

Chicago Arts Partnerships in Education's

Portfolio Development Project

Principal Investigator's Report

By

Dr. Lawrence Scripp, Principal Investigator

With

Sarah Sutherland, Research Associate

Josh Gilbert, Research Assistant

Center for Music and the Arts in Education, Inc.

Submitted to the Federal Department of Education's Arts Education Model
Development and Dissemination (AEMDD) Program

May 8, 2015

Table of Contents

1. Introduction to the CAPE Portfolio Development Project Principal Investigator's Report	3
2. Data Analysis Methodology: Establishing the Basis for Meaningful Control-Treatment School Arts/Arts Integration Learning and Academic Performance Comparisons	5
3. Five-Phase Analysis of Student ISAT Academic Test Data	7
4. Investigating Three Measures of Arts and Arts Integration Learning	12
5. The Examination of Treatment School Teacher PD and Performance Variables and Their Links to Student Arts and Academic Learning Outcomes During the Final Year of the Project	20
6. Linking the Chain of Evidence I: Direct Pairwise Correlations Between Teacher PD and Student Academic Performance Outcomes	23
7. Linking the Chain of Evidence II: Direct Pairwise Correlations Between Teacher PD and Student Arts Learning Performance Outcomes	26
8. Linking the Chain of Evidence III: Direct Pairwise Correlations Between Student Arts Learning and Academic Performance Outcomes	29
9. Determining the Strongest Links: Stepwise Regression Testing for Most Significant Teacher, Student, and Family Demographic Predictors of Academic Achievement	30
10. Conclusions	38

1. Introduction to the CAPE Portfolio Development Project Principal Investigator's Report

The CAPE Portfolio Development (PDP) Project was designed as an important experiment in the professional development of grades 4-6 visual arts and music arts specialists in Chicago Public Schools (CPS). While it has been shown in previous AEMDD projects that CPS arts cluster schools in general—and arts integration teaching artist residencies in particular—enhance academic performance (Scripp & Paradis, 2014; see pairresults.org for details), this project hypothesizes that incorporating high quality arts plus arts integration portfolios into arts specialist teaching and assessment practices will further optimize the impact of arts learning on academic achievement.

Thus, the research and evaluation question investigated in this project is

To what extent did the development of arts and arts integration classroom portfolio systems — guided by veteran CAPE teaching artists in visual and musical arts — enhance both arts learning *and* the impact of arts learning on academic performance in high minority, low economic status schools?

The investigation of this question will be based on the analysis of the multiple factors that together will represent a possible 'chain of evidence' needed to identify causal links between high quality teacher professional development and student learning outcomes. There are four main data links in this sequential chain as depicted in the Table 1:

Table 1: Multivariate Outcomes “Chain of Evidence” Analytic Framework

I. Arts Teacher Preparation PD Outcome Variables →	II. Arts Teacher Performance Outcome Variables →	III. Student Arts Learning Outcome Variables →	IV. Student Academic Performance Outcome Variables
IA. Arts Teacher PD Attendance	IIA. Arts Teacher Quantity of Student Portfolio Work	IIIA. Student Quality of Portfolio Work Ratings	IVA. Student Final Year Combined Academic Performance Test Score
IB. Arts Teacher PD Reflection/Self-Assessment Survey	IIB. Arts Teacher Classroom Observation Ratings	IIB. Student Portfolio Conferences Performance Assessment Ratings	IVB. Student Baseline to Final Year Combined Academic Performance Test Score
IC. Arts Teacher Self-Esteem/Confidence from PD Exit Survey	IIC. Arts Teacher Portfolio Conference Performance Assessment	IIIC. Student Performance Assessment Interview Ratings	
ID. Arts Teacher Combined PD Outcome Variable			

From left to right, these four columns represent a complex sequence of interrelated factors that may or may not ultimately influence student academic achievement. Taken as a whole, this model represents the various links in a chain of evidence that could predict:

- In what ways teacher PD factors (column 1) could influence teacher performance outcomes (column 2), student arts learning outcomes (column 3), and/or student academic outcomes (column 4).
- In what ways teacher performance outcomes (column 2) could influence student arts learning outcomes (column 3) and/or student academic outcomes (column 4).
- In what ways student arts learning outcomes (column 3) could influence student academic outcomes (column 4).

Another possibility is that each column of factors may only affect the adjacent column, suggesting a chain of factors that only predicts quality in the next step of the teacher-student learning sequence: teacher PD factors (column 1) could influence teacher performance outcomes (column 2), teacher performance outcomes could influence student arts learning outcomes (column 3), student arts learning outcomes (column 3) could influence student academic learning outcomes (column 4).

Then again, results may also prove that some teacher PD factors (column 1) predict student performance in the arts (column 3) and/or student academic outcomes (column 4).

Findings from this report will provide statistically significant evidence that, over the three years of project implementation, teacher PD outcomes influenced student arts and arts integration outcomes, and academic learning outcomes substantially. First, students of arts specialists—highly rated for their arts plus arts integration portfolio practices in collaboration with teaching artists in treatment ¹schools—gradually outpaced student academic and arts learning outcomes in control schools over time. While the measure of academic improvement was incremental from year to year, the overall positive pattern of academic improvement is unmistakable by the end of the project. Furthermore, although several factors influenced student learning, stepwise regression techniques revealed that PDP teacher participation in professional development and positive assessment of their PD experiences in particular predicted student Illinois Student Achievement Test (ISAT) scores when comparing baseline to final year results. By the final year of the project, it was the quality of arts plus arts integration student portfolio work along with teacher positive attitudes about PDP practices that are more deeply linked with academic achievement compared to than any other student learning or demographic factor other than “student family income” in the treatment schools.

Figure 13 presented at the end of this report delineates all significant relationships among the variables listed in Table 1 just discussed². The following sections of the report detail the methods by which conclusions about causal links between the various teacher and student learning outcome variables were drawn both in Table 1 and in Figure 13.

* * *

¹ N.B. The word “treatment” or “control” will not be capitalized except when referring titled as heading or when referring to a specific variable.

² The reader is strongly advised to refer to this figure throughout the report when multivariate analysis is being discussed in detail.

2. Data Analysis Methodology: Establishing the Basis for Meaningful Control-Treatment School Arts/Arts Integration Learning and Academic Performance Comparisons

The first step in validating the analytic methods is to understand to what extent the control and treatment school grade level longitudinal cohort data are equivalent, proportional, and can be fairly compared. The following data displays provide a measure of equivalency between the treatment and control school student population random sampled cohorts.

Accommodating disproportionate student cohort populations in control and treatment longitudinal data comparisons

By design, the number of control and treatment school student cohorts was exactly equivalent. However, by the second year of the project it was clear that two control schools would no longer participate in the project, thus making the data set disproportionate. Faced with the prospect of asymmetrical data sets, the investigators decided that, because (a) the reduced control school sample still had sufficient statistical power for determining its relationship to the variables shared between the two data sets, and (b) the treatment school sample would need to remain large in order to analyze “within-group” comparisons with respect to data only collected in the PDP schools, therefore (c) that analyzing a disproportional number of students in each cohort—though not ideal—was the best strategy for determining factors in the treatment school data that could account for differences between the control-treatment school comparisons.

Figure 1: Disproportional Number of Control and Treatment Schools Students

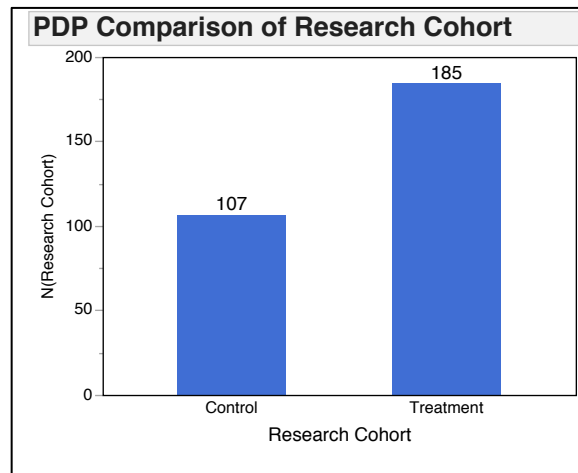


Figure 1 shows that there are 78 fewer control school students than treatment school students in the longitudinal cohort.

Comparable Demographic Factors

Although CAPE was not able to maintain equal numbers of treatment and control school students throughout the experiment, the profiles of four out of five student demographic factors in both longitudinal cohorts displayed in Table 2 were functionally equivalent.

Table 2: Control vs. Treatment School Demographic Data Percentages

	Control	Treatment
Gender	54.5% Female 45.5% Male	57.9% Female 42.1% Male
Free/Reduced Lunch (family income)	89.8% Free/Reduced 10.2% No	92.8% Free/Reduced 7.2% No
IEP Services	85.9% No 14.1% Yes	86.3% No 13.7% Yes
ELL Status	96.0% No 4.0% Yes	95.2% No 4.8% Yes
Ethnicity	45.5% Black, Non-Hispanic 52.5% Hispanic 2.0% Other	58.1% Black, Non-Hispanic 40.7% Hispanic 1.2% Other

The fifth demographic factor, ethnicity, though not equivalent, reveals that both control and treatment schools have comparable percentages of minority population students, though control school cohorts contain slightly more black students and treatment school cohorts contain comparably more Hispanic students. Because there are virtually no white students in either cohort, this project brings a particular focus onto the effect of arts and arts integration portfolios on minority students in Chicago.

It was the judgment of the researchers that despite the unequal number of total students in each cohort, the unusually high degree of equally distributed demographic factors between the two longitudinal groups provided the basis for a fair comparison.

Accounting for prior levels of student academic achievement

Because initial academic performance significantly predicts future academic performance, the longitudinal samples were randomly selected from three levels of baseline academic data collected before the PDP project began.

Sorting the longitudinal cohorts according to academic status prior to the beginning of the project also provided a precise metric for determining a degree of equivalency between the sample student cohorts. In order to achieve balanced randomly selected student cohorts, all students were classified as High (H), Average (A), or Low (L) academic achievers before the beginning of the PDP program.

A balanced tertile distribution within the normal distribution plot of the 2010-2011 ISAT Combined Average Scores was used to determine the categorical boundaries for each of the three HAL cohorts:

- H $x \geq 215$
- A $195 < x < 215$
- L $x \leq 195$

These cutoffs resulted in an identical distribution of the combined treatment and control group students chosen for the study:

- H 86
- A 80
- L 86

By virtue of this process, the final averaged academic scores for the each level of the HAL cohort in both control and treatment schools produced virtually indistinguishable 2010-2011 ISAT Combined Averaged Scores.

Table 3: Comparison of baseline ISAT scores according to Pre-designated Control- HAL cohorts

Pre Project Designation	Complete data set	Control	Treatment
High (H)	230.65	230.68	230.62
Average (A)	205.67	205.60	205.70
Low (L)	176.40	176.63	176.27

Summary Point 1: Though the PDP control and treatment student cohorts were asymmetrical in number due to the withdrawal of control schools from the project, the student demographic factors were commensurate with regard to gender, family income, ethnicity, ELL status, IEP services, and in terms of the distribution of High, Average, and Low (HAL) academically rated students randomly selected at the beginning of the project.

* * *

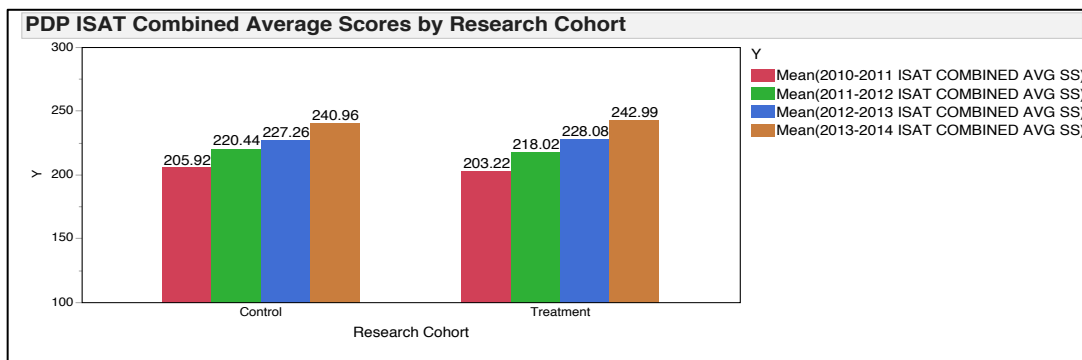
3. Five-Phase Analysis of Student ISAT Academic Test Data

Five analytic frameworks focused on academic achievement during the three years of project implementation determined that the PDP treatment schools gradually outperformed the control schools and by the final year of the project this pattern of improvement became statistically significant.

3a. Phase 1 Overall Control-Treatment (C-T) Student ISAT Test Score Comparisons

Comparing Control and Treatment school cohort academic performance serves as a first step in measuring the efficacy of PDP project. As shown in Figure 2 below, Math and Reading test scores for the Illinois Standard Achievement Test (ISAT) reveal that, spanning the years of the project implementation (baseline to third year of implementation), both control and treatment school student cohorts improved incrementally each year. From the viewpoint of each annual report, the treatment schools scores were never significantly higher than the control schools.

Figure 2: Control-Treatment School Cohort ISAT Test Score Comparisons from Baseline to Final Year of the PDP program



3b. Phase 2 C-T School "Gain score" Analysis of ISAT Scores

The gain score analysis, however, provides evidence for the significant difference between the two student cohorts. Looking more closely at the pattern of test score results longitudinally, Table 4 reveals that the treatment school cohort mean score starts out below the mean score of the control schools,³ yet as the program proceeded, the treatment cohort mean scores incrementally met and then surpassed the mean scores of the control schools by the third year of PDP. In Table 4 we see the year-by-year data previously displayed in bar chart format, with added information regarding the gain scores in column 5.

Table 4: Year-by-Year C-T School Mean Score Differences in Student ISAT Scores

	Control	Treatment	Mean Difference	Treatment School Gain score	t Prob
Baseline 2010-2011 ISAT Combined Average Means	205.92	203.22	-2.7083	—	Prob > t = 0.4128
2011-2012 ISAT Combined Average Means	220.44	218.02	-2.4226	+0.2857	Prob > t = 0.4306
2012-2013 ISAT Combined Average Means	227.26	228.08	0.8224	+3.3998	Prob > t = 0.7682
2013-2014 ISAT Combined Average Means	240.96	242.99	2.0252	+1.2028	Prob > t = 0.4909

^t= positive trend; * = significant (p value <.05); ** = very significant (p value < .01)

While none of the contiguous year mean scores are significantly different in Table 4, the fifth column data analyzed in Table 5 below shows that the average difference in the change in gain scores between the ISAT scores from 2010-2011 (baseline) to the 3rd year of implementation in 2013-2014 is statistically significant.

Table 5: Averaged Individual Student Gain Scores Between Baseline and Final Year Program Implementation

	Control	Treatment	Mean Difference	t Prob
Difference between Baseline 2010-2011 to Final Year 2013-2014 ISAT Combined Average Means Delta	35.30	39.68	4.376	Prob > t = 0.0408*

^t= positive trend; * = significant (p value <.05); ** = very significant (p value < .01)

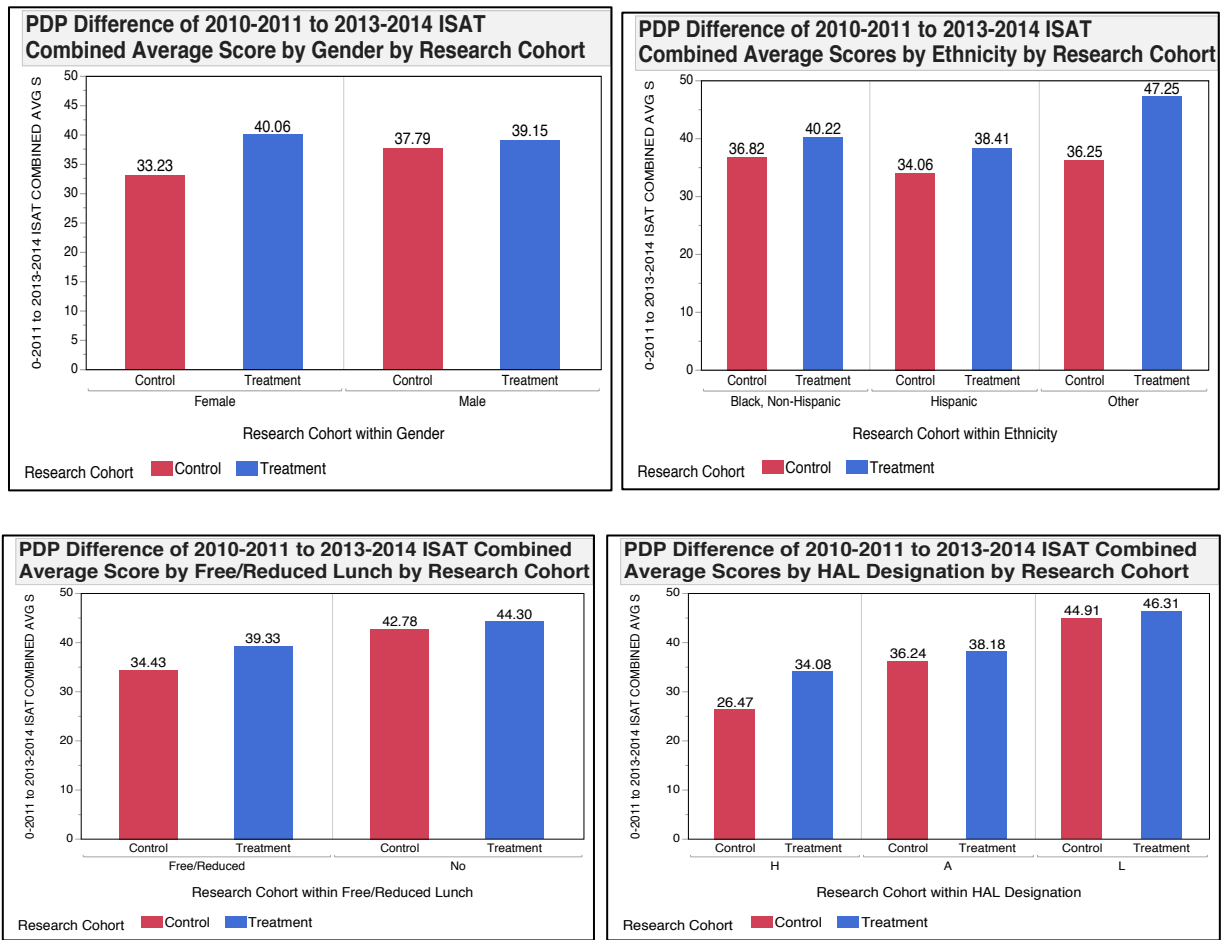
³ N.B. Although the HAL cohorts were matched by dividing the student populations scores into three equal parts, the overall Control-Treatment (C-T) baseline ISAT scores for the two cohorts were not identical.

Summary Point 2: There are positive, statistically significant differences in C-T ISAT test gain scores that indicate PDP Treatment Schools as a whole outperformed Control School cohorts when comparing baseline and final year data.

3c. Phase 3 ISAT C-T School Gain score Comparisons According to Demographic Factors

The gradual emergence of statistically significant differences between the control and treatment schools is heightened further by looking into the pattern of gain scores among the student demographic factors. The examination of baseline to final year ISAT scores in Figure 3 shows that treatment schools outperform the control schools from the viewpoint of Gender, Ethnicity, Free/Reduced Lunch (family income) and previous HAL (academic history) classification, suggesting that the gains in the Treatment Schools apply to virtually the whole spectrum of students (all “blue bar” treatment schools are higher than the “red bar” control school average gain scores).

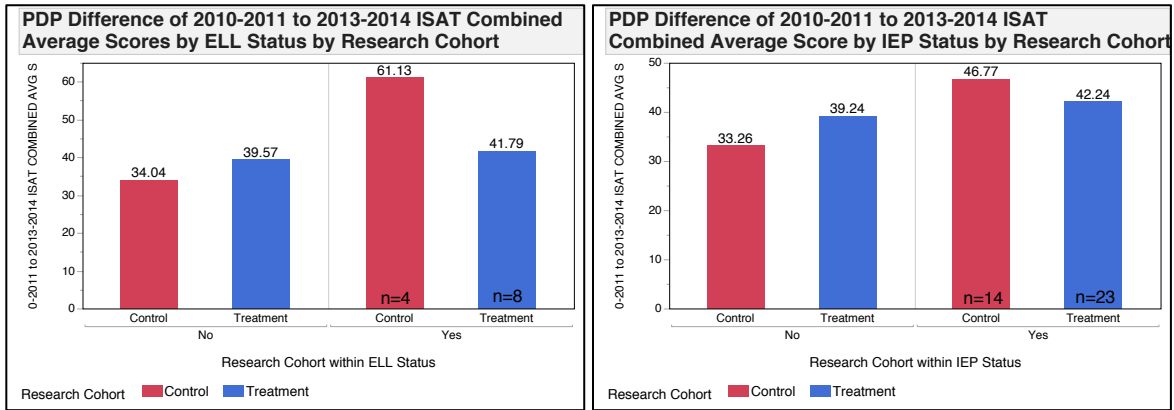
Figure 3: ISAT Baseline-Final Year Comparisons by Gender, Ethnicity, Family Income, HAL Levels of Prior Academic Achievement



It should be noted, however, that ISAT gain scores for two small sample demographic cohorts—English Language Learners (ELL status) and Individual Education Plan (IEP served) students—favor the Control Schools. Perhaps because language or learning

challenged students may not have had equal access to the portfolio process, it appears that a small number of language or learning challenged learners in treatment schools do not benefit from the PDP program in the same way all other student demographic populations do. Yet, because of the small sample size of these demographic categories (right columns in both data displays), conclusive measures of statistical significance cannot be determined.

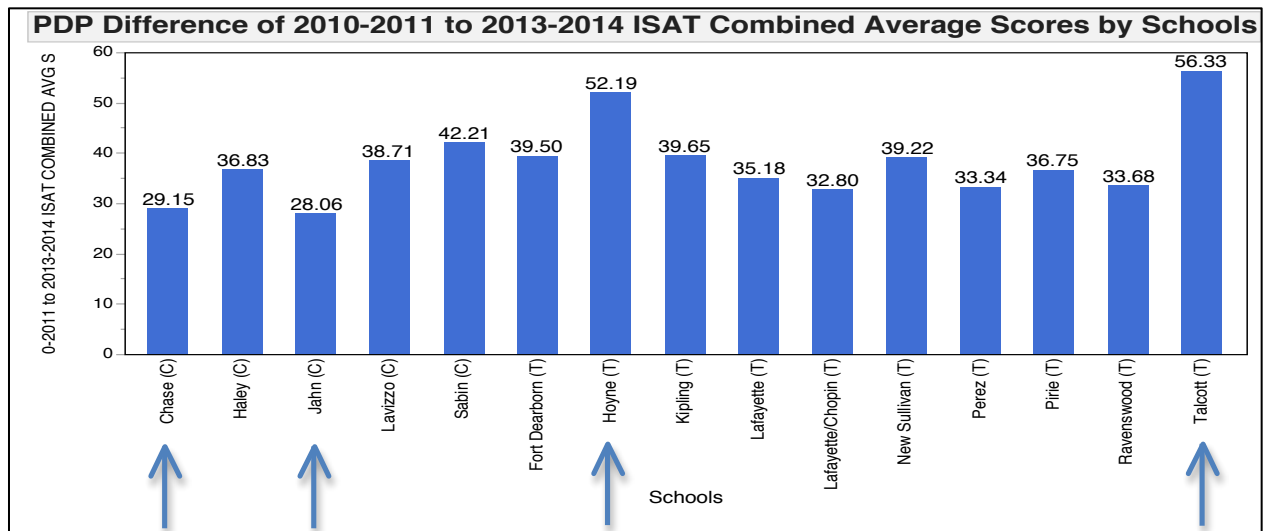
Figure 4: C-T ISAT Comparisons According to ELL and IEP Student Classifications



3d. Phase 4 Individual School Outlier Analysis

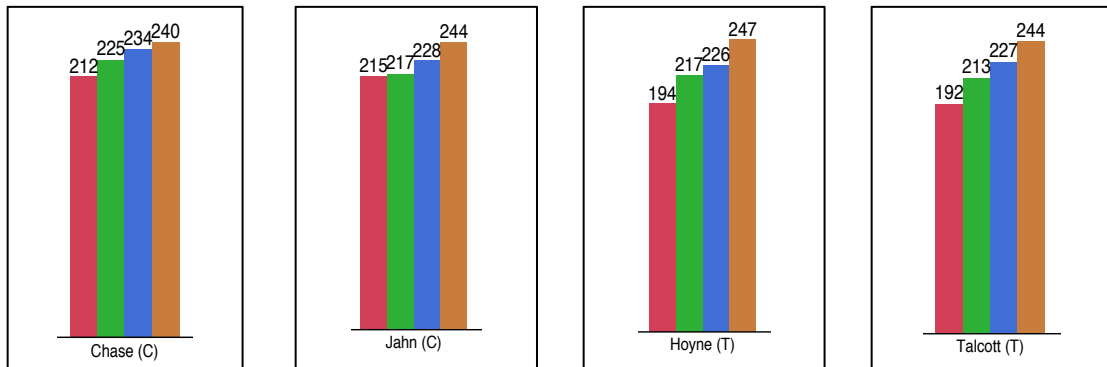
The pattern of gain scores by the separated control and treatment schools provides additional evidence indicating that the PDP project has a relatively uniform positive effect on the Treatment Schools. Figure 5 below reveals that four schools show distinctly different degrees of gain score changes over time. It appears that the control schools, labeled by (C), have two outlier “low increase schools”—Chase and Jahn—and the Treatment schools, labeled by (T), have two outlier “high increase schools”—Hoyne and Talcott. The individual school outliers indicate unusual improvement in two of the Treatment Schools, and the unusual lack of improvement of two of the Control Schools.

Figure 5: Separate School Control (C) – Treatment (T) ISAT Gain Score Comparison



Note in Figure 6 that the “low increase” control schools started with higher averaged ISAT scores and finished around the other schools’ scores (Chase 212 to 240, Jahn 215 to 244), while the “high increase” treatment schools started with lower average ISAT scores and increased the greater distance to finish at or above most the other schools’ scores (See C-T Figure 6: Outlier School ISAT Profiles (Hoyne 194 to 247; Talcott 192 to 244).⁴

Figure 6: Outlier School ISAT Profiles



3e. Phase 5 ISAT Meets/Exceeds (MEX) CPS District Benchmark Analysis

Revealing patterns of C-T school differences similar to the ISAT score analysis, Table 6 demonstrates incremental increases in the percentage of MEXs of the treatment over the control school cohorts by the second year of PDP project implementation.⁵

Table 6: Annual Differences in Percent Students Who Meet or Exceed (MEX) CPS ISAT Benchmarks

	Academic warning / Below	Meets/Exceeds
Baseline 2010-2011 ISAT MEX	C: 30/92=32.6% T: 50/160=31.2%	C: 62/92=67.4% T: 110/160=68.8%
2011-2012 ISAT MEX	C: 30/95=31.6% T: 49/154=31.8%	C: 65/95=68.4% T: 105/154=68.2%
2012-2013 ISAT MEX	C: 67/96=69.8% T: 99/161=61.5%	C: 29/96=30.2% T: 62/161=38.5%
2013-2014 ISAT MEX	C: 55/91=60.4% T: 84/151=55.6%	C: 36/91=39.6% T: 67/151=44.4%

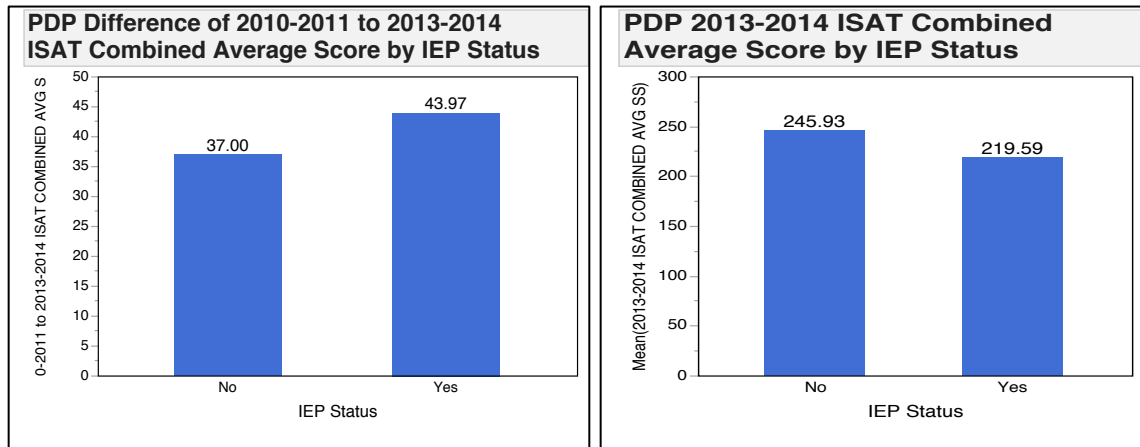
⁴ See Appendix Figure I.1 for the matched pairs analysis of the difference between 2010-2011 to 2013-2014 ISAT combined average scores by each longitudinal cohort.

⁵ Note that the overall lowering of the percentage of MEX students between year 1 and year 2 of PDP is due to changes in calibration of the MEX benchmarks by the CPS.

Anomalous IEP Student ISAT Performance in both PDP Control and Treatment Schools

Figure 7 reveals that, although IEP services in all schools did significantly narrow the ISAT performance gap between IEP and Non-IEP students from the baseline to the third year of project, these same students, on the average, still perform at a staggering rate of 26 points behind those students without IEPs.⁶

Figure 7: ISAT Gain score Differences between IEP and Non-IEP Students in Treatment Schools



Because of the statistically significant difference in the ISAT mean scores (see Appendix A: I.2 and I.3), the rest of this report will not include the IEP students to give us a more accurate picture of the overall effects of the program.

Summary Point 3: Analyses of student demographic factors, C-T outlier school ISAT profiles, and school district ISAT benchmark data provide additional evidence for the gradual yet significant effect of the PDP project on academic performance. The ongoing investigation of the effect of PDP on student learning in this report will be conducted without including data from the relatively small number of IEP outlier students in the overall analysis.

* * *

4. Investigating Three Measures of Arts and Arts Integration Learning

In the final year of the project, the CMAIE researchers administered three instruments designed to measure the impact of arts integration portfolio development on individual student arts learning:

- (1) The Arts Plus Arts Integration Performance Assessment Interview (PAI) administered to both control and treatment school cohorts

⁶ See Appendix Figures 1.2 and 1.3 for statistically significant differences between IEP and Non-IEP students.

- (2) The quantitative and qualitative assessment of individual student Portfolio Artifact Analysis (PAA) work samples in the treatment schools
- (3) The Arts Integration Portfolio Conference (AIPC) Performance Assessment Protocol designed to elicit student and teacher reflective understanding of the PDP learning outcomes based on discussion and interpretation of individual student AIP work samples

These tools were designed to engage students and teachers to reflect separately on their teaching and learning experiences related to arts integration units. The primary purpose of these tools was to provide an authentic assessment vehicle for students' level of understanding of arts and arts integration learning in the context of describing, discussing, and demonstrating aspects of their own and their peers' work.

The PAI and the AIPC provided opportunities for rating one-on-one discussion between the student and interviewer that revealed conceptual understanding, artistic process, content meaning, personal response, aesthetic criticism, and metacognition.

A secondary purpose of the AIPC was to give teachers an opportunity to articulate their views on the mission and goals of the PDP project and then to reflect on their observations of student performance in the AIPC in relation to their previous statements.

The student portfolio work was evaluated to determine the application of their knowledge and understanding in various artistic, musical, and writing projects throughout the school year. Quantitative assessment of individual student portfolio work products established a baseline measure of teacher support for PDP teaching practices. The qualitative assessment of student portfolio artifacts brought forth evidence of student interpretive understanding of individual, collaborative, and peer arts integration learning processes, products, and culminating events documented in their portfolios.

The validity of the analysis of AIPC and PAI 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 (see Appendix B: 1.1) conducted by an outside facilitator, video documentation and written transcription of each entire session, and an outside scoring team trained to rate each child's and teacher's level of response according to a common scoring rubric.

The PDP Student "Level of Complexity" Scoring System Shared Across The Three Instruments

The comparable student rating system deployed by the CMAIE team enabled the researchers to determine categorical differences in the "sophistication of response" across diverse performance tasks and work products specific to each unit of the program. Based on "skill theory" frameworks devised by Kurt Fischer⁷, the response ratings reflect

⁷ A theory of cognitive development: The control and construction of hierarchies of skills. Fischer, Kurt W. Psychological Review, Vol. 87(6), Nov 1980, 477-531

categorically different levels of cognitive complexity in the realm of artistic and integrative thinking.

The scoring system for interview transcripts in PAI, PAA, and AIPC instruments is based on a common 5-Level qualitative scale:

Table 7: PDP Common Student "Level of Response" Rating Scale for Three PDP Individual Student Learning Assessment Instruments

<p>Level NR (No Numerical Score): No Relevant Response Irrelevant or indiscernible response; silence</p>
<p>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.</p>
<p>Level 2: 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.</p>
<p>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.</p>
<p>Level 4: Systemic Understanding Substantial detail and specificity. Causal statements. Compare and contrast relationships. Critical perspective, highly complex, multiple relationships.</p>

Coding descriptions and response exemplars displayed in Appendix B demonstrate how the PDP portfolio scoring system works and provides an array of portfolio work samples that show how the assessment of artistic quality and reflective understanding of student work was made possible through the portfolio assessment processes developed in PDP project classrooms.

The results of the student arts and arts learning outcomes and their relation to standardized measures of academic achievement now follow.

* * *

Arts/Arts Integration Outcomes Measure 1: Control-Treatment School Student Arts Plus Arts Integration Performance Assessment Interview (PAI) Administered During the Final year of the project

The individual student Performance Assessment Interview (PAI) ratings (see Appendix B: 1.1) reveal important differences in the levels of understanding of arts and arts learning processes between the control and the PDP treatment schools.

Differences in mean scores displayed in Figure 9 indicate that the treatment school students' understanding of arts making processes and arts integration learning practices are significantly higher than those of the control school students.

Figure 9: Comparison of C-T PAI Mean Score Differences by the Final Year of the Project

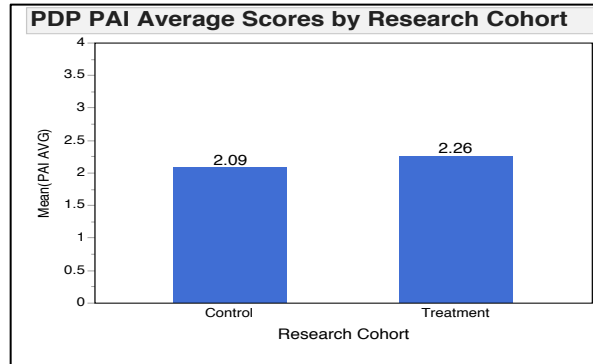


Table 8: Determination of Statistical Significant Differences of C-T PAI Score Comparisons

	Control	Treatment	Mean Difference	t Prob
PAI Average Score	2.08509	2.25944	0.174351	Prob > t = 0.0027**

t = positive trend; * = significant (p value <.05); ** = very significant (p value <.01)

Further demographic analyses reveal that the PDP treatment students outperform the control students regardless of Gender and Ethnicity:

Figure 10: Differences in PAI scores Distributed Equally According to Student Gender

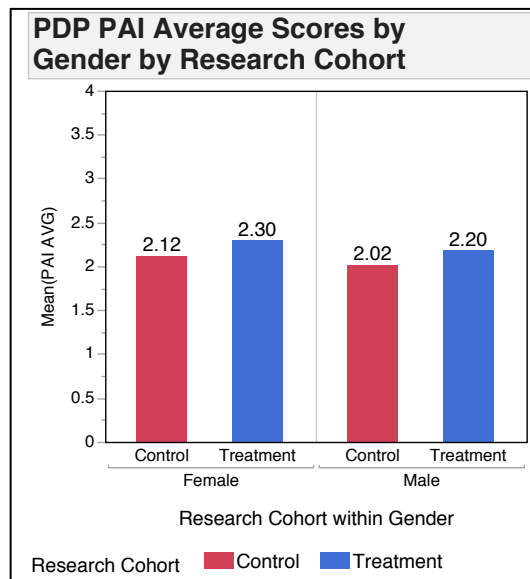
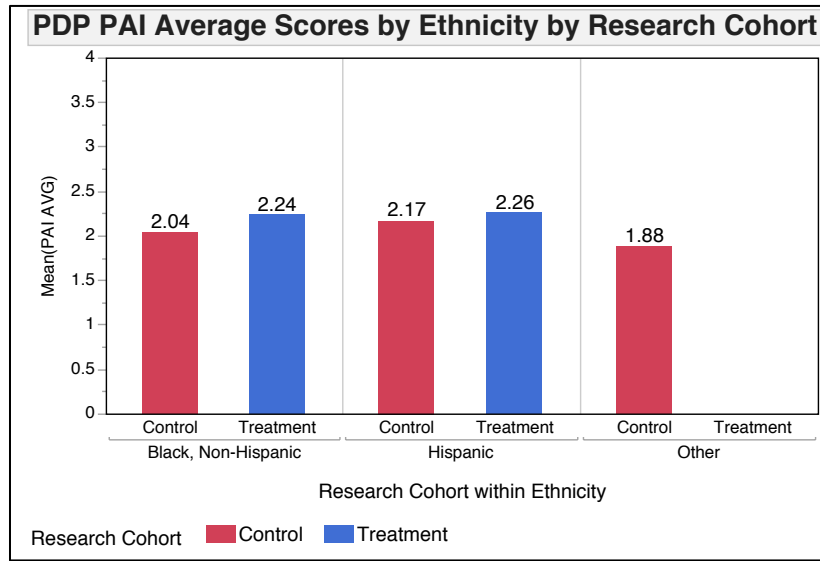
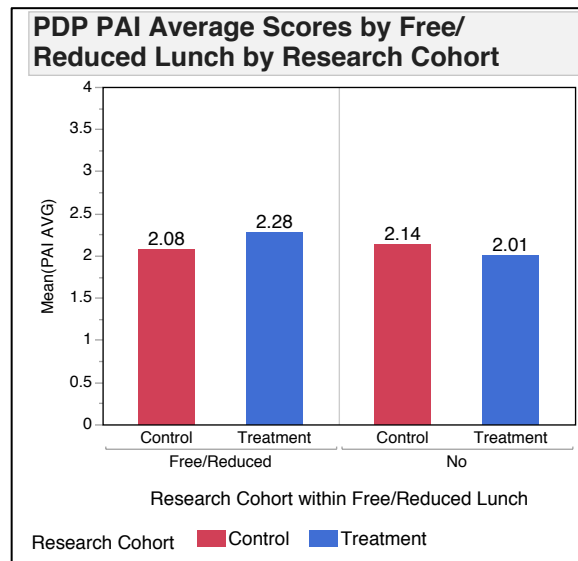


Figure 11: Differences in PAI scores Distributed Equally According to Student Ethnicity



According to analysis of PAI results based on whether or not students received Free/Reduced Lunch provisions (Figure 12), it appears that the PDP Treatment schools' low-income students showed a greater understanding of arts and arts integration processes than did the treatment school higher income students. This finding was the reversed in the control schools, where low-income students trailed both the treatment low-income students *and* the control higher income students.

Figure 12: Comparison of C-T PAI Scores by Family Income



As with the ISAT scores, the pool of ELL students is not large enough to make conclusive inferences regarding the impact of PDP.

C-T Performance Assessment Interview (PAI) Distribution of Scores According to PDP Expectations by the Final Year of the PDP Project

Benchmarks for standards of PAI scores in the final year of PDP were determined by tertile classification within the normal distribution of student performance ratings.

Meets/Exceeds (MEX) profile analysis of the PAI data in Table 8 reveals that treatment school students were far more likely to exceed the benchmark standards of arts learning and far less like to rate below these benchmarks when compared to the controls.

Table 8: C-T Differences in Benchmarks for PAI Ratings

Benchmark Categories	Below	Meets	Exceeds
Control	n=8/22 36.4%	n=11/22 50.0%	n=3/22 13.6%
Treatment	n=4/37 10.8%	n=18/37 48.7%	n=15/37 40.5%

Summary Point 4: Analysis of the Individual student Performance Assessment Interview (PAI) revealed positive evidence for the effect of PDP on the treatment school students, thereby suggesting preliminary evidence for possible causal links between student understanding of art works and art-making processes, the PDP teacher professional development program in general, and increased ISAT scores reported earlier.

* * *

Arts/Arts integration Outcomes Measure 2: Treatment School Arts Portfolio Artifacts Assessment (PAA) During Final Project Year

This variable was created to assess the quantity and quality of student portfolio work. Portfolio work samples collected in the final year of the PDP project were analyzed for (a) “number of artifacts” as an indicator of teacher level of support for the PDP project and (b) “quality of student” work products rated according to the rubrics presented in Appendix C: 1.2.

Table 9 specifies the relative distribution of student’s ability to successfully maintain an arts/arts integration portfolio system, a primary objective of the PDP teacher professional development program. Fifty-seven out of fifty-nine of the teachers’ students met or exceeded expectations for a successful PDP student portfolio system, a statistic that indicates that all teacher’s met or exceeded their responsibility to create a portfolio system for virtually all students in the treatment schools.

Table 9: Student Portfolio “Quantity of Artifacts” Distribution of Ratings according to Level of PDP Expectations by the Final Year of the Project

	Below ≤ 20	20 < Meets > 40	40 ≥ Exceeds
Quantity of Student Portfolio Artifacts	n=2/59 3.4%	n=23/59 39.0%	n=34/59 57.6%

Table 10 specifies the relative distribution of students' ability to produce high quality arts plus arts integration work products, another primary goal of the PDP teacher professional development program. The spectrum of student work ratings reveals that while virtually all PDP teachers had provided the opportunity for students to create portfolio work and most students (69.5%) met or exceeded PDP expectations, many students (30.5%) had difficulty creating detailed or multidimensional artistic work.

Table 10: Treatment School Student Portfolio "Quality of Artifacts" Distribution of Averaged Rating According to Level of PDP Expectations by the Final Year of the Project

	Below <= 2.0 (general, diffuse, single dimensional)	2.0< Meets >2.3 (multiple single dimensions, some detail)	2.3 >= Exceeds (toward inter-relational perspectives highly detailed)
Quality of Student Portfolio Artifacts	n=18/59 30.5%	n=30/59 50.9%	n=11/59 18.6%

Arts/Arts integration Outcomes Measure 3: Treatment School Arts Integration Student Portfolio Conference (AIPC) Assessment Results

This variable was created to rate the quality of individual student performance during their participation in facilitated portfolio conference protocol (Appendix B: 1.2). Performance ratings were based on the quality of description and dialogue with the facilitator and peers based on examples of student portfolio work discussed throughout the AIPC protocol. Students were rated for quality of response indicators according to the same rubric used to score the PAI responses (Appendix C: 1.1).

Table 11 specifies the distribution of treatment school students' leveled ability to reflect on the quality of arts plus arts integration work products, another primary objective of the PDP teacher professional development program. The spectrum of treatment school student portfolio conference response ratings reveals that, contrary to the quality ratings of the stand-alone portfolio work samples in the previous table, a large majority of students (82.6 %) met or exceeded expectations for critical thinking and reflective understanding of meaningful arts and arts integration learning processes, based on the interpretation of their own and their peer portfolio work products.

Table 11: Treatment School Student Portfolio Conference Performance Assessment Distribution of Averaged Ratings According to PDP Expectations by the Final Year of the Project

	Below <= 2.0 (general, diffuse, single dimensional)	2.0< Meets >2.3 (multiple single dimensions, some detail)	2.3 >= Exceeds (toward inter-relational perspectives highly detailed)
Student Portfolio Work Averaged Ratings	n=12/69 17.4%	n=26/69 37.7%	n=31/69 44.9%

Summary Point 5: Analysis of the individual student Arts Plus Arts Integration Portfolio “Number of Artifacts” in the treatment schools provided positive evidence of PDP teacher professional development outcomes by the final year of the project. The profile of the “Quality of Artifacts” ratings in the portfolios and student “Level of Response” ratings distilled from their Portfolio Conference performance assessment tasks provided evidence of the impact of portfolio practices on treatment students by the final year of the project.

* * *

Pairwise Inter-correlations Between All Three Student Learning Outcome Variables

Researchers employed multivariate “patterns and degree of correlation” analysis techniques to test for the degree of association among all treatment school student learning variables. The data summarized in Table 12 suggest that a statistically significant degree of association exists between:

- (a) “Quality of Student Portfolio Artifacts” and “Quality of Student Response in Portfolio Conference” Average Scores [positive trend]
- (b) “Quantity of Student Portfolio Artifacts” and the “Quality of Student Portfolio Artifacts” Average Scores [weak, yet statistically significant correlation]
- (c) “Student Portfolio Conference” and the “Student Performance Assessment Interview” Average Response Scores [strong, statistically significant correlation]

Table 12: Pairwise Correlations Among Treatment School Student Arts Learning Variables by Arts Teacher Type

Variable 1	Variable 2	Complete Spearman r	Complete Prob > p	Music Spearman r	Music Prob > p	Visual Arts Spearman r	Visual Arts Prob > p
Quantity of Student Portfolio Artifacts	Quality of Student Portfolio Artifacts	0.2649	0.0426*	N.S.	N.S.	N.S.	N.S.
Quantity of Student Portfolio Artifacts	Student PC Response Average Score	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
Student Portfolio Number of Artifacts	PAI Average Score	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
Quality of Student Portfolio Artifacts	Student PC Average Score	0.2510	0.1045	N.S.	N.S.	0.4226	0.0634 ^t
Student Portfolio Quality of Artifacts	Student PAI Average Score	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
Student PC Average Score	PAI Average Score	0.4744	0.0053**	N.S.	N.S.	0.8246	0.0010*

N.S. = non significant; ^t = positive trend; * = significant (p value <.05); ** = very significant (p value <.01)

Summary Point 6: Analysis of Inter-correlations among the PDP Student Learning Outcomes in treatment schools suggest that *quantity* is linked to *quality* of portfolio products, *quality* of portfolio artifacts predicts *quality* of reflection in portfolio conferences, and individual and interactive group interview reflective comments are strongly linked together by similar levels of understanding of arts and arts integration learning processes. Incidentally, the division of teacher types reveals that visual arts instructors were far more effective than their music teacher counterparts in bringing about these associations by way of their participation in PDP. Later on in statistical analysis, however, it was determined that these small sample teacher type differences did not significantly influence fundamental impact of PDP on student learning outcomes.

* * *

5. The Examination of Treatment School Teacher PD and Performance Variables and Their Links to Student Arts and Academic Learning Outcomes During the Final Year of the Project

In order to map the entire chain of evidence for the impact of PDP on the measures of student learning described earlier, teacher data was collected, validated, and reliably quantified by the research team. These data are organized into two categories: (a) teacher outcome variables linked to their participation in professional development events and (b) data collected and coded as a result of teacher performance assessment tasks and protocols.

Description and Numbering of Seven Teacher Outcome Variables

I: Four Arts Teacher Professional Development Outcome Variables

1A: Arts Teacher Attendance Data. Based on the number of exit surveys filled out, these data represent a basic measure of teacher engagement in PDP professional development events throughout the three years of the project administration. Attendance Data reveal that generally that 8 of 10 arts specialist teachers attended more than 50% of PD events, while 2 of 10 teachers attended less than 50% of the PDs offered.

1B: Arts Teacher Self-Assessment Pre-Post Survey Data. Based on averaged pre-post agreement responses (never – sometimes – most of the time – all of the time) to questions about support for arts integration learning practices in the classroom, the averaged results from all survey questions consistently reveal significant differences in teacher responses to the survey questions about:

- the depth of engagement with PDP practices in the classroom interactions,
- the maintenance of both student and teacher portfolio systems,
- the connections of PDP work to both arts & academic work, and
- the focus on providing opportunities for student reflection and self assessments.

Consistent with the goal of the treatment school classrooms, there are no low performing teacher outliers in the treatment schools (3/9 teachers in top tertile; 6/9 teachers in middle tertile; 0/9 teachers in bottom tertile).

IC: Arts Teacher Professional Development Event Exit Survey Average Score (Appendix D: 1.1). These data are based on self-esteem and confidence ratings based on averaged self-report ratings (1–2–3–4) that measure the degree of teacher understanding and confidence with PDP arts integration learning goals, content and process standards, teaching strategies, and their application to classroom practices throughout the three years of project administration. Though there are some differences among the teachers, the overall range of PDP survey responses register at a uniformly high level. That is, 10/10 overall teachers survey responses averaged in the top quartile of possible scores by the end of the project.

ID: The Arts Teacher Combined PD Outcome Variable is a composite teacher outcome variable created by averaging comparable level rubric ratings from the three previous teacher PD outcome measures, the Arts Teacher Observation Averaged Score, Teacher Quantity of Student Portfolio Work, and the Arts Teacher Portfolio Conference Performance Assessment Average Score, which follow.

II. Three Teacher Performance Assessment Outcomes

IIA: Arts Teacher Quantity of Student Work. As discussed previously, this teacher factor is measured in the final year of PDP by the number of portfolio artifacts collected from each treatment school student and is averaged by each classroom to represent each arts teacher's commitment to create and sustain an individual arts/arts integration student portfolio system according to the goals and PD practices modeled in this project. Although the averaged number of portfolio artifacts is high in terms of the PD standards by the final year of the project (8 of 9 arts teachers meet or exceed expectations of the PDP project), the distribution of averaged student number of artifacts nonetheless is used to rank order the teachers in terms of their students' ability to generate portfolio work products.

IIB: Arts Teacher Observation Averaged Score. This measure was based on expert ratings of teacher-student engagement and reflection during PDP classroom activities as described in the Teacher Observation Protocol (Appendix D: 1.3). Averaged ratings (1-2-3-4) encompass interactive factors (with an equal focus on teacher and student behaviors) such as exchange of questions, curiosities, big ideas, explicit attention to learning transfer, discussion of choices, creative processes and student-centric artistic behaviors such as active experimentation, imaginative ideas, multiple modes of expression, improvisation, "what if" questions, reflection on goals, self-assessment, respect for others, and collaboration.

Unlike previous teacher PD outcome measures, the averaged teacher observation ratings scored directly after a classroom visit revealed that most PDP arts teachers failed to display ideal behaviors during their classroom observations: 0 of 10 teacher observation exceeded the goals of the project, 1 of 10 teachers met PDP

standards of observed classroom behavior, and 9 of 10 teachers performed at below standard level of classroom practices assumed to represent ideal classroom culture for arts integration in the arts classroom.

IIC. Arts Teacher Portfolio Conference Performance Assessment Average Score. The structure of the teacher portion of the AIPC protocol (Appendix B: 1.2) shows that the arts teachers were challenged to describe the goals and practices of PDP and to interpret & assess their students' previous voiced discussion of their work in the earlier part of the portfolio conference. The transcripts of these conversations were coded and successfully scored by the CMAIE research team only in the final year of the project. Transcribed teacher responses during the AIPC were scored for levels of relevance, detail and perspective using the scoring rubric system presented in Appendix C: 1.1. Because the reflection process was based on high quality portfolio work chosen by the teacher and the students and that the index of inter-rater reliability was high—over 97% of the ratings were within the acceptable range of agreement and all problematic examples were scored twice and averaged between two independent scorers—the researchers were confident that arts teacher rank ordered averaged ratings represent a valid and reliable measure of teacher *reflective understanding* of the contribution of arts integration portfolios to student learning in the PDP classroom. The range of Arts teacher AIPC ratings were normally distributed throughout the spectrum of teacher levels (5 of 9 teachers in top tertile and 4 of 9 teachers in middle tertile).

Pairwise Inter-correlations Among All Teacher PD and Performance Assessment Outcome Variables

Similar to the inter-correlational analysis of student performance variables, PDP researchers employed multivariate “patterns and degree of correlation” analysis techniques to test for the degree of association among all treatment school teacher professional learning variables. In this case, however, results suggest that out of 21 permutations of teacher PD and performance outcomes, virtually no statistically significant degree of association exists except in the case of two paired variables:

- (d) IA: Teacher PD Attendance and IIB: Teacher Observation Ratings
- (e) IB: Teacher Self-Assessment Pre-Post Survey and IC: Teacher Self Esteem/Confidence PD Event Exit Survey

In Table 13 the first pair suggests that attendance in PD events did predict teacher observation ratings moderately well, particularly in the case of the music teachers. The second pair suggests that the teacher self-assessment and attitude surveys are linked in terms of content area. The third pair is a calculation of how the impact of teacher success in generating productive and rich portfolios and how that enhances their ability to articulate and demonstrate the goals and impact of the PDP program in their classrooms. The lack of correlation among the remaining permutations of paired variables suggest these combinations of variables were either relatively independent of each other (as indicated by negative or random correlations) or were, by design, already correlated significantly with the composite teacher ratings variable.

Table 13: Three Significant Correlations Out Of 22 Treatment School Arts Teacher PD and Performance Variables

Variable 1	Variable 2	All Arts Spearman p	All Arts Prob > p	Music Spearman p	Music Prob > p	Visual Arts Spearman p	Visual Arts Prob > p
IA: Teacher PD Attendance	Teacher Observation Ratings	0.3402	<0.0001**	0.4178	0.0002**	-0.6783	<0.0001**
IC: Teacher Self-Assessment Ratings	Teacher Self-Esteem /Confidence Ratings	0.2989	0.0002**	N.S.	N.S.	0.5472	<0.0001**
IIA: Teacher Quantity of Student Portfolio Work	Teacher PC Performance Assessment Ratings	0.2880	0.0314*	N.S.	N.S.	0.5282	0.0080**

N.S. = not statistically significant; * = significant (p value <.05); ** = very significant (p value < .01)

Summary Point 7: Only 3 of 22 paired teacher PD/performance assessment variables were strongly and positively inter-correlated: that is, (a) strong teacher PD attendance appears to predict high quality teacher observation ratings (and vice versa), (b) high levels of self-esteem or confidence implementing PDP teaching practices corresponds to high level self assessment ratings, and (c) a high level of classroom student portfolio work productivity (i.e., number of student work artifacts) predicts to a certain extent the teachers' level of sophistication of response during the PDP portfolio conference protocol. The first case suggests a causal link between teacher training and high quality arts/arts integration teaching practices; the second case suggests that significant overlap exists between the two separate teacher survey instruments such that a strong self-perception of success with the program is tied closely with high levels of confidence in incorporating the program into teacher classrooms; the third case suggests that the arts teacher's ability to document a large amount of student work artifacts in the portfolios predicts higher levels of articulation about the program's goals and the impact of the program on the quality of student arts/arts integration work. Overall, the lack of correlation among a large majority of the teacher variables suggests that the teacher data collection instruments represented relatively independent measures of teacher effectiveness in PDP.

* * *

6. Linking the Chain of Evidence I: Direct Pairwise Correlations Between Teacher PD and Student Academic Performance Outcomes

Once all teacher and student variables have been described and validated in isolation of one another or their interdependency with one another, the next step in the "chain of evidence" evaluation is to search and test for "patterns and degree of correlation" between: (a) the seven categories of teacher PD and performance outcome variables, (b) the three

categories of student arts learning variables, and (c) the two measures of student performance on standardized academic tests.⁸

6A. Correlation Between Teacher Variables and Student Academic Outcomes: Baseline – Final Year ISAT Test Gain Scores

Table 14 reveals that only two out of seven teacher factors significantly relates to ISAT Combined (reading and math) Average Gain Scores from the baseline to the final year of implementation. That is, only IA: Teacher Attendance in PD Events and (2) IIA: Teacher Quantity of Student Portfolio Work (i.e., number of portfolio work artifacts) correlated significantly with student academic performance.

The direct influence of teacher attendance on long-term academic gains—particularly in the case of the visual arts teachers—provides an essential evidentiary link between PDP arts/arts integration teacher training and the student learning gains that differentiated the treatment from the control schools by the final year of the project reported in section 4 (Arts/Arts Integration Outcomes Measure 1). The teacher ability to produce a higher quantity of documented student learning artifacts in their student portfolios that corresponded to increases in student test scores appeared to be more likely the case in music classrooms than with visual arts.

Table 14: Correlation of 7 Teacher Variables with Student ISAT Gain Scores from Baseline To Final Year of the PDP

Teacher Variable Correlation with Baseline to Final Year (2011-13) Final Year ISAT Combined Average Gain Scores	Complete Spearman r	Complete Prob > p	Music Spearman r	Music Prob > p	Visual Arts Spearman r	Visual Arts Prob > p
4 Teacher PD Outcome Variables						
IA: Arts Teacher PD Attendance	0.1871	0.0391*	N.S.	N.S.	0.4364	0.0008**
IB: Arts Teacher Pre-Post Survey Self-Assessment	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
IC: Arts Teacher Exit Survey (Self-Esteem/ Confidence with PDP)	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
1D: Combined Arts Teacher PD Outcome Variable	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
3 Teacher Performance Assessment Outcome Variables						
IIA: Arts Teacher Quantity of Student Portfolio Work (# artifacts)	0.3223	0.0175*	0.5265	0.0020**	N.S.	N.S.
IIB Arts Teacher Classroom Observation Rating	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
IIC: Arts Teacher Portfolio Conference Performance Assessment	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.

N.S. = not statistically significant; † = positive trend; * = significant (p value <.05); ** = very significant (p value < .01)

⁸ N.B. See Appendix F for complete summary chart of all teacher and student outcome variables. See Final Figure 13 Correlation-Regression Multivariate Map for a flow chart representation of all principal inter-related variables.

* * *

6B. Correlation Between Teacher Variables and Student Academic Outcomes: Final Year ISAT Test Scores

The association of teacher variables with academic achievement in the culminating year of the project was investigated to determine which factors influence academic learning in the culminating year of the arts teachers' portfolio practices. Table 15 indicates that a strong significance exists between the ISAT Combined Average score and the Teacher Self-Esteem and Confidence with arts integration practices (distilled from the IC: Teacher Exit Survey results), particularly in the case of music teachers who had less previous familiarity with portfolio documentation practices than did the visual arts teachers.

Weak but statistically significant *negative* correlations between student ISAT scores and Teacher performance ratings in the AIPC Portfolio Conference performance assessments and Teacher Observation ratings during the final year of PDP suggest teacher understanding of portfolio conference student performance was not yet sufficiently addressed in the teacher PD program.

Table 15: Strong, Significant Correlations Exist Between One Arts Teacher Outcome Variables and Final Year Student ISAT Test Scores

Teacher Variable Correlation with 2012-13 Final Year ISAT Combined Average Score	Complete Spearman r	Complete Prob > p	Music Spearman r	Music Prob > p	Visual Arts Spearman r	Visual Arts Prob > p
4 Teacher PD Outcome Variables						
IA: Arts Teacher PD Attendance Data	N.S.	N.S.	-0.4347	0.0002**	N.S.	N.S.
IB: Arts Teacher Pre-Post Survey Self-Assessment	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
IC: Arts Teacher Exit Survey (Self-Esteem/ Confidence with PDP)	0.2472	0.0046**	0.3994	0.0008**	N.S.	N.S.
1D: Combined Arts Teacher PD Outcome Variable	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
3 Teacher Performance Assessment Outcome Variables						
IIA: Arts Teacher Quantity of Student Portfolio Work (# artifacts)	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
IIB Arts Teacher Classroom Observation Rating	-0.1739	0.0496*	N.S.	N.S.	N.S.	N.S.
IIC: Arts Teacher Portfolio Conference Performance Assessment	-0.1938	0.0284*	-0.2419	0.0522 ^t	N.S.	N.S.

N.S. = not statistically significant; ^t = positive trend; * = significant (p value <.05); ** = very significant (p value < .01)

Summary Point 8: Teacher arts integration professional learning outcomes in PDP, investigated for their influence on student academic learning, have determined that teacher participation, positive self assessment, and response to portfolio conference protocols focused on the impact of PDP on student work are highly associated with treatment school academic gains. These results provide evidence that explains why PDP students improved at a greater rate than did the control schools as presented in section 3 of this report.

* * *

7. Linking the Chain of Evidence II: Direct Pairwise Correlations Between Teacher PD and Student Arts Learning Performance Outcomes

Because evidence exists that teacher PD outcomes are linked with student academic success, possible correlations between PD outcomes and arts learning outcomes can be explored to determine whether a chain of evidence can be drawn through the student arts learning variables in ways that may be linked to either or both teacher professional learning and student academic test performance data.

7A. Pattern and Degree of Correlation Between Teacher Variables and Final Year Student Arts/Arts Integration Performance Assessment Interview (PAI)

Results from correlation analysis suggest that there is no direct evidence of statistical correlation between any combination of teacher and student arts learning outcomes.

Summary Point 9: The lack of any significant correlation between any of the teacher PD or performance assessment data and final year student PAI results suggest that the arts teachers' responses to the PDP professional development program and to their ability to development productive arts/arts integration portfolio systems in their classroom were more likely to bolster academic rather than arts student learning outcomes. It appears that it is the PAI performance ratings—and not teacher PD or performance assessment variables—that correspond to student performance in the various forms of PDP project arts learning, such as student portfolio conference or PAI performance assessment ratings.

* * *

7B. Pattern and Degree of Correlation between Teacher Variables and Final Year Student Quality of Portfolio Work Artifacts

Table 16 reveals evidence that both variables IIA: Teacher Quantity of Portfolio Artifacts and IIC: Teacher Portfolio Conference Performance Assessment Ratings correlate significantly and positively with IIIA: Quality of Student Portfolio Work by the final year of the project. The predictive power of these two variables did not surprise the researchers because: (a) a higher number of student artifacts is the result of high levels of teacher support for the portfolio practices in the classroom and (b) the high quality of student artifacts should be linked with higher ratings of teacher reflection on student achievement goals in PDP as demonstrated by the higher level of sophistication of their portfolio conference interview ratings

Conversely there are also two instances of significant, yet negative correlations between IIA: Student Quality of Portfolio Work Ratings and both (a) IA: Teacher PD Attendance Data and (b) IIC: Teacher Portfolio Conference results. These rather surprising paired correlation results may be due to the focus and timing of the data collection. That is, by the final year of the project, teacher attendance in PD events may be regarded as more

supplemental than central to the quality of student work, and that the observed classroom teaching practices were less germane to evidence of arts learning than was the level of the teachers' ability to promote and understand the implications of high quality student portfolio work.

Table 16: Pattern and Degree of Significant Correlations between 7 Teacher Variables and Student Portfolio *Quality of Portfolio Work* Artifacts by the Final Year of the Project

Teacher Variable Correlations with Student Quality of Portfolio Work	Complete Spearman r	Complete Prob > p	Music Spearman r	Music Prob > p	Visual Arts Spearman r	Visual Arts Prob > p
4 Teacher PD Outcome Variables						
IA: Arts Teacher PD Attendance Data	-0.3605	0.0054**	N.S.	N.S.	-0.4663	0.0108*
IB: Arts Teacher Pre-Post Survey Self-Assessment	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
IC: Arts Teacher Exit Survey (Self-Esteem/ Confidence with PDP)	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
1D: Combined Arts Teacher PD Outcome Variable	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
3 Teacher Performance Assessment Outcome Variables						
IIA: Arts Teacher Quantity of Student Portfolio Work (# artifacts)	0.2649	0.0426*	N.S.	N.S.	N.S.	N.S.
IIB Arts Teacher Classroom Observation Rating	-0.3639	0.0058**	-0.3293	0.0657 ^t	N.S.	N.S.
IIC: Arts Teacher Portfolio Conference Performance Assessment	0.2880	0.0314*	N.S.	N.S.	0.5282	0.0080**

N.S. = not statistically significant; ^t = positive trend; * = significant (p value <.05); ** = very significant (p value < .01)

Summary Point 10: Table 16 provides evidence that by the final year of the project, the *quality* of student portfolio work depends less on continued exposure to teacher PD training or observable changes in classroom teaching practices, but rather relies more on the success of the arts teacher in generating high quality student work that, in turn, can be linked to the increasingly sophisticated metacognitive perspective on student learning revealed in the teacher portfolio conference ratings.

* * *

7C. Pattern and Degree of Correlation between Teacher Variables and Final Year Student Portfolio Conference Performance Assessment Response Ratings.

Table 17 indicates that a different kind of teacher PD outcome measure, IC: Teacher Exit Survey Ratings, a self report attitude variable that focused on issues of teacher self esteem and confidence as it pertained to the their ability to support arts/arts integration portfolio practices in their classroom, positively correlates with IIB. Student Portfolio Conference Response Ratings. This teacher attitude variable appears far more important to the music

teachers who, in comparison to visual arts teachers, were far less likely to have initiated or sustained systematic documentation and assessment of student work in their classrooms prior to the project.

Negative correlation suggests that a high degree of **IA: Teacher Attendance** did not correspond to positive student performance during the portfolio conference performance tasks with PDP practices. This result suggests that by the end of the project, teacher effectiveness may depend more on confidence developed through personal experience with PDP arts/arts integration practices than on attending more PD sessions.

Table 17: Strong, Significant Correlations between 7 Teacher Variables and Student Portfolio Conference Response Ratings by the Final Year of the Project

Teacher Variable Correlations with Student Portfolio Conference Response Ratings	Complete Spearman ρ	Complete Prob > ρ	Music Spearman r	Music Prob > ρ	Visual Arts Spearman ρ	Visual Arts Prob > ρ
4 Teacher PD Outcome Variables						
IA: Arts Teacher PD Attendance Data	-0.2412	0.0510 ^t	-0.3739	0.0321*	-0.3073	0.0820 ^t
IB: Arts Teacher Pre-Post Survey Self-Assessment	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
IC: Arts Teacher Exit Survey (Self-Esteem/ Confidence with PDP)	0.3132	0.0104*	0.4607	0.0070**	N.S.	N.S.
1D: Combined Arts Teacher PD Outcome Variable	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
3 Teacher Performance Assessment Outcome Variables						
IIA: Arts Teacher Quantity of Student Portfolio Work (# artifacts)	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
IIB Arts Teacher Classroom Observation Rating	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
IIC: Arts Teacher Portfolio Conference Performance Assessment	N.S.	N.S.	-0.3698	0.0373*	0.4813	0.0046**

N.S. = not statistically significant; ^t = positive trend; * = significant (p value <.05); ** = very significant (p value < .01)

Summary Point 11: Relatively few measures of arts teachers' professional development or performance outcomes were linked positively to student academic or arts learning outcomes. Nonetheless, the patterns and degree of correlation over time revealed that specific teacher variables mattered greatly with regard to (a) arts learning for its own sake and (b) arts integration for the sake of PDP's eventual impact on ISAT test scores. Thus, IIA.: Teacher Quantity of Student Portfolio Work and IIC: Teacher Portfolio Conference Performance Assessment Ratings strongly influences IIIA: Quality of Student Arts/Arts Integration Portfolio Work, while IC: Teacher Exit Survey results strongly linked to students' understanding of high quality arts learning and the possible impact of arts integrated learning on academic learning as demonstrated by IIB: Student Portfolio Conference Performance Assessment Ratings.

The next section of the report focuses on the correlation of student arts learning outcomes with academic outcomes.

* * *

8. Linking the Chain of Evidence III: Direct Pairwise Correlations Between Student Arts Learning and Academic Performance Outcomes

The previous sections explored the correlational links that existed between the seven arts teacher PD and performance assessment factors and the two student standardized test results and the three student arts learning outcomes. The correlations and connections between arts learning outcomes and ISAT standardize academic performance outcomes are investigated below.

* * *

8A. Patterns and Degree of Correlation Between Student Quality of Portfolio Work Artifacts and ISAT Academic Achievement Test Scores

Table 18 shows that a strong and highly sophisticated correlation exists between IIIA: Student Quality of Portfolio Work Ratings and the IVB: Student ISAT Final Year Combined Average Scores.

Table 18: Correlation of Student Quality of Portfolio Work and Academic Performance

IIIB: Student Quality of Portfolio Work Ratings and their Correlation with ISAT Academic Achievement Tests	Complete Spearman r	Complete Prob > p
IVB: Student Baseline-to-Final-Year ISAT Combined Academic Performance Average Gain Scores (2010-2014)	N.S.	N.S.
IVA: Student Final Year ISAT B=Combined Academic Performance ISAT Combined Average Scores (2013-2014)	0.4489	0.0005**

N.S. = not significant ^t = positive trend; * = significant (p value <.05); ** = very significant (p value < .01)

Summary Point 12: A very strong degree of correlation exists between IIIA: Student Quality of Portfolio Work Ratings and academic achievement, suggesting that the successful implementation of PDP in music or visual arts classrooms optimizes the effect of arts learning within portfolios on academic achievement. The reason this effect was not obtained in the baseline-to-final-year academic gain scores is probably due to several factors: (a) qualitative assessment of portfolio work was not conducted until the final year of the project, (b) the academic gains for the PDP treatment schools was not achieved significantly until the third year of the project, and (c) the impact of arts integration skills in the context of arts learning classrooms were not shown to have any connection to prior academic performance.

* * *

8B. Patterns and Degree of Correlation Between Student Portfolio Conference (PC) Performance Assessment Ratings and ISAT Academic Achievement Test Scores

Correlation analysis determined that no significant associations exist between the IIIB: Student PC Performance Assessment Ratings and IVA: Student Final Year ISAT Test Scores or the IVB: Student ISAT Baseline to Final Year Gain Scores.

Summary point 13: Unfortunately, researchers were not able to determine the degree and pattern of this correlation in the earlier years of the project because of unreliable implementation of the Student PC protocol that led to uncorrectable scoring and coding problems.

8C. Patterns and Degree of Correlation Between Student Performance Interview (PAI) Response Ratings and ISAT Academic Achievement Test Scores

Statistical analysis revealed that student IIIC: PAI Ratings do not correlate significantly with either the IVA: ISAT Final Year ISAT Combined Academic Performance Test Scores or the IVB: ISAT Baseline-to-Final-Year Test Gain Scores.

Summary Point 14: Although the IIIC: PAI Ratings are statistically isolated from the rest of the other PDP program factors, these data are significantly linked with the IIIB: Student Portfolio Conference Performance Assessment Ratings. In addition, the link between the Student PAI and Portfolio Conference Performance Assessment Ratings validates both instruments as measures of arts/arts integration teaching and learning in arts integration learning environments.

* * *

9. Determining the Strongest Links: Stepwise Regression Testing for Most Significant Teacher, Student and Family Demographic Predictors of Academic Achievement

CMAIE researchers employed multivariate “patterns and degree of correlation” analysis techniques to test for the degree of association among all treatment school teacher professional learning variables.

The two regression models investigated in this report focus on sorting out which of the 7 teacher PD and performance outcome variables, 3 student arts/arts integration learning outcome variables, and 5 student family demographic variables best fit the shape and trajectory of the (a) IVA: Student ISAT Final Year Combined Academic Test Score Data and the IVB: Student ISAT Baseline-to-Final Year Combined Academic Test Gain Scores. The teacher PD, student arts/arts integration and/or demographic variables that best fit the academic outcome data fit thus become the principal predictors of academic achievement.

9A. PDP Stepwise Fit for IVB. Baseline-to-Final Year Student ISAT (2011—2014) Combined Academic Test Gain Scores

Stepwise Regression methods focused on IVB: Student ISAT Baseline-to-Final Year Combined Academic Performance Gain Scores (2011-2014) resulted in identifying the most prominent predictors of academic achievement in the treatment schools.

The factor-by-factor stepwise fit for the difference in ISAT scores between the 2010-2011 and the 2013-2014 academic years in Table 19 reveals that the IA: Teacher Attendance in PD workshops and the concomitant increased ability of teachers to produce a greater IIA: Quantity of Student Portfolio Work Artifacts over the span of the project are by far the two strongest and statistically significant factors that explain the difference in levels of ISAT achievement from the baseline to final year of the project⁹.

Table 19: Stepwise Regression Modeling Fit for IVB. Student ISAT Baseline-to-Final Year ISAT Combined Academic Achievement Average Scores

Major Predictors of Academic Achievement Gain Scores	F ratio (Effect size)	Prob > F	R ² (Degree of explained variance per variable)	Cumulative R ² (whole model explained variance)
IIA. Teacher Quantity of Student Portfolio Work Artifacts (determined by the number of portfolio work samples collected in the final year of the project)	30.558	0.000004**	0.3499	0.3499
Teacher PD Attendance (determined by the submission of exit surveys)	14.251	0.00151**	0.2965	0.6464

N.S. = not significant ^t = positive trend; * = significant (p value <.05); ** = very significant (p value < .01)

* * *

9B. Regression Fit for Final Year 2013-2014 ISAT Combined Average Scores

The stepwise regression fit for the IVA: Final Year (2013-2014) Student ISAT Combined Academic Test Scores resulted in identifying the one prominent and three relatively ancillary predictors of academic achievement in the treatment schools during the culminating year of PDP project implementation.

In Table 20, the stepwise fit for the IVA: Final Year ISAT Combined Average Scores reveals that the IIIA: Student *Quality* of Portfolio Work Artifacts is the most significant factor in predicting academic achievement. Three other factors: (a) IA: Teacher PD Attendance, (b) IB: Teacher Pre-Post Self-Assessment Survey Ratings and (c) Free/Reduced Lunch Benefits

⁹ See Appendix E 1.1 for regression effect estimates and complete step history details.

(Family Income) Demographic Classification Data all influence academic achievement significantly, though with far less explanatory power as summarized in the table below.

Table 20: Stepwise Regression Factor Fit to Student ISAT Final Year Combined Academic Test Scores

Major Predictors of PDP Final Year Academic Achievement Test Scores	F ratio (Effect size)	Prob > F	R ² (Degree of explained variance per variable)	Cumulative R ² (whole model explained variance)
IIIA: Student Quality of Portfolio Work Ratings (Scored by the CMAIE Research Team)	22.182	0.00041**	0.3067	0.3067
Free/Reduced Lunch Benefits Student Demographic Classification Data (Family Income)	7.167	0.019*	0.1230	0.4297
IB: Teacher Pre-Post Survey Self-Assessment Ratings (Arts teacher ability to implement PDP in the classroom)	10.829	0.00585**	0.0872	0.5169
1A: Teacher PD Attendance (determined by the submission of exit surveys)	9.824	0.00791**	0.0739	0.5908

N.S. = not significant † = positive trend; * = significant (p value <.05); ** = very significant (p value < .01)

Summary Point 15: Stepwise regression analyses focused on multiple PDP teacher and student outcome variables have succeeded in determining the principal factors leading to academic achievement. Four strategies for measuring academic progress were investigated: (a) longitudinal view of patterns of academic achievement now show that teacher committed participation in high quality PD programs and the consequent proliferation of portfolio work is most highly associated with differences in academic gain score comparisons and final year results between matched pair control and treatment school academic schools, (b) the pattern and degree of pairwise correlation between teacher PD factors and student arts learning outcomes, (c) the pattern and degree of pairwise correlation between arts learning factors and academic test scores, and (d) the combination of all teacher PD outcomes, student arts/arts integration learning outcomes and student/family demographic traits were measured in the context of one another through regression analysis to determine statistically both the significance and the degree of influence on academic achievement. In sum, “baseline to final year” academic progress was most clearly linked with long-term participation in Teacher PD training sessions and the quantity of student work produced; “final year results” were most clearly linked with the quality of student portfolio work, positive ratings on teacher self assessment surveys, students’ family income status, and continued engagement in PD services.

In the final section of the report, all the correlation and regression links are mapped together to express the complexity and flow of successful arts integration program development in urban public elementary school learning environments.

* * *

9C. The Full Chain of Evidence: A Summary Correlation-Regression Flow Chart Map Depicting the Relationships Between the Teacher & Student Variables and their Hierarchically Ordered “Pattern and Degree of Association” with Student Academic Longitudinal Gain Scores and Final Year Test Results.

Multivariate analysis has provided a useful though somewhat circuitous route toward explaining the development of portfolio practices in Chicago Public School arts classrooms and its impact on both arts and academic learning. The “chain of evidence” approach takes into consideration a sequence of 7 teacher and 5 student outcome variables that have, up to this point, established the basis for arts integration PD training and program development criteria aimed at increasing both arts and academic learning. The longitudinal cohort academic outcomes have developed over time to the point that researchers can make control-treatment school comparisons, can construct a flowchart of factors that show how teacher PD response outcomes lead to teaching outcomes and how teacher performance outcomes lead to new forms of student learning, and can show that all of the factors contribute to academic achievement.

Table 21 is a summary of all significant and positive correlations and regression factors that account for the success of the academic contrast with control schools, the rising level of sophistication of student portfolio work, and the reflective thinking in both teacher PD sessions and student portfolio performance assessment protocols that provide indications of the growing of arts integration teaching and learning practices by grade 6 in the final year of the project.

Table 21: Summary of All Correlation and Regression Factors as the Basis for the Final PDP Correlation-Regression Multivariate Analysis Table

Four Teacher PD Outcome Variables		
IA: Arts Teacher PD Attendance Data	Correlation with IVB: Student Baseline to Final Year Combined Academic Performance Average Test Score (Prob > p = 0.0391)	Significant regression factor of IVA: Student Final Year Combined Academic Performance Average Test Score (Prob > F = 0.00791) (r ² = .0739) Significant regression factor of IVB: Student Baseline to Final Year Combined Academic Performance Average Test Score (Prob > F = 0.00151) (r ² = .2965)
IB: Arts Teacher Pre-Post Survey Self-Assessment	N.S.	Significant regression factor of IVA: Student Final Year Combined Academic Performance Average Test Score (Prob > F = 0.00585) (r ² = .0872)
IC: Arts Teacher Exit Survey (Self-Esteem/ Confidence with PDP)	Correlation with IIIB: Student Portfolio Conference Performance Assessment Ratings (Prob > p = 0.0104) Strong correlation with IVA: Student Final Year Combined Academic Performance Average Test Score (Prob > p = 0.0046)	N.S.
1D: Combined Arts Teacher PD Outcome Variable	Strong Correlation with IIA: Arts Teacher Quantity of Student Portfolio Work (# of artifacts) (Prob > p =<0.0001)	N.S.

Three Teacher Performance Assessment Outcome Variables		
IIA: Arts Teacher Quantity of Student Portfolio Work (# of artifacts)	<p>Correlation with IIIA: Student Quality of Portfolio Work Ratings (Prob > p = 0.0426)</p> <p>Correlation with IVB: Student Baseline to Final Year Combined Academic Performance Average Test Score (Prob > p = 0.0175)</p> <p>Strong correlation with 1D: Combined Arts Teacher PD Outcome Variable (Prob > p = <0.0001)</p>	<p>Significant regression factor of IVB: Student Baseline to Final Year Combined Academic Performance Average Test Score (Prob > F = 0.0000369) (r² = .3499)</p>
IIB: Arts Teacher Classroom Observation Ratings	N.S.	N.S.
IIC: Arts Teacher Portfolio Conference Performance Assessment	<p>Correlation with IIIA: Student Quality of Portfolio Work Ratings (Prob > p = 0.0314)</p>	N.S.
Three Student Arts Learning Assessment Outcome Variables		
IIIA: Student Quality of Portfolio Work Ratings	<p>Correlation with IIA: Arts Teacher Quantity of Student Portfolio Work (# of artifacts) (Prob > p = 0.0426)</p> <p>Correlation with IIC: Arts Teacher Portfolio Conference Performance Assessment (Prob > p = 0.0314)</p> <p>Strong correlation with IVA: Student Final Year Combined Academic Performance Average Test Score (Prob > p = 0.0005)</p>	<p>Significant regression factor of IVA: Student Final Year Combined Academic Performance Average Test Score (Prob > F = 0.00041) (r² = .3067)</p>
IIIB: Student Portfolio Conference Performance Assessment Ratings	<p>Correlation with IC: Arts Teacher Exit Survey (Self-Esteem/ Confidence with PDP) (Prob > p = 0.0104)</p> <p>Strong correlation with IIC: Student Performance Assessment Interview Ratings (Prob > p = 0.0053)</p>	N.S.
IIIC: Student Performance Assessment Interview Ratings	<p>Strong correlation with IIIB: Student Portfolio Conference Performance Assessment Ratings (Prob > p = 0.0053)</p>	N.S.
One Student Demographic Factor		
Student Demographic Factor: Free/Reduced Lunch (Gender, HAL classification, Ethnicity)	N.S.	<p>Regression factor of IVA: Student Final Year Combined Academic Performance Average Test Score (Prob > F = 0.019) (r² = .1230)</p>

(continued on next page)

Two Student Academic Assessment Outcome Variables		
<p>IVA: Student Final Year Combined Academic Performance Average Test Score</p>	<p>Strong correlation with IC: Arts Teacher Exit Survey (Self-Esteem/Confidence with PDP) (Prob > p = 0.0046)</p> <p>Strong correlation with IIIA: Student Quality of Portfolio Work Ratings (Prob > p = 0.0005)</p>	<p>Significant regression factors:</p> <p>IA: Arts Teacher PD Attendance Data; (Prob > F = 0.00791) (r² = .0739)</p> <p>IB: Arts Teacher Pre-Post Survey Self-Assessment of PDP practices (Prob > F = 0.00585) (r² = .0872)</p> <p>IIIA: Student Quality of Portfolio Work Ratings (Prob > F = 0.00041) (r² = .3067)</p> <p>Significant Regression Student Demographic Family Income Factor: Free/Reduced Lunch (Prob > F = 0.019) (r² = .1230)</p>
<p>IVB: Student Baseline to Final Year Combined Academic Performance Average Test Score</p>	<p>Correlation with IA: Arts Teacher PD Attendance Data (Prob > p = 0.0391)</p> <p>Correlation with IIA: Arts Teacher Quantity of Student Portfolio Work (# of artifacts) (Prob > p = 0.0175)</p>	<p>Significant regression factors:</p> <p>IA: Arts Teacher PD Attendance Data (Prob > F = 0.00151) (r² = .2965)</p> <p>IIA: Arts Teacher Quantity of Student Portfolio Work (Prob > F = 0.0000369) (r² = .3499)</p>

N.S. = not significant ^t = positive trend; * = significant (p value <.05); ** = very significant (p value < .01)

* * *

Correlation and Regression Analysis Findings in the context of the PDP “Chain of Evidence” Flow Chart

Using data in Table 21 as the foundation, Figure 13 (mentioned previously in section 1 of this report) summarizes the causal links in the chain of evidence that flows from a sequence of evidence from:

- I. Teacher PD Outcomes to
- II. Teacher Performance Assessment to
- III. Student Arts/Arts Integration Learning to
- IV. Academic Tests Gain Scores and Final Year results.

Significant paired correlations (thin dotted lines):

- Delineate the association of “teacher PD attendance” and “quantity of student portfolio work” on the long-term academic gain scores, a finding that shows that support for teacher development of arts integration had a direct influence on the teachers’ ability to promote an expansive documentation of student work that had a substantial effect on academic performance
- Trace the influence of teacher “quantity of student portfolio work” and “portfolio conference performance assessment ratings” on student “quality of student portfolio work,” a finding that substantiates that not only was the PDP program fully developed into the arts learning classrooms after three years, but that the quality of teacher reflective understanding of the goals and impact of the program was tied to levels of quality student work.

- Demonstrate the importance of survey data that suggest how high levels of teacher “self-confidence” with PDP arts integration portfolio practices can likely lead to students’ ability to express increasingly sophisticated levels of reflective understanding of their own work and its connections to both arts and arts integration teaching and learning as indicated by “student portfolio conference performance assessment ratings”.

Highly significant paired correlations (thin solid lines) extend the line of evidence by showing that:

- “Quality of student portfolio work” connected previously with teacher outcomes is also significantly linked to “final year ISAT academic achievement” levels, a finding that suggests that, as PDP arts/arts integration practices matured in the music and visual arts classrooms, so did its influence on academic achievement. Thus PDP not only optimized academic and arts learning compared to the matched control school longitudinal cohorts as described in the first part of this report, but also became the intermediary causal chain of factors that proceeded from teacher outcomes to high quality student arts learning outcomes that in turn predicted level of academic achievement.
- The “teacher self-confidence” PD outcome ratings positively associated with “student demonstration and reflection” portfolio conference ratings (described previously in section 9), also relate strongly to “final year academic test scores”. This chain of evidence suggests that teachers’ confident attitudes about their own competence with arts/arts integration portfolio practices is linked substantially to student academic performance.
- A positive profile of averaged teacher PD and performance assessment outcomes corresponds with a high degree of certainty to a high amount of student portfolio work. This finding is another indication of the success of the PDP professional development program taking root in arts learning classrooms in ways that support increases in academic achievement over time.
- There is a close relationship between student “performance assessment interview” and “portfolio conference” ratings, indicating that these two measures provide validated alternative assessments of student’s understanding of arts/arts integration learning processes and their possible impact on academic performance.

Stepwise regression techniques were used in PDP analysis to determine which variables emerged as leading predictors of student academic achievement in comparison to other competing factors, including student demographic data. Significant and highly significant regression factors (thin and thick solid arrows) in Figure 13 indicate that:

- Long-term baseline-to-final year test academic score results are predicted primarily by “teacher participation in the PDP professional learning events” and the “ability of teachers to get students to generate substantial amount of student portfolio work.” As indicated in Table 21, no other variables come close to that level of influence.
- The culminating year academic results are predicted primarily by “student quality of portfolio work” and to a lesser extent by teacher PD attendance, teacher self reports regarding the success of their classroom arts/arts integration portfolio practices. The relative importance of student family income factors reminds us of the difficulty of any education intervention to transcend the influence of poverty.