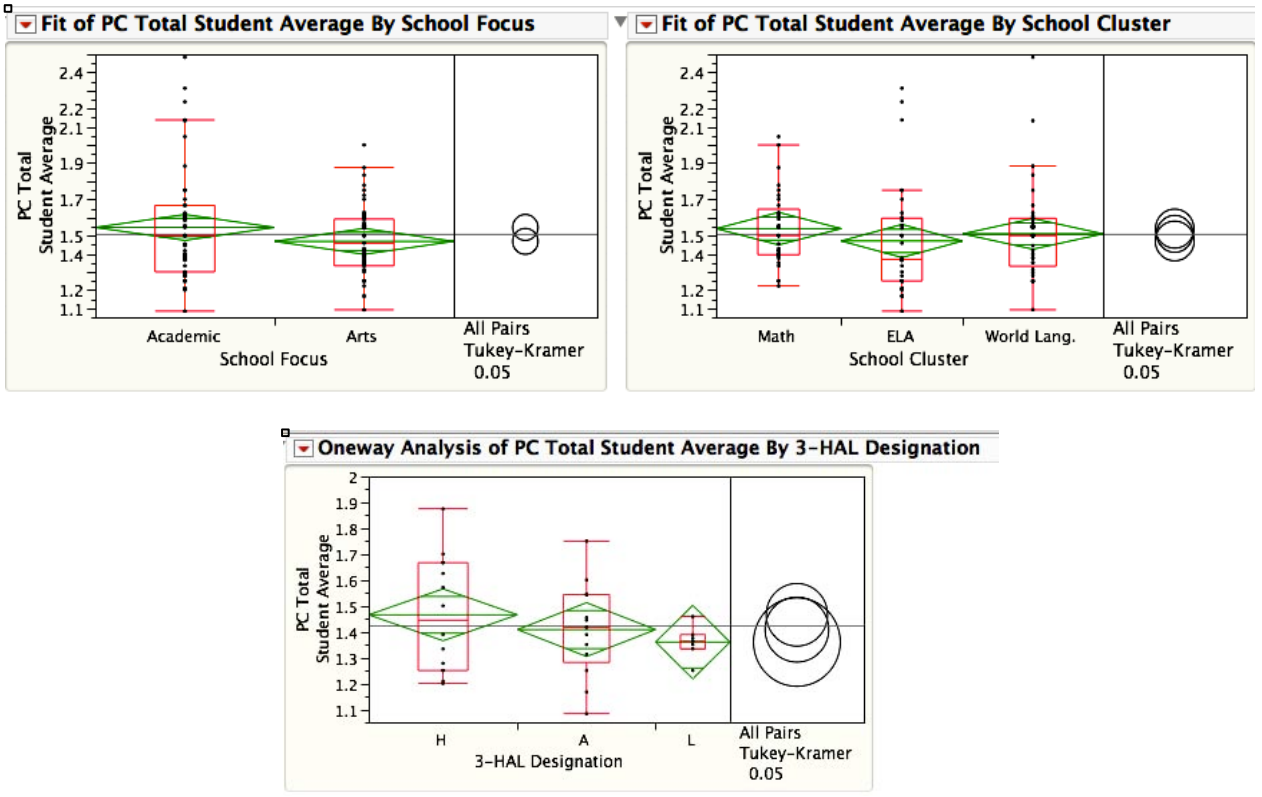
3E-1 Inquiry Question 1: *To what extent are there significant differences in PAIR Treatment School Student Portfolio Conference student ratings within treatment school types and across HAL classified learners?*

The interview content analysis in the previous section demonstrated that portfolio conference ratings in PAIR treatment schools are an important measure of these students' reflective understanding of arts integration learning processes. The following analyses explores= this hypothesis by determining of the level of statistically significant differences, or lack thereof, of the student portfolio conference ratings according to school focus and to the pre-designated high–average-low (HAL) student academic achievers.

3E-1: *Data Display Indications.* The first three displays of Student Portfolio Conference Performance Treatment School ratings indicate no significant statistical differences among PAIR Treatment Academic and Arts Focused Schools, PAIR Magnet Cluster Schools Types, nor

among pre-designated HAL academic rated students. *These findings provide evidence that the portfolio conference ratings serve as an important measure of these students' cognitive and meta-cognitive understanding of their arts integration work that is not predicted by the particular focus or academic priority of the school, and, most remarkably, a measure that is not strongly predicted by previous academic designation of the students.*

3E-1 Figure I: Total Student Portfolio Conference Averaged Ratings By PAIR Focus School (arts vs. academic), PAIR Magnet Cluster Schools Types (Math, ELA, World Language), and 3-HAL Student Designation across all PAIR schools.



3E-1 ***Summary, Emerging Themes***: All ANOVA comparative statistics were determined to be non-significant**. These data suggest that** *Student Portfolio Conference outcome data suggest that the PAIR treatment school has taken effect across all schools and student cohorts with virtually equal effect.* However, it is likely that there is insufficient statistical power due to limited sample size. Thus researchers were able only to detect weak trends in the data that favored academic schools and expected correlation of HAL designations with portfolio conference performance.

Speculation about sample size notwithstanding, *it appears more likely that the PAIR portfolio conference assessments represent an alternative measure of arts integrated learning that has taken effect equally among the various school and student cohorts*. This interpretation of the data is also supported by the testimony of many teachers cited in the previous section of this report. When asked to reflect on their own students' performance, many classroom teachers expressed surprise that the level of conventional academic performance in their class did not correspond with performance in the conference. That is, many teachers were greatly surprised to observe Low-rated HAL students in portfolio conferences performing in ways that were virtually indistinguishable from the Average or High-rated learners.

* * *

3E-2 Inquiry Question 2: ***To what extent do PAIL individual classroom averaged student work ratings predict individual student Portfolio Conference Ratings?***

The next step in this analysis is to explore the relationship of the student portfolio conference performance ratings with respect to the quantity and quality ratings given to the individual classroom student work averaged ratings.

3E-2 ***Data Display Indications:*** The table below reveals that the portfolio conference ratings are for the most part unrelated to all analyses of PAIL classroom quantity or quality indicators. All ANOVA comparative statistics were determined to be either non-significant (Wilcoxon / Kruskal-Wallis Rank Sums 1-way Test, ChiSquare Approximation NS) or even negatively correlated.

3E-2 Table 1: PAIR Student Portfolio Conference Interview Rating Correlations with classroom PAIL portfolio work quantity and quality ratings

Principal Variable: Student Portfolio Conference Averaged Ratings	by PAIL Variable	Spearman Correlation Statistic	p value
Student Portfolio Conference Averaged Ratings	PAIL Student Avg. QUANTITY Ratings (Classroom)	-.23	< .05
Student Portfolio Conference Averaged Ratings	PAIL Student Avg. QUALITY Ratings (Classroom)	NS	NS
Student Portfolio Conference Averaged Ratings PAIR ACADEMIC Focus Schools	PAIL Student Avg. QUANTITY Ratings (Classroom)	-.45	< .001

[table continued on next page]

Student Portfolio Conference Averaged Ratings PAIR ACADEMIC Focus Schools	PAIL Student Avg. QUALITY Ratings (Classroom)	.13	NS
Student Portfolio Conference Averaged Ratings PAIR ARTS Focus Schools	PAIL Student Avg. QUANTITY Ratings (Classroom)	NS	NS
Student Portfolio Conference Averaged Ratings PAIR ARTS Focus Schools	PAIL Student Avg. QUALITY Ratings (Classroom)	NS	NS

Furthermore, the statistically significant *negative* correlations between portfolio conference ratings and *quantity* of portfolio work produced in classrooms serve to reinforce further the finding that the averaging of student work documentation according to each teacher cohort failed to predict to any degree their individual portfolio conference.

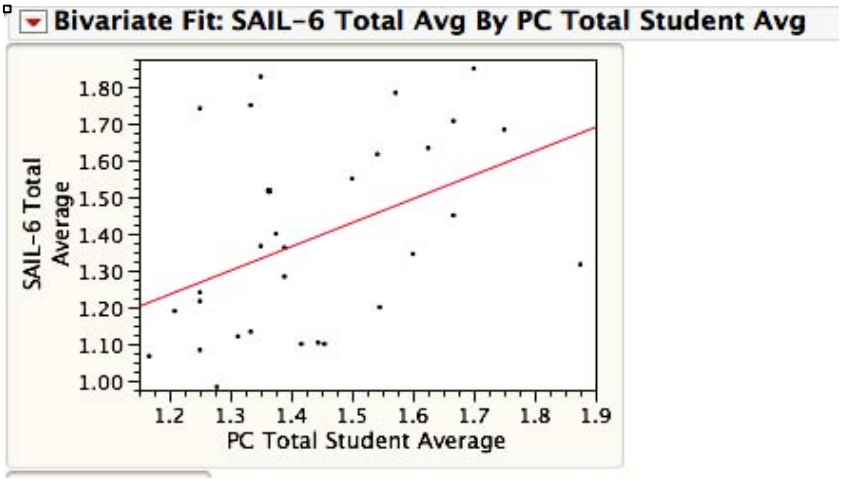
3E-2 ***Summary, Emerging Themes***: There is no statistically significant correlation between PAIL classroom student work ratings and individual student portfolio conference ratings. Because the quantity and quality of the PAIL work documentation was not sufficiently standardizes across classrooms or schools, student work could not be distinguished or compared at the individual student level. Because the quality of PAIL work samples were rated only at the classroom level, the correlation between these two factors could not be determined on an individual basis. *Since the PAIL work documentation is largely due to degree of teacher response to PAIR professional development objectives, it could not serve as a viable measure of the individual effort or understanding of arts integration learning by the students. Thus it should not be surprising that the association of PAIL work to Portfolio Conference responses was statistically negligible, regardless of their school, classroom teacher, or teaching artist affiliation*. Had the PAIR project resulted a systemic collection of highly differentiated individual student PAIL samples, analytic processes would have been more able to detect the existence of a statistical association between student quantity or quality work samples and portfolio conference responses on the individual student level. CAPE has recognized this design flaw in the study, and has since emphasized a more standard and individualized system of student portfolio work development in follow-up studies.

* * *

3E-3 Inquiry Question 3: ***To what extent do PAIR Treatment School Student Portfolio Conference ratings predict individual student SAIL Interview response ratings?***

3E-3 ***Data indications:*** Individual student SAIL interview response ratings, unlike the previous measures, are highly associated with portfolio conference response ratings. The following charts show that the SAIL interview and performance tasks are highly related to the student portfolio conference ratings, depending on the type of school focus (arts or academic) and content knowledge emphasis (Math, Writing, or World Language-Cultural studies). The first chart below displays a very strong overall relationship between the SAIL student interview and the portfolio conference ratings [F Ratio 6.61 Prob > F .02; Spearman Correlation .45 $p < .01$].

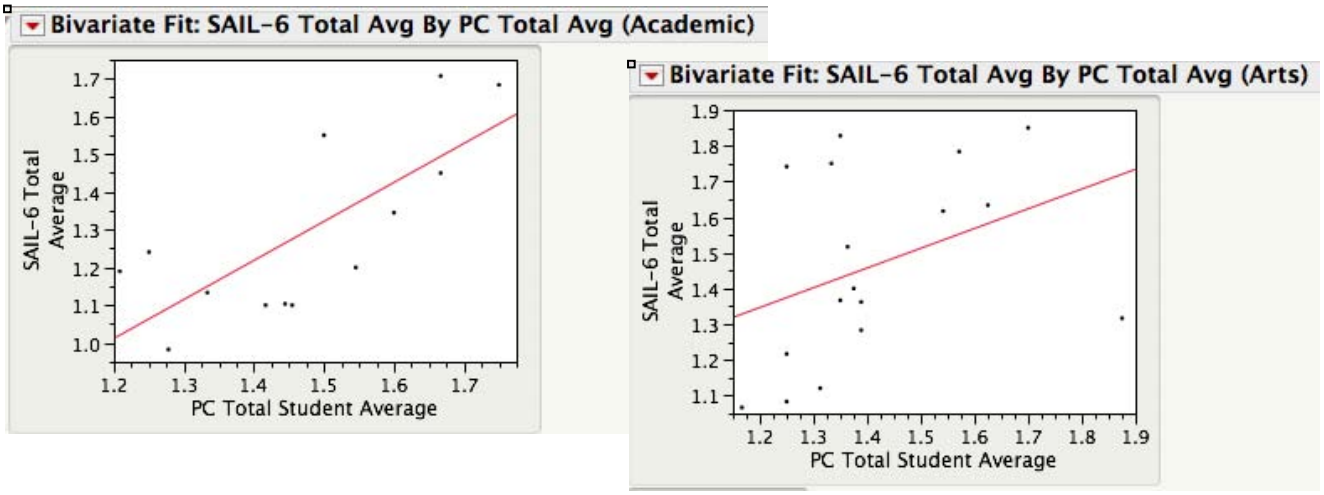
3E-3 Figure 1: Analysis of Treatment School Individual Student Portfolio Conference Ratings Associations with Student SAIL Arts Integration Interview Response Ratings.



The strong fit between these two student learning outcomes suggests the articulation of arts integration principles and the level of performance in both the SAIL and portfolio conference ratings are significantly related. The strong association between SAIL interview ratings and the portfolio conference performance assessment ratings suggests also that these students' have gained the ability to demonstrate reflective understanding of their arts integration learning processes both within and outside the presence of their own personal work. This result validates the primary goals of the PAIR project: that is, the development of young students' meta-cognitive awareness and concrete articulation of the arts integration learning processes is a result of three years of engagement in the PAIR project.

In contrast with the previous determination that the Arts Focused school students who rated higher in terms of the degree of sophistication of their reflective comments on their arts integration work during the SAIL interviews (section 3A of this report), it appears that the degree of association between SAIL and Portfolio Conference ratings is significantly higher in Academic focused PAIR schools (Spearman correlation .68 $p < .01$) compared to Arts focused schools; Spearman correlation .46 $p < .07$) [see scatterplot charts below].

E-3 Figures 2-3: Bivariate Fit of Student SAIL Average Scores by Portfolio Conference Ratings by Academic Focus Schools (ANOVA F-Ratio 14.28, Prob F < .003) and Arts Focus Schools (ANOVA NS))



The finding that PAIR Academic Focus school student portfolio conference ratings are more likely to predict SAIL scores than are Arts Focus school ratings, suggests that the portfolio conference in academic schools provides a broader or more cohesive focus on arts *plus* arts integration learning, compared to the *arts only* focus schools.

In addition, the relationship between SAIL and Portfolio Conference factors were relatively much higher in the Writing (ELA) cluster schools (Spearman r .86, $p < .02$) compared to either the World Language (weak statistical trend) or Math schools (not significant).

3E-3 *Summary, Emerging Themes*: Results from the student SAIL interview and Portfolio Conference performance assessment tasks suggest that a growing understanding of arts integration learning can be demonstrated by these two validated assessment instruments

independently over time. The high degree of correlation between ratings from these two assessment measures by the third year of the project suggests that an even higher level of interaction is taking place between both arts and arts integration learning processes over time. *The relatively higher degree of correlation between SAIL and Portfolio Conference ratings in Academic Focus arts integration schools suggests that the portfolio conference protocol captures the striking ability of non-arts magnet school students to articulate meta-cognitive connections between arts making processes and their integration with core academic.*

* * *

3E-4 Inquiry Question 4: ***Are there significant links between Treatment School Student Portfolio Conference Ratings and Classroom Teacher Ratings?***

Researchers hypothesized that teaching practices in PAIR may be significantly linked with student portfolio conference responses because (a) student portfolio conference performance depended very much on classroom teacher support for the development of the portfolio work itself and (b) classroom teachers were rated individually for their performance in the portfolio conference with respect to their articulation of project goals and their ability to describe the complexities of arts integration learning. Two classroom teacher factors are explored below for their potential influence on student portfolio conference performance: 1) Years of teacher participation in PAIR and 2) Teacher response ratings in the portfolio conference.

3E-4 *Data Display Indications*: The influence of years of teacher project participation on both teacher and student portfolio conference performance can be understood best as significant statistical trends in the chart below [3E-4 Table 1]. Note that the amount of teacher years in the project is inversely related to the student grade level (Grade 4 teacher = 3 years; Grade 5 teacher = 2 years; Grade 6 teacher = 1 year) and inversely related to the years of student experience in the PAIR project.

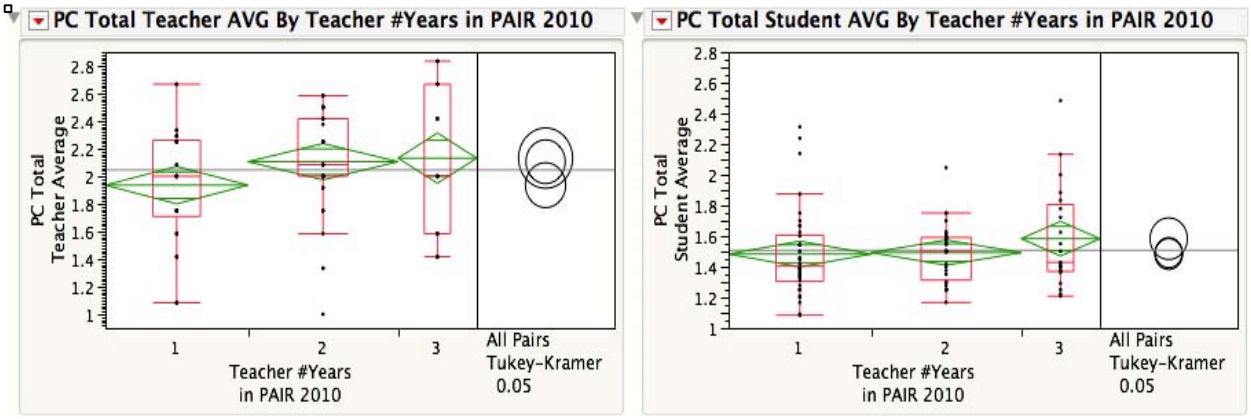
3E-4 Table 1: Teacher and Student Years of Participation by the Final Year of the PAIR Project

Years of Partticipation	Teacher Years =3	Teacher Years =2	Teacher Years =1
Student Years = 1	GRADE 4		
Student Years = 2		GRADE 5	
Student Years = 3			GRADE 6

Several findings based on Years of Teacher Participation in PAIR can be taken from the data display below [3E-4 Figure 1]. First, the teachers in general are rated significantly higher than the elementary school students in terms of level of sophistication, detail, and perspective. This finding was expected and confirms the validity of the scoring rubric for detecting hierarchical levels of sophistication in reflective comments[1]. The chart on the left outlines a statistical trend that shows that teachers with more experience in the project were rated better than those with only one year of experience. This trend confirms the value that the professional development provides teachers over the course of the project. The right chart however is perplexing: the youngest students with the least experience with the project appear to be demonstrating meta-cognitive understanding of their arts integration work at a higher level than the older, more experienced students.

3E-4 Figure 1: Teacher and Student Portfolio Conference Performance Averages According to Years of Teacher Participation in the PAIR project

(Grade 4 teacher = 3 years; Grade 5 teacher = 2 years; Grade 6 teacher = 1 year)

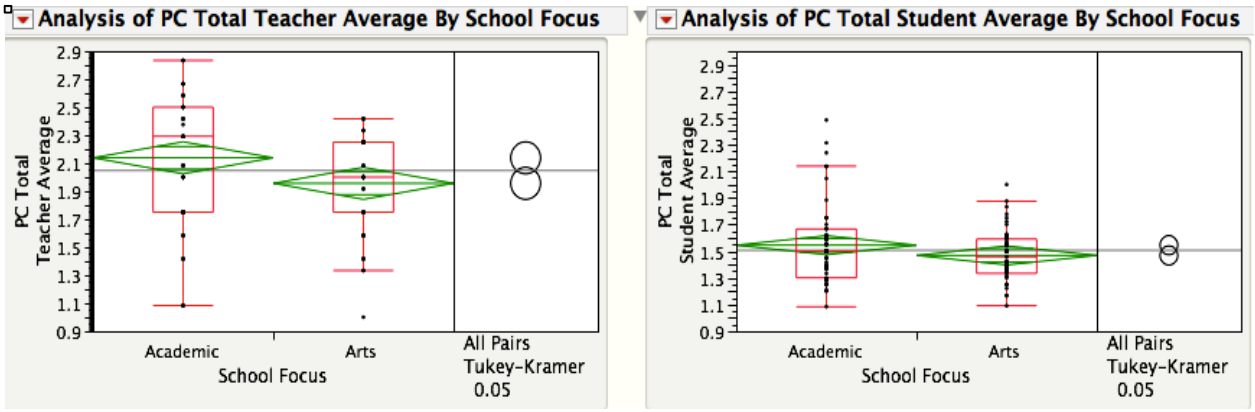


It is possible that the ongoing professional development of the fourth grade teachers over three years of the project development – the 'Fourth Grade Effect' that Gail Burnaford describes in Part I of this report – promoted the level of their most recent students reflective understanding of their arts integration work to a level commensurate with, if not beyond, the students who had three years participation in the project? This possibility will be explored in more depth in next chapter of this report [3F].

[1] It should be noted that, in rare instances, teachers were rated at the same level of articulation as some of the students in the conference transcriptions.

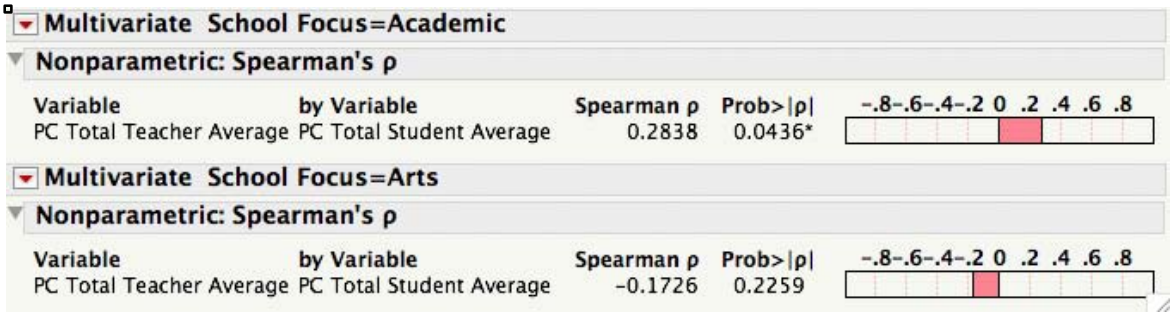
Findings concerning the influence of the school focus or academic discipline priority on Teacher and Student Portfolio Conference ratings can be drawn from the next several figures. As indicated earlier in this chapter with the student cohorts [Sections 3E-1 and 3E-3], the primarily academic-focused PAIR Schools outperformed the primarily arts-focused schools in both teacher [left chart below: Wilcoxon / Kruskal-Wallis Tests (Rank Sums) 1-way Test, Chi Square Approximation $p < .01$] and student ratings [positive, though not statistically significant trends].

3E-4 Figure 2: Teacher and Student Portfolio Conference Performance Averages according to School Focus (Arts vs. Non Arts)

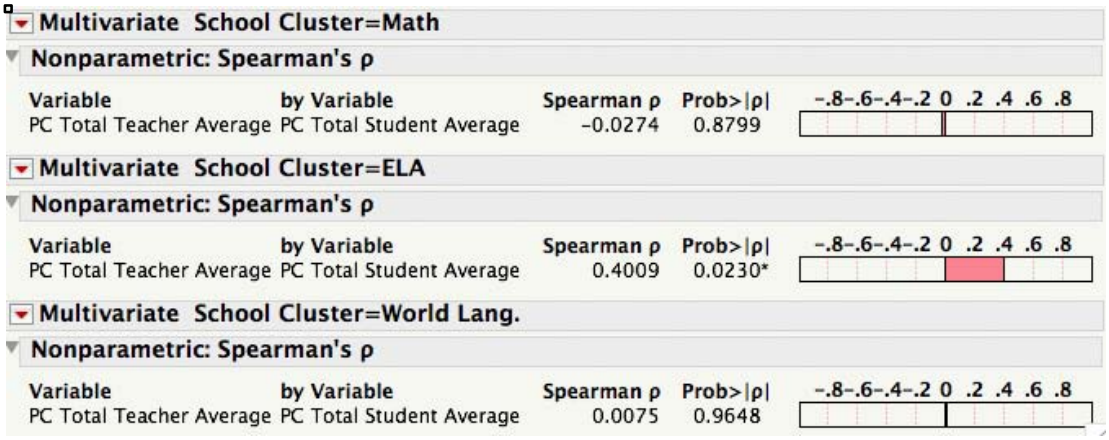


The correlation tables below [3E-3 Figures 3-4] profiles the influence of the teacher portfolio conference ratings with respect to student portfolio conference performance. These charts list statistically significant relationships only in the academic focused schools and in the context of the Writing Cluster Schools in particular.

3E-4 Figure 3: Pearson Correlation Table measures the degree of association between Teacher and Student Portfolio Conference Ratings by Focus Schools (Arts vs. Non Arts)



3E-4 Figure 4: Pearson Correlation Table measures the degree of association between Teacher and Student Portfolio Conference Ratings by Magnet School Cluster Designation [Math, Writing (ELA), and World Languages]



Multivariate School Cluster=Math

Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|	-.8 -.6 -.4 -.2 0 .2 .4 .6 .8
PC Total Teacher Average	PC Total Student Average	-0.0274	0.8799	

Multivariate School Cluster=ELA

Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|	-.8 -.6 -.4 -.2 0 .2 .4 .6 .8
PC Total Teacher Average	PC Total Student Average	0.4009	0.0230*	

Multivariate School Cluster=World Lang.

Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|	-.8 -.6 -.4 -.2 0 .2 .4 .6 .8
PC Total Teacher Average	PC Total Student Average	0.0075	0.9648	

3E-4 ***Summary, Emerging Themes:*** This section examined the possible effects of teacher ratings on student level of understanding of arts integration learning during their portfolio conference performance. Since there was no significant evidence of teacher PAIL student work documentation influencing SAIL interview scores, this chapter looked at the years of project professional development and direct participation of teachers in the portfolio conferences as a way to gauge the support students received from their teachers. Researchers surmised that students whose teachers were highly rated in the portfolio conferences would better be able to reflect on the meta-cognitive aspects of arts learning during the student portfolio conference sessions.

Relatively weak statistical trends in the data analysis suggest how crucial the teacher support may be. Because the very youngest students equaled or excelled the older, more experienced students and that teachers with higher ratings of their ability to articulate project goals and results tended to predict higher student ratings, the professional development and resulting changes in classroom practices may well be critical to optimizing the impact of the PAIR program. The emergence of the Writing Schools as a school where the association between teacher portfolio conference ratings is strongest supports earlier findings that the value of arts integration may be most clearly detected with PAIR units that most closely with fundamental concepts shared between arts and academic disciplines.

* * *

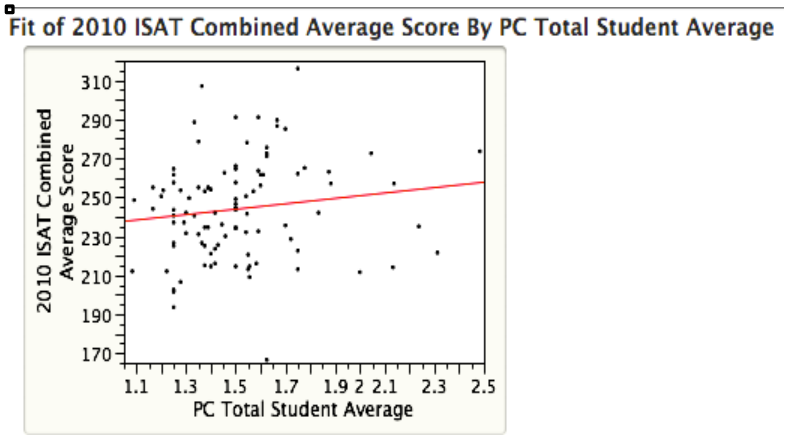
3E-5 Inquiry Question 5: ***Are there significant links between student Treatment School Student Portfolio Conference Ratings and ISAT standardized Test Scores? How do results vary according to school types?***

The final Inquiry question of this chapter examines the association of student portfolio conference findings on standardized measures of academic achievement.

3E-5 *Data Display Indications*: The following scatterplots and correlation tables indicate that there are statistically significant links between Student Portfolio Conference Ratings and combined Math and ELA ISAT test scores depending on School Focus, Primary Academic focus, and pre-designated student HAL academic achievement cohorts.

The first display [3E-5 Figure 1 below] plots the statistically significant relationship between averaged Student Portfolio Ratings and ISAT academic achievement scores in the PAIR Treatment School longitudinal cohort focus schools [Spearman correlation .43 $p < .02$]. This relatively weak general effect is not dependent on the type of Focus School (Arts vs. Non Arts) and only seems to be strong in the Math Cluster Schools.

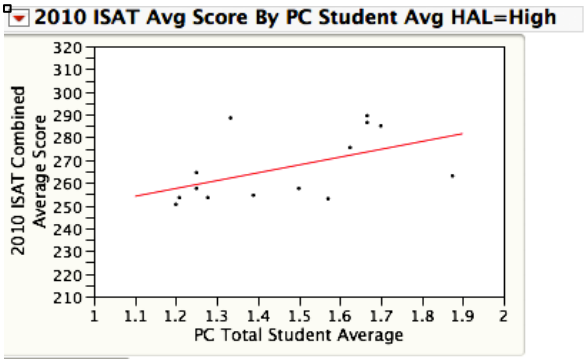
3E-5 Figure 1: Degree of Association of Combined Student Portfolio Conference Ratings and ISAT Combined Academic Test Scores



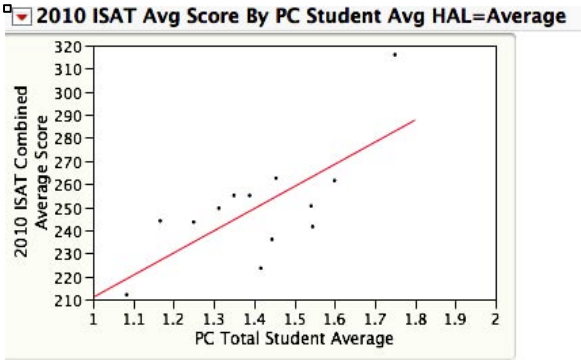
Subsequent analyses of HAL-rated student statistics [3E-5 Figure 2] reveal a startling pattern of association between the portfolio conference ratings and ISAT scores that suggests that the

general ISAT-Portfolio Conference Rating correlation is statistically strong for previously designated High [Spearman .57 p<.03] and moderately strong Average [Spearman .52 p <.07] Academic rated students and wildly unpredictable for the Low Performing students.

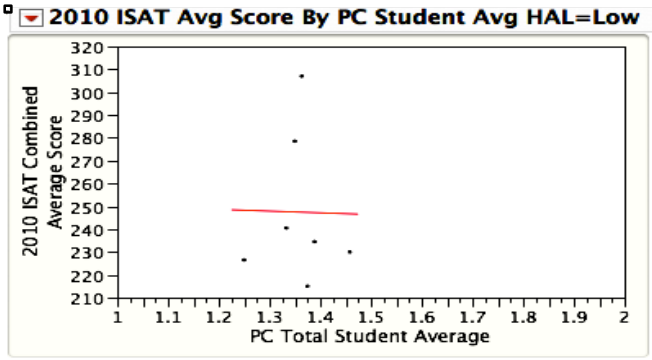
3E-5 Figure 2: Degree of Association of Student Portfolio Conference Ratings and ISAT Combined Academic Test Scores According to Pre-designated Academic Achievement Levels (High, Average, and Low)



[Linear Fit F ratio: 4.09 p < .07]
[Spearman .57 p < .03]



[Linear Fit F ratio: 11.34 p < .01]
[Spearman .52 p < .07]

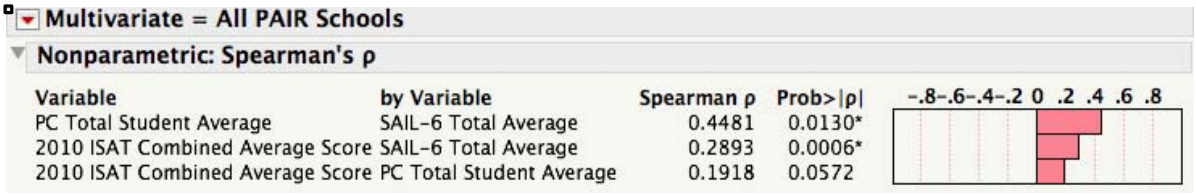


[not statistically significant]

The correlation tables below show how the employment of two inter-related measures of student arts integration learning – the SAIL interview and the PAIR Portfolio Conference protocol - have provided effective tools for developing profiles of student arts integration learning achievement. These data displays also capture the overall significant relationship of the SAIL and Portfolio Conference Assessments with academic test scores.

The first data display [3E-5 Figure 3] below establishes the primary significant trend: the inter-relationship among the SAIL and Portfolio Conference data indicates that together, these two measures are highly significant predictors of student academic achievement.

3E-5 Figure 3: Degree of Correlation among Student Portfolio Conference Ratings, Student SAIL Interview Response Ratings, and ISAT Combined Academic Test Scores



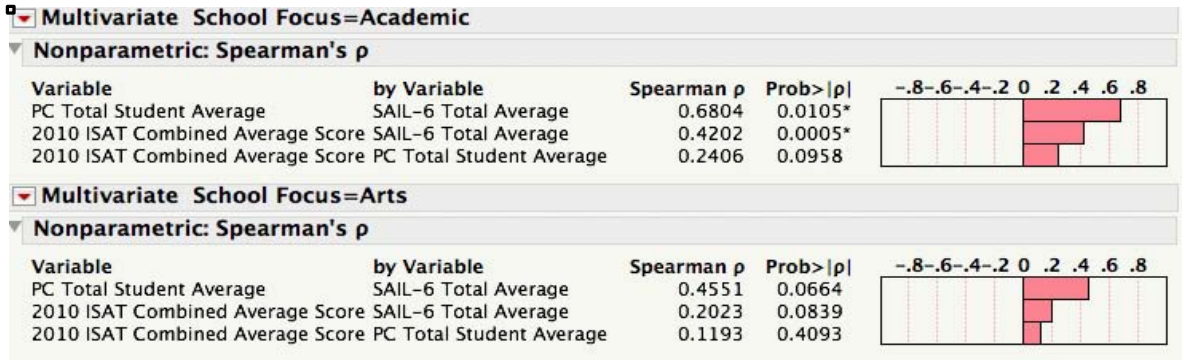
Multivariate = All PAIR Schools

Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|
PC Total Student Average	SAIL-6 Total Average	0.4481	0.0130*
2010 ISAT Combined Average Score	SAIL-6 Total Average	0.2893	0.0006*
2010 ISAT Combined Average Score	PC Total Student Average	0.1918	0.0572

The next data display [3E Figure 4] demonstrates differences in the 'degree of correlation' between portfolio conference performance assessment rating and academic achievement when comparing Arts and Academic Focus arts integration programs. These data suggest that the development of meta-cognitively rich, reflective understanding of how arts integration learning impacts academic achievement is significantly more likely to occur in PAIR Academic Focus schools in comparison to PAIR schools that focus primarily on arts learning.

3E Figure 4: Differences in the Degree of Correlation between Student Arts Integration and Academic Outcomes when comparing Arts Focus to Academic Focus PAIR Treatment Schools



Multivariate School Focus=Academic

Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|
PC Total Student Average	SAIL-6 Total Average	0.6804	0.0105*
2010 ISAT Combined Average Score	SAIL-6 Total Average	0.4202	0.0005*
2010 ISAT Combined Average Score	PC Total Student Average	0.2406	0.0958

Multivariate School Focus=Arts

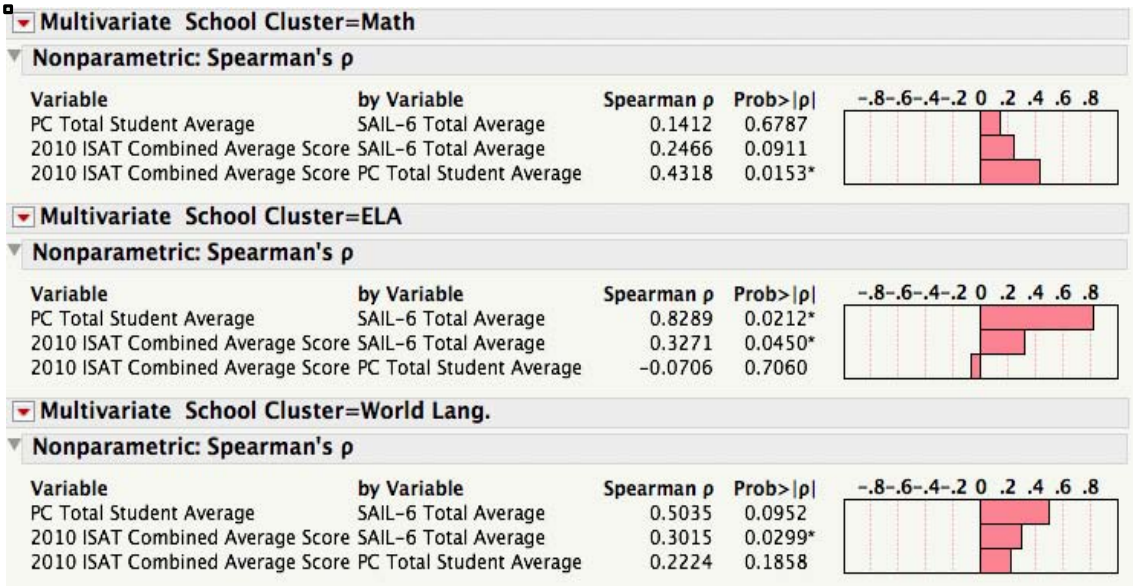
Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|
PC Total Student Average	SAIL-6 Total Average	0.4551	0.0664
2010 ISAT Combined Average Score	SAIL-6 Total Average	0.2023	0.0839
2010 ISAT Combined Average Score	PC Total Student Average	0.1193	0.4093

The final three data displays [3E Figure 5] reveal interesting possible differences among the three case study PAIR Magnet Cluster School groups. Due to small sample size, these data patterns can only be considered statistical trends, ye it does appear that there may have been

striking differences in the impact of PAIR according to the particular academic focus of the school. Math Focus schools, for example, benefitted most broadly by PAIR in general and portfolio conference ratings that were by far the most predictive of ISAT scores among all the school types. ELA schools and World Language School students were much more connected to ISAT results through the SAIL Interview data compared to the portfolio conference ratings.

3E-5 Figure 5: Differences in the Degree of Correlation between Student Arts Integration and Academic Outcomes when Comparing the Particular Academic Focus of PAIR Treatment Cluster Schools (Math, ELA, World Language)



Multivariate School Cluster=Math

Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|
PC Total Student Average	SAIL-6 Total Average	0.1412	0.6787
2010 ISAT Combined Average Score	SAIL-6 Total Average	0.2466	0.0911
2010 ISAT Combined Average Score	PC Total Student Average	0.4318	0.0153*

Multivariate School Cluster=ELA

Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|
PC Total Student Average	SAIL-6 Total Average	0.8289	0.0212*
2010 ISAT Combined Average Score	SAIL-6 Total Average	0.3271	0.0450*
2010 ISAT Combined Average Score	PC Total Student Average	-0.0706	0.7060

Multivariate School Cluster=World Lang.

Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|
PC Total Student Average	SAIL-6 Total Average	0.5035	0.0952
2010 ISAT Combined Average Score	SAIL-6 Total Average	0.3015	0.0299*
2010 ISAT Combined Average Score	PC Total Student Average	0.2224	0.1858

From the viewpoint of the portfolio conference facilitators and observers, the primary indications of PAIR program effect on student learning demonstrated by the Math schools may be due to the degree of learning transfer that was stimulated by connecting math – a subject areas most removed from the arts in elementary schools – with dance choreography and visual art. This particular path of arts processes, as was indicated in the portfolio conference student responses in math schools, may have made the most compelling case for arts integration learning transfer.

3E-5 *Summary, Emerging Themes*: The statistical analysis of the portfolio conference data indicates that a significant link has been established with the SAIL interview response ratings and ISAT test score data. *Similar to the associations between the ISAT and SAIL arts integration interview ratings, the significant degree of ISAT - Portfolio Conference student*

learning outcome correlations can be understood as evidence of the causal link that can exist among multiple subject areas when teachers and teaching artists intentionally focus on learning processes and concepts that between arts and academic content and skills. In addition, the results of the Portfolio conference assessment outcomes confirm the value of arts integration teaching as an approach to engaging young children in the investigation, discussion and demonstration key core academic skills through arts learning processes that result in measurable levels of meta-cognitive understanding of integrative learning practices.

* * *

In the next chapter [3F], all teacher professional development factors will be reviewed comprehensively for their possible impact on all measures of student learning.

3F: Statistical Analysis of Teacher and Student Intersection Outcomes

Methods and Detailed Statistical Analyses of the Intersections among PAIR Teacher Professional Development Outcomes, Student Learning Outcomes, and Student and School Demographic Factors

3F: Table of Contents

3F: Teacher and Student Outcome Intersection Analysis: *Methods and Detailed Statistical Analyses of the Intersections among PAIR Teacher Professional Development Outcomes, Student Learning Outcomes, and Student and School Demographic Factors*	1
3F: Table of Contents	1
3F: List of Tables, Figures and Appendices	4
3F: Introduction	7
Formulating and testing the PAIR teacher-student learning intersection hypothesis	7
The multivariate analytic framework	8
Review of teacher professional development and student learning factors in previous data analysis	9
3F-1: Categorizing Possible Intersections Between Teacher Professional Development and Student Learning Outcome Variables in PAIR Treatment Schools	12
3F-2: Methodological Implications for Multivariate Analysis of the Eleven Teacher-Student Program Outcomes	14
3F-2: Seven PAIR Teacher Professional Development Outcome Variables (A: I-VII)	15
3F-2 Teacher PD Outcome Variable A-I: *Teacher Key Effect Ratings*	15
3F-2 Teacher PD Outcome Variable A-II: *Teacher Years of Participation in the PAIR Project*	16
3F-2 Teacher PD Outcome Variable A-III: *Teacher PAIR Professional Development (PD) Session Attendance.*	17
3F-2 Teacher PD Outcome Variable A-IV: *Teacher Year-End Curriculum and Survey [YECS] Key Questions*	17
3F-2 Teacher PD Outcome Variable A-V: *Classroom PAIL Student Work Sample Ratings*	19
3F-2 Teacher PD Outcome Variable A-VI: *Combined Teacher Professional Development Ratings*	20
3F-2 Teacher PD Outcome Variable A-VII: *Teacher Portfolio Conference Interview Response Ratings*	20

3F-2 Student Survey Response and Performance Assessment Outcome Variables (B: I-IV)	21
3F-2 Student Survey Response Outcome Variable B-I: *Student Survey Response Ratings*	22
3F-2 Student Learning Outcome Variable B-II: *Student Snapshots of Arts Integration Learning (SAIL) Interview and Performance Assessment Response Ratings*	23
3F-2 Student Learning Outcome Variable B-III: *Student Portfolio Conference Interview Response and Performance Assessment Ratings*	24
3F-2 Student Learning Outcome Variable B-IV: *Student Academic Performance Ratings*	25
3F-3: Statistical Evidence of Teacher-Student Performance Outcome Intersections	26
A taxonomy of statistical analysis methods	26
3F-3 Teacher Outcome Exhibit I: *Four Exemplary Teacher Effect Designations* and their Association with Teacher Professional Development and Student Learning Outcomes	29
Statistical trends connecting 'Teacher Effects' to student survey and learning outcomes	33
3F-3 Teacher Outcome Exhibit II: *'Years of Teacher Participation in PAIR'* and its Influence on Teacher and Student Outcomes	36
3F-3 Teacher Outcome Exhibit III: *The 'Averaged Teacher Attendance' Ratings in PAIR Professional Development Sessions* and their Association with Teacher Survey Ratings, Teacher Professional Development Outcomes, and Student SAIL Interview Ratings	38
3F-3 Teacher Outcome Exhibit IV: '*Teacher Year-End Curriculum and Survey (YECS)' Ratings* and their Association with other Teacher Professional Development and Student Learning Factors	40
YECS survey questions and the ISAT academic test scores	42
3F-3 Teacher Outcome Exhibit V: '*Quantity' and 'Quality' Ratings of PAIL Classroom Student Work Samples* and their Association with Teacher Performance in Portfolio Conference, Student Survey Response and Learning Outcomes	45
3F-3 Teacher Outcome Exhibit VI: '*Combined Teacher Professional Development Factor Ratings'* and their Association with Teacher Portfolio Conference, Student Survey and Student Arts Integration Interview Response Outcomes	47
Teacher Impact on Student ISAT Scores: Patterns and Degrees of Correlation between 'Combined Teacher PD Factors' and 2010 ISAT academic performance scores by 'Years of Teacher Participation in PAIR'	48
3F-3 Teacher Outcome Exhibit VII: '*Teacher Portfolio Conference Ratings'* and their Association with Student Survey Response and Learning Outcomes	51
The influence of teacher portfolio performance outcomes on arts integration student learning	52

3F-3 Student Outcome Exhibit VIII: '*Student Survey Response Ratings'* and their Association with Teacher Professional Development Factors and Student Learning Outcomes, Controlling for Student Demographic Factors	54
3F-3 Student Outcome Exhibit IX: *Student 'Snapshots of Arts Integration Learning' (SAIL) Interview Ratings'* and their Association with Teacher Professional Development Factors and Student Learning Outcomes, Controlling for Student Demographic Factors	56
3F-3 Student Outcome Exhibit X: *'Student Portfolio Conference' Performance Assessment Ratings* and their Association with Teacher Professional Development and Student Learning Outcomes, Controlling for Student Demographic Factors	59
3F-3 Student Outcome Exhibit XI: *Student Academic Performance Ratings* and their Association with PAIR Professional Development, Student Survey, and Student Arts & Arts Integration Outcomes, Controlling for Student Demographic Factors	63
The importance of student academic performance data analysis to the success of the PAIR arts integration project	63
3F-3 Exhibit XI Teacher Professional Development – Student Academic Intersection Data Indications	64
3F-3 Exhibit XI Student Arts Integration Learning – Student Academic Performance Intersection Data Indications	66

* * *

3F: *List of Tables andnFigures*

3F-1 Table 1:	Eleven Treatment School Teacher-Student Longitudinal Cohort Learning Variables	13
3F-1 Table 2:	Six Teacher Key Effect Categories	15
3F-2 Table 1:	KEY Teacher Year-End Curriculum and Survey (YECS) Response Ratings	18
3F-2 Table 2:	KEY Classroom Student PAIL (PAIR Arts Integration Research) Work Sample Ratings	19
3F-2 Table 3:	KEY Student Survey (SS) Response Factors	23
3F-2 Table 4:	KEY Snapshots of Arts Integration Learning (SAIL) Interview Factors	24
3F-3 Table 1A:	Seven Categories of PAIR Teacher Professional Development-Student Learning Outcome Intersection Variables	27
3F-3 Table 1B:	Four Categories of Student Outcomes	28
3F-3 Table 1C:	Categories of Student and Treatment School Demographic Factors	28
3F-3 Exhibit I Table 1:	Mean Score Comparisons of Six Teacher Professional Development Outcomes by Teacher Effect Profile Categories (Exemplary vs. Not Exemplary)	30
3F-3 Exhibit I Figure 1:	Mean Score Differences between Exemplary vs. Non-Exemplary Teacher 'Content Expertise' and 'Documentation to Learn' Effects on 'Teacher Portfolio Conference Ratings'	31
3F-3 Exhibit I Figure 2:	Mean Score Differences between Exemplary vs. Non-Exemplary Teacher 'Fourth Grade Effect' and 'Documentation to Learn' Effects on Teacher PD Combined Factors (super variable) and 'Teacher Portfolio Conference' Ratings	32
3F-3 Exhibit I Figure 3:	Mean Score Differences between Exemplary vs. Non-Exemplary Teacher 'One or More Effects' on 'Teacher Portfolio Conference Ratings'	33
3F-3 Exhibit I Table 2:	Mean Score Comparisons of Four Student Learning Outcomes by Teacher Effect Profile Categories (Exemplary vs. Non-Exemplary Effect)	34
3F-3 Exhibit I Figure 3:	Influence of Mean Score Differences of Teacher 'Document to Learn' Effect on Student Portfolio Conference ratings and the "Healy Effect' on 'Combined Student Survey Responses'	34
3F-3 Exhibit II Figure 1:	Mean Score Differences in Combined Teacher Professional Development Variables (3F-2: A-VI) and Teacher Portfolio Conference Ratings (3F-2: A-VII) According to 'Years of Teacher Participation in PAIR	37

(table continued on next page)

3F-3 Exhibit III Figure 1:	Statistically Significant Associations of 'Averaged Teacher Survey', 'Averaged Classroom Student Work' Ratings, 'Teacher Portfolio Conference' Ratings and Student SAIL Interview Response' Ratings according to Levels of 'Teacher Attendance in PAIR' PD Sessions	39
3F-3 Exhibit IV Figure 1:	Bivariate Fit of Averaged 'Teacher Survey' and Averaged 'Student Arts Integration Learning' (SAIL) Interview Ratings	41
3F-3 Exhibit IV Figure 2:	Correlation Table Tracing the Degree of Association Among the Three Primary 'Averaged Teacher Survey' Ratings and 'Student Arts Integration Learning' (SAIL) averaged Response Ratings	41
3F-2 Exhibit IV Figure 3:	Correlation Table Indicating Random Association between 'Combined Teacher Year-End Survey' Averaged Responses and Combined Averaged Test Scores, Grades 4-6 by the final year of PAIR	42
3F-2 Exhibit IV Figure 4:	Correlation Table Indicating a Positive Association Trends between two of the three Teacher 'Year-End Curriculum and Survey' Responses and the Final Year ISAT Averaged Combined Test Scores, Grades 4-6	42
3F-2 Exhibit IV Figure 5:	Correlation Table Indicating Positive trends between Averaged, Combined Teacher 'Year-End Survey' Responses and ISAT Academic Achievement scores in Grades 4 by the final year of PAIR	43
3F-3 Exhibit IV Figure 6:	Correlation Table Indicating a Shift in Direction and Degree of Association Between Averaged Teacher 'Year-End Curriculum and Survey' Survey Responses and Final Year ISAT Test Scores, Grades 4-6	44
3F-3 Exhibit V Figure 1:	Teacher PC Ratings by PAIL Student Work Quality and Quality Ratings	46
3F-3 Exhibit VI Figure 1:	Bivariate Fit between 'Combined Teacher PD Factors' and 'Teacher Portfolio Conference Performance'	47
3F-3 Exhibit VI Figure 2:	Bivariate Fit between 'Combined Teacher PD Factors' and 'SAIL Student Interview' Ratings	48
3F-3 Exhibit VI Figure 3:	Negative, Statistically Insignificant Correlation of 'Teacher Combined PD Factors' and Student ISAT Scores for Teachers with One Year of Participation in PAIR	49
3F-3 Exhibit VI Figure 4:	Positive, Yet Statistically Insignificant Relationship Between Teacher 'Combined PD Factors' and Student ISAT Scores for Teachers with Only Two Years of Participation in PAIR	49
3F-3 Exhibit VI Figure 5:	Positve, Statistically Significant Correlation of 'Combined Teacher PD Factors' and Student ISAT Scores for Teachers with Three Years of Participation in PAIR	50
3F-3 Exhibit VII Figure 1:	Statistical Association Between 'Teacher Portfolio Conference' Ratings and 'Student SAIL Interview' Ratings in Grade 6 (Final Year of PAIR)	52

3F-3 Exhibit VII Figure 2:	Statistical Association Between 'Teacher Portfolio Conference' Ratings and 'Student Portfolio Conference Ratings' in Grade 6 (PAIR Final Year)	52
3F-3 Exhibit VII Figure 3:	Correlation Table Indicating the Degree of Association Among 'Teacher Portfolio Conference,' 'Student SAIL Interview', and 'Student Portfolio Conference' Ratings	53
3F-3 Exhibit VIII Figure 1:	Bivariate Fit of 'Averaged Student Survey' Response Ratings by Student ISAT Averaged Reading and Math Standardized Test Scores	55
3F-3 Exhibit IX Table 1:	Correlation Table Indicating Degree of Association Between 'Student SAIL Interview' Ratings and 'Combined Teacher PD Factors'	57
3F-3 Exhibit IX Figures 1:	Bivariate Fit of 'Student SAIL Interview Ratings with 'Combined Teacher PD Factors' and 'Teacher Portfolio Conference Ratings'	57
3F-3 Exhibit IX Figure 2:	Bivariate Fit between 'Student SAIL Interview Ratings' and 'Student Portfolio Conference Ratings'	58
3F-3 Exhibit IX Figure 3:	Bivariate Fit between 'Student SAIL Interview' Ratings and Combined Student Averaged Academic Scores'	58
3F-3 Exhibit X Figure 1:	Bivariate Fit between Grade 6 Arts Integration 'Student Portfolio Conference' and 'SAIL Interview' Ratings	60
3F-3 Exhibit X Figure 2:	Bivariate Fit of Only Grade 6 'Student Portfolio Conference' and 'Teacher Portfolio Conference' Ratings	61
3F-3 Exhibit X Figure 3:	Bivariate Fit of Combined (Grades 4-6) Student and Teacher Portfolio Conference Ratings in the Final Year of PAIR	62
3F-3 Exhibit X Figure 4:	Regression Factor Analysis Table Indicating Principal Predictors of Student Portfolio Conference Performance from among all Teacher Professional Development, Student Learning and Demographic Factors	62
3F-3 Exhibit XI Figure 1:	Correlation Table indicating Overall influence of Combined Teacher PD Factors on Student Academic Test Scores, and Progressively Positive Effect of Years of Teacher Participation on Student Academic Test Scores	65
3F-3 Exhibit XI Figure 2:	Comparative Bivariate Fit of Three Student Arts Integration Learning Outcomes with Student Academic Achievement Ratings	67
3F-3 Exhibit XI Figure 3:	Correlation Table indicating the 'Pattern and Degree of Correlation' of Combined Student Academic Achievement Scores in relation to Student Survey Response and Two Arts Integration Learning Outcomes	68
3F-3 Exhibit XI Figure 4:	Regression Factor Analysis Table Indicating the Principal Predictors of Student Academic Achievement from among all Teacher, Student, and Student Demographic Outcome Variables (Excepting 'Family Income')	69

3F: Introduction

Formulating and testing the PAIR teacher-student learning intersection hypothesis

In the PAIR project there was an underlying hypothesis that the systemic incorporation of high quality arts integration teaching practices would (a) positively affect student learning outcomes in both arts and academic subject areas and (b) result in a highly predictive degree of association between the individual teacher and student outcomes. Statistical evidence for the determination of significant links among *individual* teacher professional development, student arts, and student academic learning outcomes were needed to test this hypothesis.

By design, the PAIR research project provided for collecting multiple measures of teacher professional development outcomes that can be tested to test for statistically significant inter-relationships among all teacher outcome ratings, among student learning outcome ratings, and the degree of association at the junction between these two types of variables. Thus this final chapter can focus explicitly on the measurable intersections between individual teaching practices and individual student learning by reviewing results form previous sections of the this report and producing new statistical evidence in order to find significant connections between individual classroom teacher practices most closely associated with differences in student outcomes.

The chart below summarizes the general path of analysis taken in this chapter:

Teacher Professional Development Ratings, → their possible impact on Teacher Performance Outcomes, → and their possible impact on Student Learning Outcomes.

The 'chain of factors' ordered from left to right functions as a logical basis for measuring for impact of PAIR at the intersection of teacher and student learning:

> (a) *If* the *positive inter-correlation of* PAIR professional learning factors (that is, productive attendance in workshops indicated by survey results, curriculum design) provides evidence of the quality and coherency of PAIR arts integration professional development program,

(b) *And a high degree of association between PAIR professional development factors and teacher performance ratings* (that is, quantity and quality of student work documentation, performance in portfolio conferences) provides evidence of the positive impact of teacher learning on teaching practices,

(c) *Then the determination of significant predictive relationships between PAIR teacher professional development or performance outcomes and various student outcomes* (e.g., student surveys, arts integration interviews, performance assessments, test scores) would provide statistical evidence that causal links exist between teacher professional learning factors and student learning success in the PAIR project.

Since positive identification of teacher effect factors and positive treatment-control school comparisons of standardized test score data were reported in Part 1 and 2 of this report, what remains to be seen in Part 3 is whether the data reported earlier can be linked significantly with the results of PAIR internal arts integration assessments and external testing of academic test scores.

The multivariate analytic framework

In order to make the case for significant intersections between teacher and student learning multiple types of data will be represented and analyzed as this section unfolds.

Categorically defined teacher professional development outcome data will be analyzed to understand differences among teacher performance due to grade level assignment, years of participation in the study, or according to exemplary teacher effect designations, as described in Part 1 of the report.

Qualitatively determined teacher and student learning factors based on arts integration survey responses, performance assessment ratings, classroom documentation of student work, and portfolio conference quality ratings will provide the opportunity to test for multiple strands of their influence on student academic learning variables.

Multivariate factor statistical methods including both 'pattern and degree of correlation' procedures and factor regression analysis procedures will be used to determine the highest

degree of statistical association among all teacher and/or student learning outcomes, and a comprehensive array of school and student demographic factors.

Review of teacher professional development and student learning factors in previous data analysis

Up to this point in the report assumptions about high quality teacher implementation of the PAIR program have been drawn from (a) control-treatment school analysis of teacher survey data and reflection sheets, (b) treatment school evaluation of teacher degree and quality of participation in the professional development aspects of the program, and c) treatment school analysis of teacher statements culled from the portfolio conference protocols that focused on their understanding of PAIR project objectives and their ability to describe the impact of the program based on their observation of student performance in the conference.

In the Teacher Impact section (Part 1) of this report, Dr. Burnaford and her colleagues summarized results from (a) control-treatment survey comparisons that focused on indicators of high quality general academic teaching practices and b) within-treatment school analyses of the quality and focus on teacher arts integration practices from surveys, interview, attendance and student work documentation data that provided evidence that the professional development objectives of the PAIR project had been achieved.

In addition to meeting program evaluation objectives, overall analysis of teacher reflections revealed four principal effects of professional development arts integration teaching practices in the PAIR program in the treatment schools:

(1) The 'teacher content knowledge effect' (based on the judgment of 'exemplary' vs. 'non exemplary' degree and articulation of the importance of content knowledge teachers demonstrated in their arts integration units),
(2) The 'document to learn effect' (indicated by exemplary teachers who were most open to PAIR documentation practices and their resulting PAIL portfolio work samples and labels),
(3) The 'fourth grade effect' (the exemplary impact of the program on teachers who worked initially with the teaching artists to create the first stage drafts of the PAIR units and participated in the project throughout the entire three year period), and

(4) The 'Healy effect' (named after the Healy World Language school whose faculty and administrators had previously made a long-term commitment to arts learning initiatives that appeared to provide optimal conditions for developing the PAIR project in their school).

Qualitative analysis of these four effects resulted in subset of PAIR teacher cohorts who rated highest for one or more of these categories of exemplary practice. In Part 3 of the PAIR report these exemplary teacher cohorts will be subsequently evaluated for their possible effect on student learning factors.

In the Student Academic Performance Outcomes section (Part 2) of this report, Scripp and his colleagues demonstrated the extent to which the student learning outcomes in the longitudinal treatment school and cross-section student cohorts compared favorably with the control groups both in terms of (a) achievement ratings from the Illinois Snapshots of Academic Test (ISAT) assessments and (b) statistical evidence of a significant re-distribution of academic ranking among pre-designated academic achievement student cohorts (High, Average, Low) by the end of the project. The statistical measure of equity stemmed from evidence that the high, medium, or low (HAL) academic performance cohorts performed better as a whole and were far less differentiated with respect to academic performance in the treatment PAIR schools compared to control schools by the end of the project. Taken together, these results suggest that, *as PAIR Treatment arts and academic focused schools engaged in high-quality arts integration professional development programs over three years* (as described in Part I), *significantly higher overall student performance levels and the ability of low performing students to close the achievement gap occurred simultaneously* (as reported in Part 2).

In the Student Arts Learning Assessments and their Intersections with Teacher-Professional Development Outcomes section (Part 3) of this report, multiple measures of student arts integration learning and their inter-correlation have already been analyzed in detail in sections 3A-E. Results from internally developed arts learning performance assessment instruments, rubrics for rating student work, interview protocols, and surveys developed or adopted for the PAIR project so far have provided valid and reliable analysis of individual student inter-related understanding of PAIR arts and arts integrated learning concepts, skills, and creative processes.

The individual student Snapshots of Arts Integration Learning (SAIL) interview (section 3A) and the Student Survey (section 3B) outcomes provided two initial indicators of reflective, meta-cognitive understanding of arts integration teaching and learning practices: (a) SAIL scores provided control-treatment comparison ratings that validated the impact of the PAIR program and were also highly correlated with ISAT test score results and (b) Student Survey data provided validation of treatment school changes in the culture and ethos of collaborative teaching, reflective thinking and presentation of arts learning the PAIR treatment schools in contrast to the control schools. Control-Treatment student cohort comparisons of these variables provided positive indications that PAIR resulted in distinct changes in the classroom culture of collaborative teaching and student reflection that appears linked with higher ratings of understanding of arts and arts integration learning processes.

Differentiation within the various types of Treatment School cohorts revealed differences in the quality of student work and student ability to reflect on their work according to individual school or student classification. Analysis of PAIL classroom rated student work samples (section 3C) and individually scored PAIR Portfolio Conference responses (sections 3D & 3E) provided criterion referenced ratings of student work and levels of student reflective understanding of arts integration learning that revealed how students worked with and understood the implications of arts integration teaching and learning.

The range of data reported in Part 3 of this report makes it possible to look for the causal links that may exist among teaching practices, the quality of student work, and eventually the student learning outcomes in the PAIR project.

After detailed inspection of the four classroom PAIL case study samples (3C) and the same four classroom Portfolio Conference case study vignettes (3D), for example, it became increasingly clear that the categorization of teachers according to the four exemplary effects described in the Teacher Impact (Part I) of the report could be linked with the qualitative analysis ratings of the student outcome data contained in the previous sections (3: A-E). The "Content Knowledge Matters Effect" and the "Documenting to Learn Effect" (collecting student work inspires higher levels and quality of teacher and student reflection) were well substantiated in the qualitative analysis of exemplary student work in the first part of this report. Teacher reflections on content

knowledge in math, reading and social studies consistently emerged out of the interpretation of the arts integration student work samples (section 3C), as did reflective comments by the teacher on the documentation processes that captured the creative processes guided by the teaching artists (section 3D). As noted previously in 3C, the finding that the number of years in PAIR predicts quality of student work documentation, for example, validates the "Fourth Grade Effect" described in Section I of this report (long term engagement in the project design and implementation impact matters).

In other cases, data may be linked from the initial teacher professional development outcomes to both student arts and academic outcomes. Once instance of this 'triangulation' of causal links could be substantiated by the World Language Cluster School relatively strong PAIL ratings re-affirming the "The Healy Effect" (as described in Part I of this report) and its link to both accelerated ISAT performance (reported in Part 2), and the possibility that these factors may be linked with various other student arts learning outcomes (reported in Part 3: A-E).

Since the preliminary analysis of these separate exemplary teaching practices and student learning outcomes completed, various hypotheses suggesting that PAIR teacher professional development factors can be modeled and understood as a chain of causal links that contribute to differences in student learning outcome can now be tested statistically.

* * *

3F-1 Categorizing Possible Intersections Between Teacher Professional Development and Student Learning Outcome Variables in PAIR Treatment Schools

Section 3F-1 Inquiry Questions: *What are the 7 types of PAIR treatment school classroom teacher professional development outcomes that may impact the 4 key student learning outcomes in this report? What evidence indicates that there are statistically significant intersections among the 11 teacher professional development and student learning outcomes?*

The table below [3F-1 Table 1] describes *seven* teacher professional development factors (A: I-VII) to be tested for their association with each other and with *four* student learning factors including survey data, two arts integration learning outcomes, and academic test scores (B: I-IV).

3F-1 Table 1: Eleven Treatment School Teacher-Student Longitudinal Cohort Learning Variables

Teacher-Student Intersection Variables	Data Source Type	Description	Modeling Type	Value Range
A. Seven PAIR Teacher Profession Development & Performance Assessment Outcome Variables [Data collected from all PAIR treatment school teachers and analyzed in the final year (3) of the project]				
A-I: **Teacher Key Effect Ratings** *all teachers designated after project completion in 2010*]	Teacher PD outcome designations	Ratings derived from interview and descriptive data (plus two multiple rating categories) [all teachers]	Categorical Rating of Teacher Quality of Response to PAIR	Each teacher effect rated as "No Effect" or "Exemplary Effect"
A-II: **Teacher Years of Participation in PAIR** [*all teachers rated after project completion in 2010*]	Teacher measure of project experience, exposure, availability of resources [inversely related to years of student years of experience in PAIR]	Levels of participation determined by teacher grade level exposure to PAIR (3 years - grade 4 teachers; 2 years - mostly grade 5 teachers; 1 year - mostly grade 6 teachers)	Numeric-Ordered Categorical	Teacher project participation rated as 1,2,3 years [inversely related to years of student participation in PAIR]
A-III. **Teacher PD Session Attendance** [*all teacher data averaged after project completion in 2010*]	Teacher PD Participation measure [*all teachers*]	Maximum 0-4 sessions per year (averaged across years in PAIR)	Numeric-Ordinal	Teacher project participation averaged rating from 0.00- 4.00
A-IV. **Teacher Year-End Curriculum and Survey [YECS] Ratings** [*all teacher data averaged after project completion in 2010*]	Teacher PD self report ratings of PAIR impact on teaching and classroom culture	Scaled response ratings of three key survey questions most differentiated from control school results	Numeric-Continuous Non-parametric Rating Scale	Three individual questions: scaled response 1-6; also average of combined factors
A-V. **Classroom PAIL Work Sample Ratings** [*all classroom data averaged after project completion in 2010*]	Classroom ratings of teacher-teaching artist support for student work documentation [*all teachers*]	Ratting determined by classroom levels of 'quantity' and 'quality' of student work (based on qualitative analysis in section IIID of this report	Numeric-Continuous Non-parametric Rating Scale	7 point scale rubric ratings
A-VI. **Combined Teacher Professional Development 'Super Variable' Ratings** [*all teacher data averaged after project completion in 2010*]	Combination of Teacher ratings from Teacher-Student Intersection Variables IV, V, VII [*all teachers*]	Ratings determined by averaging Variables IV PD attendance), V (YECS key factor survey responses), VII (averaged classroom ratings of student sample work)	Numeric-Continuous Non-parametric Rating Scale	Standardized 7 point scale rubric ratings in all three variables were averaged
A-VII. **Teacher Portfolio Conference Interview Response Ratings** [*all teacher reflection and observation data averaged after project completion in 2010*]	Teacher reflection ratings on program goals, impact and observation of student performance [*small sample teachers*]	Ratings determined by level of sophistication in responses	Numeric-Continuous Non-parametric Rating Scale	7 point scale rubric ratings averaged across PC sections 1 and 3

(table continued on next page)

B. Four PAIR Student Survey Response and Performance Outcome Variables [All data collected from the treatment school longitudinal cohort students and analyzed in the context of the final year (3) of the project]				
B-I. **Student Survey Response Ratings** [*2010 student longitudinal cohorts*]	Survey ratings indicate student perception of classroom teaching practices that most distinguish treatment from control school classroom learning environment factors	Students rate change in treatment classroom culture by presence of 'student reflection' 'collaborative teaching and learning' 'participation in arts learning,' 'role of arts integration' based on their classroom experience.	Numeric-Continuous Non-parametric	Five categories of Averaged Agreement Responses (1.00 – 8.00]
B-II. **Student SAIL Interview and Performance Assessment Ratings** [*2010 6th grade student longitudinal cohorts*]	Student reflection and perception on the nature of arts learning in relation to academic and arts integration teaching and learning	Ratting determined by individual student levels of detail, relational and systems thinking	Numeric-Continuous Non Parametric	7 point scale rubric ratings
B-II: **Student Portfolio Conference Interview Response Ratings** [2010 *small sample student longitudinal cohorts*]	Student demonstration on reflection on the nature of arts integrated learning	Ratting determined by individual student levels of detail, relational and systems thinking	Numeric-Continuous Non Parametric	7 point scale rubric ratings
B-IV: **Student Combined ISAT Academic Performance Ratings** [2010 *all student longitudinal &follow-up cohorts*]	Student understanding of language literacy skills and math concept and problem-solving skills	Rating determined by averaging individual student Reading and Math scores	Number-Continuous	Scaled Scores

* * *

3F-2: Methodological Implications for Multivariate Analysis of the Eleven Teacher-Student Program Outcomes

The Teacher-Student Rating variables are varied in scope and statistical power depending on the model type, properties and range of the data structures listed above. Each of the sections below present a brief summary of each of the eleven variables and their implications for data analyses in terms of the degree of association—both within and across— the *seven* teacher (A: I-VII) and the *four* student (B: I-IV) survey, interview, or performance assessment responses.

3F-2: Seven PAIR Teacher Professional Development Outcome Variables (A: I-VII)

The first seven variables capture the range of teacher professional development factors that provided descriptive, yet measurable data related to the quality, coherency and cumulative impact of professional development services provided during the PAIR project. These teacher outcomes range from simple calculations of the degree of teacher exposure or participation in PAIR (variables AII-AIII) to the scaled ratings of the *scope* and *quality* teacher documentation of student work or reflective understanding about the impact of arts integration program practices on student learning (AIV-AVII). The attribution of the exemplary classroom teaching effects to treatment school teacher cohorts (A-I) and averaged professional development factor profiles (A-VI) also may provide statistically powerful evidence of the internal links among teacher outcomes that can be tested for their degree of influence on various factors of student learning.

* * *

3F-2: Teacher PD Outcome Variable A-I: *Teacher Key Effect Ratings*

The exemplary teacher effect ratings were adopted directly from the qualitative analysis performed by Dr. Gail Burnaford in Part I of this report. This designation identifies cohorts of teachers who were cited for their exemplary teaching practices according to four effects resulting from analysis of teacher interview data, quality of participation in the PAIR professional development program and evidence of high quality teaching and documentation practices. The multiple effect categories represent a cohort of teachers who were rated as exemplary in more than one type of effect.

3F-2 Table 2: Six Teacher Key Effect Categories

Teacher Key Effects:	
Teacher Effect 1:	Content Knowledge Effect
Teacher Effect 2:	Document to Learn Effect
Teacher Effect 3:	Fourth Grade (Long-Term Ownership) Effect
Teacher Effect 4:	Healy (School Culture) Effect
Teacher Multiple Effect 5:	More than One Effect
Teacher Multiple Effect 6:	One or More Effects

Multivariate statistical analyses eventually will reveal to what extent these teacher effect designations serve as valuable qualitative criteria for interpreting the impact of the PAIR program in Part I of this report with respect to quantitative ratings of professional development outcomes (AIII-VII) and student survey, interview, and performance assessment ratings that focus on arts learning, arts integration learning, and/or academic performance ratings.

* * *

3F-2 Teacher PD Outcome Variable A-II: *Teacher Years of Participation in the PAIR Project*

This variable is determined by the amount of years a teacher participated in the PAIR project. Common to all teachers were an opporunity to attend five professonal development workshops per year and twenty plus days spent collaobrating with two CAPE teaching artists on the development and implmentation of PAIR unit in two different art forms.

This variable is used to evaluate the possible impact of teacher time spent in the project on various professional development and student learning outcome variables. Evidence that teacher survey and performance task responses, or that student learning outcomes increase significantly according to years of teacher experience in PAIR can be used to validate the effectiveness of the PAIR professional development program. This evidence would suggest also that long-term, ongoing experience in the PAIR program is necessary to achieve optimal program impact results. As presented earlier in section IIIC of this report, teacher initial and long-term experience in the project was describes as the 'fourth grade effect' because of the relatively more intensive training and opporutnities for curriculum refinement that was afforded most to the fourth grade teachers in the PAIR project. Earlier reports that fourth grade students final year cohorts academically outperformed the other grade levels whose teachers had less time for teacher professional development time suggested evidence of the fourth grade effect by the final year of the project. In consequent analyzes, effects of teacher 'years in PAIR' can be tested on some variables that allow for comparisons across all three years of project implementation.

* * *

3F-2 Teacher PD Outcome Variable A-III: *Teacher PAIR Professional Development (PD) Session Attendance.*

This variable is determined by the amount of professional development sessions attended by a teacher within one academic year of the PAIR project. Common to all teachers was the opporunity to attend five professonal development workshops each year of PAIR. The amount of PD session events attended for teachers with more than one year in the project were averaged according to the amount of their respective participation years in the PAIR project.

Like the 'years in the project' factor, this variable can be used to evaluate the impact of teacher professional development sessions on various professional development and student performance outcome variables. Evidence that either or both outcomes increase significantly according to the degree of PD teacher attendance would suggest that teacher commitment to professional development sessions is necessary to achieve optimal results in the PAIR project.

Since the program evovled throughout every year of the project, each teacher grade level cohort had to assimilate new aspects of the program, to reflect on past practice, and to collaborate on the production of updated PAIR curriculum and assessment practices during these PD sessions. The design of these sessions was to provide teacher learning experiences that would sustain arts integration experiences for sixth grade students in the later years of the project, even though sixth grade classroom teachers had received relatively less years of project participation than their fourth and fifth grade colleagues.

* * *

3F-2 Teacher PD Outcome Variable A-IV: *Teacher Year-End Curriculum and Survey [YECS] Key Questions*

Teacher 'Years of Participation' and 'Degree of Attendance' in PD sessions suggest that teachers benefitted more or less by the degree of time spent on professional development aspects of the PAIR program. The teacher Year End Surveys (YECS) provides a primary measure of the *qualitative* difference on the impact of the PAIR professional development program on individual classroom collaborative teaching practices.

As reported earlier in Part I of this report, most of the comparisons of the 'Teacher YECS survey' responses were not significantly different when comparing the treatment to control schools. However, three key teacher survey questions emerged as the primary indicators of impact on PAIR treatment school classroom practices and culture. A high level of agreement that working with the teaching artist by "Brainstorming Ideas" for curriculum development, "Getting Consensus" on the PAIR unit, "Modeling the Process" of collaborative teaching, and/or by combining all three of these factors most effects the quality of PAIR program implementation in the classroom project. Thus, individual differences in the degree of response to these PAIR program survey questions presumably provide a good indication of each individual teacher's commitment to, and satisfaction with, the three most distinctive qualities of the PAIR teacher-teaching artist collaborative practices.

As seen in differences in the control-treatment teacher survey comparisons and the degree and quality of collaboration between the classroom teacher and teaching artists, classroom changes due to quality of teaching artist collaboration establishes the fundamental value of 'pairing' of academic and arts focused schools through the joint effort of classroom teachers and artists in this project. These difference also may explain why the teaching artist relationship in PAIR treatment schools may have contributed significantly to better and more equitable student learning performance outcomes compared to control schools by the final year of the project.

3F-2 Table 1: KEY Teacher Year-End Curriculum and Survey (YECS) Response Ratings

Key Teacher Variable Ratings Relevant to the Quality and Focus of Collaboration between the Classroom Teacher and the Teaching Artist	
Teacher YECS Variable 1:	Brainstorming, Planning with Teaching Artist on PAIR Units
Teacher YECS Variable 2:	Reaching Consensus on PAIR Unit planning
Teacher YECS Variable 3:	Modeling the Process of Arts Integration Collaboration with Teaching Artist in Classrooms
Teacher YECS Composite Variables:	The combined average of the previous three factors.

* * *

3F-2 Teacher PD Outcome Variable A-V: *Classroom PAIL Student Work Sample Ratings*

Classroom Work Sample PAIL ratings, described in section IIIC, indicate the overall quantity and quality levels of documentation of student work in each treatment school classroom. All PAIR teachers had the opportunity to attend professional development sessions in years 2-3 of the project in order to learn how to optimize the process of collecting, organizing, and annotating each student's work sample into PAIL folders.

The quantity and quality of the PAIL collections of student work thus reflected each teacher's degree of commitment to PAIR student work products as indications of high-quality arts integration teaching and learning. A high PAIL classroom rating thus indicates a high level of each teacher's commitment and ability to create systemic, high quality documentation of student work that can also influence other student learning outcomes. If classroom PAIL ratings are significantly assocated with student arts or academic performance, results will support the inference that teacher committement to high quality documentation practices, a key teacher effect described above, can result in higher levels of student learning outcomes.

3F-2 Table 2: Classroom Student PAIL (PAIR Arts Integration Learning) Work Sample Ratings

Individual Classroom Ratings of Pail Student Work Samples Organized by PAIR Treatment School Teachers	
PAIL Variable 1:	Quantity (Scope, Range) of Student Work Samples [4-point Rubric)
PAIL Variable 2:	Quality (Detail, Sophistication) of Student Work Samples [4-point Rubric)
PAIL Composite Variable 3:	The combined average of the previous two factors.

* * *

3F-2 Teacher PD Outcome Variable A-VI: *Combined Teacher Professional Development Ratings (Average of Teacher Attendance, YECS Three Key Response Ratings, PAIL Ratings)*

This composite variable averages the combination of three key independent teacher professional development ratings. This variable was created in order to detect the combined effect of teacher professional learning outcomes based on (a) participation in PAIR PD sessions, (b) response to key teacher-teaching artist collaboration factors of arts integration curriculum development and implementation, and (c) the production and organization of PAIL student work samples resulting from the PAIR units. These variables could be amalgamated because these three sub-factors involved all PAIR teachers in the project and the outcome ratings were determined by similarly constructed standardized rubric rating scales by the research team.

This combined PD factors variable was created for the purpose of (a) providing a more expansive continuum of general professional development ratings in order to rank order teachers according to those had been impacted the most comprehensively and (b) to increase the statistical power of the sum of these qualities to better detect statistical associations with similar scaled student outcome measures.

* * *

3F-2 Teacher PD Outcome Variable A-VII: *Teacher Portfolio Conference Interview Response Ratings*

In part one of this report the Teacher Portfolio Conference response ratings were used to determine, among other factors, qualities of teacher reflection that determined valuable criteria for the four teacher effects described above. From the perspective of this chapter, teachers Portfolio Conference responses can be used to rate each individual teacher's ability to provide a clear description of (a) PAIR program goals, (b) examples of the impact of PAIR on classroom learning environment and its integration across academic disciplines, and (c) student demonstration of understanding of PAIR arts integration learning processes and their connections across the curriculum during the portfolio conference sessions. The teacher interview response performance rubric ratings, as described in section 3D of this report, were derived from coding teacher portfolio conference responses for their level of detail, articulation, perspective, relational thinking, and meta-cognitive understanding of the PAIR project.

High teacher portfolio conference ratings would indicate a teacher's relatively deep understanding of the goals, processes and anticipated outcomes of arts integration teaching and learning processes. A strong, positive relationship between teacher portfolio conference ratings and student arts and/or academic learning outcomes would indicate the extent to which the level of teacher's reflective, meta-cognitive understanding of the project affects student achievement. As the final teacher professional development factor this variable essentially provides a performance assessment rating of the level of observation and reflection on the nature, range and significance as they understand the development of arts integration practices in their classroom over the course of the project.

* * *

3F-2 Student Survey Response and Performance Assessment Outcome Variables (B: I-IV)

The next four variables capture the range of student perception and performance outcome data that were developed and/or collected over the three years of the study. These outcomes range from surveys designed to capture student perception of the increased presence of arts integration practices in the classroom (variables B-I) to the rating interview responses focused in the relationship of arts, academic of archived student work (B-II) to the scaled ratings of reflective comments about peer work in portfolios conferences (B-III) and to the state standardized measurements of academic achievement (B-IV).

The 2010 data collection is the central focus of the statistical results of the PAIR project due to the fact that statistically significant effects of the treatment school teacher and student cohorts emerged only during the final year of the project. As noted before, it has been hypothesized that the fourth grade teachers - those who experienced the most proflonged professional development and the most opportunity to work the first draft PAIR initiatives in their classrooms with the teaching artists – may be performing at a higher level of arts integration project implementation. By year three of PAIR, the teacher 'fourth grade effect' may also predict higher performance by their students by the final year of the project.

Comparing student learning ratings across three years of PAIR revealed a cumulative 'sixth grade effect' on student performance despite the fact that the grade 6 teachers had the least amount of professional development training and experience in the final year of the project. Arguably, student learning results from the control-treatment school comparisons woiuld have been far more pronounced by year three of PAIR had these teachers received professsional development and program support services from the beginning of the project.

There are other factors that contribute to the sixth grade effect. Because the PAIL student work products samples and the Portfolio Conference performance assessment tasks were developed in the last part of the project, the quality of intensive professional development support for PAIR teachers in year three was probably higher for the sixth grade teachers. Thus the continuity of instruction with teaching artists over the three years of the project and their support to PAIL work samples, revised SAIL tasks, and newly developed portfolio assessment tasks may have been the deciding factors for achieving statistically significant treatment school student learning outcomes by the project's culumination.

* * *

3F-2 Student Survey Response Outcome Variable B-I: *Student Survey Response Ratings*

As a result of prior analysis reported in Section 3B of this report, the Student Survey Response factors have been grouped into three primary composite variables that best captured the impact of teacher PAIR treatment school classroom practices in contrast to the control school classrooms. These include positive change in presence and importance of:

(1) student reflection in teaching practice;

(2) presentation and exhibition of art work in the classroom;

(3) arts integration processes and practice.

In addition, the creation of a composite variable averaging the three primary factors of student survey responses listed above could be used to detect the association of differences in student perception on other teacher or student project outcomes. A significant degree and pattern of association with positive change in student perception of these three factors would provide evidence to substantiate the teacher survey findings or the impact of various teacher professional

development outcomes. Student survey responses may also establish connections between student awareness of key differences in treatment school classrooms and student performance outcomes.

3F-2 Table 3: KEY Student Survey (SS) Response Factors

	Key Rating of Changes in PAIR Classrooms by the final year of the project
SS Variable 1:	Presence and Importance given to Student *Reflection*
SS Variable 2:	Presence and Importance of *Presentation* or Exhibition of Student Arts or Arts Integration Work
SS Variable 3:	Presence and Importance of *Arts Integration* Classroom Practices
SS Composite Variable:	Combining Averaged Reflection, Presentation, and Arts Integration Student Survey responses

Student Survey Response ratings, like all other major student learning outcome variables, can be subjected to regression factor analysis to determine statistically the extent to which certain variables can be understood as causal links in the chain of teacher and student learning outcomes.[1]

* * *

3F-2 Student Learning Outcome Variable B-II: *Student Snapshots of Arts Integration Learning (SAIL) Interview and Performance Assessment Response Ratings*

The individual student SAIL ratings, described in Section 3A of the report, represent a major measure of student understanding of arts and arts integrated teaching and learning. The SAIL protocol challenged each student to describe and provide examples of arts learning processes in two disciplines, to elaborate further on how academic discipline-based learning may be similar to the those in the two arts disciplines, and to elaborate further on nature of arts integration teaching and learning strategies they experienced in their classroom PAIR units.

[1] See section later sections in 3F and 3G of this report for extensive step-wise regression factor analysis of all student outcomes.

Since SAIL assessments were designed to serve as a leading student performance indicators of the differences between control and treatment school students' understanding of arts integration processes, the teacher outcome variables (described above) that are determined to be most highly associated with SAIL treatment school student outcomes could be seen as the significant *intermediary* factors contributing to high quality arts integration program results reported earlier in Parts 2 and 3 of this report.

3F-2 Table 4: KEY Snapshots of Arts Integration Learning (SAIL) Interview and Performance Assessment Factors

	Key Factors of by the final year of the project
SAIL Variable 1:	Averaged response ratings in with respect understanding of two Arts Disciplines featured in that particular Treatment School
SAIL Variable 2:	Averaged response ratings in with respect to understanding parallel between Academic and the two Arts Disciplines featured in that particular Treatment School
SAIL Variable 3:	Averaged response ratings in with respect to understanding arts integration practices and its impact on Academic emphasis in that particular Treatment School
SAIL Composite Variable:	Combining Averaged Arts, Academic, and Arts Integration SAIL ratings

* * *

3F-2 Student Learning Outcome Variable B-III: *Student Portfolio Conference Interview Response and Performance Assessment Ratings.*

The individual student portfolio conference ratings, described in Section IIID-E of the report, is a second major measure of student understanding of arts and arts integrated teaching and learning. The Portfolio Conference protocol challenged students to demonstrate understanding of arts learning processes in two disciplines based on their own and peer work. This highly interactive context provides many opportunities to rate students' ability to perform and reflect on work based on tasks that grow out of the work samples provided in the portfolio.

The student performance rubric ratings, similar to the teacher response ratings, were derived

from coding portfolio conference responses for their level of detail, articulation, perspective, relational thinking, and meta-cognitive understanding of their working processes and the academic content emphasized in their PAIR units.

As indicated previously, the portfolio conference ratings were only collected in the treatment schools. Thus, the student portfolio conference ratings are a leading indicator of the treatment school students' understanding of arts integration processes. The teacher professional development outcome variables (described above) that are determined to be most highly associated with Student Portfolio Conference treatment school student outcomes would be seen as the strongest PAIR program factors contributing to the treatment school students ability to articulate cognitive, social-emotional, and meta-cognitive understanding of arts integration learning outcomes in the context of sample arts integration student work products.

* * *

3F-2 Student Learning Outcome Variable B-IV: *Student Academic Performance Ratings*

The individual student state standardized test scores, described in Part 2 of this report, are the major measure of student academic performance for PAIR treatment and control schools. Fortunately, the Illinois Snapshots of Academic Test (ISAT) standard scores were criterion-referenced to expectations for grade level performance. Thus the ISAT ratings represent an externally validated measure of academic progress that is used here to focus on the comparable impact of the PAIR project on control and treatment schools over time.

In this project, the student ISAT ratings are also positioned as the final outcome variable of the eleven intersection factors described above. Thus the PAIR teacher impact *and* student learning intersections variables (described above) determined to be most highly associated with external student learning outcomes (ISAT) could then be seen as the strongest PAIR 'intermediary' or 'through-line' factors that led to enhance academic achievement in schools that employed high-quality arts integration teaching and learning practices.

* * *

3F-3: Statistical Evidence of Teacher-Student Performance Intersection Outcomes

In this third section of 3F, statistical evidence for the degree of association among *seven* teacher-classroom intersection variables and the *four* student learning outcomes in PAIR Treatment Schools will be organized by five focus questions:

3F-3 Inquiry Questions:

1) What are the teacher professional development [PD] factors that are most strongly linked with teacher PD outcome ratings?

2) What are the teacher impact variables that can be linked with student survey responses to changes in their arts integration classrooms?

3) What are teacher PD outcome variables that are most strongly linked with student SAIL and Portfolio Conference Performance Arts and Arts Integration Learning Outcome Assessments? To what extent does the degree or pattern of association of these variables provide evidence that the PAIR teaching factors positively predict PAIR internal measures of student arts integration learning outcomes?

4) What are teacher professional development variables that most strongly linked with student performance outcomes on external standardized ISAT Academic Tests? To what extent do these links suggest that PAIR arts integration teaching and learning objectives positively impact achievement test performance and equity factors?

5) What evidence has emerged from this project that suggests that teacher-student learning intersections impact the quality and impact of the PAIR approach to arts integration teaching and learning?

A taxonomy of statistical analysis methods

Each of the *seven* types of teacher impact variables will be tested for their degree of association with the *four* student response and performance learning outcomes [see table below]. The three major statistical processes used to determine the statistically significant ($p < .05$) degree of association among the variables include analyses of:

(1) The *bivariate linear fit* among two or more variables (using ANOVA F statistic and t values for measuring significant mean differences); chains of these factors can be examined in terms of comparative pattern or degree of correlation with respect to predictive value for teacher and student outcomes.

(2) The *degree and pattern of correlation* measures of continuous, non-parametric ratings (ANOVA F for multivariate and Spearman r values for paired variables); these comparisons will provide more powerful statistic evidence of association and the possibility of both direct and 'intermediary' predictors of student outcomes.

(3) *Step-wise regression processes* that model the influence of multiple variables on a singular focus on primary teacher professional development or student learning outcome variables in order to (a) sift out primary predictors from isolated or confounding factors of correlation, (b) rank-order competing variables among the many possible predictors of primary teacher and student learning outcomes and (c) determine the cumulative level 'explained variance' (r^2 value) of all the predictor variables in the regression equation (chain of competing factors).

After the models most predictive of student achievement are determined from the program variables, demographic factors can be added to the overall model to test the strength of association of predictive variables controlling for the influence of student demographic factors (gender, income status, ethnicity, ELL designation, HAL academic designation) and categorical differences among the treatment schools (teaching artist assignment, type of school type (arts vs. non arts), school primary academic focus (World Language, Writing (ELA), or Math).

3F-3 Table 1A: Seven Categories of PAIR Teacher Professional Development-Student Learning Outcome Intersection Variables

A. PAIR Teacher Professional Development Outcome Variables						
A-I: **Teacher Key Effect Ratings** [all teachers]	A-II: **Teacher Years of Participation in PAIR** [all teachers]	A-III: **Teacher PD Session Attendance Ratings** [all teachers]	A-IV. **Teacher Year-End Curriculum and Survey [YECS] Ratings** [2010 all teachers]	A-V: **Classroom PAIL Work Sample Ratings** [2010 all teachers]	A-VI: **Teacher Combined 'Super Variable" *PD* Ratings** [2010 all teachers]	A-VII. **Teacher Portfolio Conference Interview Response Ratings** [2010 small sample teachers, all grades]

(Tables continued on next page)

3F-3 Table 1B: Four Categories of Student Outcomes

B. Four Categories of Student Survey Response and Task Performance Outcome Variables			
B-I:	B-II:	B-III:	B-IV:
Student Survey (SS) Response Ratings	Student SAIL Arts integration Interview and Performance Assessment Ratings	Student Portfolio (PC) Conference Interview Response Ratings	Student ISAT Combined Academic Performance Ratings
[2010 student longitudinal cohort]	[2010 6th grade student longitudinal cohorts]	[2010 small sample student longitudinal cohorts]	[2010 all student longitudinal & follow-up cohorts]

3F-3 Table 1C: Categories of Student and Treatment School Demographic Factors

Student Learning *Demographic* Variables	School and Teacher *Categorical* Variables
• Gender • Family Income (Free & Reduced Lunch) • Ethnicity • ELL Classification • HAL Academic Achievement Pre-Designation	• Arts vs. Non Arts Schools • Cluster School Designation (Writing, Math, or World Languages) • Teaching Artist (Art Form) Designation

The following eleven PAIR Teacher and Student Exhibits will report detailed statistical findings from each of the *seven* teacher professional development outcome categories (Exhibits I-VII) and the four student outcome categories (Exhibits I-IV). These teacher outcome exhibits will provide evidence of the *coherency and impact* of the PAIR treatment school program factors as indicated by the pattern and degree of inter-correlation among teacher survey, interview and performance assessment factors. The student outcome exhibits will provide evidence of the *influence* of PAIR teaching practices on student learning outcomes as indicated by the pattern and degree of association among teacher-student survey, interview and performance assessment data.

* * *

3F-3 TEACHER OUTCOME EXHIBIT I: *Four Exemplary Teacher Effect Designations* and their Association with Teacher Professional Development and Student Learning Outcomes

The analysis of the first of seven teacher outcome variables is based on four specific categories of teacher impact on student learning outlined in Part I of the PAIR report *[Burnaford, PAIR Report Part I: Teacher Impact pp. 56-68]*:

- The Content Expertise Effect: *Pedagogical Content Knowledge Matters*
- The Documenting to Learn Effect: *Collecting Student Work Inspires Reflection*
- The Fourth Grade Effect: Curriculum: *Designers Have Ownership Through Long-term Project Participation*
- The Healy Effect: *Initiatives Build on Other Initiatives in a School*

The validation of these categories is based on whether or not PAIR treatment school teachers "*demonstrated significant differences from their peers on specific variables*, as reported on the surveys, the *coded comments from portfolio conferences*, the *coded open-ended responses on the surveys*, and the *reported pedagogy identified on student work labels."* Since these variables are believed by Burnaford and her colleagues to "show promise with respect to impact of the program and possible relationship to student achievement," they were analyzed below for their degree of association with other teacher impact factors (Exhibits II-VII) and with the four major categories of student achievement (Exhibits VIII-XI).

3F-3 Exhibit I Data Indications: *Two of the teacher effects influence most strongly the coherence and quality of the PAIR teacher professional development outcomes.* Results from the table below [Exhibit I Table 1] reveals the predominate positive effect of the 'Fourth Grade Effect' (long-term teacher engagement) and the 'Healy Effect' (school wide teacher engagement) on teacher participation in PD events, highly engaged collaboration with teaching artists, and the quality and quantity of student work documentation. Nonetheless, the primary predictors of teacher portfolio conference interview response – 'Content Expertise Effect' and the "One of More Teacher Effects – are highly related to the teachers' ability articulate arts integration

teaching practices and student learning outcomes persuasively in the final year teacher-student portfolio conferences.

3F-3 Exhibit I Table 1: Mean Score Comparisons of Six Teacher Professional Development Outcomes by Teacher Effect Profile Categories Designation (Exemplary vs. Not Exemplary)

A1: Teacher Effect Profile Category	A-II: Teacher Years in PAIR	A-III: Teacher PD Session Attendance Ratings	A-IV. Teacher Year- End Curriculum and [YECS] Survey Ratings	A-V: Classroom PAIL Work Sample Ratings	A-VI: Teacher Combined Factors of PD Ratings	A-VIII. Teacher Portfolio Conference Interview Response Ratings
	Mean Differences ANOVA F ratio - Probability Value					
Content Expertise Effect	NA	ns	ns	ns	ns	F ratio 20.2 $p < .0001$
Document to Learn Effect	NA	ns	ns	F ratio 6.1 $p < .01$	ns	ns
Fourth Grade Effect	NA	F ratio 25.8 $p < .0001$	F ratio 4.8 $p < .02$	F ratio 10.2 $p < .002$	F ratio 10.2 $p < .002$	ns
Healy Effect	NA	F ratio 67.3 $p < .0001$	F ratio = 38.0 $p < .0001$	ns	F ratio 45.6 $p < .0001$	ns
One or More Effects	NA	ns	ns	ns	F ratio 8.3 $p < .004$	F ratio 15.9 $p < .0001$
More Than One Effect	NA	ns	ns	ns	ns	ns

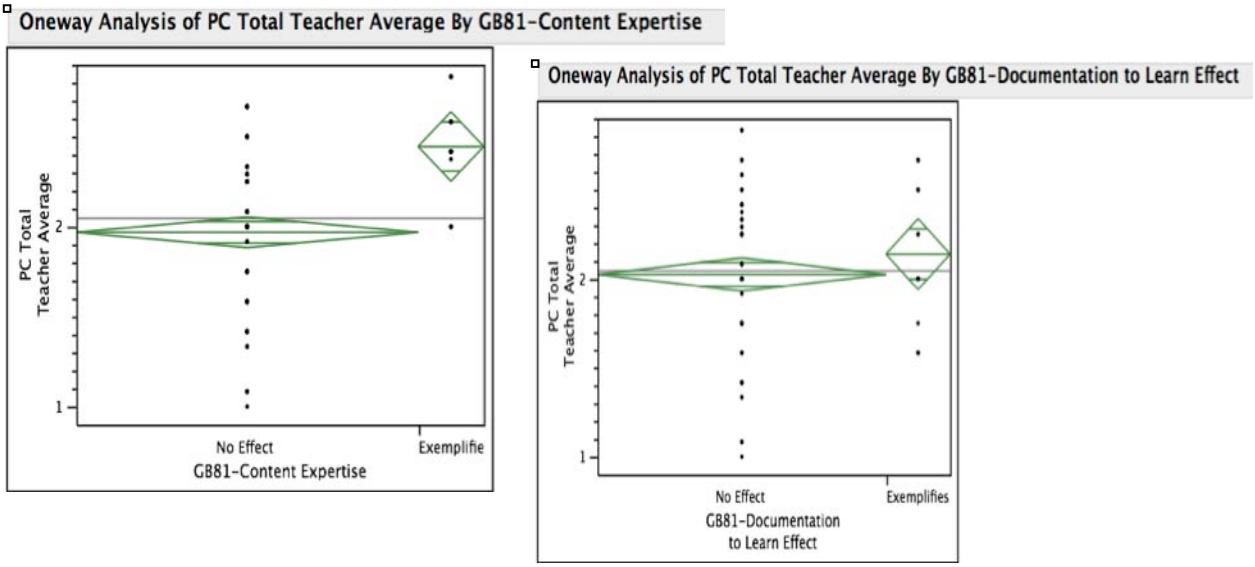
NA = Not Applicable due to small uneven grade distribution and small sample size
ns = Mean score differences are not statistical different

More detailed analysis of the PAIL samples (A-V) (not included in the table above) clarifies the difference between factors of quantity from averaged quality of the student work, another teacher outcome variable. The *quantity and scope of PAIL student work* was primarily predicted by the "Documentation to Learn Effect" (F ratio 26.1, $p < .0001$) while the *quality of student work* most strongly with the "Healy Effect" (F Ratio 22.2, $p < .0001$). Only the 'Fourth Grade Effect' was significantly linked with *both quantity* (F Ratio 4.8, $p < .03$) *and quality* (F Ratio 9.0, $p < .003$)

of student work. The high predictive value of the general effect of exemplary effects ('One or More Effects') indicates the general cohesive impact of the teacher effect variables on the ability of teachers to generate and organize high quality student work artifacts.

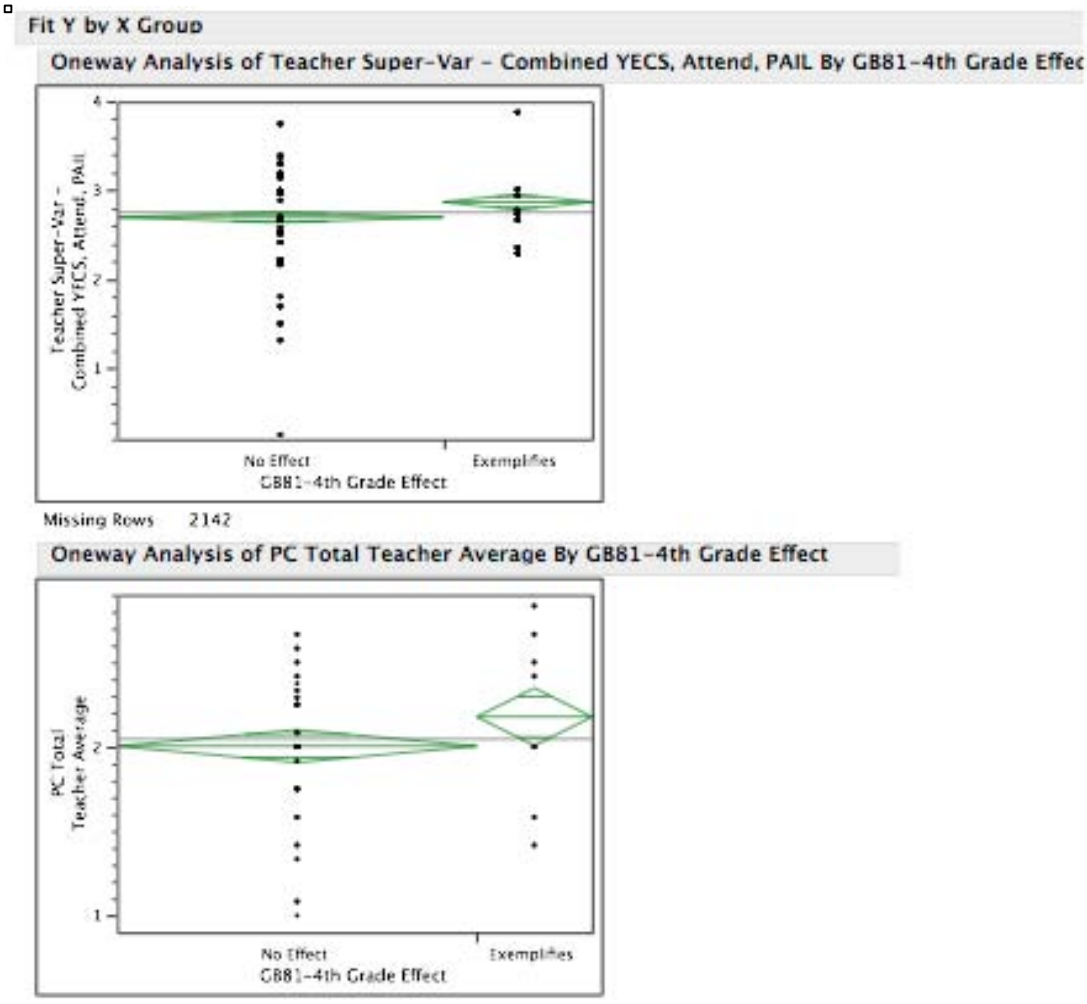
While the tables of means comparisons above provide clear evidence for the validity of the teacher effects analysis, the range and shape of these data plots reveals even more about the significance of these differences. In the figure below [3F-3 Exhibit I Figure 1], both data displays show that the exemplary teacher effects not only are defined by higher mean scores, but also by the absence of low-level responses. That is, none the exemplary 'content expertise' teachers or 'document to learn' teachers scored poorly on the teacher portfolio conference performance assessments.

3F-3 Exhibit I Figure 1: Mean Score Differences between Exemplary vs. Non-Exemplary Teacher 'Content Expertise' and 'Documentation to Learn' Effects on 'Teacher Portfolio Conference Ratings'



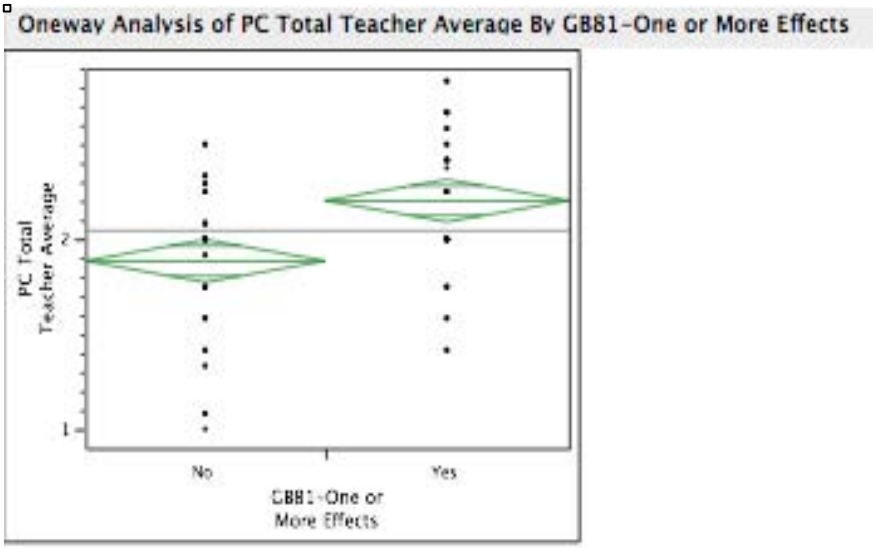
Similarly, the next two data displays [3F-3 Exhibit I Figure 2] reveal that the 'fourth grade effect' teachers who demonstrate ownership and depth of experience with the PAIR project through high teacher PD survey and portfolio conference ratings are a cohort of teachers who only score relatively high on both their survey self-reports (concerning teaching artist collaboration) and their ability to articulate the values and student impact of PAIR in their classroom during the portfolio conferences.

3F-3 Exhibit I Figure 2: Mean Score Differences between Exemplary vs. Non-Exemplary Teacher 'Fourth Grade Effect' and 'Documentation to Learn' Effects on Teacher PD Combined Factors (super variable) and 'Teacher Portfolio Conference' Ratings



Another example of means comparison plots [3F-3 Exhibit I Figure 3 below] shows how it may be possible that, whatever the strong suit of the teacher is, the ability of teachers to 'exemplify one or more' of the Teacher Effects Ratings predict a comparitively higher level of performance during facilitated teacher portfolio sessions. This distribution of Teacher Effects also suggests that different teacher effects appear to contribute to higher ratings regarding the teacher's ability to articulate project goals and methods, and to more keenly interpret student's reflective understanding of arts integration learning during the teacher-student portfolio conferences.

3F-3 Exhibit I Figure 3: Mean Score Differences between Exemplary vs. Non-Exemplary Teacher 'One or More Effects' on 'Teacher Portfolio Conference Ratings'



Statistical trends connecting 'Teacher Effects' to student survey and learning outcomes

Though Certain Exemplary Teacher Effects appear to associate deeply with teacher professional development outcomes, none of the 'Teacher Effect' exemplary ratings were connected strongly with student survey responses or learning outcome factors, thus demonstrating the limitations of the teacher effect analysis. The positive, yet weak trends in these student survey data demonstrated in the table below [3F-3 Exhibit I Table 2] suggest that (a) teachers at schools who have had long-term experience with school reform initiatives (the 'Healy Effect') are most closely associated with student perceptions of changes in classroom practices and (b) that teachers who demonstrate the 'Document to Learn' effect are most closely associated with student performance in their portfolio conferences.

[table on the next page]

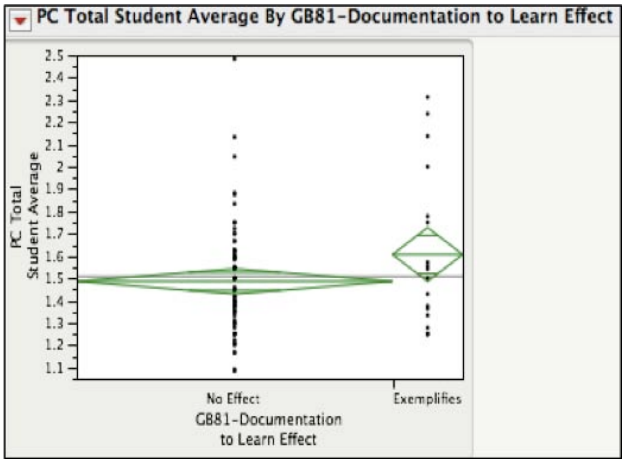
3F-3 Exhibit I Table 2: Mean Score Comparisons of Four Student Learning Outcomes by Teacher Effect Profile Categories (Exemplary vs. Non-Exemplary Effect)

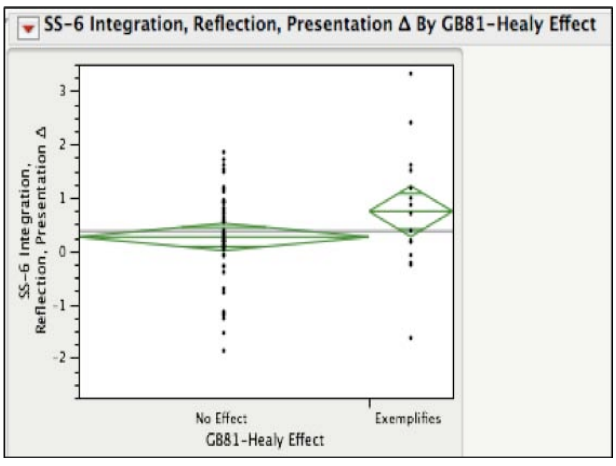
AI: Teacher Effect Profile Category	B-I: Student Survey (SS) Response Ratings	B-II: Student SAIL Interview and Performance Assessment Ratings	B-III: Student Portfolio Conference (PC) Interview Response Ratings	B-IV: Student ISAT Combined Academic Performance Ratings
Content Expertise Effect	ns	ns	ns	ns
Document to Learn Effect	ns	ns	[F ratio 3.2, P < .08]*	ns
Fourth Grade Effect	ns	ns	ns	ns
Healy Effect	[F ratio 3.2, P < .08]*	ns	ns	ns
One or More Effects	ns	ns	ns	ns
More Than One Effect	ns	ns	ns	ns

ns = Positive mean score differences are not statistical different
*= Positive trend, through not statistically significant

The data plots below [3F-3 Exhibit I Figure 3], linked to the table above, show how important statistical trends can be detected even though they do not meet the most rigorous standards of statistical significance.

3F-3 Exhibit I Figure 3: The Influence of Mean Score Differences of Teacher 'Document to Learn' Effect on Student Portfolio Conference ratings and the "Healy Effect' on 'Combined Student Survey Responses' ('Arts Integration,' 'Time for Reflection,' 'Student Arts Learning Presentations')





* * *

3F-3 Exhibit I: Summary and Emerging Themes: *Qualitatively-defined categories of Exemplary Teaching Effects can significantly predict teacher professional development variables, but only weakly influence student performance outcomes.* It appears that the predominance of the 'Fourth Grade' (long-term teacher PD) and 'Healy' (school wide arts integration practices) effects with regard to most of the other teacher impact ratings provides an important key to the quality and coherence of the PAIR program in the treatment schools. That is, those teachers who started at the ground-level stages of the project as it developed and who received the highest amount of ongoing guidance in the PAIR project (the fourth grade teacher cohort), and those teachers with more comfort with school reform initiatives in general (Healy effect teacher cohort) are more likely to participate in the professional development events, to reflect positively on their collaborative arts integration units with the teaching artists, and to produce and organize a better scope and quality of student work. To the degree that these teacher outcomes can predict student academic or artistic achievement, the 'fourth grade effect' criteria for exemplification may represent important underlying factors in determining the quality and impact of future dissemination of arts integration work.

The relatively inconclusive measures of association between teacher effects on student survey and portfolio conference ratings indicate that the determination of the exemplary 'teacher effects' does not yet provide a direct, strong predictor of student arts integration learning or academic achievement outcomes. However, keeping in mind the relatively small teacher sample size and that the teacher effect cohorts are categorical rather than scaled data, the Key Teacher Effects may yet serve as frameworks for understanding qualitative differences in the quality of teacher implementation of the PAIR project.

In sum, the Teacher Effect categories that were determined by teachers' self-reflective statements and their alignment with professional training objectives reported earlier Part 1 of this report provided (a) useful criteria for understanding the overall success of the teacher professional development outcomes in the project, (b) evidence that the exemplary effects of influence of 'years of teacher participation in the project' can impact the quality of teacher performance in the project, and (c) that these effects may function also as 'intermediary' factors in multivariate analyses that represent a critical link in the path of association of teacher professional development outcomes that can directly predict student learning.

* * *

3F-3 TEACHER OUTCOME EXHIBIT II: '*Years of Teacher Participation in PAIR'* and its Influence on Teacher and Student Outcomes

The 'Years of Teacher Participation' variable was used to test for the influence of ongoing opportunities for professional development and project participation on teacher and student performance. The 'fourth grade effect' is named for the traits of exemplary fourth grade teachers, the teachers who were trained from the very beginning of the three-year project and were the most experienced and potentially most dedicated cohort of PAIR teachers. In contrast to the Fourth Grade Effect, the Years of Experience variable represents a mixture of exemplary and non-exemplary teachers who served in one of three grade level cohorts.

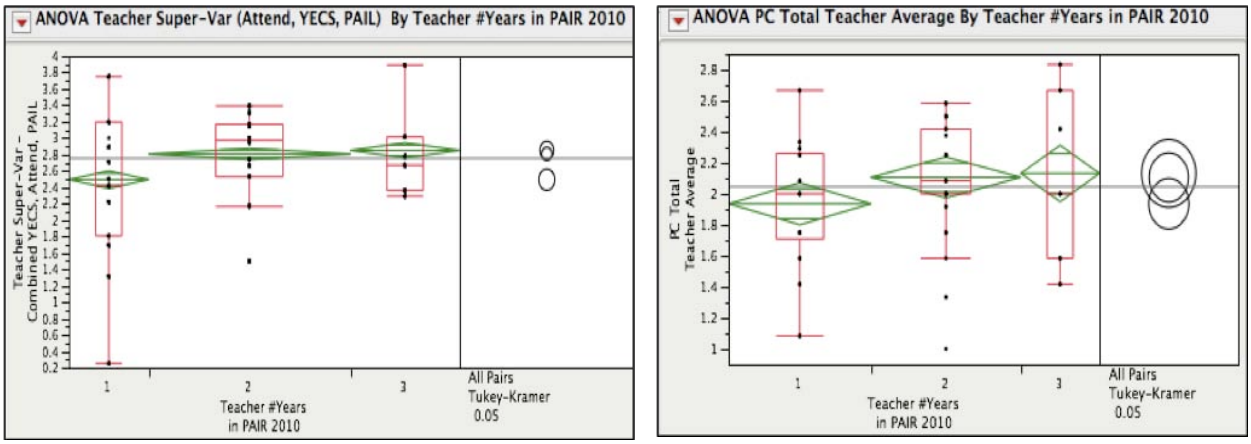
3F-3 Exhibit II Data Indications: *Years of Teacher Participation weakly predicts the quality of teacher professional development ratings.* Although most mean differences in the other Teacher Impact Ratings tend to reflect more positive results over time, Years of Teacher Participation could only be used as a categorical variable for cohort analysis depending the structure of the data and the requirement that the dataset be equally distribute over the three years of the project.

The strongest influence of teacher 'Years of Participation in PAIR' surfaced only in the analysis of its significant fit with Teacher Professional Development Combined Factors (super variable) ratings. As the charts below indicate, teachers who were in the project two or more years were statistically distinguished from those with only one year in the project [ANOVA F Ratio 14.5, $p < .0001$]. To a lesser extent this same pattern obtained with respect to Teacher Portfolio Conference Ratings as well [ANOVA F Ratio 2.1; $p < .12$].

Later on, as indicated in 3F Exhibit VI Figure 1 below, the Years of Teacher Participation revealed an even more profound effect on the cumulative impact of teacher participation in PAIR on student state standardized academic test score results.

[figure on next page]

3F-3 Exhibit II Figure 1: Mean Score Differences in Combined Teacher Professional Development Variables (3F-2: A-VI) and Teacher Portfolio Conference Ratings (3F: A-VII) According to 'Years of Teacher Participation in PAIR'



3F-3 Exhibit II Summary and Emerging Themes: *The Impact of Years of Teacher Professional Development variable had limited effect on teacher professional development outcomes.* Even though there are positive mean differences according to the 'Years of Teacher Participation' with respect to levels of student performance in Student SAIL and Portfolio Conference Ratings, 'Teacher Years in PAIR' is not a statistically significant factor on its own right, especially when compared to other categorical variables such as Teacher Effects, which are qualitatively rather than quantitatively defined.

Furthermore, these results make the point that ratings of the 'Fourth Grade Effect' predicts qualitative better than quantitative measures based on 'years of participation in the PAIR project.' The exemplary quality of professional development matters a great deal more than it does to simply rating teachers by 'years of participation'. Nonetheless, the categorical variable Years of Teacher Participation can be employed to examine changes in degree of correlation between PD and student outcomes over time when the data are drawn from all three grade levels and all teachers in the final year of the project (e.g., Teacher outcome factor A-IV in exhibit IV and the intersection between teacher factor AVI and student factor BIV discussed in Exhibit VI).

* * *

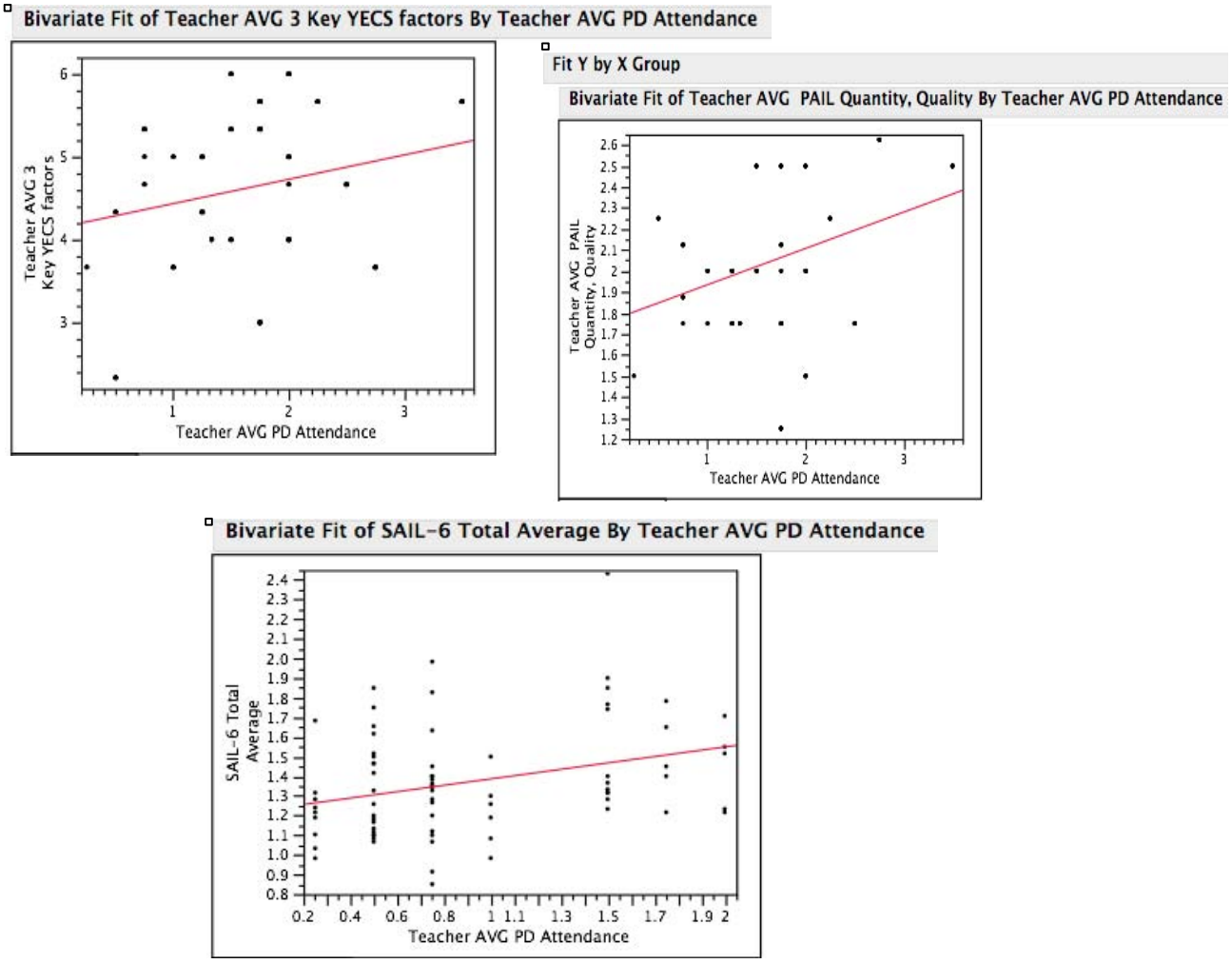
3F-3 TEACHER OUTCOME EXHIBIT III: *The 'Averaged Teacher Attendance' Ratings in PAIR Professional Development Sessions* and their Association with Teacher Survey Ratings, Teacher Professional Development Outcomes, and Student SAIL Interview Ratings

In contrast to Years of Teacher Participation in PAIR categories, Teacher Averaged Attendance ratings capture to some extent the relative depth of engagement with the content of the PAIR professional development sessions. High levels of teacher attendance provided multiple opportunities for teachers to design, critique and revise their new arts integration practices. Professional Development Session Surveys administered during these sessions provided repeated opportunities for teachers to reflect (a) on their time spent in *collaboration and planning* with the teaching artists, (b) on the extent to which they were able to *share their work* with their fellow teachers, (c) on how they were *deepening their curriculum* by incorporating PAIR arts integration practices into their primary focus non-arts content area, and (d) on their planning and implementation of *student learning documentation* processes. Thus the cumulative time spent attending professional development events suggests that high attending teachers are much more likely to be focused on building a reflective awareness of their PAIR implementation. By linking these attendance data to other professional development and student learning outcomes we will be better able to judge the impact of PAIR on teachers who attended these sessions.

3F-3 Exhibit III Data Indications: *Analysis of teacher attendance data revealed multiple statistically significant avenues of association with respect to other teacher and student outcomes.* The three data displays in the figure below [3F-3 Exhibit III Figure 1] trace the robust relationships between high averaged attendance in teacher professional development events and positive year-end survey ratings about collaborative practice [F ratio 23.7, $p < .0001$; $r = .18$ $p < .001$], quantity and quality of PAIL classroom work samples [F ratio 54.4, $p < .0001$; $r = .26$, $p < .0001$], and student SAIL Arts Integration Interview ratings [F ratio 8.0; $r = .26$, $p < .0001$].

[figures on next page]

3F-3 Exhibit III Figure 1: Statistically Significant Associations of 'Averaged Teacher Survey', 'Averaged Classroom Student Work' Ratings, 'Teacher Portfolio Conference' Ratings and Student SAIL Interview Response' Ratings according to Averaged Annual Levels of 'Teacher Attendance in PAIR' PD Sessions



3F-3 Exhibit III Summary and emerging themes: *Teacher attendance data can be linked directly to both teacher and student performance outcomes.* These attendance data results reveal a surprisingly broad impact on both program quality and performance outcomes. Not only do these data directly validate the efficacy of the professional development focus and its programmatic impact, but they also imply that these PD events had an indirect, yet transformative positive impact on both teachers and student performance.

* * *

3F-3 TEACHER OUTCOME EXHIBIT IV: '*Teacher Year-End Curriculum and Survey (YECS)' Ratings* and their Association with other Teacher Professional Development and Student Learning Factors

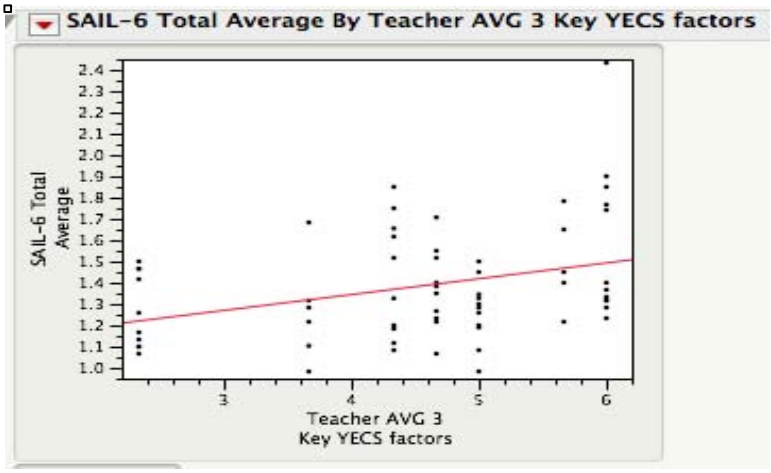
3F-3 Exhibit IV Data Indications: *Teacher 'Year-End Curriculum and Survey' responses related to positive ratings of their collaboration with teaching artists predicted various forms of student learning.* The teacher 'Year End Curriculum and Survey' (YECS) averaged ratings provided comparisons that were previously used for control-treatment school differences. As reported earlier, the final year of YECS Survey data revealed that there were highly significant differences between treatment and control group teacher cohorts regarding the highly statistically significant relationships between teachers and teaching artists *brainstorming to plan*, *coming to consensus in planning,* and *modeling discussion and feedback in front of students in the classroom.* Because these factors draw a particularly salient distinction among the PAIR control and treatment teachers, these same factors can now be tested for their predictive power on student learning.

The three YECS questions analyzed here represent the three most distinctive factors of the control-treatment comparisons. These questions happen also to represent a core indicator of the PAIR goal to foster collaboration between the classroom teacher and teaching artist. Thus strong indications of sharing of ideas, finding consensus in the unit planning, and modeling integrative teaching and learning in the classroom would validate the success of the project professional development objectives.

The scatterplot below [3F-3 Exhibit IV Figures 1] displays the intriguing intersection between the averaged factors of collaborative classroom teacher and teaching artist connection and the students' ability to articulate their understanding of arts integration learning processes. As the chart indicates, the averaged combined ratings of these three variables are highly associated with the averaged student SAIL Interview data ratings [F ratio 7.7; $p < .007$; Spearman r statistic .30, p value < .002].

(figure on next page)

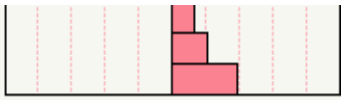
3F-3 Exhibit IV Figure 1: Bivariate Fit of Averaged 'Teacher Survey' and Averaged 'Student Arts Integration Learning' (SAIL) Interview Ratings



YECS data from the final year of the longitudinal sample also reveal the individual degree of correlation between each key teacher survey question and the averaged SAIL student ratings [Exhibit IV Figures 2 below]. Positive correlations indicate in the chart below that when the teacher and teaching artist feel they are able to effectively model the instruction of these units collaboratively, they are more likely to influence positively the 6th grade students' ability articulate the nature of arts and arts integration learning in their interviews.

3F-3 Exhibit IV Figure 2: Correlation Table Displaying the Degree of Association Among the Three Primary 'Averaged Teacher Survey' Ratings and 'Student Arts Integration Learning' (SAIL) Averaged Response Ratings

SAIL-6 Total Average	YECS-T-A Both Brainstorm	0.1392	0.2727
SAIL-6 Total Average	YECS-T-A Get Consensus	0.2127	0.0916
SAIL-6 Total Average	YECS-T-A Model Process	0.3862	0.0016*

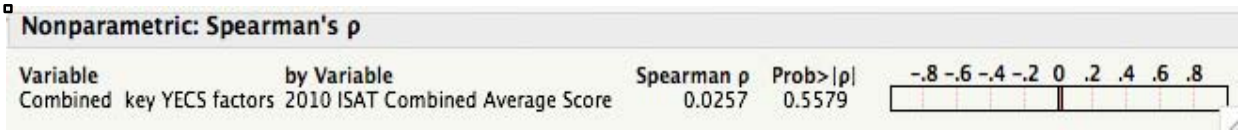


In addition, the sub-analysis of HAL student cohort survey responses indicate far stronger association with the factors depicted above for those students pre-designated as *Average* academic learners [YECS 'model process' ratings with SAIL interview ratings; r = .45, p value < .04] or as *Low* academic learners [YECS 'get consensus' or 'model process' ratings with SAIL ratings [r = .41, p value < .06 and r = .40, p value < .08 respectively].

YECS survey questions and the ISAT academic test scores

The averaged combined, averaged YECS survey data are generally are not strongly related to any student learning outcome other than the SAIL interview outcomes, including the combined math and reading state standardized test scores as seen in the chart below [3F-3 Exhibit IV Figure 3].

3F-2 Exhibit IV Figure 3: Correlation Table Indicating the Random Association between 'Combined Teacher Year-End Survey' Averaged Responses and Combined Averaged Test Scores, Grades 4-6 by the final year of PAIR

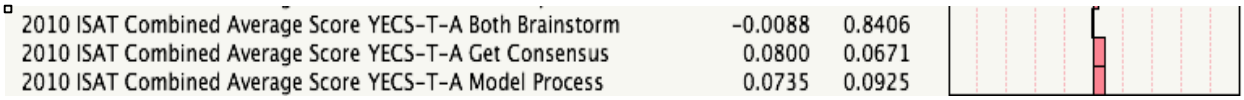


Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|	-.8 -.6 -.4 -.2 0 .2 .4 .6 .8
Combined key YECS factors	2010 ISAT Combined Average Score	0.0257	0.5579	

However, when the averaged YECS analysis focuses on the three questions separately (see chart below), positive statistical trends results, although not strong enough to rise to the benchmark of statistical significance. The sub-correlations charted below [Exhibit IV Figure 4] suggest that 'getting consensus' on the arts integration curriculum units (Spearman r = .08, p < .07) and "modeling the process" of collaborative arts integration instruction in classrooms (Spearman r = .07 p < .09) tend to predict academic achievement in the PAIR project. These findings suggest that the level of teacher positive self-reflections on collaborative curriculum design and implementation factors – and not the brainstorming of ideas – tends to be associated with overall classroom academic performance.

3F-2 Exhibit IV Figure 4: Correlation Table Indicating a Positive Association Trends between two of the three Teacher 'Year-End Curriculum and Survey' Responses and the Final Year ISAT Averaged Combined Test Scores, Grades 4-6

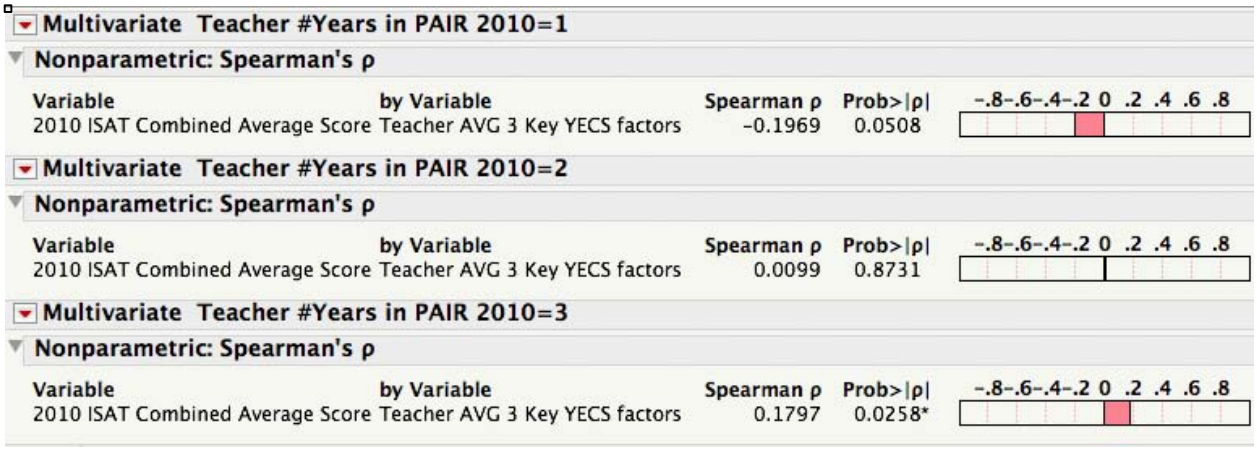


2010 ISAT Combined Average Score YECS-T-A Both Brainstorm	-0.0088	0.8406
2010 ISAT Combined Average Score YECS-T-A Get Consensus	0.0800	0.0671
2010 ISAT Combined Average Score YECS-T-A Model Process	0.0735	0.0925

Furthermore, when breaking down the averaged YECS variable ratings according to 'Years of Teacher Participation in PAIR' by the final year of the project [see 3F-3 Exhibit IV Figure 5

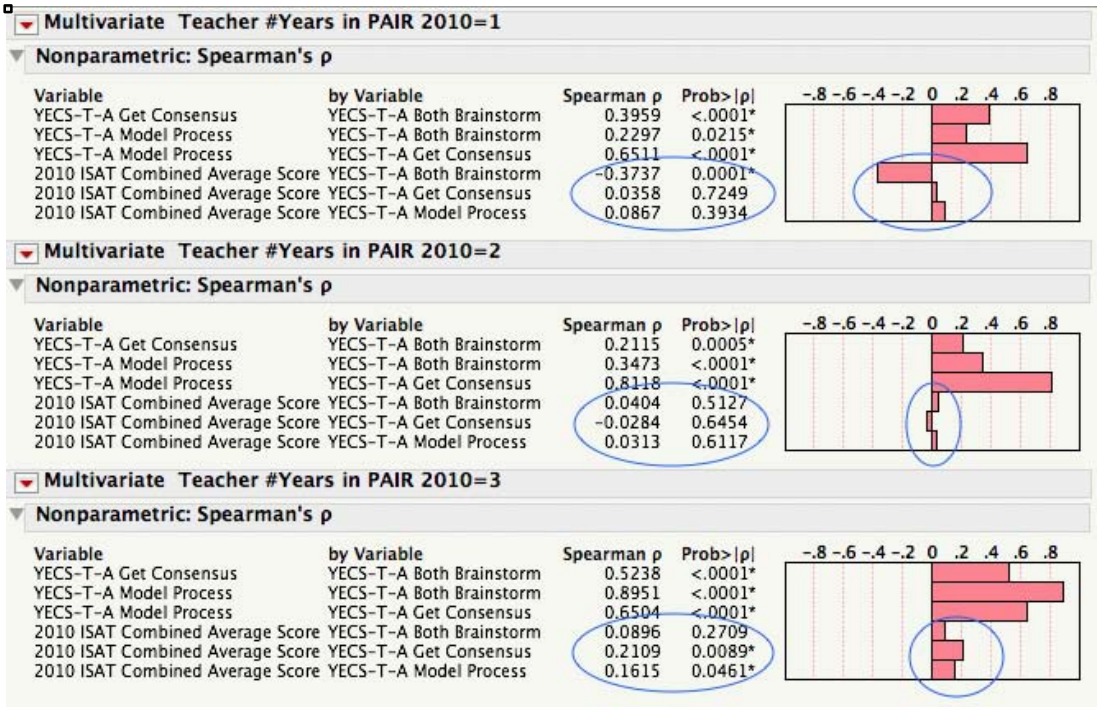
below], results indicate that positive, significant statistical trends are obtained *only* by the fourth grade students who had the most experienced PAIR teachers during the final year of the project. That is, the pattern of correlation between the YECS and ISAT data over time indicates that, although the 6th grade longitudinal student cohorts may outperform the control school cohorts [as reported in Part 1 of the report] *despite* having less experienced PAIR teachers (as indicated by negative correlations with YECS ratings in grade 6), the 'fourth grade effect' students succeed in standardized tests *with the benefit* of guidance from their more experienced teachers (as indicated the statistically significant 'degree of association' between individual survey questions and ISAT student test scores indicated in the figure below.

3F-2 Exhibit IV Figure 5: Correlation Table Indicating Positive trends between Averaged, Combined Teacher 'Year-End Survey' Responses and ISAT Academic Achievement scores in Grades 4 by the final year of PAIR



In the analysis below [3F-3 Exhibit IV Figure 6] we see an extraordinary positive shift in the degree of correlation according to the increasing 'years of teacher participation' in the project. As the teacher become more experienced with the project, random results or weak statistical trends with the aggregated data increasingly indicates positive, statistically significant evidence of the connection of teacher ratings of arts integration collaborative practices in the YECS surveys and student ISAT test results in those teacher's classrooms. When looking at the shifts in degrees of correlation from 1 to 3 years of teacher participation in the PAIR project, significant *negative* correlations turned into weak positive correlations, and then finally strong, statistically significant positive correlations.

3F-3 Exhibit IV Figure 6: Correlation Table Indicating a Shift in Direction and Degree of Association Between Averaged Teacher 'Year-End Curriculum and Survey' Survey Responses and Final Year ISAT Test Scores, Grades 4-6



Multivariate Teacher #Years in PAIR 2010=1

Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|
YECS-T-A Get Consensus	YECS-T-A Both Brainstorm	0.3959	<.0001*
YECS-T-A Model Process	YECS-T-A Both Brainstorm	0.2297	0.0215*
YECS-T-A Model Process	YECS-T-A Get Consensus	0.6511	<.0001*
2010 ISAT Combined Average Score	YECS-T-A Both Brainstorm	-0.3737	0.0001*
2010 ISAT Combined Average Score	YECS-T-A Get Consensus	0.0358	0.7249
2010 ISAT Combined Average Score	YECS-T-A Model Process	0.0867	0.3934

Multivariate Teacher #Years in PAIR 2010=2

Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|
YECS-T-A Get Consensus	YECS-T-A Both Brainstorm	0.2115	0.0005*
YECS-T-A Model Process	YECS-T-A Both Brainstorm	0.3473	<.0001*
YECS-T-A Model Process	YECS-T-A Get Consensus	0.8118	<.0001*
2010 ISAT Combined Average Score	YECS-T-A Both Brainstorm	0.0404	0.5127
2010 ISAT Combined Average Score	YECS-T-A Get Consensus	-0.0284	0.6454
2010 ISAT Combined Average Score	YECS-T-A Model Process	0.0313	0.6117

Multivariate Teacher #Years in PAIR 2010=3

Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|
YECS-T-A Get Consensus	YECS-T-A Both Brainstorm	0.5238	<.0001*
YECS-T-A Model Process	YECS-T-A Both Brainstorm	0.8951	<.0001*
YECS-T-A Model Process	YECS-T-A Get Consensus	0.6504	<.0001*
2010 ISAT Combined Average Score	YECS-T-A Both Brainstorm	0.0896	0.2709
2010 ISAT Combined Average Score	YECS-T-A Get Consensus	0.2109	0.0089*
2010 ISAT Combined Average Score	YECS-T-A Model Process	0.1615	0.0461*

3F-3 Exhibit IV Summary and Emerging Themes: *Teacher Survey data focused on the successful collaboration of core classroom teacher-teacher artist practices can positively influence the quality of arts, arts integration and academic learning*. By the third year of the PAIR, the averaged combined teacher survey ratings are positively associated with the grade 6 student arts integration learning outcomes, thus establishing a fundamental measure of the successful intersection of teacher professional development ratings with the intended positive effect on student arts and arts integration learning. Furthermore, when broken down by two separate response factors – teacher consensus in planning and implementing arts integration units with a teaching artist – the correlations with student arts integration interview and performance assessment ratings are even more robust.

Yet, because these strong correlations could only be tied to the student longitudinal cohort of students whose teachers had only participated intensively in the project for one year (due to the design of the project), the effect of degree of years of teacher project implementation experience could not be determined.

The correlation of teacher survey responses with student ISAT academic scores, however, can be measured across all grade levels at the final year of the project. This analysis was possible because both the YECS and the ISAT tests were given to all participants in the final year of the project. However, because this association was only statistically significant in the final year of the project for the fourth grade students who had the more experienced PAIR teachers, *these findings do suggest that only the cumulative degree of teacher experience with PAIR teaching artists can directly influence academic achievement as measured by standardized tests.*

Taken together, YECS survey results suggest that there may be an underlying transformational impact that years of teacher professional development have on the student SAIL interview ratings and, after three years of experience with PAIR, on the student ISAT academic achievement scores. As the PAIR collaborative process ratings become positive and increasingly significant over time as illustrated here, *these data suggest that the classroom teacher's ratings of the intensity and quality of the collaborative process with the teacher artist – both in curriculum planning and implementation – increasingly predict levels of student academic achievement in proportion to their commitment to long-term professional development in arts integration practices.*

* * *

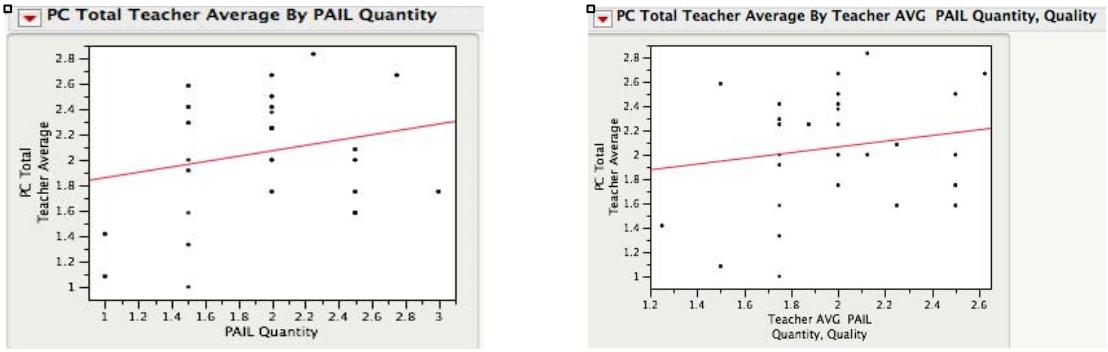
3F-3 TEACHER OUTCOME EXHIBIT V: '*Quantity' and 'Quality' Ratings of PAIL Classroom Student Work Samples* and their Association with Teacher Performance in Portfolio Conference, Student Survey Response and Learning Outcomes

As previously mentioned, the generation and organization of PAIL student work samples were begun late into the project and were for the most part in a state of evolution. Not all teachers were able to manage the process well and the standards for collecting and interpreting student work were not entirely clear especially to the teachers who were not able to attend certain professional development workshops in the final year of the project. As a result of this program implementation flaw, the work samples within the classrooms were not sufficiently *individualized* and the *work samples across classrooms were not sufficiently standardized* to make rank ordered comparisons at the individual student level. Therefore, the research team

rated the samples by classroom, thus turning an intended individual student outcome variable into a teacher/classroom teacher professional development outcome variable.

3F-3 Exhibit V Data Indications: *The Quality of PAIL Classroom Ratings is at best a weak predictor of Teacher Portfolio Conference Observation and Reflection.* Due to factors described above, the PAIL ratings are not highly predictive of any variable except for weak statistical trends that connect teacher performance during student-teacher portfolio conferences with PAIL Student Work Quantity [F 4.1, p < .05] and Quality [F 2.5, p < .11] ratings. These trends are depicted in the figure below:

3F-3 Exhibit V Figure 1: Teacher PC Ratings by PAIL Student Work Quality and Quality Ratings



3F-3 Exhibit V Summary and Emerging Themes: *Due to project design flaws, the PAIL student work ratings were not statistically powerful enough to determine their relationship with other teacher and student learning outcomes.* The classroom ratings of PAIL student work, as discussed previously, were influenced to some extent by the degree and quality of teacher participation in the PAIR project. However, the creation of PAIL student work documentation processes intensified only during the later stages of the project. Thus, teachers who had been active in the project from its beginning may have had more experience, but they did not have sufficient practice with documenting work samples from their classrooms to produce optimal results.

CAPE has since corrected for this research design flaw in subsequent studies so that researchers can plan to evaluate the intersection of individual student work samples in relation to teacher professional development outcomes in the future.

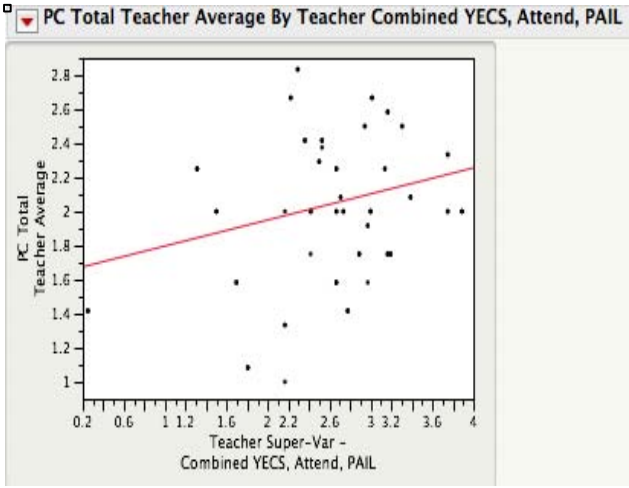
* * *

3F-3 TEACHER OUTCOME EXHIBIT VI: '*Combined Teacher Professional Development Factor Ratings'* and their Association with Teacher Portfolio Conference, Student Survey and Student Arts Integration Interview Response Outcomes

IIIF-3 Exhibit VI Data Indications: *The 'Combined Professional Development ('Super-variable') Factors significantly predict Teacher Portfolio Conference Ratings, SAIL Student Interview Performance Ratings, and ISAT Test scores when controlling for Teacher Years in PAIR.* The strategy of averaging three teacher PD factors into one variable was to create statistically standardized indexes of each teacher's level of project participation, ability to document and annotate student work samples, and the ability to reflect on professional development outcomes through survey responses. These combined indexes represented balanced profiles of teacher professional development outcomes that, when averaged together, provided strong statistical measures of its relationship to multiple teacher and student project outcomes as indicated by the following data displays [3F-3 Exhibit VI Figures 1-2].

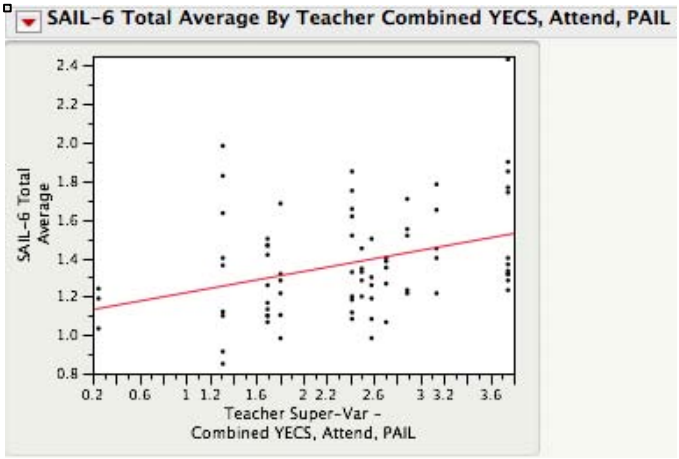
3F-3 Exhibit VI Figures 1-2: **Bivariate Fit between 'Combined Teacher PD Factors' and 'Teacher Portfolio Conference Performance', 'SAIL Student Interview' Ratings**

The first data display [3F-3 Exhibit VI Figure 1] to the right captures the positive statistical trend of association between Combined Teacher Professional Development Factors and Teacher Portfolio Conference performance ratings [F 7.6, $p < .007$; $R = .17$, $p < .10$].



(figure on next page)

The next data display [3F-3 Exhibit VI Figure 2] below captures the positive statistically significant relationship between 'Combined Teacher PD Factors' and 'Student SAIL Interview' performance ratings [F 10.1, p < .002; r = .32, p < .005].



Teacher Impact on Student ISAT Scores: Patterns and Degrees of Correlation between 'Combined Teacher PD Factors' and 2010 ISAT academic performance scores by 'Years of Teacher Participation in PAIR'

The data displays below capture the positive statistical trends between 'Combined Teacher PD Factors' and student ISAT academic performance ratings also displays strong evidence for the cumulative impact of three years of teacher participation in PAIR. The three charts below [Exhibit VI Figuree 3-5] shows that, *even though treatment school student test scores were generally higher than control schools by year three of the project, combined teacher professional development ratings predicted stduent standardized test scores only after three years of teacher paprrtcipation in PAIR.*

The first data display [Exhibit VI Figure 3] shows that no postiive relationship exists between combined professional development factors and ISAT scores for teachers with one year of experience with PAIR.

(figure on next page)

3F-3 Exhibit VI Figure 3: Negative, Statistically Insignificant Correlation of 'Teacher Combined PD Factors' and Student ISAT Scores for Teachers with One Year of Participation in PAIR

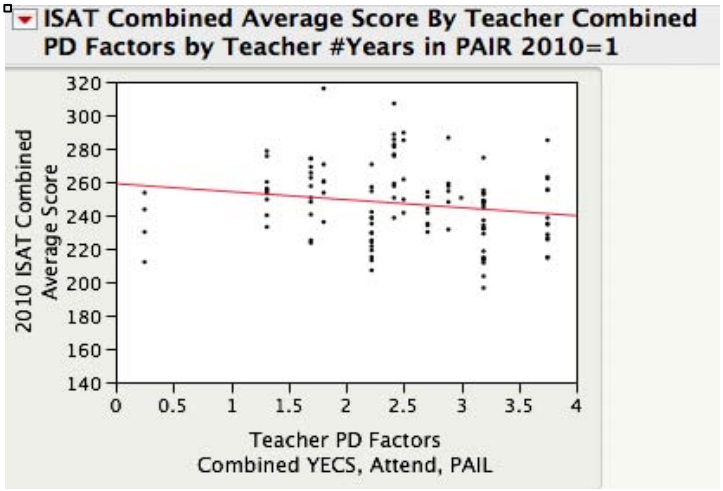
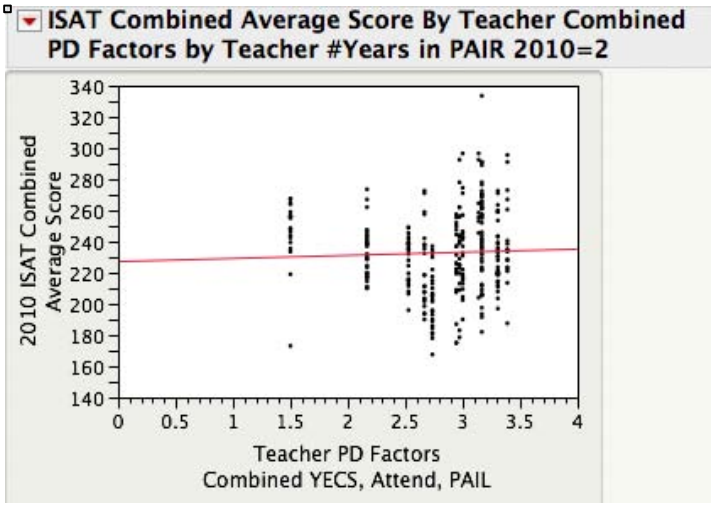


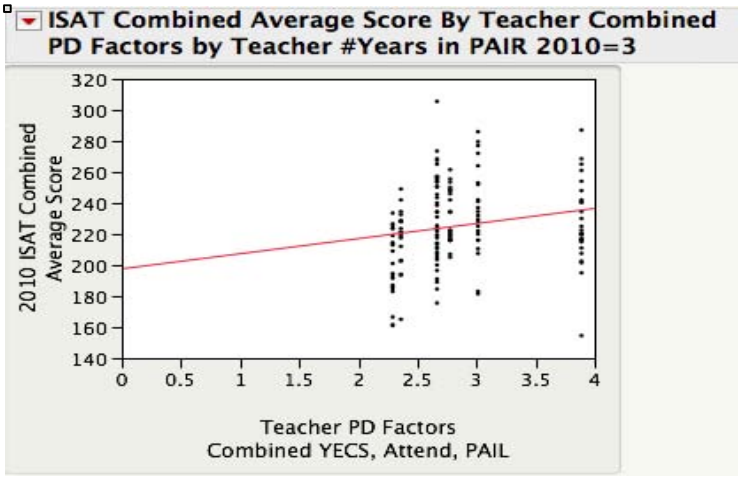
Exhibit VI Figure 4 below indicates indicate that, though statistically insignificant, the relationship of combined professional development ratings to student achieement is far more positive after two years of teacher participation in PAIR compared to results after one year.

3F-3 Exhibit VI Figure 4: Positive, Yet Statistically Insignificant Relationship Between Teacher 'Combined PD Factors' and Student ISAT Scores for Teachers with Only Two Years of Participation in PAIR



After three years in the PAIR, however, the intersection of Grade 4 teacher PD ratings and ISAT scores is highly significant [F ratio = 5.6, $p < .02$; $r = .27$, $p < .0006$]. The figure below [3F-3 Exhibit VI Figure 5] shows the convergence of higher and more consistently rated PD outcomes and their association with student achievement ratings suggests that, with time, the PAIR averaged professional development ratings substantially predict positive academic achievement as well a better understanding of arts integration learning reported earlier.

3F-3 Exhibit VI Figure 5: Positve, Statistically Significant Correlation of 'Combined Teacher PD Factors' and Student ISAT Scores for Teachers with Three Years of Participation in PAIR



3F-3 Exhibit VI Summary and Emerging Themes: *The averaged 'Combined Teacher Professional Development' ratings strongly predict student Interview and Portfolio Conference ratings, and increasingly predict student ISAT test scores as teachers became more experienced with the PAIR project.* The combined teacher outcome variable was constructed as a measure of inter-correlated composite professional development factors that, because of added statistical power, would more reliably detect effects otherwise not captured by each factor separately. The YECS, Attendance, and PAIL student work variables, could be combined because of rating intervals, shared data ranges and the shape of distributed ratings.

Results from the Teacher Outcome Factors analysis show that the composite teacher PD ratings, representing an aggregate of YECS other factors such as teacher attendance in professional development session and the ratings of classroom student work sample ratings, is strongly

associated with two primary arts integration student learning outcomes. In addition, this variable also became increasingly predictive of student academic test scores as teachers became more experienced with the PAIR project.

The finding of a particularly significant relationship between 'Years of Teacher Participation in PAIR' and the connection of 'Combined Teacher Professional Development Ratings' to student academic outcomes hold broad implications for future research on the impact of arts integration programs on student academic test scores. The challenge that emerges from these data is that it may take three full years of teacher professional development before arts plus arts integration schools are consistently able to both (a) outperform comparison schools in academic achievement and (b) establish an unassailable, highly predictive relationship between professional development outcomes and academic achievement scores at the classroom level.

In retrospect, had there been more student SAIL interview and PAIR portfolio conference data collected in all grade levels and throughout all three years of this project, it may have been possible to detect earlier and more systemic evidence that could substantiate the connection between changes in teaching practices due to professional learning and improvement in student arts integration or academic learning outcomes.

* * *

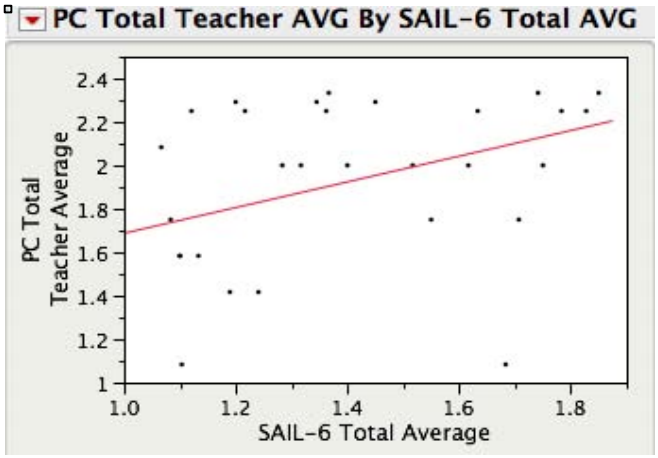
3F-3 TEACHER OUTCOME EXHIBIT VII: '*Teacher Portfolio Conference Ratings'* and their Association with Student Survey Response and Learning Outcomes

3F-3 Exhibit VII Data Indications: *By the third year of PAIR 'Teacher Portfolio Conference Performance Rating's strongly influenced "Student SAIL Interview' ratings and 'Student Portfolio Conference' ratings, but not Student ISAT academic performance ratings.* In this exhibit, the 'Teacher Portfolio Conference' performance ratings variable emerges as a major predictor of student arts integration outcomes. That is, each individual teacher's level of articulation of PAIR goals and description of student arts integrated learning predicts student interview and portfolio assessment ratings, but not student survey or student academic outcomes.

The influence of Teacher Portfolio Performance Outcomes on Arts Integration Student Learning

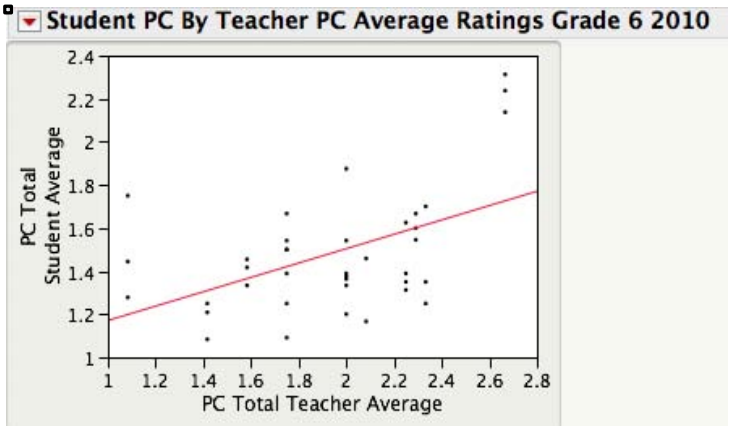
The first chart below [3F-3 Exhibit VII: Figures 1] depicts the statistically significant degree of association between Teacher Portfolio Conference rating and student SAIL interview scores [F ratio = 4.8, $p < .04$; Spearman r = .43, $p < .02$].

3F-3 Exhibit VII Figure 1: Bivariate Fit Between 'Teacher Portfolio Conference' Ratings and 'Student SAIL Interview' Ratings in Grade 6 (Final Year of PAIR)



Similar to the previous chart, the positive intersection between *teacher* and *student* portfolio conference ratings, depicted below [3F-3 Exhibit VII Figure 2], is highly statistically significant for 6th grader by the final year of the PAIR project [F ratio = 11.4, $p < .002$; Spearman r = .38, $p <. 02$].

3F-3 Exhibit VII Figure 2: Statistical Association Between 'Teacher Portfolio Conference' Ratings and 'Student Portfolio Conference Ratings' in Grade 6 (Final Year of PAIR)



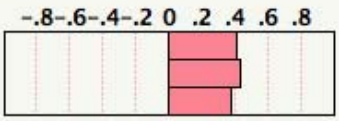
The correlation table figure below [3F-3 Exhibit VII Figure 3] summarizes the non-parametric correlation measures showing again that the internal measure of PAIR teacher and student performance outcomes are highly inter-related.

3F-3 Exhibit VII Figure 3: Correlation Table Indicating the Degree of Association Among 'Teacher Portfolio Conference,' 'Student SAIL Interview', and 'Student Portfolio Conference' Ratings

Teacher Portfolio Conference Correlation with Student SAIL and PC Ratings Grade 6 2010

Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|
PC Total Student Average	SAIL-6 Total Average	0.4088	0.0308*
PC Total Teacher Average	SAIL-6 Total Average	0.4312	0.0220*
PC Total Teacher Average	PC Total Student Average	0.3767	0.0197*



3F-3 Exhibit VII Summary, Emerging Themes: *The PAIR teacher and student performance outcomes are highly positively inter-correlated confirming that the teacher portfolio conference ratings predict student performance in arts integration learning and not student academic achievement.* Teacher portfolio conference ratings represent the last in a series of teacher performance outcomes unique to this project. The conferences were piloted in the second year of the project and presented in professional development sessions to forecast how researchers would rate teacher and student understanding of the goals and outcomes of arts integration teaching and learning. The conferences were administered comprehensively only during the final year of the project as a way to measure (a) the culminating effect of PAIR professional development and (b) its contribution to all student learning survey response and performance outcomes.

As previously reported in this section of the report [3F-3], the Teacher Effects (Factor AI, Teacher Outcome Exhibit I) and the averaged 'Combined Professional Development' variables (Factor AVI, Teacher Outcome Exhibit VI) are most highly predictive of the 'Teacher Portfolio Conference' ratings - especially for teachers who had been implementing the PAIR project for three years. As reported in this exhibit, the 'Teacher Portfolio Conference' ratings correlated most strongly with the 6th grade 'Student SAIL Interview [3F-3 Teacher Outcome Exhibit V] and 6th grade 'Student Portfolio Conferences' [3F-3 Teacher Outcome Exhibit VII], suggesting that

the *teacher conference performance outcomes generally most benefitted the students who had the intensive three-year experience with PAIR.* Taken together, *data analysis of the first seven exhibits provides a chain of evidence that confirms the significant impact of PAIR teacher professional learning factors on both teacher and student arts integration performance outcomes.*

Reported next are four Student Learning Outcome Exhibits (Exhibits VIII-XI) that focus on the direct impact of PAIR teacher factors on the measures of student learning and the degree and pattern of correlation among these student outcome factors.

* * *

3F-3 STUDENT OUTCOME EXHIBIT VIII: '*Student Survey Response Ratings'* and their Association with Teacher Professional Development Factors and Student Learning Outcomes, Controlling for Student Demographic Factors

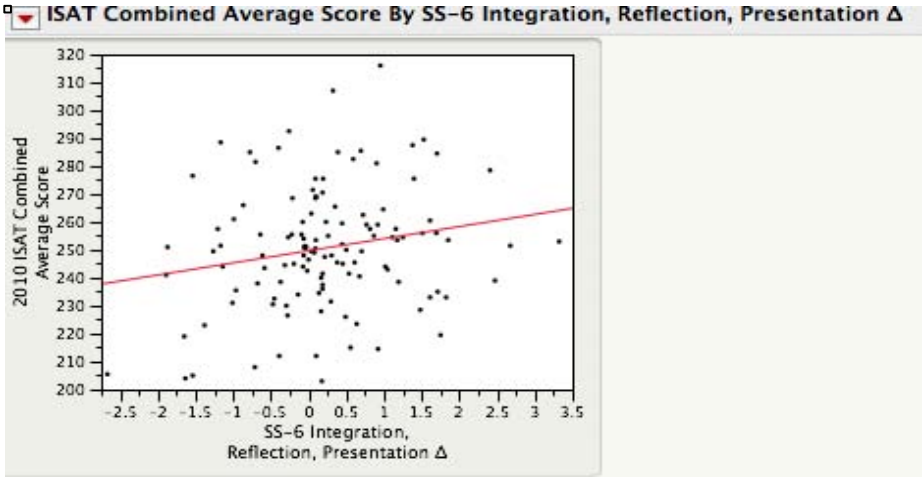
3F-3 Exhibit VIII Data Indications: *There is a negligible impact of the seven teacher professional development outcomes on 'Student Survey Response Ratings' and there appears to be a limited influence of 'Student Survey Responses' on the three Student Outcomes.* The longitudinal student surveys were designed to capture the perception of classroom practices and culture as the study progressed. By the final year of the project, control-treatment school comparisons differed most with regard to averaged questions focused on the presence of opportunities for students to engage in integrative teaching and learning experiences, reflection on the learning process, and the opportunity to present arts integration work in their classroom.

The averaged combined response rating of these three factors produced the clearest indication of comparative change in PAIR classroom practices. Yet the data analyses revealed no direct significant relationship between the treatment school students' perceptions of change in classroom culture and any of the seven teacher professional development factors. That is, the general experience of PAIR — more than the particular impact of teacher PD factors — influenced the student perception of change in their classroom culture.

However, the student survey data is connected significantly to their arts and academic learning outcomes, particularly with regard to academically challenged or minority students. Regression factor analysis reveals a mild degree of association between change in the students' student survey and student portfolio conference ratings [$r^2 = .28$, $p < .07$]. The only other significant demographic factors, student HAL designations and family income, tended to be *negatively correlated* with student survey data, *thus demonstrating that the perception of changes due to arts integration classroom practices were more important to pre-designated low academic or low income PAIR Treatment School students.* Ethnicity demographic designations also were related highly with the student survey response ratings, thus demonstrating a greater sensitivity to change in teacher practices by the minority status students, though this statistical effect accounted for less than 7% of the explained variance in the survey responses.

The display below [3F-3 Exhibit VIII Figure 1] shows that there is also a positive relationship between combined student survey ratings and student academic achievement. The bivariate fit of student survey results and combined reading and math scores reveals that a positive statistical trend exists between student perception of growth in the experience of arts integrated classroom practices and the students' ability to perform on academic tests [F ratio = 5.0, $p < .003$; Spearman $r = .17$, $p < .07$].

3F-3 Exhibit VIII Figure 1: Bivariate Fit of Averaged 'Student Survey' Response Ratings by Student ISAT Averaged Reading and Math Standardized Test Scores



3F-3 Exhibit VIII: *Summary and Emerging Themes*: The relative isolation of student survey responses from virtually all teacher professional development factors was unexpected. It would seem that the relatively strong ratings of changed teaching practices due to PAIR as indicated by teacher professional development survey responses would likely to predict a stronger rating of change in classroom practices by the students. Yet this was not the case.

Nonetheless, there appears to be a statistical link between perception of change in classroom practices and student portfolio conference performance ratings. That is to say, the students' ability to demonstrate and articulate arts integration learning appears to correlate with student survey responses by students most sensitive to the increased presence of arts integration practices in their classrooms.

Furthermore, the student survey ratings also are positively associated with ISAT test scores. Apparently PAIR students who did become more aware of the collaborative presence of the teaching artist and peer collaboration in the arts integration units in their classrooms and its influence on their portfolio performance assessment tasks, also tended to do better on standardized tests by the third year of the project.

* * *

3F-3 STUDENT OUTCOME EXHIBIT IX: *Student 'Snapshots of Arts Integration Learning' (SAIL) Interview Ratings'* and their Association with Teacher Professional Development Factors and Student Learning Outcomes, Controlling for Student Demographic Factors

3F-3 Exhibit IX Data Indications: *Strong statistical connections exist between the SAIL Interview Response Ratings and both teacher professional development and student learning outcomes.* As indicated in the correlation matrix below [3F-3 Exhibit IX Table 1**]**, for example, SAIL ratings are linked directly to several formative professional development variables such as attendance, survey responses, and teacher portfolio conference performance. *It appears conclusive that, in general, PAIR treatment school students' understanding of arts and arts integration processes result from exposure to those PAIR teachers who were highly-rated for*

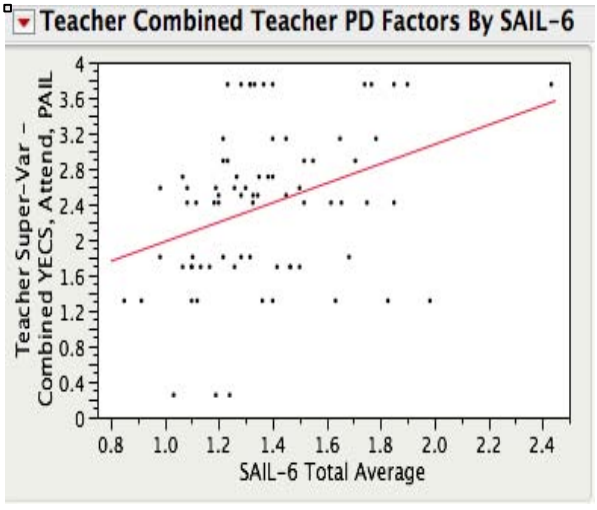
positive response to CAPE professional learning experiences, and, conversely, the students who demonstrate a lack of arts integration learning are linked with teachers with low professional development ratings.

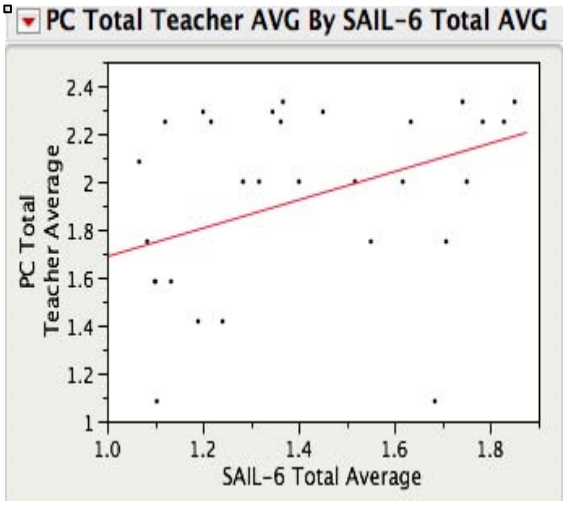
3F-3 Exhibit IX Table 1: Degree of Association Between 'Student SAIL Interview Ratings' and 'Combined Teacher Professional Development Factors'

SAIL-6 Total Average	Teacher AVG PD Attendance	0.3013	0.0082*
SAIL-6 Total Average	Teacher AVG 3 Key YECS factors	0.3002	0.0159*
SAIL-6 Total Average	Teacher Super-Var - Combined YECS, Attend, PAIL	0.3218	0.0046*
SAIL-6 Total Average	PC Total Teacher Average	0.4311	0.0174*

In the two scatterplots below [3F-3 Exhibit IX Figure 1**]**, the high degree of bivariate fit between SAIL scores and Combined Teacher PD factors [F ratio = 10.1, p < .002; Spearman r = .32, p < .005] or SAIL scores and Teacher Portfolio Conference ratings [F = 5.1, p <. 03; Spearman r = .43, p < .02] provides clear evidence of the strong association among these variable, despite being limited by sample size and the occurrence of ties in student scores in both variables.

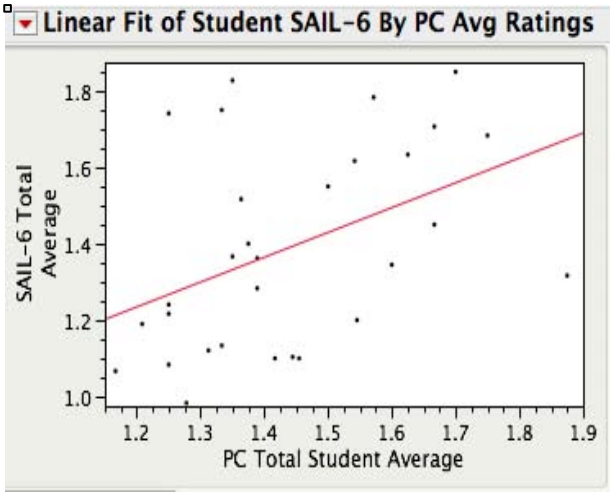
3F-3 Exhibit IX Figure 1: Bivariate Fit of 'Student SAIL Ratings' with 'Combined Teacher Professional Development Factors' and 'Teacher Portfolio Conference Ratings'





It appears also [3F-3 Exhibit IX Figure 2**]** that the SAIL ratings are aligned strongly with student portfolio conference performance ratings [F ratio = 6.6, $p < .02$; Spearman r = .45, $p < .01$].

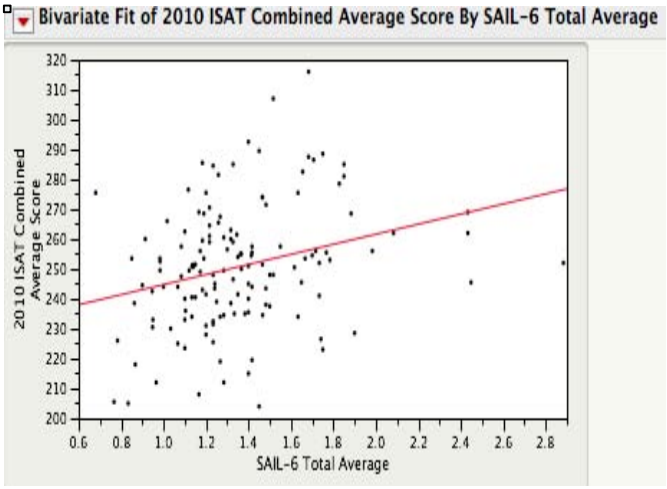
3F-3 Exhibit IX Figure 2: Bivariate Fit between 'Student SAIL Interview Ratings' and 'Student Portfolio Conference Ratings'



Taken together, results depicted by these three figures suggest that the SAIL ratings are linked causally to PAIR teacher professional development outcomes.

Furthermore, since SAIL test are not focused explicitly on academic performance in reading or math, the highly statistically significant association between SAIL and student ISAT standardized test scores [F ratio = 11.5 $p < .001$; Spearman r = .29, $p < .001$] indicate that a causal link also exists between these two variables.

3F-3 Exhibit IX Figure 3: Bivariate Fit between 'Student SAIL Interview' Ratings and 'Combined Student Averaged Academic Scores'



3F-3 Exhibit IX Summary, Emerging Themes: *Analysis of Student 'Snapshots of Arts Integration Learning' (SAIL) assessment data is highly related to multiple measures of teacher professional development and student learning outcomes.* Students whose teachers (a) attended PD more PAIR PD sessions and (b) learned more successfully how to integrate arts teaching and learning practices into their curriculum by collaborating more effectively with teaching artists, (c) responded to SAIL Interviews with more a more sophisticated understanding of both arts and arts integration learning processes.

The SAIL data provided an independent, internally consistent measure of the impact of PAIR on student arts and academic learning. The close association of SAIL ratings with ISAT test scores suggest that the higher order thinking and creatively skills that many teachers thought children developed in their PAIR units actually do predict performance on math and language test scores. Thus, *the SAIL assessment data functions as the powerful intermediary variable that provides both a measure of PAIR's program development integrity and impact*. In this scenario, the SAIL assessment ratings become the nexus point in multivariate analysis that illuminates the link between (a) the evaluation of PAIR teacher professional development outcomes and their impact on PAIR student arts integration learning outcomes and (b) the evaluation of PAIR student arts integration learning outcomes and their impact on state standardized academic achievement test scores [see also IIIG Summary Figure 5 in the next chapter].

* * *

3F-3 STUDENT OUTCOME EXHIBIT X: *'Student Portfolio Conference' Performance Assessment Ratings* and their Association with Teacher Professional Development and Student Learning Outcomes, Controlling for Student Demographic Factors

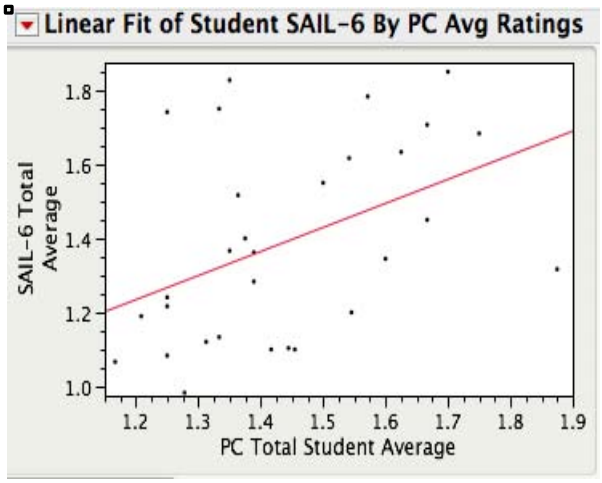
3F-3 Exhibit X Data Indications: *PAIR 'Student Portfolio Conference' ratings provided an arts integration performance assessment — in parallel with the SAIL Interview responses — that also correlated significantly with academic achievement.* In PAIR, the Student Portfolio Conferences provided an alternative strategy for assessing students' understanding of arts integration learning that is highly correlated with the previously reported SAIL student arts integration assessments. Unlike the SAIL Interview response data, the portfolio conference

ratings were based on PAIR students' ability to articulate and demonstrate their reflective understanding of arts integration learning in the context of presenting their own work samples. By design, the portfolio conference sessions were highly interactive, involved peers, and were based on the examination and demonstration of each student's own arts integration work.

Both Student Portfolio Conference and SAIL assessments were conducted in the final year of the project. Whereas the SAIL assessments covered only the 6th graders who participated in the PAIR for three years, the Student PC Ratings presented here were sampled across grades 4-6 in the final year of the project. Thus the Portfolio Conference data provides a look at arts integration learning across grade levels, but with smaller sample size per grade level than was the case with the grade 6 SAIL assessments.

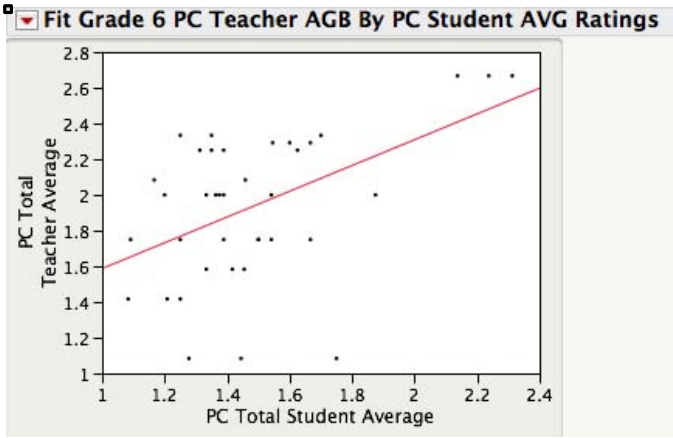
Nonetheless the strong association between Student Portfolio Conference and SAIL ratings demonstrate that arts integration assessments to be mutually reinforcing. As displayed below [3F-3 Exhibit X Figure 1] the linear fit of Student Portfolio Conference with SAIL ratings [F ratio = 6.6, $p < .02$] and the resulting non-parametric correlation between the factors [Spearman r = .45, $p < .01$] were both highly statistically significant. *Thus evidence from this data display suggests strongly that PAIR students who are most articulate about the general principles of arts integration tend also to be the students who can better demonstrate arts integration connections in the context of their own and the peer's work in portfolio conferences.*

3F-3 Exhibit X Figure 1: Bivariate Fit between Grade 6 Arts Integration 'Student Portfolio Conference' and 'SAIL Interview' Ratings



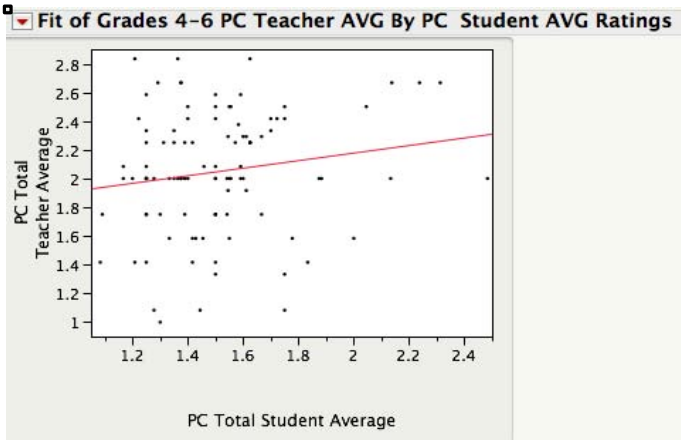
When looking only at the grade 6 cohort ratings from the primary PAIR longitudinal cohort, the degree of linear fit with Teacher Portfolio Conference data [3F-3 Exhibit X Figure 2 below] suggests a strong, highly statistically significant link with the Student Portfolio Conference ratings [F ratio = 11.4, $p < .002$; Spearman r = .38, $p < .02$]. *This data display shows that PAIR students whose teachers are articulate about the general principles of arts integration and the evaluation of arts integration student learning, are more likely to be the students who can better demonstrate arts integration connections in the context of their own and the peer's work in portfolio conferences.*

3F-3 Exhibit X Figure 2: Bivariate Fit of Only Grade 6 'Student Portfolio Conference' and 'Teacher Portfolio Conference' Ratings



However, the degree of linear fit among *combined* grade level 4-6 Student and Teacher Portfolio Conference ratings in the final year of the PAIR (3F-3 Exhibit X Figure 3 below) results in a relatively weak positive statistical trend [F ratio = 2.7, $p < .10$; Spearman r = .15, $p < .12$] due to statistically insignificant correlations in Grades 4 and 5. *This finding is important because it argues strongly the necessity for long-term engagement in PAIR before the positive relationship between 'Student Portfolio Conference' and 'Teacher Portfolio Conference' can be established. Results from these data suggest the three-year mark in PAIR ('sixth grade effect') is the tipping point for determining the close relationship between teaching and learning in the evolution of PAIR arts integration practices.*

3F-3 Exhibit X Figure 3: Bivariate Fit of Combined Grade Level (4-6) Student and Teacher Portfolio Conference Ratings in the Final Year of PAIR



Furthermore, the regression factor analysis matrix below [3F-3 Exhibit X Figure 4] identified both 'Student SAIL Interview' ratings [$r^2 = .60$, $p < .04$] and 'Combined Teacher Professional Development Ratings' [$r^2 = .37$, $p < .003$] and as by far the most significant, positive predictors of 'Student Portfolio Conference' ratings, controlling for all other teacher PD, student learning and student demographic factors. Together these two variables account cumulatively for approximately 98% of the variance in Student Portfolio Conference ratings.

3F-3 Exhibit X Figure 4: Step-wise Regression Factor Analysis Table Indicating Principal Positive Predictors of Student Portfolio Conference Performance from among all Teacher Professional Development, Student Learning and Student Demographic Factors

Regression Factor Analysis: Student PC by Teacher PD, SAIL & Demographic Factors

Step	Parameter	Action	"Sig Prob"	Seq SS	RSquare
1	SAIL-6 Total Average	Entered	0.0411	0.105179	0.5992
2	Teacher Super-Var - Combined YECS, Attend, PAIL	Entered	0.0026	0.064567	0.9670
3	3-HAL Designation{H-A}	Entered	0.0326	0.004782	0.9943
4	Ethnicity{Black-White&Asian}	Entered	0.0700	0.000872	0.9992
5	2010 ISAT Combined Average Score	Entered	0.0776	0.000134	1.0000

These statistical results make clear the role 'Student Portfolio Conference' as an intermediary arts integration learning outcome variable in parallel the SAIL Interview outcomes. However, when analyzing the degree of association between the Student Portfolio Conference outcomes and ISAT test scores, there exists only a modest positive correlation [F ratio = 6.6, $p < .02$; Spearman $r = .19$, $p < .06$] between these two factors in contrast to the stronger relationship

between the SAIL and ISAT test score results. The next and final section of this chapter reveals that, when comparing all competing regression factors (including student demographic data) that can predict ISAT scores, the portfolio conference ratings emerge as the most significant predictive variable.

3F-3 Exhibit X Summary and Emerging Themes: *Grade 6 Student Portfolio Conference and the previously-reported student SAIL Interview ratings both demonstrate that two validated forms of arts integration learning are both positively linked with PAIR professional development factors and are highly predictive of academic achievement.*

Whereas SAIL interview protocols challenged students to describe the nature and importance of arts-integrated learning in contrast with domain specific learning, the Portfolio Conference elicited student description and demonstration of arts-integrated learning based on their work and in the context of interactive discussion with their peers. These two measures of student arts integration learning proved to be highly inter-correlated and associated significantly with academic achievement. Taken together, these two arts integration assessments developed in PAIR provide evidence of nexus point in the chain of causal factors that proceeds from (a) individual teacher professional development outcomes that (b) positively predict individual levels of student arts integration learning, and that (c) consequently predict measures of academic achievement.

A final summary of the through-line hypothesis of autonomous, yet inter-dependent causal factors that prescribe the coherency of PAIR's arts integration practices and their impact on student academic achievement will be discussed further in the following section of this chapter (3F-3 Exhibit XI) and in the overall summary chapter (3G).

* * *

3F-3 STUDENT OUTCOME EXHIBIT XI: *Student Academic Performance Ratings* and their Association with PAIR Professional Development, Student Survey, and Student Arts & Arts Integration Outcomes, Controlling for Student Demographic Factors

The importance of student academic performance data analysis to the success of the PAIR arts integration project

The student Illinois Snapshot of Academic Test (ISAT) academic performance data provide an essential benchmark for judging the success of any public school in Chicago during the course of the PAIR project. Investment in arts integration curriculum innovation is not sustainable if a school fails to meet minimum academic standards of achievement or improvement in standardized tests. Therefore it was incumbent on the researchers to evaluate each element of the PAIR arts integration program for its positive contribution to student academic achievement.

As reported in Part 2 of this report, significant gains in academic scores occurred in PAIR schools in contrast to CPS district averages and comparisons with other CPS arts magnet schools. By the third year of PAIR implementation, control-treatment school comparisons made it clear that the general effects of rigorous arts integration professional development and teaching practices (described in Part 1 of this report) positively predicted both statistically significant academic gains, especially for students categorized as low-performing students at beginning of the project.

The multiple variable teacher-student intersection analysis methods detailed in Part 3 of the whole report focused on the thorough examination of (a) seven independent strands of teacher professional learning outcomes both for indications of the internal consistency of the treatment school programs and for their separate and combined contribution to four forms of student arts and academic learning outcomes, and (b) each strand of student arts integration survey, interview and performance assessment response both for indication of the internal validity of arts integration learning and for their degree of association with external measures of academic achievement.

In this final chapter of Part 3 of the PAIR report, all teacher professional development and all student survey, student learning and student demographic factors will be examined to determine which parts of the PAIR program had the most direct influence on the ISAT combined reading and math scores.

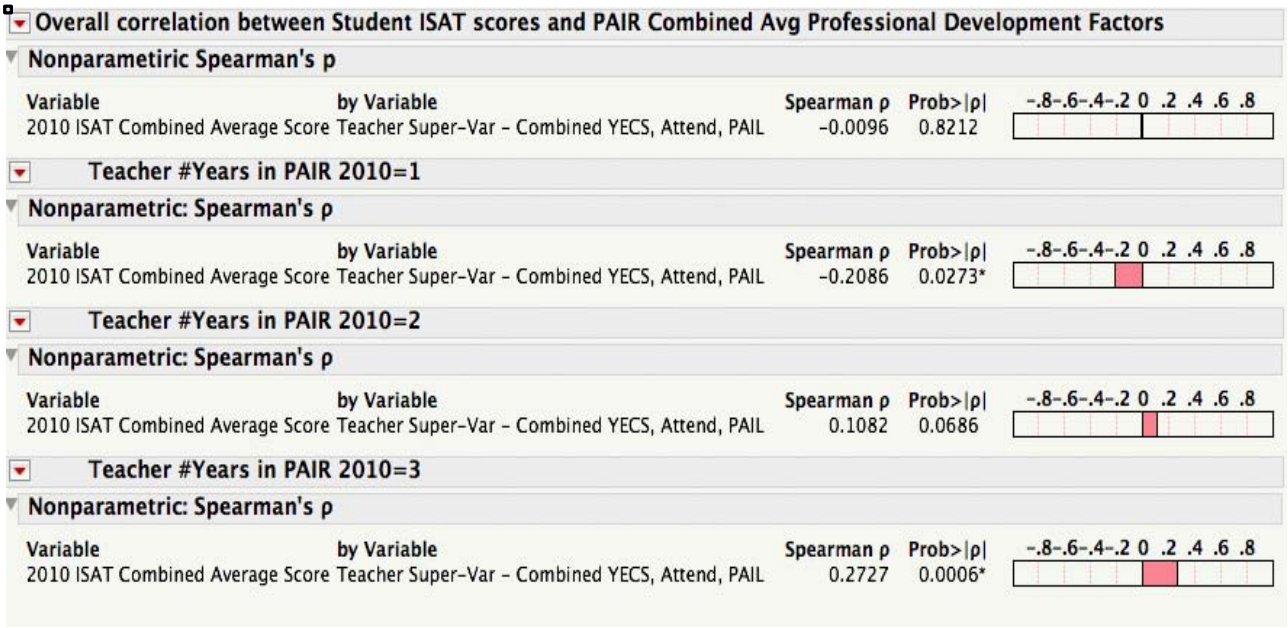
3F-3 Exhibit XI Teacher Professional Development – Student Academic Intersection Data

Indications: *PAIR Teacher professional development factors increasingly predict student academic achievement over time ('Fourth Grade Effect').* As previously indicated in Part 3, the level of PAIR teacher professional development outcomes predicted significant positive

difference in the quality of student arts learning. Yet there was no single professional development factor in PAIR found to be strongly associated with student academic test scores. *However, when parsing the data by years of teacher experience in PAIR, positive correlations between combined professional development factors and student achievement are clearly aligned with levels of experience of teachers in the PAIR project.*

The top line in the data display below (3F-3 Exhibit XI Figure 1] shows that the *overall* correlation (grades 4-6) between 'Combined Professional Development Factors' and student academic performance to be negligible. Subsequent correlations reveal that sixth grade teachers with only one year of experience with PAIR professional development training did not *positively* influence student ISAT achievement scores with respect to their PD ratings [r = –.21, p < .03] and fifth grade teachers with two years experience mildly influenced ISAT scores according to their PD ratings [r = +.11, p < .07]. However, fourth grade teachers with three years experience strongly predicted academic performance in proportion to their combined professional development outcome ratings [r = +.27, p < .001].

3F-3 Exhibit XI Figure 1: Correlation Table Indicating the Overall influence of Combined Teacher PD Factors on Student Academic Test Scores, and the Progressively Positive Effect of Years of Teacher Participation on Student Academic Achievement Ratings



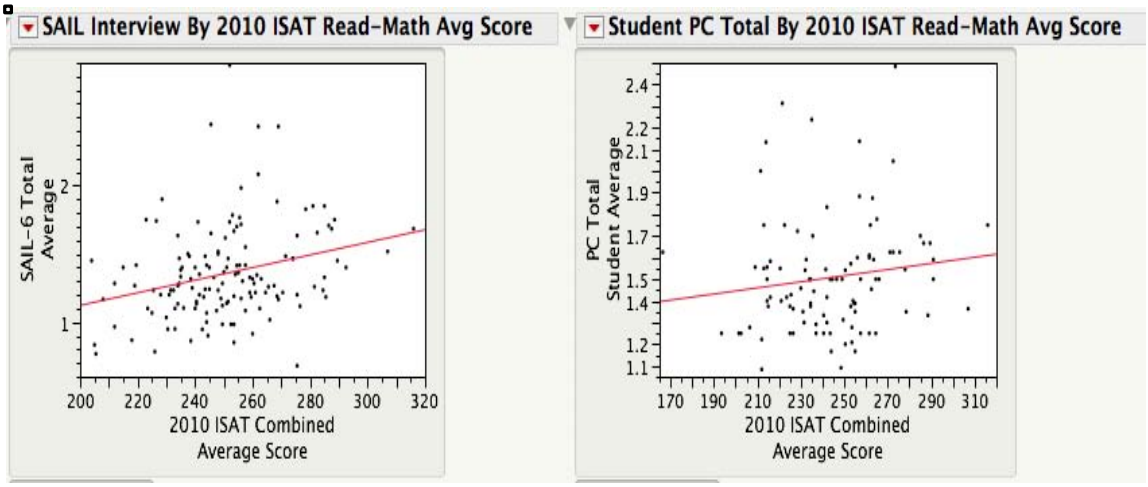
The chart above thus confirms the "fourth grade effect" hypothesis first formulated in Part 1 of this report—that *positive PAIR teacher professional development outcome ratings had progressively stronger influence on student learning proportionate to the years of high quality participation with the PAIR professional development program.* Thus, the fourth graders who had three-year veteran PAIR teachers with high professional development outcome ratings were most likely to achieve the highest academic scores in their grade level.

3F-3 Exhibit XI Student Arts Integration Learning – Student Academic Performance Intersection Data Indications: *There are three leading associative links between PAIR program arts integration outcomes and academic achievement: 'Student Survey responses, 'Student SAIL Interview' ratings, and 'Student Portfolio Conference' ratings'.* Through analyses of 'degrees of correlation' and regression factor analysis among these student arts learning intersection variables, researchers were able to make judgments regarding the possible causal links among these three factors and student academic outcomes.

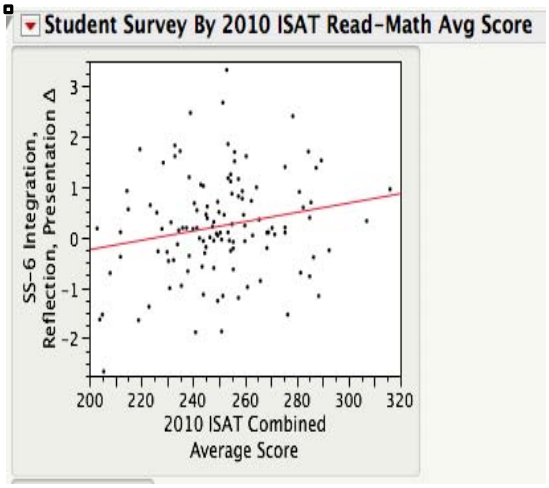
The three data displays below [displayed in 3F-3 Exhibit XI Figure 2] compare the relative degrees of association between ISAT scores and the three internally assessed student arts learning factors: (a) 'Student Survey' response ratings [F ratio = 4.99, $p < .03$; Spearman $r = .17$, $p < .07$], (b) 'Student SAIL Interview' ratings [F ratio = 11.47, $p < .001$; Spearman $r = .29$, $p < .001$], and (c) 'Student Portfolio Conference' interview and performance assessment ratings [F ratio = 1.94, $p < .17$; Spearman $r = .19$, $p < .06$].

(3F-3 Exhibit XI Figure 2 on next page)

3F-3 Exhibit XI Figure 2: Comparative Bivariate Fit among Three Student Arts Integration Learning Outcomes with Student Academic Achievement Scores



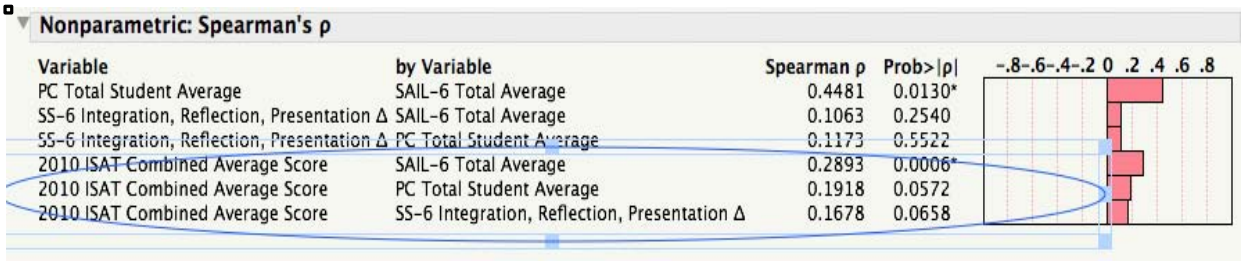
(the third scatterplot display on next page)



The top three lines of the correlation table below [3F-3 Exhibit XI Figure 3] reveal that the 'Student Survey Response' data are independent of the two highly-correlated student learning factors ['SAIL interview and 'Student Portfolio Conference'). Yet the three correlation statistics at the bottom of the chart show that all three factors correlate significantly with the student academic performance scores. This triangulation of predictive factors of academic achievement suggests that student survey data that reflects changes in the students' perception of a PAIR classroom culture – the emergence of arts integration practices, the presentation of arts learning work, and the increase in time spent on reflection – is congruent with data from arts integration

learning assessments that capture levels of student understanding of arts integration practices in these classrooms. Taken together these data suggest an important through-line of causal factors that link changes in classroom culture to student arts and academic learning.

3F-3 Exhibit XI Figure 3: Correlation Table indicating the 'Pattern and Degree of Correlation' of Combined Student Academic (averaged Reading and Math) Achievement Scores in relation to Student Survey Response and Two Arts Integration Learning Outcomes



Nonparametric: Spearman's ρ

Variable	by Variable	Spearman ρ	Prob>|ρ|
PC Total Student Average	SAIL-6 Total Average	0.4481	0.0130*
SS-6 Integration, Reflection, Presentation Δ	SAIL-6 Total Average	0.1063	0.2540
SS-6 Integration, Reflection, Presentation Δ	PC Total Student Average	0.1173	0.5522
2010 ISAT Combined Average Score	SAIL-6 Total Average	0.2893	0.0006*
2010 ISAT Combined Average Score	PC Total Student Average	0.1918	0.0572
2010 ISAT Combined Average Score	SS-6 Integration, Reflection, Presentation Δ	0.1678	0.0658

Follow-up regression factor analyses reveal that when the relative fit of the three triangulating factors is measured against academic achievement, the 'Student Portfolio Conference' ratings demonstrate the strongest linear fit with combined reading and math test scores [see 3F-3 Exhibit XI Figure 3 below]. Furthermore, '*Student Portfolio Conference' averaged ratings are the strongest predictor of academic score even when student demographic variables (that include gender, family income, ethnicity, ELL, and High, Average, or Low (HAL) academic standing previous to the PAIR project) are factored into the regression equation.*

Although the statistical significance of this correlation is relatively low [$p < .09$] when pitting the relatively small sample size of treatment school 'Student Portfolio Conference' data against all competing factors this variable is, by far, the most prominent predictor of academic achievement among all other variables and this factor explains a relatively high degree of variance in the academic achievement ratings [$r^2 = .35$).

Furthermore, when removing just one demographic factor from the regression analysis – that is, the 'Family Income' variable determined by free and reduced lunch status and its negative statistical correlation with academic scores[2] – Student Portfolio Conference ratings assume a

highly statistically significant and remarkably strong positive predictor of academic achievement [$p < .01$, $r^2 = .62$] in comparison to all other teacher professional development, student learning, and student demographic factors of the 6 most prominent factors selected from the regression analysis [see Exhibit XI Table 3].

3F-3 Exhibit XI Figure 4: Stepwise Regression Factor Analysis Table Indicating the Principal Predictors of Student Academic Achievement from among all Teacher, Student, and Student Demographic Outcome Variables (Excepting 'Family Income'[3])

Step	Parameter	Action	"Sig Prob"	Seq SS	RSquare
1	PC Total Student Average	Entered	0.0116	4307.808	0.6211
2	SS-6 Integration, Reflection, Presentation Δ	Entered	0.1378	862.1787	0.7454
3	PC Total Teacher Average	Entered	0.1513	643.1797	0.8381
4	3-HAL Designation[A-H]	Entered	0.1943	423.7852	0.8993
5	SS-6 Integration and Presentation Δ	Entered	0.1417	396.3576	0.9564
6	Teacher AVG PAIL Quantity, Quality	Entered	0.0077	297.7544	0.9993

3F-3 Exhibit XI Summary and Emerging Themes: *The detailed correlation and regression factor analyses of teacher, student learning, and student demographic variables establishes that PAIR student arts integration learning outcomes do predict academic achievement.* 'Snapshots of Arts Integration' (SAIL) Interviews and the interactive 'Student Portfolio Conference' performance assessments together provide a bridge to academic achievement. In addition, those students who responded to survey questions by indicating positive changes in the classroom arts integration practices also tended to do better in the academic achievement tests, suggesting that degree of positive changes in teaching practices can significantly impact student achievement scores.

Although teacher professional development outcomes do not directly predict student academic achievement, path analysis results showed that students who had teachers with the most experience with PAIR *and* achieved relatively high ratings of averaged professional development outcomes scored higher than control school cohorts on the ISAT tests by the final year of the project.

Altogether these findings confirm that, although positive teacher professional development factors strongly predict student arts integration learning outcomes, it is the student arts learning

outcomes that most directly predict academic success. From this viewpoint, the sequence of analyses of the PAIR outcome variables through the twelve exhibits examined in this chapter now delineates a path of positive teacher-student arts integration intersections that in due course lead to excellence and more opportunities for equity in academic achievement in CPS schools.

* * *

The results of all teacher and student outcomes and their intersections presented in Part 3 sections A-F will be summarized in section 3G, the final chapter of the comprehensive PAIR Report.

* * *

3G: PAIR Comprehensive Report FINAL SUMMARY

A Synopsis of the PAIR report through the Lens of Five Teacher Professional Development – Student Learning Outcome Intersection Data Flow Charts

3G: *Table of Contents*

3G: PAIR Comprehensive Report FINAL SUMMARY	1
3G: Table of Contents	1
3G: List of Tables and Figures	2
3G Introduction: The *Summary Figure Sequence*	3
3G-1: A Synopsis of Inter-related Aspects of PAIR Individual Teacher Professional Development Outcomes	4
3G-2: A Synopsis of PAIR Teacher Professional Development Outcomes – Student Learning Outcome Intersections	9
3G-3: A Synopsis of Inter-related Aspects of PAIR Individual Student Learning Outcome Factors	13
3G-4: A Synopsis of PAIR Student Outcome Regression Factors	17
3G-5: A Synopsis of Correlation-Regression Analyses of Combined PAIR Teacher-Student Intersection Factors	22
3G-6 Final PAIR Whole Report Conclusions, Implications for Future Research	26
Analysis of teacher-student learning intersections in relation to teacher interview data	26
The value of PAIR comprehensive data collection and multivariate analysis methods	28
Implications for future research	31

* * *

3G: *List of Tables and Figures*

3G-1 SUMMARY Figure 1:	A Flow-Chart Map of Inter-Related PAIR Individual Teacher Professional Development Rating Factors	6
3G-1 *Reference* Table 1:	Mean Difference and Correlations Statistics for All PAIR Inter-Related Teacher Professional Development Rating Factors	7
3G-1 SUMMARY Figure 2:	A Flow-Chart Map of All PAIR Teacher Professional Development – Student Learning Outcome Intersections.	10
3G-2 *Reference* Table 2:	Mean Difference and Correlation Statistics for All PAIR Intersecting Teacher Professional Development – Student Outcome Rating Factors	11
3G-3 SUMMARY Figure 3:	A Flow-Chart Map of PAIR Inter-Related Student Outcome Factors	14
3G-3 *Reference* Table 3:	Correlation Statistics Among All PAIR Student Outcome Rating Factors	15
3G-4 SUMMARY Figure 4:	A Flow-Cart Map of PAIR Teacher Professional Development and Student Regression Factors Most Associated with Student Learning Outcomes	18
3G-4 *Reference* Table 4:	Regression Correlation Data Analysis Results from All PAIR Professional Development, Student Learning, Student Demographic Factor Intersections	19
3G-5 SUMMARY Figure 5:	A Correlation-Regression Map of PAIR Combined Teacher-Student Outcome Intersection Factors	23

* * *

3G Introduction: The *Summary Figure Sequence*

In this final section of the PAIR report five data displays are used to summarize the *degree and paths of association* among the principal factors of teacher professional development and student learning outcomes in the PAIR treatment schools. These inter-related data patterns represent the statistically significant links (a) among teacher professional development categories and outcome factors [A I-VII][1], (b) among student learning outcome variables [BI-IV], and (c) the intersections between both sets of variables

The data displays (3G 1-5: Summary Figures 1-5) are ordered according to the hypotheses of the project. A *synopsis of data indications* will provide guidance toward interpreting each corresponding *summary figure*. Taken together the *summary figure sequence* will depict all teacher variables and the strength of correlation among them as an indication of the coherence of the program development, and the extent to which the PAIR teacher professional development outcomes predict student survey responses or learning outcomes. The key at the bottom of each of the *summary figures* (from left to right) explains the solid and dotted vectors in terms of strength of association, the geometric shapes in terms of comprehensiveness of the data sample, and the dark to light shading scheme in terms of sources of data (teacher professional development to classroom to student learning outcome variables). Follow-up *reference tables* provide the precise statistics associated with each *summary figure*.

* * *

[1] The teacher and student factors are fully described in PAIR Report section 3F-3: A1-VIII, BI-IV.

3G-1: A Synopsis of Inter-related Aspects of PAIR Individual Teacher Professional Development Outcomes (AI-VIII)

Since the teacher professional development (PD) factors are sequentially ordered, the correlations between factors will be interpreted as causal links that trace the impact of teacher professional development outcomes on variables that resulted from prior events. It is reasonable to assume, for example, that high levels of *individual teacher attendance ratings* in PAIR professional development sessions [A-III] could influence key factors of program quality as indicated by teacher *year-end teacher survey* ratings [3A-IV], which may lead also to higher *quality of classroom rated student work samples* [A-V], and eventually may result in higher levels of sophistication in *teacher reflection* comments made during the Teacher-Student portfolio conferences [A-VII].

3G-1 Summary Figure 1 Data Indications: *PAIR program effectiveness is indicated by the significant degree of association among all teacher professional development [PD] factors.* Because several teacher PD factors[2] (AI-VI) are significantly inter-correlated to some degree, Summary Figure 1 below confirms the overall cohesiveness of the sequential aspects of the PAIR professional development program. That is, high levels in *teacher attendance* in PD sessions and certain exemplary *teacher effects* both predict positive performance ratings in teacher portfolio conference sessions. In addition, these strongly inter-correlated ratings among various factors of professional development implementation provide the foundation for eventually determining the relative impact of PAIR on the principal student learning outcomes variables of the project.

In particular, dark arrows emanating from the Attendance [A-III] factor box indicate that averaged annual *attendance* significantly predicts the teacher professional development survey outcomes and is the most significant contributor to the combined teacher PD factors variable. This determination provides critical evidence that the treatment schools benefitted from PAIR in proportion to their engagement in arts integration training and not just by *years of participation in the project*, a variable that is not linked with teacher performance outcomes. Furthermore, teachers who embody the quality of various key

[2] Note that teacher Combined and Averaged PD Factor A-VI variable is by definition correlated to various degrees with professional development factors AIII-V,; overall, it is most highly correlated with the Teacher Attendance variable.

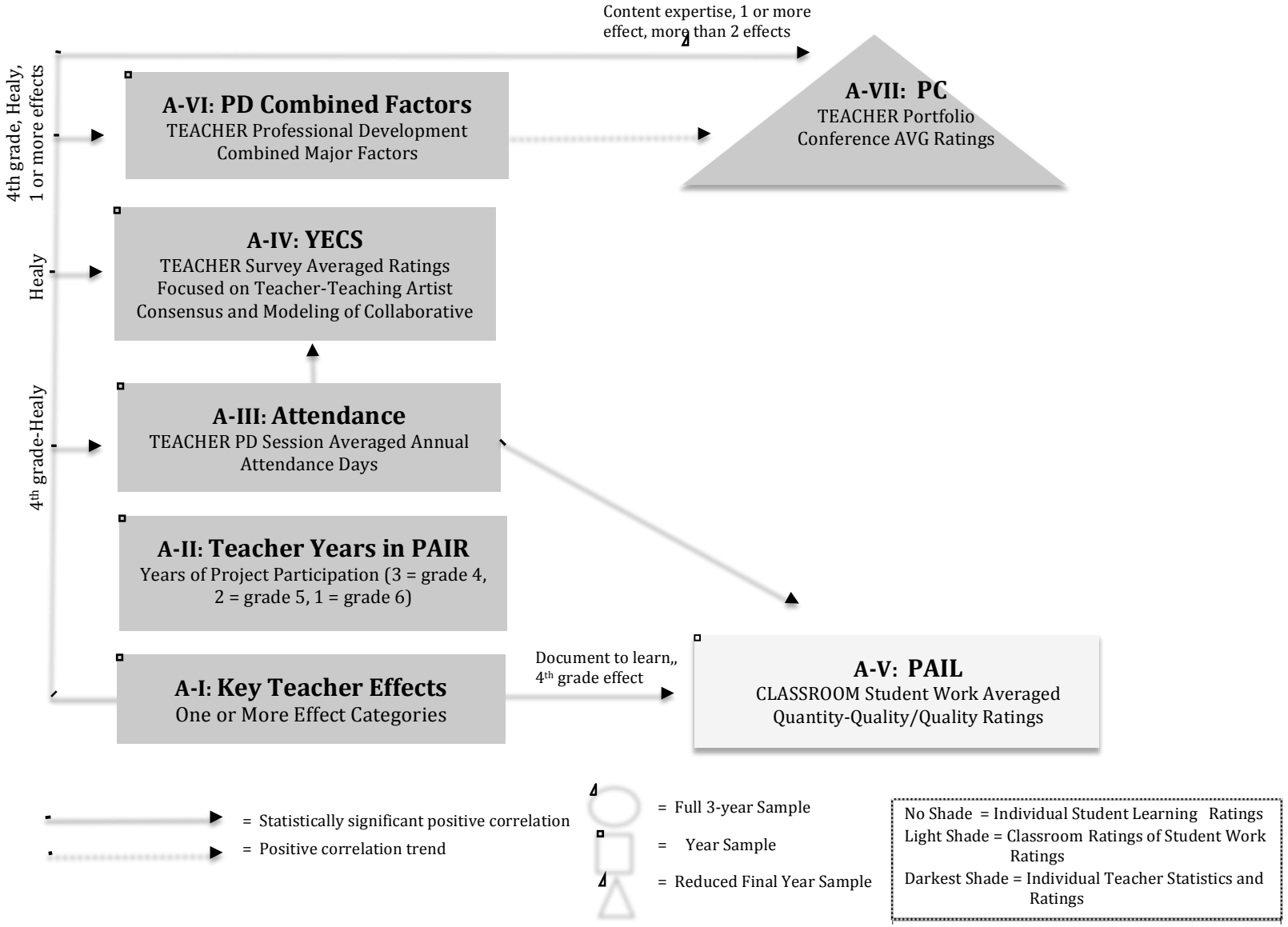
teacher effect categories, as previously reported, do perform better with respect to virtually all teacher PD outcomes, depending on the particular effect labeled in the figure below. Thus, mean score differences between teachers who do or do not exemplify the "content expertise" effect, for example, are most likely to be perform better during the portfolio conference sessions. Mean score differences between teachers who do or do not exemplify the "fourth grade or Healy" effect were most likely to attend PAIR sessions, rate higher on reflection on key teacher-teaching artists dynamics of the PAIR project, deliver higher quality student work samples, and rate highly on the combined teacher PD factors[3].

Overall, this first correlation map display below shows that *teacher performance in portfolio conferences* represents a fundamental measure PAIR program's success since virtually all teacher professional development factors are highly predictive of this teacher performance assessment task[4].

(3G-1 Summary Figure 1 on next page)

[3] See detailed correlation charts at the end of this chapter and section IIIF for precise correlation statistics for both averaged and individual strands of teacher professional development variables.
[4] (see 3D-E for full description of the teacher-student Portfolio Conference protocol, scroing rubrics, and qualitative analysis.

3G-1 SUMMARY FIGURE 1: A Flow-Chart Map of Inter-Related PAIR Teacher Professional Development Rating Factors. Multivariate map of the statistically significant inter-relationships (determined by mean score differences or pair-wise correlations) among all teacher professional development outcome variables by the final year of the PAIR project.



3G-1 Reference Table 1: Mean Difference and Correlations Statistics for All PAIR Inter-Related Teacher Professional Development Rating Factors

	A-II: **Teacher Years in PAIR**	A-III: **Teacher AVG PD Session Attendance Ratings**	A-IV: **Teacher AVG Year- End Key Curriculum and Survey [YECS] Ratings**	A-V: **Classroom PAIL Work AVG Quantity-Quality Student Work Sample Ratings**	A-VI: **Teacher Combined AVG Attendance, YECS, PAIR Ratings**	A-VIIIL: **Teacher AVG Portfolio Conference Interview Response Ratings**
Mean Differences t Test ratio, p Probability Confidence Value; ANOVA F ratio, p Probability Confidence Value						
AI: Teacher Effect: Content Expertise Effect	NA	ns	ns	ns	ns	t ratio=5.9, P<.0001; F ratio = 20.2 p <.0001
AI: Teacher Effect: Document to Learn Effect	NA	ns	ns	t test ratio=3.7, p<.0003; F ratio = 6.1 p <.01	ns	ns
AI: Teacher Effect: Fourth Grade Effect	NA	t ratio=4.3, p<.0001; F ratio = 25.8 p <.0001	ns	t test ratio=3.0, p<.003; F ratio = 10.02 p <.002	t test ratio=3.5, p<.005; F ratio = 10.2 p <.002	ns
AI: Teacher Effect: Healy Effect	NA	t ratio=6.7, P<.0001; F ratio = 67.3 p <.0001	t ratio=8.0, p<.0001; F ratio = 38.0 p <.0001	ns	t test ratio=6.0, P<.006; F ratio = 45.6 p <.0001	ns
AI: Teacher Effect: One or More Effects	NA	ns	ns	ns	t test ratio=2.8, p<.006; F ratio = 8.3 p <.004	t test=4.0, p<.0001; F ratio = 15.9 p <.0001
AI: Teacher Effect: More Than One Effect	NA	ns	ns	ns	ns	t test=3.1, p<.002; F ratio = 7.9, p<.006
ANOVA F ratio statistic, Probability value; Spearman correlation r statistic- Probability Value						
AII: Years of Teacher Participation	NA	ns	ns	ns	ns	ns
AIII: Teacher PD AVG Attendance	NA	NA	F ratio=23.7, P<.0001; r = .18, p<.0001 (model process)	F ratio =.54.4, p<.0001; r=.26, p<.0001 (Quantity-Quality)	NA (correlated by variable definition)	ns
AIV: Teacher AVG 3 Key YECS Survey Ratings	NA	NA	NA	ns	NA (correlated by variable definition)	ns
AV: Teacher AVG PAIL Student Work Ratings	NA	NA	NA	NA	NA (correlated by variable definition)	ns
AVI: Combined Attend, YECS, PAIL Ratings	NA	NA	NA	NA	NA	F ratio 7.6, p<.007; r = .17, p>.10
AVII: Teacher PC Ratings	NA	NA	NA	NA	NA	NA

3G-1: Summary, Emerging Themes

The importance of this first summary figure cannot be overestimated. Without compelling evidence of inter-correlating links among these teacher variables, it would be difficult to argue that the PAIR teacher training strands were significantly defined and cohesively implemented among the treatment schools. The extraordinary degree of association among these factors indicates also that teachers who exemplified PAIR Teacher Effects such as taking ownership of program development processes, focusing on documentation of student learning, or emphasizing content expertise in arts integration teaching and learning, validated the success of the PAIR professional development program.

Moreover, the path to excellence in overall teacher development is pervasive, varied and web-like. The various kinds and combinations of exemplary Teacher Effects, strong indications of arts integration practices indicated in the YECS teacher survey responses, and high levels of attendance in teacher professional development sessions all mattered to the development of the PAIR program in the treatment schools throughout the years of project development. Therefore the determination of a high degree of inter-correlation among sequence of teacher professional development variables outlined here provides the first step in presenting conclusive statistical evidence that the PAIR professional learning outcomes directly influenced student arts and academic performance.

* * *

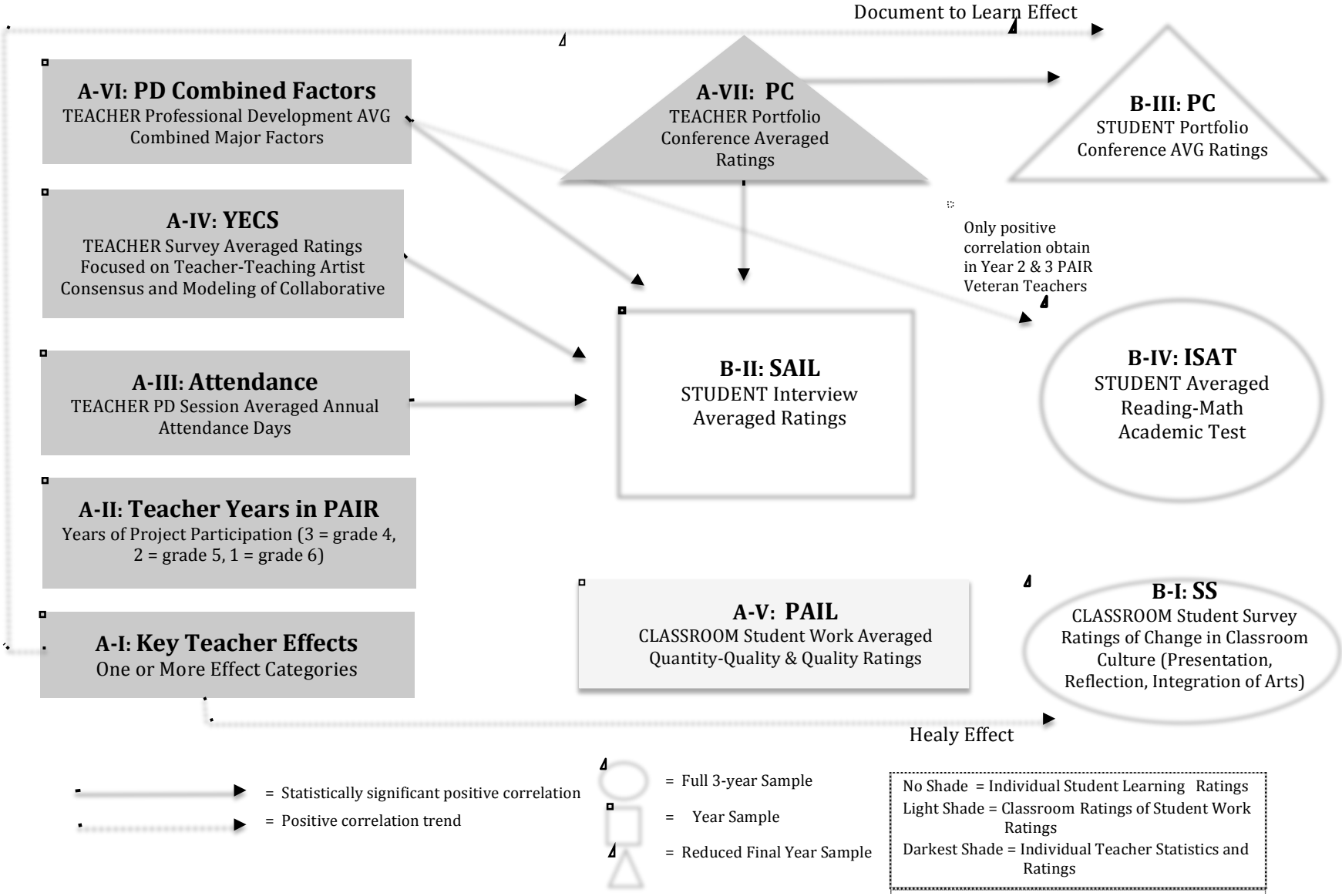
3G-2: A Synopsis of PAIR Teacher Professional Development Outcomes – Student Learning Outcome Intersections

3G-2 Data Indications: *Several PAIR program development factor ratings have significant impact on student learning factors.* The second figure below displays the probable impact of PAIR teacher professional development outcomes on student learning. The relationships among the variables are depicted by solid (highly significant) and dotted (statistical trend) vectors that portray a relatively discrete focus between five teacher PD factors and three student learning outcomes that emerged by the end of the final year of the PAIR project. These data indicate that:

- Teacher Attendance, Survey, and Combined Factors variable [AIII-IV, VI] ratings are highly associated with the Student Arts Integrated Learning Interview ratings [B-II];
- Teacher Portfolio Conference Ratings [A-VII] and the exemplary teacher Document to Learn Effect to a lesser extent predict Student Performance Conference Ratings;
- Positive statistical trends characterize the exemplary teacher Healy Effect on the student survey arts integration change ratings;
- Combined Teacher PD Factor ratings [A-VI] correlates positively with standardized measures of individual student averaged reading and math scores, yet only with teachers in their second or third year in the project.

(3G-2 Summary Figure 2 on next page)

3G-2 SUMMARY FIGURE 2: A Flow-Chart Map of All PAIR Teacher Professional Development – Student Learning Outcome Intersections. A multivariate correlation map of the statistical trends and statistically significant associations among teaching professional development and student learning outcome variables by the final year of the PAIR project.



3G-2 Reference Table 2: Mean Difference and Correlation Statistics for All PAIR Intersecting Teacher Professional Development – Student Outcome Rating Factors

	B-I: **Student Survey (SS) Response Ratings**	B-II: **Student SAIL Interview and Performance Assessment Ratings**	B-III: **Student Portfolio Conference (PC) Interview Response Ratings**	B-IV: **Student ISAT Combined Academic Performance Ratings**
	Mean Differences t Test ratio, p Probability Confidence Value; ANOVA F ratio, p Probability Confidence Value for each Student Learning Outcome			
AI: Teacher Effect: Content Expertise	ns	ns	ns	ns
AI: Teacher Effect: Document to Learn	ns	ns	t test ratio=1.4, p<.09; F ratio=3.2, p<.08]	ns
AI: Teacher Effect: Fourth Grade	ns	ns	ns	ns
AI: Teacher Effect: Healy	t test ratio=1.5, p<..08; F ratio=3.2, p<.08	ns	ns	ns
AI: Teacher Effect: One or More Effects	ns	ns	ns	ns
AI: Teacher Effect: More Than One Effect	ns	ns	ns	ns
ANOVA F ratio statistic, Probability value; Spearman correlation r statistic- Probability Value				
AII: Years of Teacher Participation	NA	NA	ns	ns
AIII: Teacher PD AVG Attendance	ns	F ratio=8.0, p< .006; r=.30, P<.008	ns	ns
AIV: Teacher AVG 3 Key Survey Ratings	ns	F ratio=7.7, p< .007; r=.30, P<.02	ns	ns
AV: Teacher AVG PAIL Student Work Ratings	ns	ns	ns	ns
AVI: Combined Attend, YECS, PAIL Ratings	ns	F ratio=10.1, p<.002; r=.32, p<.005	ns	Overall: ns Year 1 Teacher: F ratio = 3.5, p<.06; r = –.21, p<.03 Year 2 Teacher F ratio = .38, ns r = .11, p<07; Year 3 Teacher F ratio = 5.6, p<.02; r= .27, p<.001
AVII: Teacher PC Ratings	ns	F ratio = 4.8, p<.04; Spearman r = .43, p<.02	F ratio = 11.4, p<.002; Spearman r = .38, p<.02	ns

3G-2: Summary, Emerging Themes. The results from the figure below indicate decisively that strong statistical links exist between specific teacher professional development and internal developed student arts integration learning factors, and to a certain extent with external measures of academic achievement. Results displayed here indicate that

(a) high levels of *attendance* in teacher PD sessions [A-III], *combined professional development factors* [A-VI], and positive *attitudes* regarding the teacher-teaching artist collaborative planning and implementation processes directly predicted ratings of student perception of how *arts integration processes* changed their classroom culture [B-I];

(b) positive teacher PD *performance outcomes* such as exemplary documentation processes, the ability to articulate program goals, and to cite evidence of student learning during facilitated Portfolio Conferences predicted student ability to demonstrate and reflect on their arts integrated learning work during student performance assessments;

(c) high levels of *combined professional development factors* [A-VI] linked positively with *scholastic performance* [B-IV] modestly for classroom teachers in their second year of PAIR, and highly significantly for third year PAIR teachers (the Fourth Grade effect).

These paths of association outlined in the data patterns support he principal hypotheses of this project: that teacher arts integration professional development outcome factors should significantly influence both student arts and academic learning performance over time.

* * *

3G-3: A Synopsis of Inter-related Aspects of PAIR Individual Student Learning Outcome Factors

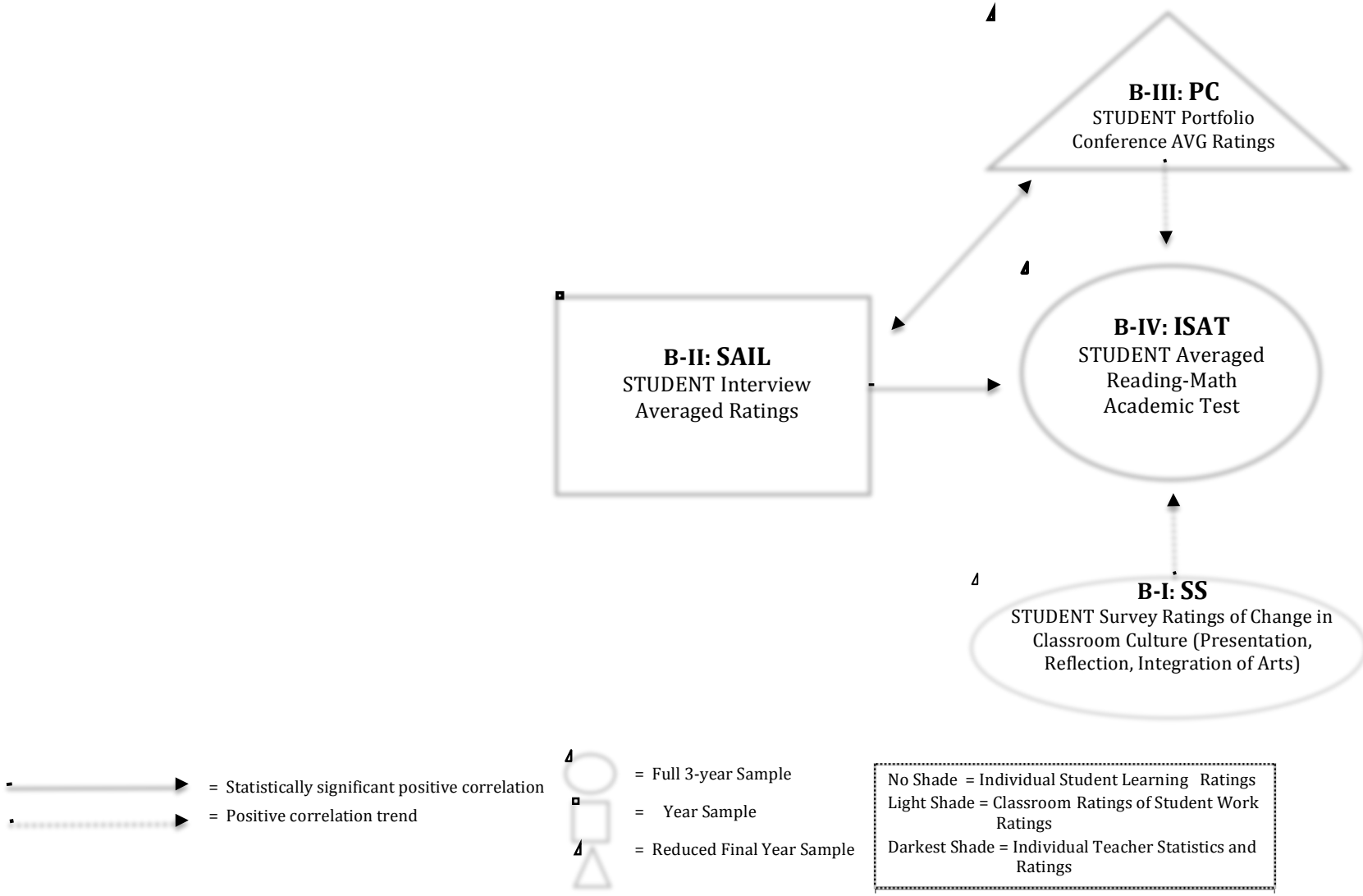
IIIG-3 Figure 3 Data Indications: *PAIR arts integration student learning outcome variables, as a whole, are significantly related to each other and, to a lesser extent, with student academic achievement test performance.* The third summary figure below explores the relationship between PAIR internal (BI-III) and external state measures of academic performance (B-IV). The relationships among the variables are depicted again by solid (highly significant) and dotted (statistical trend) vectors that portray clearly positive relationships among student survey results, two internal performance assessments and one standardized test of academic achievement implemented in the final year of the PAIR project. The vectors in the data display show that:

- Student measures of arts integration learning [BII-III] are very highly correlated, a finding that validates both the reflective (interview) and interactive (portfolio conference) student learning assessments that are focused on student understanding of the goals, practices and impact of both arts and arts integrated teaching and learning.
- Student SAIL arts integration interview ratings, and, to a lesser extent, student portfolio conference ratings are linked with academic achievement.
- Students' perception of change in arts integration classroom practices, as indicated by student survey responses in B-I, are mildly correlated with academic performance [B-IV].

[3G-3 Summary Figure on next page]

3G-3 SUMMARY Figure 3: A Flow-Chart Map of PAIR Inter-Related Student Outcome Factors.

Multivariate correlation map of statistical trends and statistically significant inter-related associations among student learning outcome variables by the final year of the PAIR project.



3G-3 Reference Table 3: Correlation Statistics Among All PAIR Student Outcome Rating Factors

	B-1: **Student Survey (SS) Response Ratings**	B-II: **Student SAIL Interview and Performance Assessment Ratings**	B-III: **Student Portfolio Conference (PC) Interview Response Ratings**	B-IV: **Student ISAT Combined Academic Performance Ratings**
ANOVA F ratio statistic, p Probability Confidence Value; Spearman correlation r statistic, p Probability Confidence Value				
B-1: **Student Survey (SS) Response Ratings**	NA	ns	ns	F ratio=5.0, p< .003; r=.17, p< .07
B-II: **Student SAIL Interview and Performance Assessment Ratings**	–	NA	F ratio=6.6, p< .02; r=.45, p< .01	F ratio=11.5, p< .001; r=.29, p< .0006
B-III: **Student Portfolio Conference (PC) Interview Response Ratings**	–	–	NA	ns
B-IV: **Student ISAT Combined Academic Performance Ratings**	–	–	–	NA

3G-3 Summary, Emerging Themes: These results suggest that a relatively high degree of association exists between all the student variables and the academic learning outcomes.

These data suggest also that the arts integration assessments can be understood as valid, alternative tools for measuring the success of arts integration teaching practices. In section II of this report we learned that the PAIR treatment schools outperformed the control school cohorts both academically and in terms of internal measures of student arts and arts integration understanding. Because there is now statistical evidence of the positive, strong interaction between individual teacher professional development and student learning ratings, we can now draw the conclusion that the PAIR program outcomes influenced the improved and more equitable academic test results demonstrated by the treatment schools in Part 2 of this report. The added significance of the student survey data is that it measures to some extent changes in the classroom culture that has become infused with arts integration practices and, in particular,

with the positive presence of teaching artist-classroom teacher collaboration. In treatment schools, students' perception of a classroom culture of increased collaboration, reflection and the increased presence of arts in the classroom can be linked also with higher achievement in that classroom.

The fact that PAIR student learning outcomes are now linked directly academic learning in the context of extensive teacher training in arts integration practices is significant in itself. A more deeply probing analysis of the pathways of achieving teacher-learning intersections in the follow-up charts will show the extent to which professional development outcomes can function as preliminary links in the path toward academic performance. Going beyond the causal links among treatment school arts integration professional development and student learning outcomes outlined in the chart below, subsequent statistical regression analyses illustrated in follow-up figures methods will be used to sift out which of the principal program outcome variables are the most powerfully associated with both student arts and academic learning outcomes.

* * *

3G-4: A Synopsis of PAIR Student Outcome Regression Factors

As described previously, regression models used for multivariate analysis are based on statistical procedures that gauge the 'goodness of fit' of many different kinds of variables according to the one principal outcome factor. Proceeding stepwise through all variables, regression analysis tests each independent variable for its degree of correlation with the one principal dependent variable in comparison to every other variable entered into the equation. Theoretically, once all the variables are measured for degree of correlation (r), effect size (r^2), and probability of error (p value), the cumulative 'explained variance' of all the variables will be close to 100%.

Var 1 + Var 2 + Var 3 + Var 4 + … = explained variance in one principal outcome variable.

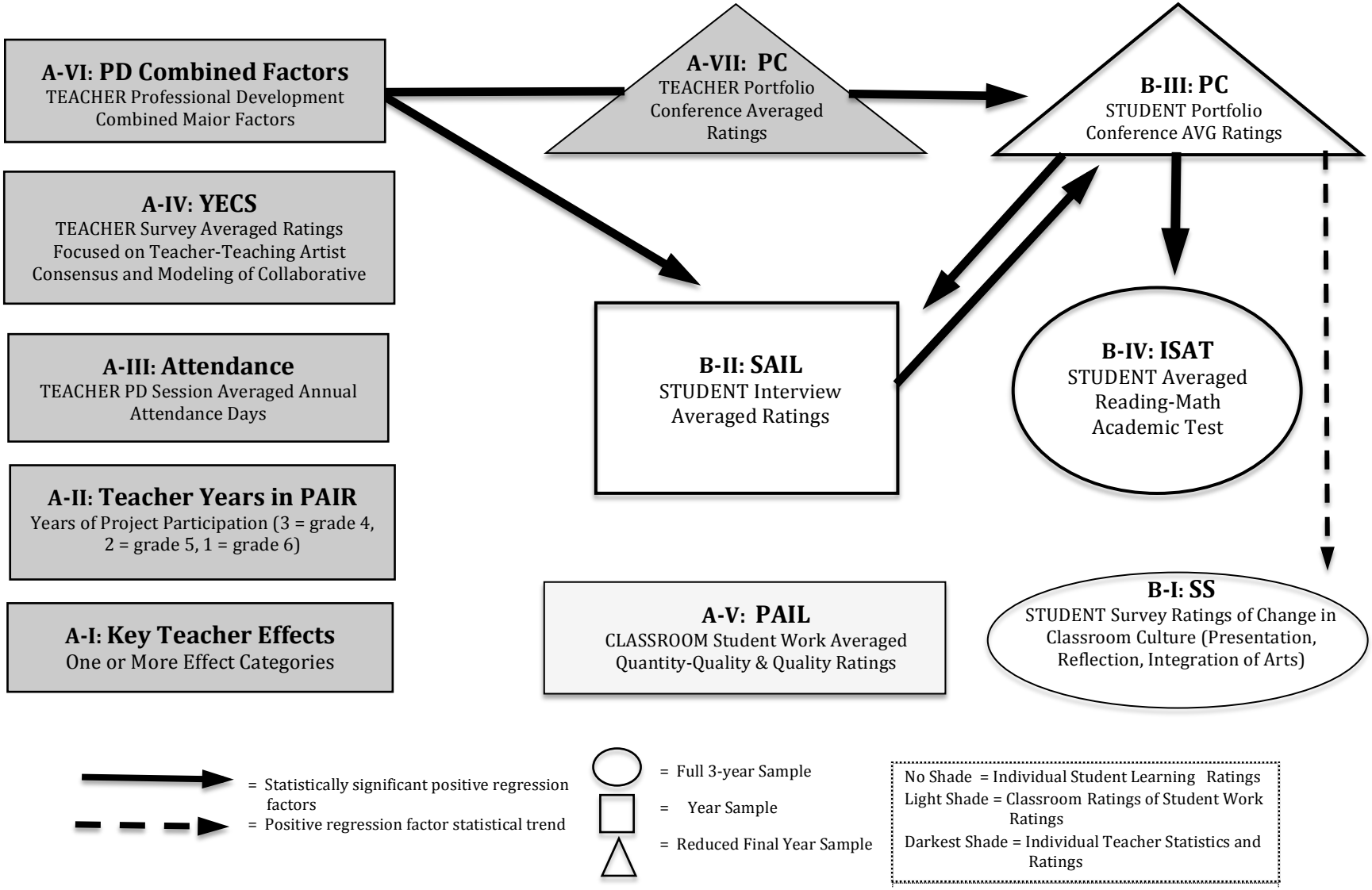
Step by step, regression analysis sifts through the many variables that may be significantly correlated when paired with the principal outcome variable, but do not have the same 'wellness of fit' or predictive power.

Using regression modeling techniques, the general linear model that measured the degree of correlation of all teacher professional development outcome variables, student learning variables, and student demographic factors were fit to each of the four major student learning outcomes of this project: student survey ratings [B-I], arts integration (SAIL) interview ratings [B-II], Student portfolio conference performance (PC) assessments [B-III], and student state standardized academic test scores [B-IV].

3G-4 Data Indications: *Combined teacher PD factors predict student arts integration learning ratings, and student arts integration learning ratings predict student academic achievement.*

The thick solid arrows directed toward the three major student outcome variables in 3G Summary Figure 4 below indicate which of the 'feeder' variables are most significantly predictive of student comprehension of arts integration learning processes and/or academic achievement test results. As indicated previously, with the sole exception of family income in the context of standardized tests, no demographic student factors (income, gender, ethnicity, ELL status, school selection, or pre-designated HAL academic status ratings) were statistically significant in comparison with the PAIR teacher professional development or student outcome factors that appear in the regression map figure below.

3G-4 SUMMARY Figure 4: A Flow-Chart Map of PAIR Teacher Professional Development and Student Regression Factors Most Associated with Targeted Student Learning Outcomes: Multivariate map of statistically significant student learning outcome regression factors linked to the student learning outcomes by the final year of the PAIR project.



3G-4 Reference Table 4: Regression Correlation Analysis Results from All PAIR Professional Development, Student Learning, Student Demographic Factor Intersections

	Four Principal Student Learning Outcome Target Variable			
Regression Factors	B-1: Student Survey (SS) Response Ratings	B-II: Student SAIL Interview and Performance Assessment Ratings	B-III: Student Portfolio Conference (PC) Interview Response Ratings	B-IV: Student ISAT Combined Academic Performance Ratings
Stepwise regression r correlation values, p probability confidence value, $r^{2=}$ explained variance statistic (percent of outcomes directly associated with the factor)				
AI: Teacher Effects (any of 6 categories)	ns	ns	ns	ns
AII: Years of Teacher Participation	ns	ns	ns	ns
AIII: Teacher PD AVG Attendance	ns	ns	ns	ns
AIV: Teacher AVG 3 Key Survey Ratings	ns	ns	ns	ns
AV: Teacher AVG PAIL Student Work Ratings	ns	ns	ns	ns
AVI: Teacher Combined Attend, YECS, PAIL Ratings	ns	**Sig Prob = .002 r^2 = .37**	**Sig Prob = .003 r^2 = .37**	ns
AVII: Teacher PC Ratings	ns	ns	ns	ns
B-1: Student Survey (SS) Response Ratings	ns	ns	ns	ns
B-II: Student SAIL Interview Ratings	ns	ns	**Sig Prob= .04 r^2 = .60**	ns
B-III: Student Portfolio Conference (PC) Interview Ratings	**Sig p = .06 r^2 = .28**	**Sig Prob = .04 r^2 = .60**	ns	**Sig Prob = .09 r^2 =. 35 [Sig Prob = .01 r^2 =. 62]***
B-IV: Student ISAT Combined Academic Performance Ratings	ns	ns	ns	ns
Student Gender	ns	ns	ns	ns
Student Family Ethnicity (Black-Asian differences)	**ns**	ns	ns	ns
Student Family Income Category	ns	ns	ns	ns
Student ELL Status	ns	ns	ns	ns
Student HAL Academic Designation	ns	ns	ns	ns
Designated Teaching Artist	ns	ns	ns	ns
School Arts/Academic Focus	ns	ns	ns	ns
School PAIR Cluster Type	ns	ns	ns	ns

***** these highly significant results are obtained when the negative effect of **income is excluded as a regression analysis factor**

3G-4 Summary, Emerging Themes: The overall regression analysis indicates that, controlling for all student and school demographic elements, (a) the Combined Factors of Teacher Professional Development [A-VI] variable is the single best predictor of student arts integrated learning, (b) the two principal student arts assessment instruments [B II-III] produced highly inter-correlated, statistically reliable measures of learning by the PAIR treatment school student cohorts, and (c) the student performance conference ratings appear to be the single best positive predictor of academic achievement. By including all demographic factors in the regression equation, the regression analysis substantiates the general hypothesis of the PAIR model development project statistically when we can conclude that

- *Controlling for all student demographic factors*, averaged teacher professional development and survey outcome ratings most powerfully predicted student arts integration learning outcomes. The fact that HAL designations did not predict arts integration ratings confirms that high achievement in arts integration and academic assessments was achieved by students previously designated as low-level academic learners;
- *Controlling for all student demographic factors (excepting only the confounding influence of family income)*, student arts integration learning ratings most powerfully predict student academic achieve test outcomes. In this case positive arts integration learning outcomes are more predictive of student academic achievement than gender, ethnicity, prior HAL or ELL classifications.

This statistical distillation of most predictive factors explains in broad terms why arts integration treatment school student cohorts outperformed the control school cohorts in academic test results by the end of the project. By rating levels of both individual teacher training and student learning, the intersection of these factors demonstrate that it is the quality ratings of arts integration professional development factors that predicts best the caliber of student arts learning and its positive association with levels of academic achievement.

Thus, the progression of these summary data maps, backed up by detailed methodological description and data analyses presented in this report, demonstrates that the positive results of control-treatment test score comparisons can be understood more precisely through the predictive power of the analysis of ratings of teacher professional development and student arts

learning factors. In this model, internal measures of PAIR arts integration teacher and student learning define a well calibrated intervention effect that can be shown to be linked significantly with external measures of academic performance.

* * *

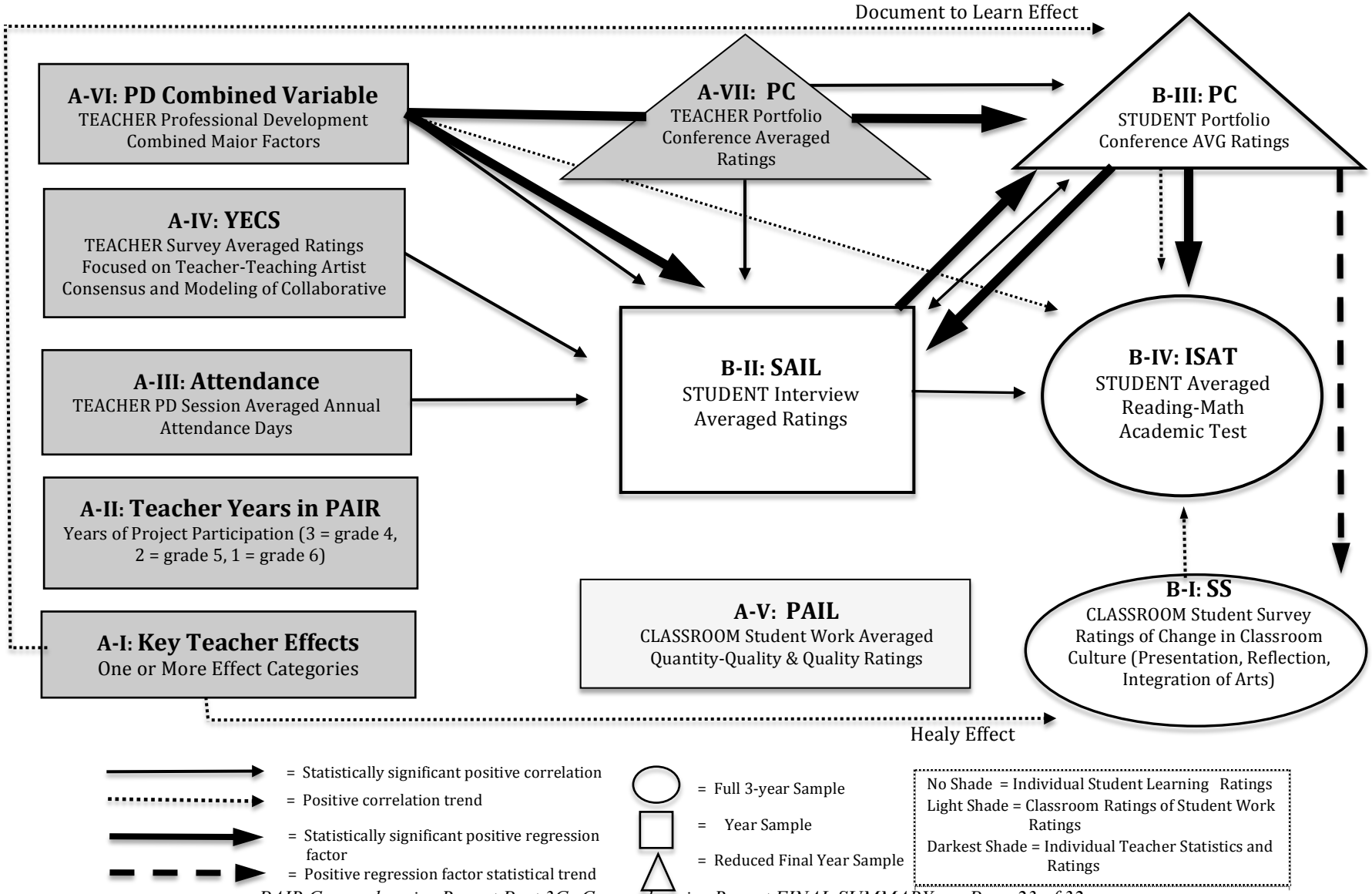
3G-5: A Synopsis of Correlation-Regression Analyses of PAIR Combined Teacher-Student Intersection Factors. This final summary map portrays the entire set of statistically significant correlation and regression factors related to student outcomes in one display.

3G-5 Data Indications: *Looking at both pair-wise correlation and multiple factor regression analyses, it appears that the averaged combined teacher professional development factors best predicts both teacher and student performance outcomes; the student interview and portfolio conference ratings considered together are highly predictive of academic achievement.*

The combination of all principal correlation and regression factors provides a two-dimensional view of the statistical patterns and degree of association directly related to the intersection of teacher and student learning outcomes. Correlations constitute degrees of association that infer either causal (one-way) links or interactions (two-way) between any two variables. The regression vectors represent the result of a hierarchical rank ordering of predictive variables directed toward a single student learning outcome variable (e.g., Teacher PD Combined Factors surfacing as a primary predictor of student SAIL or Portfolio Conference ratings as indicated in the figure below). Taken together the data display below provides two indications of the degree and pattern of association among all teacher-student learning intersections. These intersections suggest that

- Distinctions between mean scores of exemplary and non-exemplary Teacher Effect influenced all measures of teacher professional development, but only mildly predicted Student Survey or Student portfolio conference ratings directly.
- Virtually all teacher professional development outcome ratings are strongly associated with student arts learning outcomes. The Combined Teacher Professional Development Factors is the only variable that strongly predicts student achievement, though most likely occurring only in classrooms whose teachers are either Year 2 or 3 veteran PAIR teachers.
- Student Learning outcomes are highly inter-correlated and predictive of student academic achievement ratings, especially in the case of Student Portfolio Conference ratings.

3G-5 SUMMARY FIGURE 5: A Correlation-Regression Map of PAIR Combined Teacher-Student Outcome Intersection Factors. Multivariate map of the statistical trends and statistically significant intersections among teacher professional development and student learning outcome variables by the final year of the PAIR project.



3G-5 Summary, Emerging Themes. Observations from this chart are organized from four perspectives:

1. Teacher professional development correlation-regression factors:
    a) In general, the exemplary Teacher Effect designations predict the qualities of the PAIR professional training program but only mildly with student learning outcomes.
    b) The Teacher PD Composite Factors variable suggests that the whole effect of PAIR teacher professional development outcomes is greater than the sum of its parts. Unlike other more narrowly construed teacher training variables, the more comprehensive professional development Combined PD Factors variable [A-VI] and the Teacher Portfolio Conference Performance Ratings [A-VII] are significantly linked with both art integration and to a lesser extent academic student learning.
2. Classroom-Rated correlation factors:
    a) The correlation between the quality ratings of classroom student work samples [A-V] does not quite meet the criteria for strong statistical significance suggesting that *individual ratings*, and not classroom averaged ratings, of student work samples might provide a better determination of the link of arts integration work to other student learning outcomes in future studies.
3. Student learning correlation and regression factors:
    a) The extraordinary strength of association between student portfolio conference and interview assessments [BII-III] demonstrates the validity of performance assessments of the treatment school students. However correlated these student learning assessments are, it appears that their degrees of association with academic differ. These differences may be due to differences in sample size. Nonetheless, the portfolio conference student ratings appear to be the most penetrating predictor of academic achievement, suggesting that the highly interactive and multimodal nature of this form of assessment may be more relevant to combining factors of student work with reflection on learning processes. That is, the interactive performance tasks in the context of demonstration and discussion of portfolio work samples and the presence

of peer feedback may prove to be a more valid and powerful link between arts integration and academic learning assessments.

4. Student Survey Response correlation-regression factors:

   a) As discussed previously, the student surveys may be key factor for linking student perception of classroom presence of arts integrated teaching and learning practices with academic achievement resulting from these practices. Otherwise, the lack of correlation between the student survey results with teacher professional development outcomes is surprising and points to a potential weakness in the project. Part I of the report did not report overwhelming differences between control and treatment teacher survey responses, yet there may yet be better ways of drawing out more powerful distinctions between arts, academic, and arts integration classroom culture practices in the minds of students in future studies.

*** * ***

3G-6 Final PAIR Whole Report Conclusions, Implications for Future Research

Documenting and measuring the intersections between teacher professional development and student learning outcomes became the focus of Part 3 of the PAIR report, commissioned by CAPE to augment the PAIR AEMDD program evaluation report submitted in March of 2011.

Part 3 of this report is needed to fulfill CAPE's mission to take seriously its role as a leading research-based arts learning organization funded by the federal government to provide extensive data analysis relevant to the impact of high quality arts integration teaching and learning.

Leadership in arts integration program development requires the determination by CAPE staff, teaching artists, community partners, and researchers to seek out and solve the complexities of assessing interdisciplinary teaching and learning in public schools. The content of this report resulted from the opportunity to use extensive multivariate analysis to probe the intersections between teacher and student learning in order to provide evidence for the impact of teacher professional development outcomes on measures of student understanding and success with arts integration learning.

This concluding section of the report is framed by teachers' subjective comments on the impact of PAIR on student learning reported in Part 3: D & E, testimony that has been since confirmed by the statistical evidence of student learning presented in the later sections of the report.

Analysis of teacher-student learning intersections in relation to teacher interview data

From the viewpoint of the researchers the validity of this report required a thorough investigation of the hunches and beliefs that teachers expressed about the potential causal links between the quality of students arts integration work and higher level thinking skills that should lead to better performance on standardized academic tests.

> **Teacher** "I find that it [PAIR] appeals to different intelligences. They [students] need to get up and move and do things with their hands. It's not just visual and listening skills. I think it helped them to develop a better understanding of the material. … I think all the students were engaged. I think they were very excited about learning. It was kind of fun to be in the classroom. And I think it...I'm *hoping* it helped develop higher level thinking skills."

The quotation above captures one of the many transformational aspects of reported anecdotally by PAIR teachers during their portfolio conference interviews (as reported in Part I of the comprehensive report). Although the newly adopted PAIR arts-based collaborative teaching practices proved to be productive and highly engaging for students according to teacher observations, the eventual statistical analysis of the intersection data lends far more credibility to the assertion that PAIR classroom units helped students develop higher level thinking skills. The teacher-student intersection data (reported in Part 3: E, F, G} suggests that teachers who engaged productively in the PAIR professional development sessions and collaborative arts integration classroom practices were most likely to articulate the increased presence of multiple intelligence based creative and critical thinking learning processes as an alternative to their past reliance on pencil and paper activities and tests.

The facilitated teacher-student portfolio conferences provided an opportunity for teachers to observe and reflect on the impact of their PAIR arts integration practices on their students' cognitive and social-personal development. As reported in the qualitative analysis sections of this report based on portfolio conference transcripts (Part 3D), teachers talked about how their dance, music, drama, or visual arts integration units encouraged student to make their own connections with learning in math, cultural studies and literature. They felt that students were more likely both (a) to find their own 'voice' in arts integration processes and (b) to reflect on a wide range subject area content objectives shared between the 'paired' academic and arts disciplines:

> "What we were looking to do on all writings, regardless of whether it's ISAT or what we're using it for, we always want the student's own voice to come through and not just mimic what it is that they've read, but we want their input, their take on things. So this [PAIR] was helping them develop confidence in their writing so that their own true personality and their voice would come through. When they have to make connections, it's harder for them to make a connection if it's not their own voice and their own real experiences."

Since PAIR provided new ways for teachers to include student input into their work and make more connections across disciplines, teachers felt that arts integration learning opportunities were more equitably distributed among their disadvantaged students. Teachers who observed ELL learners in their portfolio conference sessions, for example, noted that arts integrated teaching

resulted in increasingly sophisticated conceptual and reflective thinking skills that appeared to aid the development of their students' English proficient writing skills:

> **Teacher**: " [My students] are English Language Learners, so during this year I didn't see, or didn't get the growth in the writing as I expected to see in my previous classes, especially the first year with the project in my English Only classroom. However, their confidence and ability to speak to others using the English language has grown significantly, and I think it's because of the fact that they can express their knowledge through other means rather than by paper and pencil. … when they have the opportunity to say it to you, they know that they can express their ideas, what they know and what they've learned…."

Furthermore, the classroom teachers' portfolio conference experiences revealed how previously low or average academically rated students were now able to reflect on their learning in ways that were not noticeably different from their higher rated students.

> **Teacher:** "I think it's [the portfolio conference] a great thing because, I mean, it brings everybody [to] the same level. … And I think that's important because you're not going to [limit expectations for] the low group or you're the high group. Who knows? You couldn't tell from that [what we observed in today's session]."
>
> **Interviewer:** You think people watching the tape and rating student responses would know that they're low, medium and high [previously rated academic students]?
>
> **Teacher**: "I don't think so. That's why I'm very glad, because all three were answering, all three had something to say… They were really putting things together with everything else. … I see that they're integrated, and you can't really tell the difference. … I don't think if you just looked at them you could say like, oh, well, this one really struggles with math and this one [doesn't].

The value of PAIR comprehensive data collection and multivariate analysis methods

Although PAIR teachers provided anecdotal evidence of change in their classroom practices and in student learning outcomes, rigorous qualitative and statistical data analyses were required to confirm the project's positive impact CPS teachers and students.

Control-treatment comparisons provided evidence of positive differences for PAIR on both measures of teacher professional development surveys that defined contrasts in teaching

practices (Part I) and student learning outcomes that demonstrated different levels and distributions of understanding of arts learning and academic performance (Parts II and III).

Control-treatment comparisons resulted in strong statistical evidence of PAIR's positive effect on standard measures of school academic performance. *Within-treatment school analyses* of the level and quality of a much wider range of teacher professional development factors provided an opportunity to statistically determine the degree and patterns of association among these factors and their intersection with student learning outcomes. Treatment school analyses provided evidence of the cohesive impact of PAIR individual teacher outcomes the individual student arts integration learning outcomes, and, controlling for virtually all demographic factors, its connection to individual student levels of academic performance.

The main sections of the report provided three lenses for distilling teacher and learning effects. The mixed methods analysis in Part 1 of this report suggests that teacher training resulted in significant, yet categorically different effects on the teachers. Assessment of teacher commitment to applied content knowledge expertise, production and analysis of student work documentation, and in-depth commitment to the development of collaborative arts integration practices made it possible sort out exemplary classroom teachers who demonstrated these traits from those who did not. Control-treatment comparisons and within-treatment school ratings indicate that the exemplary teacher effects became clearly identifiable and statistically significant predictive factors of student learning by the final year of the project.

Part 2 control-treatment comparisons of averaged state academic test results revealed significant positive differences between (a) arts cluster magnet schools compared to other Chicago Public schools, (b) PAIR treatment arts-academic compared to control arts-academic magnet schools, (c) primarily arts focused magnet schools compared to primarily academic focused magnet schools, and (d) world language cluster schools compared to math and writing cluster schools. In addition, the tracking of initially categorized high, average, and low academically rated students at third grade reveals that PAIR schools demonstrated a far more equitable distribution of improved academic performance than did the control schools by sixth grade.

Part 3 of the report provided *within treatment school* in-depth statistical analysis of the internal measures of individual student work samples, student performance assessments ratings, and tests

of the predictive relationship between rank-ordered teacher professional development ratings and student learning outcomes.

Overall, these data provided statistical evidence of the positive impact of PAIR on student learning over time. Each newly developed student learning measure was examined separately for its contribution to understanding the impact of PAIR on individual measures of student performance over time. Detailed quantitative analysis of student outcome ratings and extensive qualitative analysis of teacher and student performance during the portfolio conferences provided insight as to why the control-treatment school standardized tests comparisons revealed positive differences in academic performance and indications of increased equity for the PAIR treatment schools.

The final summary figures [3G Summary Figures 1-5] revealed strong evidence of positive relationships among teacher-student intersections based on the degree and pattern of inter-correlation among variables. The hierarchical rating of the most significant correlations through stepwise regression factor analysis focused on determining which teacher professional development variables were most predictive of the major student learning outcome variables.

The multivariate analytic methods showed that the teacher professional development and student learning outcomes could be mapped separately or as intersections according to the statistical measures of association within and across the teacher-student outcome rating variables. The overall degree and pattern of statistically significant correlation among multiple teacher professional development factors provided conclusive evidence of treatment school program *cohesiveness* over time. Similarly, overall degree and pattern of statistically significant correlation among student learning outcome factors provided conclusive evidence of treatment school program *direct impact* on two different types of arts integration assessments.

By examining the relationship of seven teacher professional development outcome ratings to four student survey and learning outcomes, evidence of the comprehensive impact of PAIR project emerged as a *chain of causal factors* that supported many aspects of the classroom teacher perceptions and reflections on the transformational impact of PAIR asserted in interview data. Tests of strength of association between teacher and student rating variables demonstrated that

professional development outcomes significantly predict student performance on two complementary measures of arts integration learning (SAIL Interview and Student Portfolio Conference ratings), controlling for all student demographic factors such as gender, ethnicity, ELL status, prior academic designation, size of school, type of magnet school, etc.). *Controlling for all demographic factors except for the confounding and surprisingly lack of influence by family income factors, that final 'through-line analysis' indicated that high levels of student arts integration learning associated with positive ratings of individual teacher professional development outcomes, predicted better performance on academic performance.*

Implications for future research

The PAIR model development project tested for the impact of high-quality classroom teacher-teaching artist led arts integration units in Chicago Public Schools. The creation of new individual rating systems for PAIR classroom teacher effectiveness and the development of new instruments for assessing meta-cognitively rich aspects of student arts integration learning outcomes made possible the extensive multivariate analysis of teacher and student intersections. In effect, the mapping of statistically significant assessment ratings of both individual teacher and individual student survey, interview, and performance assessment outcomes constituted a comprehensive window onto the impact of PAIR interventions in public schools dedicated to building connections between arts and academic learning practices as a tool for school excellence and equity.

Now that this groundbreaking study has (a) provided a validated model for measuring the quality of arts integration program and quality of teacher professional development outcomes, (b) devised practical, reliable ways of rating the impact of PAIR teaching practices on individual students' arts integration learning, *and* (c) established a statistically significant predictive association between arts integration learning and more equitable and diverse pathways to academic improvement, *policy makers now can responsibly expect – or even require - schools to provide arts integration teaching and learning practices as a central strategy for promoting integrative learning that benefits the comprehensive education of every public elementary school student.*

Thus, future researchers needs no longer ask *if,* but *to what extent* qualitatively validated and quantifiably evaluated measures of arts integration program development can predict higher and more equitable indications of student learning. Schools that choose to create high quality arts integration programs based the principles, tools and practices of the PAIR reported here will now be able to rigorously assess for themselves the impact of creative, multiple arts integration teaching and learning practices on multiple aspects of school performance. In future education policy discussions, results from teacher professional training and student outcomes ratings systems reported here can provide a basis for building whole school, district or state-wide accountability systems that will be necessary for ensuring the responsible implementation of sustainable high quality arts integrated teaching, learning, and assessment practices in American public elementary schools.

* * *