

Arts Achieve, Impacting Student Success in the Arts: Preliminary Findings After One Year of Implementation

Tara M. Mastrorilli, Susanne Harnett, Jing Zhu

Abstract

The *Arts Achieve: Impacting Student Success in the Arts* project involves a partnership between the New York City Department of Education (NYCDOE) and five of the city's premier arts organizations. *Arts Achieve* provides intensive and targeted professional development to arts teachers over a three-year period. The goal of the project is to improve the quality of arts teachers' instruction through in-service professional development on the use of balanced (formative and summative) assessment, leading to increases in students' arts achievement. Starting in the 2011-2012 school year, arts teachers formed art discipline-based professional learning communities (PLCs) to work together, using a process of inquiry and action research that focuses on reviewing student data and examining impact on current instructional practice. Additionally, each arts teacher was paired with a facilitator from the arts organizations to support them over the course of the project. The specific professional development activities included: on-site consultancies, assessment retreats, inter-visitations, and an online community. *Arts Achieve* also provides participating arts teachers with resources to support this work, such as units of study and technology bundles.

To measure the impact of the *Arts Achieve* project on arts teachers and students, Metis Associates designed a cluster randomized control trial study, whereby 77 schools were assigned to treatment or status-quo control conditions by arts discipline (dance, music, theater, visual arts) and school level (elementary, middle, high). In the planning year of the project, Benchmark Arts Assessments were developed in each arts discipline and school level to measure students' arts achievement. Findings from Year 1 indicate that, while there were not statistically significant differences between the growth of treatment and control teachers, the students of treatment teachers demonstrated significantly greater growth in arts achievement from the students of control teachers. The results suggest that a more sensitive tool for detecting change in teachers is needed. Successes and challenges of project implementation are discussed, and potential areas for additional inquiry in the coming years of the grant also are recommended.

Introduction

Teaching and learning in the arts, like any subject, requires the use of assessment to allow teachers to reflect on students' progress and adjust their instruction to meet students' needs. Assessment also provides students with feedback, which allows them to reflect on their own learning. More specifically, current research highlights the importance of balanced assessment (formative and summative) in instruction to support student learning and increase student achievement (Black & Wiliam, 1998; Hattie & Temberley, 2007; Kluger & DeNisi, 1996). In general, however, the arts lack access to high quality assessments to inform arts teachers about their students' knowledge and skills in a particular art form. The *Arts Achieve: Impacting Student Success in the Arts* project was conceptualized to address this issue and, in doing so, improve teacher instruction and student learning in the arts.

In spring 2010, the *Arts Achieve* project was funded by two United States Department of Education (USDOE) grants: the New York City Department of Education (NYCDOE) Office of Arts and Special Projects (OASP) received an Arts Education Model Development and Dissemination (AEMDD) grant and Studio in a School (STUDIO) received an Investing in Innovation (i3) grant to implement the program. The goal of *Arts Achieve* is to improve the quality of arts teachers' instruction through intensive and targeted professional development on the use of balanced assessment, leading to increases in students' arts achievement. To evaluate the efficacy of the *Arts Achieve* project, the project partners hired Metis Associates, a national evaluation and consulting firm, to conduct an experimental study, currently underway. At the end of the first project year, which was used as a planning year, 77 NYC public schools were assigned to treatment or status-quo control conditions using a stratified (art discipline and school level) random assignment process. In this article, we report data on the preliminary effects of the *Arts Achieve* project on arts teachers' instructional practices, as well as on students' arts achievement, after one full year of implementation.

Theoretic Framework

The logic model, which is displayed in Figure 1, illustrates the theoretical framework of the *Arts Achieve* project. The project is based on the theory that, when arts teachers participate in professional development, including being members of professional learning communities (PLCs), that focuses on action research and the use of balanced assessment strategies, the quality of their arts instruction will improve. In turn, the project hypothesizes that enhanced arts instruction will lead to improved student achievement in the arts. The paragraphs below provide an overview of research on these components, which have guided the work of *Arts Achieve*.

Arts Achieve draws on current research around professional development for educators, which emphasizes the need for PLCs to develop the collective capacity of staff to work together in order to improve teacher practice and student learning. Through PLCs, teachers work together in teams, learning from each other in environments that are both school-based and job embedded (DuFour, Eaker, DuFour, 2005). The development of PLCs in the arts community is particularly needed, given that arts teachers can be the only staff members in their schools teaching in their content area, and many have reported that they do not feel integral to the faculty (Burnaford, 2009).

Several research studies provide robust support regarding the impact of PLCs on teacher practices, school culture, and student achievement. For example, Louis and Marks (1998) conducted a multi-site study in 24 schools on the impact of PLCs. Through classroom observations and interviews with teachers, they documented the presence of authentic pedagogy (i.e., pedagogy that emphasizes higher order thinking, the construction of meaning through

conversation, and the development of depth of knowledge that has value beyond the classroom) and examined the connection between the quality of classroom pedagogy and the existence of core characteristics of a PLC. This study found that the presence of a professional community in a school contributes to higher levels of social support for achievement and higher levels of authentic pedagogy. Bolam, McMahon, Stoll, Thomas, & Wallace (2005) examined survey data from 393 schools and interview-based case study data from 16 school sites. Both the survey and case study data revealed that when teachers work in PLCs, there are fundamental shifts in the ways that they approach their work, increasing their collaboration, reducing their feelings of isolation, and improving their overall morale.

There is also strong evidence that when teachers work in PLCs, their students experience increased achievement. Eight recent studies (Berry, Johnson, & Montgomery, 2005; Bolam et al., 2005; Hollins, McIntyre, DeBose, Hollins, & Towner, 2004; Louis & Marks, 1998; Phillips, 2003; Strahan, 2003; Supovitz, 2002; Supovitz & Christman, 2003) all found positive relationships between teachers' participation in PLCs and student academic achievement. In these studies, results of student achievement gains varied with the strength of the PLC in the school (Bolam et al., 2005; Louis & Marks, 1998) or with the specific focus of the efforts of teams or small communities of teachers (Supovitz, 2002; Supovitz & Christman, 2003).

Underlying all of the work of the PLCs is an emphasis on action research and use of appropriately balanced assessment strategies to review and discuss student work and teacher practice. Action research involves teachers engaging in inquiry and reflection on their current practice and student work. In contrast to one-day professional development sessions, professional development that involves action research is more participant-driven and incorporates inquiry and reflection that occurs over a period of time. Zeichner (2003) notes that these components are in alignment with the standards and guiding principles for professional development as set by national organizations (e.g., the American Federation of Teachers, the National Councils of Teachers of English and Mathematics) and academic scholars (e.g., Darling-Hammond & McLaughlin, 1996).

Studies have found that teachers who engage in action research report that they have higher self-esteem and confidence levels (Dadds, 1995; Loucks-Horsely, Hewson, Love, & Stiles, 1998), develop self-analysis skills that are applied to their teaching (Day, 1984), and become more aware of how they impact their students (Allen, Shockley, & Baumann, 1995). In a review of studies on four action research professional development programs, Zeichner (2003) reported that engaging in action research helps teachers create a more student-centered environment, in which the teachers focus on listening to and observing their students to influence instruction. He explained further that teachers begin to see the point of view of their students and allow them to have more input in the classroom.

In order to inform teachers' action research, balanced assessment is a critical tool to gather evidence on student progress. Current research emphasizes the need for teachers to use assessment as a tool to help gather student performance data and assess how they should target their instruction to meet students' needs (Gewertz, 2010; Stiggins, 2010). Assessment also provides feedback to students in order for them to shape their understanding and improve their learning. Formative assessment helps students answer the question, "How am I doing?" as they are learning new material, while summative assessment answers the question, "How did I do?" at the end of learning a unit (Shute, 2008; Stiggins, 2005).

Recent meta-analyses have documented the effectiveness on the use of assessment practices on student achievement (Black & Wiliam, 1998; Hattie & Temberley, 2007; Kluger &

DeNisi, 1996). Black and Wiliam (1998) examined 250 studies from the research addressing a range of student factors and teacher instructional practices, including formative assessment strategies. They concluded that formative assessment has a more profound effect on learning than do other typical educational interventions, finding effect sizes between 0.4 and 0.7. Moreover, they concluded that assessment practices have a stronger effect on low achieving students than on high achieving ones, as they are instrumental in developing meta-cognitive skills and enhancing motivation. Hattie and Temberley's (2007) research described the results of 12 meta-analyses that included 196 studies and 6,972 effect sizes. Notably, they found that the average effect size for the use of formative feedback was 0.79.

Arts Achieve Project Implementation

The *Arts Achieve* project includes a partnership among the NYCDOE OASP, five of the city's premier arts organizations, and Metis Associates, the project's evaluator. The arts organizations and their particular arts focus for the project are as follows: STUDIO (lead partner, visual arts); ArtsConnection (theater); the Dance Education Laboratory at the 92nd Street Y (dance); the Weill Music Institute at Carnegie Hall (music), and the Cooper Hewitt National Design Museum (technology). Figure 2 provides a description of each of the partner organizations. Prior to the first year of project implementation, the NYCDOE and the partner organizations developed and piloted 12 Benchmark Arts Assessments, one in each arts discipline (dance, music, theater, and visual arts) and school level (elementary, middle, and high). The Benchmark Arts Assessments measure students' arts content knowledge and performance skills based on local arts standards set in the NYCDOE *Blueprints for Teaching and Learning in the Arts*.¹ The development of the Benchmark Arts Assessments is described further in the Methods section.

The *Arts Achieve* project provides professional development to arts teachers over a three-year period to help them engage in action research and to learn to use the data from the Benchmark Arts Assessments and ongoing formative assessments. Starting in the 2011-2012 school year, the arts teachers formed art discipline-based PLCs to work together toward the goal of improving their practice and student learning. Teaching artists from the arts organizations were partnered with the participating arts teachers to work together over the course of the project. In contrast to an artist-in-residence program, the teaching artist's role is that of a facilitator, who helps support the arts teacher with project work, as well as plan for instruction and provide assistance in the classroom. The *Arts Achieve* in-service professional development includes:

- *On-site Consultancies*: Each facilitator visits his/her partner school about twice a month over the course of the school year, for a total of 22 visits over the course of the school year. During each of these visits, the facilitator and arts teacher work directly with two of the teacher's targeted art classes² and participate in a separate joint instructional planning meeting. While the specific activities during each visit are unique to the needs of the arts teacher, the facilitator's time in the classroom may include observing, modeling, or co-teaching. During the planning time, the arts teacher and facilitator may discuss the teacher's action research and frequently review student work, examine formative and summative assessment data, and discuss the instructional implications of data.
- *Assessment retreats*: The facilitators and arts teachers meet three times over the course of the school year as a full group to focus on the appropriate use of formative and summative assessment strategies and the use of data from the

Benchmark Arts Assessments to inform instruction. The assessment retreats are facilitated by Dr. Heidi Andrade, an expert in formative assessment practices. These meetings also provide an opportunity for participants to meet within their discipline and grade level specific PLCs.

- *Inter-visitations*: Two school inter-visitations per year provide additional opportunities for participants to meet within their PLCs. During these inter-visitations, the arts teachers deepen the work of the PLCs by discussing student work and sharing best practices around instruction and use of formative assessment strategies. The initial visit was to a model school, while all subsequent visits have been to other treatment schools.
- *Online Community*: To support ongoing collaboration and share ideas and resources, a social-networking site was created for the project using Ning. This online community allows the *Arts Achieve* participants to share student work by posting pictures or videos, to upload units of study or links to websites, to network with other teachers in the project, to make announcements, and to archive project documents.

In addition, the *Arts Achieve* project provides participating arts teachers with resources to support this work. Resources include *Blueprint*-aligned exemplary units of study, links to websites and other sources, their students' results from the Benchmark Arts Assessments, and technology bundles. The technology bundles include three iPads and a projector to use in their classrooms. The arts teachers participated in a full-day of training on how to operate and incorporate the iPads into their instruction. Through the incorporation of these additional resources, teachers will have increased access to arts content, formative and summative assessment strategies, and instructional strategies designed to reach students through multiple venues.

The *Arts Achieve* project also provides the participating facilitators with additional training each year of the project. In the first year, three days of training were held—two prior to the start of implementation in the schools and one during the school year. In subsequent years, additional training days have been provided. During these trainings, the facilitators are provided with techniques and strategies on how to work with their partner arts teachers, apply formative assessment in the classroom, and use student data from formative assessments and the Benchmark Arts Assessments to improve instructional practice.

Purpose of Study

The literature on the positive impact of formative assessment, action research, and PLCs on teacher instruction and student achievement, as described above, provides ample evidence to support the *Arts Achieve* framework. However, the literature is limited in that it does not provide evidence on specific impact on arts teacher instruction or student achievement in the arts. The purpose of this study is to extend and deepen this previous research, and to examine whether, and to what extent, providing teachers with high quality and intensive professional development can positively impact arts teachers and their students.

Overall, the *Arts Achieve* study addresses the following three research questions:

1. What is the nature of *Arts Achieve* implementation? What are the successes and challenges of *Arts Achieve* implementation?
2. What is the impact of the *Arts Achieve* project on arts teachers' *Blueprint* knowledge and instructional practices, including their use of the *Blueprint* standards and their interpretation and use of formative and summative assessment data?

3. What is the impact of the *Arts Achieve* project on students' arts achievement? Are there differential impacts of the *Arts Achieve* project by arts discipline and school level?

Methods

To evaluate the impact of the *Arts Achieve* project, Metis Associates designed a cluster randomized control trial, whereby schools were assigned to either the treatment or status-quo control condition. This design allows for comparisons between teachers and students who participate in the treatment with those who do not. This section describes how the sample was selected, the instruments used to collect teacher and student background characteristics and outcomes, and the analyses used to address the study's research questions.

Sample

The *Arts Achieve* sample was selected using a two-stage process. First, NYCDOE public schools that met the basic eligibility requirements were recruited and then randomly selected to participate. To identify eligible schools, information about their arts programs was gathered through the 2009-2010 NYCDOE *Annual Arts Education Survey*.³ To be eligible to participate, schools had to be part of the NYCDOE public school system and have an existing arts program in at least one arts discipline. Additionally, eligible *elementary* schools had to: 1) serve *all* grades 3-5; 2) offer at least 30 hours of instruction in one of the arts disciplines to all grades 3-5; and 3) have a certified or a cluster teacher assigned to teach the arts. Eligible *middle* schools had to: 1) serve *both* grades 7 and 8; 2) offer at least one full year of instruction in an arts discipline to eighth-grade students; and 3) have an arts teacher on staff certified to teach the arts. Eligible *high* schools had to: 1) serve *any* grade 9-12; 2) offer at least one full year of instruction in an arts discipline to students in any grade 9-12; and 3) have an arts teacher on staff certified to teach the arts.

Two informational meetings in winter 2011 were held to describe the program and to recruit volunteer schools to participate in the study. Schools recruited for participation were blocked on two characteristics--arts discipline (dance, music, theater, and visual arts) and school level (elementary, middle, and high schools)--thereby creating 12 blocks. In cases where there were more eligible schools than needed, schools were randomly selected to participate.

In the second stage of the selection process, schools were randomly assigned to the treatment or control group. Within each arts discipline block, eight schools were selected at the elementary level, six at the middle school level, and six at the high school level. At the elementary school level, five schools were assigned to the treatment group and three schools were assigned to the control group. At the middle and high school level, three schools were assigned to the treatment group and three schools were assigned to the control group.

Table 1 displays the number of participating schools, arts teachers, and students who participated in the first implementation year. The data are categorized by arts discipline, school level, and group. A total of 77 schools, 43 treatment and 34 control, participated in the first year of implementation. Because of shortages in the number of eligible dance and theater programs, three schools were in more than one block. One school with both elementary and middle school grades offered instruction in dance and was randomly assigned to the control group at both levels. A middle school that offered instruction in both theater and visual arts was randomly assigned to the theater treatment group and then randomly assigned to the visual arts control group. In this case, the school was pulled out as a visual arts control school. One high school with both dance and theater programs was randomly assigned to both the dance and theater treatment groups. Lastly, one theater treatment elementary school and one theater control high school elected not to participate prior to the start of implementation.

The treatment was targeted mainly toward the school's arts teacher in the arts discipline upon which the school was blocked. Within the treatment schools, the arts teachers' targeted classes in the treatment and control schools were selected based on grade (grade 5 in elementary, grade 8 in middle school,⁴ and grades 9-12 in high school); year-long instruction (all target classes had to participate in year-long arts instruction); and (if multiple classes met these eligible requirements) scheduling convenience. The classes selected for assessment in the control schools were selected using the same criteria. Arts teachers in the treatment schools participated in all the *Arts Achieve* professional development, whereas the arts teachers in the control schools participated only in evaluation activities.

In general, one arts teacher per school participated in the treatment, although there are some schools with more than one participating arts teacher. A total of 79 arts teachers participated in the project, 44 arts teachers from the treatment schools and 35 arts teachers from the control schools. A total of 4,066 students received yearlong instruction in art within the arts teachers' targeted classes, including 2,046 students in the treatment schools and 2,020 students in the control schools. Table 2 displays the demographic profile of participating students by group.

Twenty-four teaching artists, who all had previous experience working in schools, participated in the project as facilitators (six dance, five music, seven theater, and six visual arts). The majority of facilitators were partnered with only one arts teacher; however, some worked with two or three teachers.

Instruments

Ongoing data collection for *Arts Achieve* consists of program documentation, surveys and focus groups with arts teachers, and the Benchmark Arts Assessments. Additionally, secondary data are collected from the NYCDOE, including students' background characteristics, and English Language Arts achievement. This section describes the *Arts Achieve* data collection procedures and instruments.

Program documentation. To measure *Arts Achieve* implementation, program documentation is collected, including professional development materials, participant attendance, and observations of the trainings. Data on the number of on-site consultancies also are collected.

Arts teacher surveys. To collect information on arts teachers' knowledge and instructional skills, Metis and program staff developed surveys and administered them to the participating arts teachers in the treatment and control schools at the beginning (pre) and end (post) of the school year. The surveys ask questions about arts teachers' characteristics, including their years of teaching experience and certification in the designated arts discipline. The surveys also use Likert-scale questions to ask about arts teachers' knowledge and use of the NYCDOE *Blueprints* and use of formative and summative assessment data. Open-ended questions were also included to learn about the successes and challenges of implementing the *Arts Achieve* work in the school. Thirty-six out of forty-four (81.8%) of the treatment arts teachers responded to both the pre- and post-surveys, while 17 out of 35 (48.6%) of the control arts teachers responded to both surveys.

Composite scores of teachers' *Blueprint* knowledge, *Blueprint* use, and formative assessment strategies were calculated using multiple items on the arts teacher pre- and post-surveys. The *Blueprint* knowledge scale consisted of five items and had a Cronbach's alpha value of 0.77 on the post-survey. The *Blueprint* use and use of formative assessment strategies post-survey scales were also found to have high internal consistencies (4 items $\alpha=0.79$, 3 items $\alpha=0.73$, respectively).

Benchmark Arts Assessments. As described above, the Benchmark Arts Assessments were developed in the planning year and were used to measure the impact of *Arts Achieve* on students' arts achievement. Prior to the first year of implementation, teams were assembled to spearhead the creation of the assessments. The teams were led by the NYCDOE Arts Directors in the four arts forms, and also were comprised of school-based arts teachers, staff from the participating arts organizations, and NYCDOE experts in test and measurement. A total of 12 assessments were developed (three [one each for 5th grade, middle school, and high school] in each of the four arts disciplines [dance, music, theater, and visual arts]). The assessments were designed to measure the extent to which students have developed the knowledge and skills that are expected in a particular art form by fifth grade, middle school, and high school, according to the NYCDOE *Blueprints*. They also have been aligned with the Common Core Capacities in English Language Arts.

The Benchmark Arts Assessments each have multiple components, including performance and written sections, and include activities that address content knowledge, transferable concepts, and skills in the designated arts discipline. The assessments allow students to demonstrate their depth of knowledge in the content area through the analysis of performances, evaluation of other masterworks, and the creation of their own works of art. Question types include short-answer, fill-in-the-blank, multiple choice, and performance tasks. The majority of the questions are scored using a four-point scaled rubric.

Prior to each administration of the assessments, *Arts Achieve* staff provided training for adjudicators in administering and scoring the assessments. The adjudicators were current or retired NYCDOE arts teachers or facilitators. Separate training was conducted for each of the arts disciplines, but each included background on the context and purpose of the Benchmark Arts Assessments and training on the rubrics to score student work. Two adjudicators each administered and scored the assessments.

In spring 2011, the assessments were piloted in a sample of NYCDOE public schools that met the same requirements as the study schools, and the psychometric properties of the assessments were analyzed. Based on the results of the pilot, the assessments were found to have acceptable levels of reliability and validity. The assessments are administered to students in participating classes in the treatment and control schools in both the beginning (pre) and end (post) of the school year. Total scores for each of the assessments use a scale of 0 to 100.

Table 3 displays the internal consistencies of each of the spring 2012 Benchmark Arts Assessments. The Cronbach's alpha values range from 0.72 on the elementary school *music-vocal* assessment to 0.89 on the high school *theater-direction* assessment, indicating that the internal consistencies are acceptable. Inter-rater reliabilities were calculated for each task scored using a rubric on each of the assessments. The Kappa values ranged from 0.13 on the high school dance assessment, group dance performance collaboration task to 1.00 on the music elementary assessment, vocal and instrumental performance tasks. The appendix presents the reliabilities for each task.

Secondary data. Data on treatment and control students' characteristics were obtained from the NYCDOE to use as covariates in the analysis models. Student characteristics include student demographics (including gender, race/ethnicity, English language learner status, special education status, free/reduced priced lunch status), and average daily attendance. Students' scores on the New York State English Language Arts (NYS ELA) exam also were collected as a measure of student academic achievement. In elementary school and middle school, the NYS

ELA assessment is administered to all students in grades three through eight in the spring of each school year.

Data Analysis

Descriptive statistics were calculated, including frequencies, means, and standard deviations, on the attendance, survey, and arts achievement data. Multiple regression analyses were conducted to measure the impact of the *Arts Achieve* project on arts teachers' knowledge and instructional practice scores, as measured by the arts teacher post-survey and students' arts achievement scores on the post-Benchmark Arts Assessments. Potential confounding factors or covariates that have a relationship with the treatment or outcome are included in each regression model to reduce threats to the study's internal validity. Covariates included in the teacher outcome models included arts teachers' years of experience, arts discipline certification, and pre-survey composite scores. In the student arts achievement models, student demographics (i.e., gender, race, English language learner status, special education status, free and reduced price lunch status, and average daily attendance) and prior achievement (i.e., spring 2011 NYS ELA exam score and pre-Benchmark Arts Assessment score) were used as covariates. All covariates included in the multiple regression models were grand mean centered. For statistically significant impact effects, Glass's delta was calculated to obtain the impact's effect size.

Results

As stated above, the *Arts Achieve* project is based on the theory that participation in the *Arts Achieve* professional development will have positive impacts on arts teachers' knowledge and instructional skills and, in turn, on students' arts achievement. This section presents preliminary implementation and impact findings of the project after one full year of implementation.

Implementation Findings

As described earlier, the participating arts teachers and their partner teaching artists, or facilitators, formed PLCs to work together to engage in action research and to learn to use data from formative assessments and the Benchmark Arts Assessments. The facilitators collaborated with their partner arts teacher to identify gaps in student learning and devise an action plan that would address these gaps. The facilitator played the role of "critical friend" to the arts teachers, and did not necessarily work with the students directly (though this may have been an element of the work). During the first year of implementation, facilitators were expected to conduct a total of 22 on-site consultancies over the course of the school year. Each consultancy included an extended planning session with the teacher to discuss strategies, examine data, debrief on previous meetings, and review student work.

Table 4 displays the average number of *Arts Achieve* on-site consultancies conducted by facilitators in Year 1. The data are presented by arts discipline and school level. Across all treatment schools, the facilitators and arts teachers conducted an average of 19.6 (SD=2.25) consultancies, which is slightly less than the expected 22. The average number of consultancies ranged from 18.0 (SD=1.53) in the music middle schools to 21.3 (SD=1.00) in the dance and visual arts high schools. Facilitators explained that scheduling the consultancies with their partner arts teachers was often a challenge due to other commitments on both of their parts. Facilitators often had restrictions on the days that they could visit due to residency commitments in other schools, and arts teachers often had other commitments during the school day as well. Additionally, the art schedule in the buildings often was disrupted due to testing, school trips, and other activities.

Despite the challenges in scheduling, the consultancies, data collected through surveys and focus groups indicate that the arts teachers and facilitators were largely successful in fostering strong partnerships. Arts teachers were highly positive overall about their facilitators and appreciated the opportunity to collaborate with other artists. For example, on an end-of-year survey, one teacher wrote, “My partner facilitator was very helpful, encouraging, reflective, intelligent, and I couldn’t imagine how she could have been more helpful. She was a great partner...[a] first rate artist, instructor, [and] mentor.” Another teacher explained that her facilitator was an excellent resource for planning instruction. The consultancy planning time allowed her opportunities to bounce ideas off a fellow artist educator. These types of opportunities, in general, were perceived as critical, given that the great majority of arts teachers have no school-based peers with whom to collaborate. The only critique that many of the participating arts teachers expressed was that they wished they had more time to meet and plan with their facilitators.

In addition to the consultancies, the Arts Achieve project allows for other opportunities for the facilitators and their partner teachers to collaborate. In Year 1, these opportunities included three full-day assessment retreats and two school inter-visitations. Additionally, there were three training days designed for facilitators only. Table 5 displays participants’ attendance at these professional development activities in Year 1. The majority of arts teachers and facilitators participated in all four assessment retreats over the course of the year; however, the great majority of arts teachers and facilitators did not attend both of the inter-visitations scheduled during the first year. While most teachers attended the first inter-visitation in fall 2011 to observe best practices in a model school, a smaller number with scheduling conflicts were able to attend the spring inter-visitation at one of their peer’s schools. Despite scheduling challenges, both facilitators and arts teachers reported that these professional development opportunities were extremely useful in developing their skills and knowledge in formative assessment practices. Arts teachers valued the time to meet in their PLCs and to continue the planning that they had worked on during their consultancies. Likewise, facilitators provided positive feedback on the professional development sessions targeted for them specifically. Initially, many felt challenged by their new role as facilitator in participating schools and struggled to develop trusting and productive relationships with their partner teachers. These sessions allowed them to form their own PLCs and to share successes and challenges in their work with teachers, as well as to learn the expectations and best practices associated with their role as facilitator.

Impact Findings

Arts teacher knowledge and instructional practice. Changes in arts teacher knowledge and instructional practice were measured through closed- and open-ended items on pre and post self-report surveys. Figure 3 displays data from the closed-ended items, including treatment and control arts teachers’ perceptions about their content knowledge and instructional practices at pre- and post-administration. The data show that after one year of implementation, arts teachers in both the treatment and control schools made some growth in their instructional knowledge and skills over the course of the school year. From pre- to post-survey administration, arts teachers in the treatment schools had slight increases in their perceptions of their *Blueprint* knowledge ($M_{pre} = 3.48$, $SD_{pre} = 0.41$; $M_{post} = 3.55$, $SD_{post} = 0.35$) and *Blueprint* use in instruction ($M_{pre} = 3.37$, $SD_{pre} = 0.96$; $M_{post} = 4.15$, $SD_{post} = 0.81$). Similarly, arts teachers in the treatment schools reported more use of formative assessment strategies ($M_{pre} = 3.99$, $SD_{pre} = 1.05$; $M_{post} = 4.95$, $SD_{post} = 0.84$) and a better understanding of the importance of reviewing and analyzing data for instruction ($M_{pre} = 6.79$, $SD_{pre} = 2.99$; $M_{post} = 7.72$, $SD_{post} = 2.30$) over the course of the school

year. However, when compared to gains made by the arts teachers in the control schools, the results of multiple regression analyses indicated that there was no statistically significant ($p > .05$) effect of treatment on arts teachers' knowledge or instructional practices, after controlling for teachers' years of experience and arts certification.

While the regression analyses did not find statistically significant differences between the treatment and control groups, data from the open-ended portions of the treatment teachers' post-surveys provide evidence to support the impact of *Arts Achieve* on teachers' knowledge and skills. One teacher explained that the project was allowing her to perfect her craft as a teacher. Another teacher explained that he was more focused. This teacher elaborated that through his review of the data, he knew where his students' weaknesses were and that they were aligned with the instructional areas on which he had not placed focus over the course of the year. He further explained that he planned to re-focus his instruction in the coming year. Another teacher summarized her learning by saying, "I am very grateful to be involved with Arts Achieve. I have learned a lot about my teaching, and am so excited that there are so many ways that I am improving my craft," and elaborated that she looked forward to continuing her growth over the years of the project.

Student arts achievement. Multiple regression analyses were conducted to measure the impact of the *Arts Achieve* project on students' arts achievement after controlling for student demographics (i.e., gender, race, English language learner status, special education status, free and reduced priced lunch status, and average daily attendance) and previous achievement (i.e., spring 2011 NYS ELA exam score and pre-Benchmark Arts Assessment score). Table 6 displays the results of the multiple regression analysis on students' arts achievement across all arts disciplines and school levels. The results indicate that, after controlling for differences in demographics and previous achievement, the participating students in the treatment schools performed 3.11 points higher on the post-Benchmark Arts Assessments than participating students in the control schools after one year of implementation. The predicted mean post-Benchmark Arts Assessment score for the average treatment students is 62.7 as compared to 59.6 for the average control student. The Glass's delta indicated that there was a small effect size of 0.18.

To determine if there were differential impacts by arts discipline and school level, separate regression models were conducted after selecting the appropriate subsample. Table 7 displays the results of the multiple regression analyses on students' arts achievement by arts discipline. After controlling for demographic and previous achievement differences, the results indicate that, across all grade levels, participating students in the treatment schools scored higher on the post-Benchmark Arts Assessments than participating students in the control schools in music, theater, and visual arts. Across all school levels, the theater schools had the largest statistically significant ($p < .05$) treatment effect with an effect size of 0.42. Participating students in the theater treatment schools performed 7.44 points higher on the post-Benchmark Arts Assessments than participating students in the control schools. The predicted mean post-Benchmark Arts Assessment score for the average treatment theater students is 60.3 as compared to 52.8 for the average control student. The music and visual arts treatment effects also were statistically significant ($p < .05$) with effect sizes of 0.26 and 0.21, respectively. The average treatment music student had a predicted mean post-Benchmark Arts Assessment score of 68.5, whereas the average control music student mean is 64.6. The average treatment visual arts student had a predicted mean post-Benchmark Arts Assessment score of 56.7, whereas the

average control music student mean is 53.85. The dance treatment effect is approaching significance, $p = .05$, with a Glass's delta of 0.11.

Table 8 displays the results of the multiple regression analyses on students' arts achievement by school level. Across arts disciplines, the results of the multiple regression analyses indicate that participating students in the treatment elementary and high schools scored higher on the post-Benchmark Arts Assessments as compared to students in the control elementary and high schools. Participating students in the treatment elementary schools performed 3.79 points higher on the post-Benchmark Arts Assessments than participating students in the control elementary schools. Participating students in the treatment high schools performed 6.66 points higher on the post-Benchmark Arts Assessments than participating students in the control high schools. The elementary and high school treatment effects were statistically significant ($p < .05$) with effect sizes of 0.25 and 0.34, respectively. The treatment effect for participating students in the middle schools was not statistically significant ($p > .05$).

Discussion

The *Arts Achieve* project was designed with the goal of improving the quality of arts teachers' instruction through intensive and targeted professional development, which, in turn, is intended to lead to an increase in student arts achievement. After the first of three years of implementation, preliminary results provide promising indication of the impact of the project on student arts achievement.

Arts Achieve was developed based on literature that connects improved student achievement with increased teacher use of formative and summative assessment and data-driven decision making practices. While evidence on the effectiveness of these practices is abundant across education, little (if any) research on its impact on arts teachers exists. The reason for this lies in the fact that there are few assessments in the arts and, as such, arts teachers tend not to use data to guide their instruction. Furthermore, arts teachers generally have not been trained in the use of data-driven decision making.

The *Arts Achieve* project began with the intense, but ultimately rewarding, process of developing arts assessments. During the first project year (which was used for planning and development), 12 Benchmark Arts Assessments were created, one for each of the four arts disciplines and three school levels. The Benchmark Arts Assessments are aligned with the NYC DOE's *Blueprints* and represent an important step in defining and measuring expectations for teaching and learning in the arts. The assessments, which largely have been found to have acceptable levels of reliability and validity, mark an important deliverable for the grant.

However, the ambitious *Arts Achieve* project aimed to do much more than simply create the assessments. The project aimed to provide intensive professional development for teachers, which includes the use of data from the Benchmark Arts Assessments, to help them shape their instruction, and ultimately improve student arts achievement.

This article provides data on implementation and impact after one full year of implementation according to the three main research questions addressed by the *Arts Achieve* study, including: 1) successes and challenges of implementation; 2) impact of the project on teachers' knowledge and skills; and 3) impact of the project on students' arts achievement.

Successes and Challenges

As described above, a major success of the project was the creation of 12 reliable and valid Benchmark Arts Assessments. Additionally, strong partnerships between arts teachers and facilitators were forged, fostering growth for both groups. The project required both groups to move outside of their comfort zone—facilitators had to play a role in schools that was often

brand new to them, while teachers had to learn to dig into data and to reflect on their instructional implications. The pairings had to develop strong, trusting relationships that would ultimately lead to productive change in student learning.

Challenges confronted were not atypical to programs implemented in urban school environments. Teachers in general—and arts teachers are no exception—have extreme demands on their time. There are constant distractions within the schools, and oftentimes, art can take a backseat to content areas that are subject to high-stakes testing. As a result, fewer than expected on-site consultancies took place over the course of the year, and a small percentage of arts teachers and facilitators were able to participate in both of the planned school inter-visitations. Furthermore, while most of the partnerships were strong, a few suffered from trust issues, teacher reluctance to change their practices, and/or personality conflicts.

Based on the findings from the first year of implementation, the project has made some adjustments to the design, including fewer in-school visits and more assessment retreat time. Additionally, the role of the facilitator has been better communicated to participants, and teachers and facilitators were given more time in the second year of implementation to strengthen their relationships, as well as to work together on action research.

Impact on Arts Teachers' Knowledge and Skills

Impact of the project on arts teachers' knowledge and skills was measured through self-report surveys, which were distributed to teachers in both the treatment and control groups on a pre and post basis. While teachers indicated in open-ended portions of the survey that the project was highly impactful to them, results of multiple regression that examined differences in growth between treatment and control teachers did not detect statistically significant differences. Given the contrast in findings, it is possible that the scales designed to measure teacher knowledge and skills are not sensitive enough to detect change over the course of the school year. It is also possible that additional implementation time is needed to detect an impact in teacher growth.

Impact on Student Arts Achievement

Findings from the analyses of student arts achievement indicated that, across all grade levels and arts disciplines, students who received yearlong arts instruction in the treatment schools demonstrated greater growth in their arts achievement than students who received yearlong arts instruction in the control schools. This finding supports the hypothesis that the work with teachers ultimately leads to improvement in student growth.

Interestingly, however, given that our theory of change is that professional development will positively impact teacher knowledge and skills that, in turn, will lead to increases in student arts achievement, the fact that there are significant positive effects for students' arts achievement but not for teacher growth, further suggests that the measure through which teacher growth is assessed is not sensitive enough to detect change. In coming years, teacher growth will be explored in greater depth, and additional tools, including qualitative methods, will be used to gather data.

It is also notable that there were differential impacts on student achievement according to arts discipline and school level. These results will be explored in greater detail in coming years to examine whether patterns persist and, if so, what factors are contributing to the differential impact.

Overall, findings from the first year of implementation of *Arts Achieve* provide highly promising results. Though it is still in progress, the project has already produced assessments that have the potential to be valuable to the field. Furthermore, the study provides strong evidence that increasing arts teachers' knowledge and skills in formative and summative assessment and

the use of data leads to improvement in student arts achievement. It has furthermore raised new questions and areas of inquiry into which the field may explore in the future.

References

- Allen, J.B., Shockley, B.B., Baumann, J.F. (1995). Gathering 'round the kitchen table': teacher inquiry in the NRRC school research consortium. *Reading Teacher*, 48, 526-529.
- Berry, B., Johnson, D., & Montgomery, D. (2005). The power of teacher leadership. *Educational Leadership*, 62, 56-60.
- Black, P. & Wiliam, D. (1998). *Inside the Black Box: Raising Standards Through Classroom Assessment*. *Phi Delta Kappan*, 80, 139-148.
- Bolam, R., McMahon, A., Stoll, L., Thomas, S., & Wallace, M. (2005). *Creating and sustaining effective professional learning communities*. University of Bristol, UK.
- 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, 1-45.
- Dadds, M. (1995). *Passionate Inquiry and School Development*. Falmer Press: London, UK.
- Day, C. (1984). Teachers' thinking intentions and practice: An action research perspective. In R. Halkes & J. Olson (Eds.) *Teacher Thinking*. Swets & Zeitlinger: Lisse, NL.
- DuFour, R., Eaker, R. & DuFour, R. (2005). *On common ground: The power of professional learning communities*. Bloomington, IN: Solution Tree
- Gewertz, C. (2010). Experts Lay Out Vision for Future Assessments. *Education Week*, 29 (23). Retrieved from http://www.edweek.org/ew/articles/2010/02/23/23assessment_h29.html?tkn=VY%5bFA3BBpVhGN9TxBWdYT%2Byv0d9k%2Fn0QJf9z&cmp=clpedweek.
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research* 77.1, 81-112.
- Hollins, E.R., McIntyre, LR, DeBose, C., Hollins, K.S., & Towner, A. (2004). Promoting a self-sustaining learning community: Investigating an internal model for teacher development. *International Journal of Qualitative Studies in Education*, 17, 247-264.
- Kluger & DeNisi (1996). The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. *Journal of Applied Psychology*, 254-284.
- Loucks-Horsely, S., Hewson, P., Love, P. & Stiles, K. (1998) *Designing Professional Development for Teachers of Science and Mathematics*. Corwin Press: Thousand Oaks, CA.
- Louis, K.S. & Marks, H.M. (1998). Does professional learning community affect the classroom? Teachers' work and student experiences in restructuring schools. *American Journal of Education*, 106, 532-575.
- Phillips, J. (2003). Powerful learning: Creating learning communities in urban school reform. *Journal of Curriculum and Supervision*, 18, 240-258.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78, 153-189.
- Stiggins, R. (2010). Essential formative assessment competencies for teachers and school leaders. *Handbook of Formative Assessment*. New York: Routledge
- Strahan, D. (2003). Promoting a collaborative professional culture in three elementary schools that have beaten the odds. *The Elementary School Journal*, 104, 127-146.
- Supovitz, J. A. (2002). Developing communities of instructional practice. *Teachers College Record*, 104, 1591-1626.
- Supovitz, J. A., & Christman, J. B. (2003). Developing communities of instructional practice: Lessons from Cincinnati and Philadelphia. *CPRE Policy Briefs*, #RB-39.

Zeichner, K. (2003). Teacher research as professional development for p-12 educators in the U.S. *Educational Action Research*, 11, 301-325.

Notes

¹ The *Blueprints* set clear standards for what students should know, understand, and be able to do in each of the four arts disciplines (dance, music, theater, and visual arts) as they move through the school system from Pre-K through 12th grade. They are based on National arts standards and support the NYS Standards for Arts Instruction. Scope and sequence of learning are identified on the *Blueprints* through five strands: Art Making, Literacy in the Arts, Making Connections, Community and Cultural Resources, and Careers and Lifelong Learning. Benchmarks are delineated at four levels – Grades two, five, eight, and 12.

² In elementary schools, fifth-grade classes are targeted. In middle schools, eighth-grade classes are targeted, though some sixth- and seventh-grade classes may be included. In high school, the targeted classes can include any grade from nine through 12.

³ Each spring all, NYCDOE schools are asked to complete the *Annual Arts Education Survey*. The survey collects information on schools' arts programs. More specifically, the survey includes questions about student participation in arts courses by arts discipline, arts activities conducted, and certified school-based arts teachers.

⁴ In some cases, if eighth grade could not be targeted, the treatment was geared toward grade six or seven.

Table 1

Number of Arts Achieve Schools, Teachers, and Students by Arts Discipline, School Level, and Group

Art Discipline	School Level	Treatment			Control		
		N Schools	N Arts Teachers	N Students	N Schools	N Arts Teachers	N Students
Dance	Elementary	3	5	212	3 ^a	3	193
	Middle	3	3	117	3 ^a	3	160
	High	3 ^b	3	98	3	3	135
<i>Dance Subtotal</i>		<i>11</i>	<i>11</i>	<i>427</i>	<i>9</i>	<i>8^c</i>	<i>488</i>
Music	Elementary	5	5	185	3	3	209
	Middle	3	3	195	3	3	289
	High	3	3	181	3	5	186
<i>Music Subtotal</i>		<i>11</i>	<i>11</i>	<i>561</i>	<i>9</i>	<i>11</i>	<i>684</i>
Theater	Elementary	4 ^d	4	238	3	3	128
	Middle	3	3	66 ^e	3	3	150
	High	3 ^e	4	103	2 ^f	2	92
<i>Theater Subtotal</i>		<i>10</i>	<i>11</i>	<i>407</i>	<i>8</i>	<i>8</i>	<i>370</i>
Visual Arts	Elementary	5	5	340	3	3	225
	Middle	3	3	153	2 ^g	2	91
	High	3	3	158	3	3	162
<i>Visual Arts Subtotal</i>		<i>11</i>	<i>11</i>	<i>651</i>	<i>8</i>	<i>8</i>	<i>478</i>
Overall Total		43	44	2,046	34	35	2,020

^a Although there are three control elementary schools with a dance program and three control middle schools with a dance program, one school is included in both groups. Consequently, there are a total of eight control schools with a dance program, not nine schools.

^b Although there are three treatment high schools with a theater program and three treatment high schools with a dance program, one school is included in both groups. Consequently, in the two groups, there are a total of five high schools, not six high schools.

^c Although there are three control elementary school dance teachers and three control middle school dance teachers, one teacher is included in both groups. Consequently, there are a total of eight control school dance teachers, not nine schools.

^d A theater treatment elementary school elected not to participate in the project prior to the start of implementation.

^e A treatment theater middle school did not provide students with yearlong theater instruction, therefore, this number represents the number of students in two of the treatment theater middle schools.

^f A theater control high school elected not to participate in the project prior to the start of implementation.

^g A visual arts control middle school was initially randomly assigned as a treatment theater middle school; therefore the school was pulled out of the study as a control school for visual arts.

To control for diffusion of treatment effects, study schools cannot participate as both treatment and control schools.

Table 2
Demographic Profile of Participating Students by Group

Student Characteristic	Treatment	Control
N Students	2,046	2,020
Female	54.8%	56.5%
Race		
White	14.9%	14.6%
African American	28.6%	31.5%
Asian	18.5%	17.7%
Hispanic	37.0%	35.7%
Other	0.93%	0.45%
English Language Learner	14.0%	8.7%
Special Education	13.9%	11.4%
Free-Reduce Priced Lunch	83.8%	79.3%
Average Daily Attendance (SD)	93.5 (10.58)	92.8 (10.21)
Spring 2011 NYSELA Proficiency ^a	50.4%	55.6%

^a Students' spring 2011 NYSELA proficiency level was not available for all participating students. There were 1,618 treatment students with spring 2011 NYSELA scores and 1,467 control students with scores.

Table 3
Internal Consistencies of the Spring 2012 Benchmark Arts Assessments

Benchmark Arts Assessment					
Art Discipline	School Level	Subtest ^a	N Cases	N Items	α
Dance	Elementary		286	16	0.85
	Middle		214	15	0.88
	High		120	19	0.84
Music	Elementary	Instrumental	88	15	0.76
		Vocal	203	15	0.72
	Middle		319	23	0.86
	High		225	26	0.83
Theater	Elementary	Playwriting	127	12	0.85
		Costume Design	85	13	0.83
	Middle	Acting - Actors	65	9	0.81
		Musical Theater - Actors	25	7	0.77
		Playwriting - Actors	18	7	0.78
		Acting - Design	1	7	–
	High	Musical Theater - Design	5	7	–
		Playwriting - Design	0	7	–
		Direction	132	10	0.89
		Acting	39	11	0.77
Visual Arts	Elementary		441	29	0.85
	Middle		228	25	0.88
	High		183	18	0.87

^a The music elementary assessment and each of the theater assessments asked students to choose between tasks on the assessments. In these cases, the internal consistencies are presented for each of the subtests.

Table 4

Average Number of Arts Achieve On-site Consultancies by Arts Discipline and School Level

Art Discipline	School Level	N Schools	Mean	SD
Dance	Elementary	3	19.4	3.91
	Middle	3	20.7	0.58
	High	3	21.3	1.53
<i>Dance Subtotal</i>		<i>11</i>	<i>20.3</i>	<i>2.72</i>
Music	Elementary	5	20.2	1.48
	Middle	3	18.0	1.00
	High	3	20.0	1.73
<i>Music Subtotal</i>		<i>11</i>	<i>19.5</i>	<i>1.63</i>
Theater	Elementary	4	19.5	1.00
	Middle	3	15.7	2.08
	High	3	19.7	3.06
<i>Theater Subtotal</i>		<i>10</i>	<i>18.4</i>	<i>2.63</i>
Visual Arts	Elementary	5	19.6	1.52
	Middle	3	20.0	1.73
	High	3	21.3	1.53
<i>Visual Arts Subtotal</i>		<i>11</i>	<i>20.2</i>	<i>1.60</i>
Overall Total		43	19.6	2.25

Table 5

Arts Teacher and Facilitator Attendance at Professional Development Training

Professional Development Trainings	Arts Teachers	Facilitators
Assessment Retreats		
Attended All Four Days of Training	56.8%	76.7%
Attended Three Days of Training	22.7%	7.0%
Attended Two Days of Training	18.2%	9.3%
Attended One Day of Training	2.3%	4.7%
Inter-visitations		
Attended All Two Visits	13.6%	18.6%
Attended One Visit	54.5%	58.1%
Facilitator Training		
Attended All Three Days of Training	—	58.1%
Attended Two Days of Training	—	16.3%
Attended One Day of Training	—	18.6%

Table 6
Student Arts Achievement Multiple Regression Results (N=2,726)

Variable	B	B SE
Constant	59.61**	0.39
Pre Arts Achievement Score	0.32**	0.02
Female	2.30**	0.54
Spring 2011 NYSELA Scale Score	0.16**	0.01
Hispanic ^a	-2.53**	0.64
African American ^a	-5.39**	0.68
Average Daily Attendance	0.30**	0.06
English Language Learner	-3.47*	1.05
Free-Reduce Priced Lunch	-1.43*	0.68
Special Education	-1.58	0.84
Grade	0.57**	0.15
Group	3.11**	0.52
R^2		0.33
F		122.07**

^a The comparison group included white, Asian, or other racial backgrounds.

* $p < .05$, ** $p < .001$.

Table 7
Student Arts Achievement Multiple Regression Results by Arts Discipline

Variable	Dance (N=683)		Music (N=821)		Theater (N=529)		Visual Arts (N=693)	
	B	B SE	B	B SE	B	B SE	B	B SE
Constant	62.31**	0.82	64.60**	0.59	52.86**	1.04	53.85**	0.66
Pre Arts Achievement Score	0.49**	0.04	0.46**	0.04	0.33**	0.04	0.38**	0.03
Female	3.60*	1.24	1.29	0.81	2.82*	1.36	2.51*	0.75
Spring 2011 NYSELA Scale Score	0.11**	0.03	0.11**	0.02	0.14**	0.03	0.15**	0.02
Hispanic ^a	-2.72	1.38	0.28	0.94	-1.25	1.83	-2.48*	0.87
African American ^a	-4.67*	1.57	-1.02	1.18	-1.29	1.65	-5.57**	0.99
Average Daily Attendance	0.17	0.11	0.55**	0.10	0.23	0.12	0.15	0.09
English Language Learner	-3.91	2.28	-2.22	1.85	-0.50	3.21	-1.66	1.26
Free-Reduce Priced Lunch	-3.23*	1.47	1.40	0.98	0.94	1.65	-0.68	1.17
Special Education	-0.39	1.86	-5.77**	1.34	0.13	1.96	0.69	1.20
Grade	0.27	0.35	0.40	0.25	0.13	0.38	-0.97**	0.22
Group	2.27	1.16	3.88**	0.86	7.44**	1.40	2.84**	0.77
R^2		0.42		0.41		0.29		0.49
F		44.93**		52.83**		20.50**		61.92**

^a The comparison group included white, Asian, or other racial backgrounds.

* $p < .05$, ** $p < .001$.

Table 8
Student Arts Achievement Multiple Regression Results by School Level

Variable	Elementary (N=1,388)		Middle (N=874)		High (N=464)	
	B	B SE	B	B SE	B	B SE
Constant	58.85**	0.47	65.16**	0.68	51.58**	1.26
Pre Arts Achievement Score	0.28**	0.02	0.23**	0.03	0.48**	0.05
Female	1.63*	0.64	1.91	1.01	6.44**	1.74
Spring 2011 NYSELA Scale Score	0.14**	0.01	0.29**	0.03	0.12*	0.03
Hispanic ^a	-1.95*	0.77	-2.02	1.11	0.43	2.41
African American ^a	-4.14**	0.83	-6.06**	1.44	-3.91	2.22
Average Daily Attendance	0.22*	0.08	0.34*	0.12	0.43**	0.10
English Language Learner	-2.85*	1.14	-6.26*	2.68	0.24	3.08
Free-Reduce Priced Lunch	0.96	0.92	-1.45	1.11	-6.17*	2.16
Special Education	-1.482	0.93	-3.14	1.79	0.77	2.77
Grade	— ^b	—	-1.52	0.80	2.38*	0.74
Group	3.79**	0.64	0.12	1.04	6.66**	1.57
<i>R</i> ²		0.32		0.33		0.38
<i>F</i>		65.35**		40.09**		26.52**

^a The comparison group included white, Asian, or other racial backgrounds.

^b Grade was not included in the elementary model, given that all students in the subsample were in fifth grade.

p* < .05, *p* < .001.

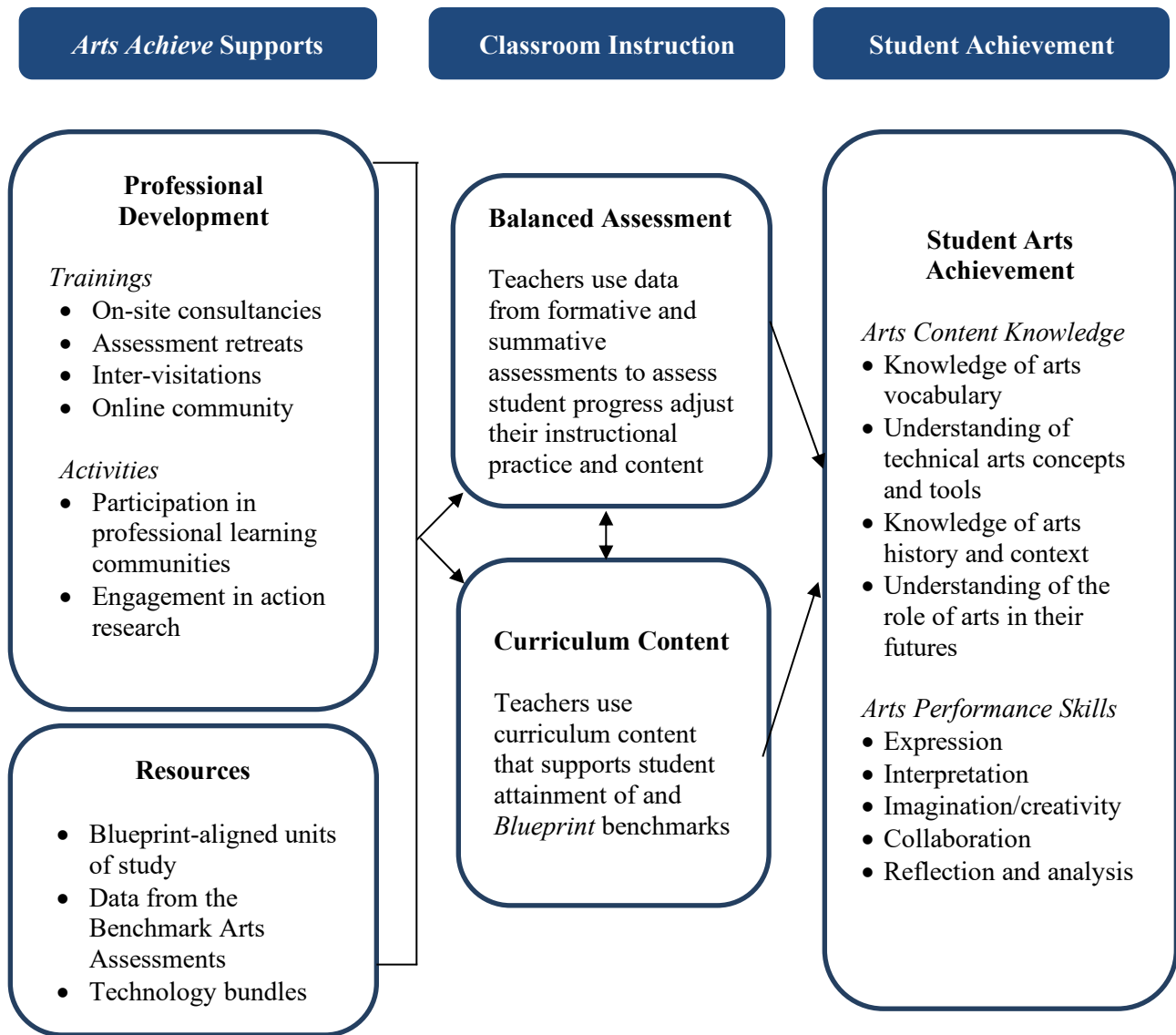


Figure 1. *Arts Achieve* Logic Model.



The New York City Department of Education (NYCDOE) consists of over 1,700 schools that serve 1.1 million students each year. The NYCODE Office of Arts and Special Projects (OASP) supports universal access to arts education and increased quality in the arts through supports for school leaders, the development of curriculum and professional development for teachers of dance, music, theater, visual art, and the moving image.



Founded in 1979, ArtsConnection provides in-school and after school programs taught by professional teaching artists in dance, music, theater, and visual arts in NYC schools.

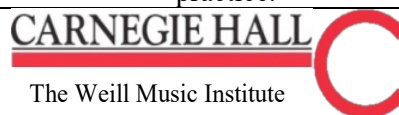


For twenty years, the 92Y Dance Education Laboratory has provided a professional development program to dance educators nationwide interested in developing their teaching practice.



Lead Partner

For 36 years our community of professional artists has been teaching visual arts to hundreds of thousands of New York City children in under-served public schools, daycare centers, and community-based organizations. We offer quality art workshops where children explore their creativity and learn the joys and benefits of making art. Studio also collaborates with and develops the ability of those who provide or support arts programming and creative development for youth both in and outside of schools.



The Weill Music Institute creates visionary programs that embody Carnegie Hall's commitment to music education. These programs inspire audiences of all ages and nurture tomorrow's musical talent, reaching more than 400,000 people each year through national and international partnerships, in New York City schools and community settings, and at Carnegie Hall.



Smithsonian
Cooper-Hewitt, National Design Museum

Smithsonian's Cooper-Hewitt, National Design Museum is the only museum in the nation devoted exclusively to historic and contemporary design. The Museum presents compelling perspectives on the impact of design on daily life through active educational and curatorial programming.



metis associates
making a meaningful difference

Metis Associates is a national research and consulting firm headquartered in New York City. Metis has over 35 years of experience in education-based evaluation, working with a wide range of organizations committed to making a meaningful difference in the lives of children, families, and communities.

Figure 2. Arts Achieve Arts Partners.

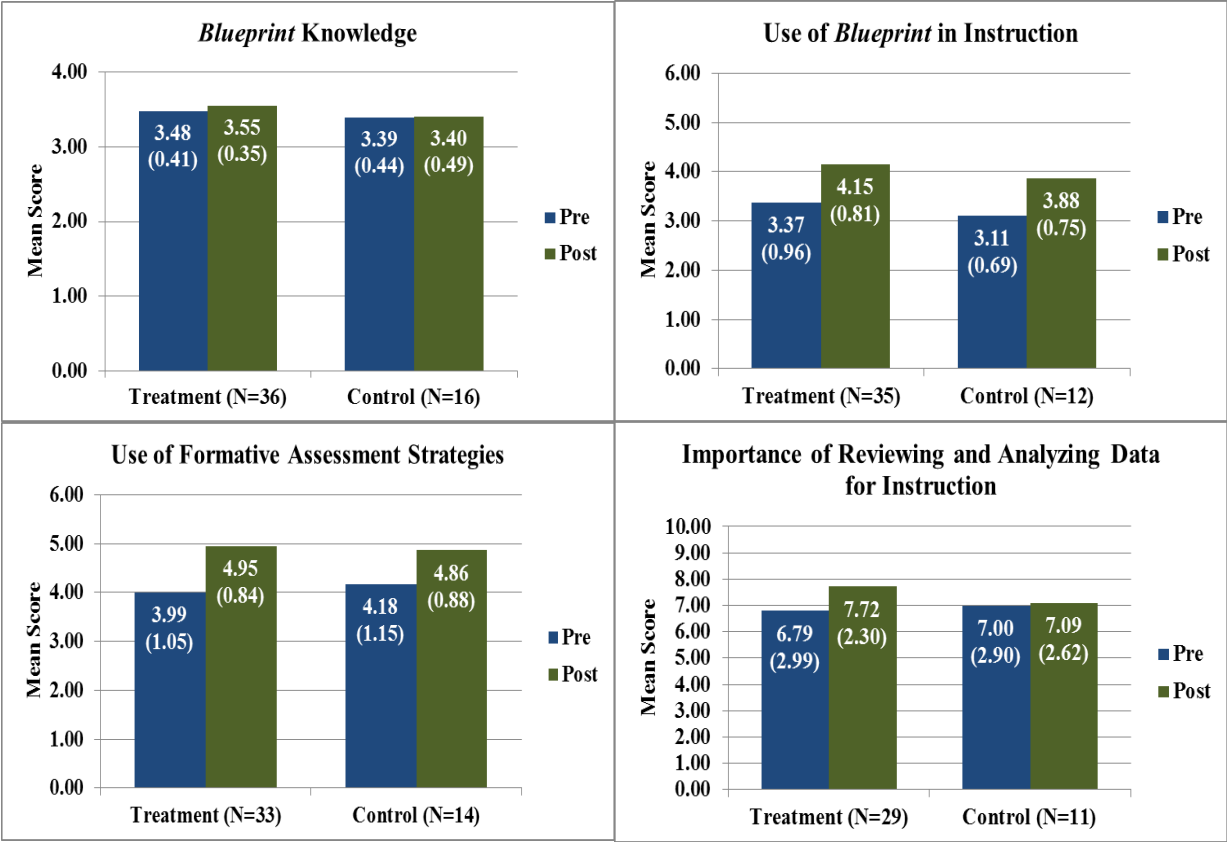


Figure 3. Arts Teachers' Perceptions of Their Content Knowledge and Instructional Practices from Pre- to Post-Survey. Note: The y-axes represent the response scale range. Standard deviations are presented in parentheses below the mean score.

Appendix

Table A1
Inter-rater Reliabilities of the Spring 2012 Fifth-Grade Benchmark Dance Assessments

Task Number	Task and Description		N Cases	Kappa
1.1	Stylistic Hallmarks in Dance: Written interpretation of a masterwork dance performance	Content	399	0.58**
1.2		Comprehension	399	0.62**
1.3		Usage	399	0.67**
2.1	Analysis of Dance Elements: Analyzing dance elements in a masterwork dance prompt	Body	398	0.74**
2.2		Dynamics	391	0.87**
2.3		Space	397	0.72**
2.4		Relationship	385	0.74**
3.1	Dance Styles and Genres: Comparing and contrasting dance in other styles and genres	Description of Selected Style	396	0.68**
3.2		Compare and Contrast	360	0.64**
4.1	Dance Performance: Performs masterwork dance prompt, choreographs solo performance, and collaborates on a duet	Movement	384	0.14**
4.2		Replication	379	0.26**
4.3		Solo Performance, Creativity	384	0.32**
4.4		Solo Performance, Performance Quality	383	0.24**
5.1	Choreographic Choices	Duet Collaboration	383	0.24**
5.2		Content	340	0.75**
5.3		Comprehension	341	0.71**
		Usage	341	0.72**

* $p < .05$, ** $p < .001$.

Table A2

Inter-rater Reliabilities of the Spring 2012 Middle School Benchmark Dance Assessments

Task Number	Task and Description		N Cases	Kappa
1.1		Choreographic Devices	271	0.64**
1.2	Dance Analysis: Recognize, identify, and categorize dance elements of a masterwork dance prompt	Body	293	0.87**
1.3		Dynamics	292	0.94**
1.4		Space	292	0.84**
1.5		Relationships	290	0.87**
2.1			Individual Accuracy of Movement	285
2.2	Movement Replication and Trio Performance: Collaborate on a trio of a masterwork dance prompt	Individual Performance Quality	285	0.55**
2.3		Group Collaboration	285	0.65**
3.1		Solo Dance Performance: Choreograph and perform solo routine using elements from masterwork dance prompt	Creativity	285
3.2	Performance Quality		285	0.50**
4	Choreographic Choices: Description of choreographic choices for solo		284	0.82**
5	Dance Styles and Genres: Compare and contrast masterwork dance prompt to dancing in another style/genre		261	0.74**
6.1	The Expression of Dance: Written response to the masterwork dance prompt and solo performance, addressing elements of dance and the relationship between movement and expression	Content	234	0.81**
6.2		Comprehension	234	0.76**
6.3		Usage	233	0.74**

* $p < .05$, ** $p < .001$.

Table A3
Inter-rater Reliabilities of the Spring 2012 High School Benchmark Dance Assessments

Task Number	Task and Description	N Cases	Kappa
1	Stage Directions: Recognize and identifying stage directions of a masterwork dance prompt	230	0.98**
2.1		220	0.89**
2.2	Dance Analysis: Recognize, identify, and analyze movements of a masterwork dance prompt	246	0.63**
2.3		228	0.67**
2.4		247	0.97**
3	The Body and Movement: Demonstrate basic knowledge of anatomy and kinesiology related to dance	240	0.82**
4	Dance as a Profession: Demonstrate knowledge of roles in the dance profession	225	0.77**
5	Dance Styles and Genre: Compare and contrast different dance styles/genres	223	0.57**
6.1	Dance as Expression: Written response of masterwork dance prompt addressing the relationship between movement and expression of themes	220	0.67**
6.2		220	0.68**
6.3		220	0.52**
7.1	Solo Dance Performance: Participate in warm-up and create a solo routine based on a masterwork dance prompt	221	0.34**
7.2		220	0.34**
7.3		221	0.38**
8.1	Group Dance Performance: Collaborate to create a group performance combining solos	194	0.47**
8.2		194	0.24**
8.3		195	0.13*
9.1	Reflecting on Dance Performance: Written reflection on dance performance	215	0.58**
9.2		213	0.60**

* $p < .05$, ** $p < .001$.

Table A4
Inter-rater Reliabilities of the Spring 2012 Fifth-Grade Benchmark Music Assessments

Task Number	Task and Description		N Cases	Kappa
1.1		Music Vocabulary	402	0.93**
1.2		Reading Music I	—	—
1.3	Completing a Composition: Demonstrating knowledge of music elements	Reading Music II	—	—
1.4		Reading Music III	—	—
1.5		Reading Music IV	425	0.97**
2	Musical Form: Identifying Musical Form		390	0.88**
3	Rhythmic Composition: Composing a short rhythm		411	0.91**
4.1	Planning a Composition: Selecting appropriate elements for a musical composition based on a short film	Instruments	423	0.83**
4.2		Dynamics	356	0.91**
4.3		Tempo	390	0.85**
5.1	Identifying Elements of Music: Listening to a music compositions and identifying the elements of music	Musical Elements	419	0.75**
5.2		Writing Skills	419	0.73**
6.1	Listening to and Comparing Musical Compositions: Listening to two music compositions and identifying elements of music	Musical Elements	411	0.86**
6.2		Writing Skills	399	0.78**
7.1	Music Performance: Performing learned instrumental or vocal piece	Instrumental	100	1.00**
7.2		Vocal	280	1.00**

Note: Tasks 1.2 through 1.5 do not have kappa values, given that they had right or wrong answers.

* $p < .05$, ** $p < .001$.

Table A5

Inter-rater Reliabilities of the Spring 2012 Middle School Benchmark Music Assessments

Task Number	Task and Description		N Cases	Kappa
1.1		Tempo	416	0.90**
1.2	Elements of Music: Listening to a music composition and identifying the elements of music	Dynamics	394	0.85**
1.3		Blending of Voices	424	0.78**
1.4		Instruments/Voices	416	0.79**
1.5		Artistic Choices	408	0.75**
2		Composition: Composing a short rhythm		422
3	Note Identification: Identifying notes of a composition		415	0.98**
4A.1	Identifying the Elements of Music – Recording A: Listening to two different music compositions and identifying the elements of music	Instruments/Voices	423	0.83**
4A.2		Melody	414	0.86**
4A.3		Harmony	401	0.87**
4A.4		Mood	418	0.76**
4A.5		Tempo	402	0.87**
4B.1	Identifying the Elements of Music – Recording B: Listening to two different music compositions and identifying the elements of music	Instruments/Voices	422	0.81**
4B.2		Melody	413	0.86**
4B.3		Harmony	398	0.86**
4B.4		Mood	418	0.77**
4B.5		Tempo	403	0.90**
5.1	Genres and Styles: Comparing and contrasting two different music compositions	Similarities	422	0.82**
5.2		Differences	423	0.83**
5.3		Writing Skills	422	0.76**
6	Reading Music: Listening to a music composition and identifying the written score		421	0.98**
7	Musical Connections: Identifying and describing roles in music		376	0.99**
8	Music Performance: Performing composed rhythm		414	0.77**

* $p < .05$, ** $p < .001$.

Table A6
Inter-rater Reliabilities of the Spring 2012 High School Benchmark Music Assessments

Task Number	Task and Description	N Cases	Kappa	
1	Notation, Rhythm, and Meter: Demonstrating knowledge of the basic elements of music composition	356	0.91**	
2	Pitch: Identifying the pitch of music notation	360	0.95**	
3.1	Dynamics and Tempo: Demonstrating knowledge of dynamics and tempo	Dynamics	365	0.96**
3.2		Tempo	365	0.98**
4	Key: Identifying the key of a musical composition	367	0.99**	
5A.1	Identifying the Elements of Music – Composition 1: Listening to and identifying elements of musical composition	Instruments/Voices	367	0.79**
5A.2		Mood	356	0.80**
5A.3		Meter	311	0.96**
5A.4		Tempo	330	0.96**
5A.5		Melody	343	0.88**
5A.6		Dynamics	293	0.98**
5B.1	Identifying the Elements of Music – Composition 2: Listening to and identifying elements of musical composition	Instruments/Voices	365	0.73**
5B.2		Mood	336	0.79**
5B.3		Meter	297	0.94**
5B.4		Tempo	325	0.94**
5B.5		Melody	336	0.81**
5B.6		Dynamics	292	0.89**
5C.1	Identifying the Elements of Music – Composition 3: Listening to and identifying elements of musical composition	Instruments/Voices	359	0.64**
5C.2		Mood	344	0.82**
5C.3		Meter	295	0.96**
5C.4		Tempo	325	0.92**
5C.5		Melody	326	0.70**
5C.6		Dynamics	297	0.91**
6.1	Genres and Styles of Music: Comparing the elements of music of the three musical composition prompts	Similarities	360	0.83**
6.2		Differences	356	0.85**
7	Writing a Review: Written reflection on one of the three musical compositions	334	0.66**	

* $p < .05$, ** $p < .001$.

Table A7
Inter-rater Reliabilities of the Spring 2012 Fifth-Grade School Benchmark Theater Assessments

Task Number	Task and Description	N Cases	Kappa
1.1		–	–
1.2	Theater Content and Vocabulary: Demonstrating knowledge of theater content and vocabulary	Theater Vocabulary I	–
1.3		Theater Vocabulary II	–
1.4		Theater Vocabulary III	–
		Theater Vocabulary IV	–
2A.1	Theater Analysis – Playwriting – Writing a short dialogue using a photograph of a play scene	Development of Character	207 0.56**
2A.2		Engaging Dialogue	207 0.60**
2A.3		Clear Conflict	207 0.55**
2B.1	Theater Analysis – Costume Design: Describing the relationship between the costumes and characters within a photograph of a play	Description of Costume	143 0.56**
2B.2		Analysis of Character Relationship Between	143 0.62**
2B.3		Character and Costume	143 0.54**
2B.4		Writing Skills	143 0.59**
3.1	Theater Performance: Acting in a two- person scripted scene	Focus and Commitment	334 0.35**
3.2		Vocal Skills	332 0.31**
3.3		Physicality	331 0.31**
3.4		Objectives	330 0.42**
4.1	Group Tableau: Group tableau performance	Reacts Physically to Imaginary Circumstances	– –
4.2		Responds Physically to Other Characters	– –
4.3		Reveals the Character through Physical Presence	– –
4.4		Reveals the Setting through Movement	– –

* $p < .05$, ** $p < .001$.

Note: Tasks 1.1 through 1.4 and Task 4.1 through 4.4 do not have kappa values, given that they had right or wrong answers.

Table A8

Inter-rater Reliabilities of the Spring 2012 Middle School Benchmark Theater Assessments

Task Number	Task and Description	N Cases	Kappa	
1A.1		Collaboration	106	0.27**
1A.2		Vocal Skills	106	0.35**
1A.3	Theater Performance – Acting: Acting in a two-person scene and incorporating director’s adjustments	Physicality	106	0.47**
1A.4		Staging	106	0.22**
1A.5		Objectives	105	0.37**
1A.6		Responds to Direction	53	0.45**
1B.1	Theater Performance – Musical Theater: Choreographing, rehearsing, and performing a musical number in a group and then incorporating director’s adjustments	Collaboration	50	0.13
1B.2		Vocal Skills	49	0.65**
1B.3		Staging	49	0.36**
1B.4		Responds to Direction	32	0.68**
1C.1	Theater Performance – Playwriting: Collaborating with a partner on completing a scripted scene and then incorporating director’s adjustments	Objectives	83	0.77**
1C.2		Dramatic Structure	83	0.49**
1C.3		Collaboration	82	0.68**
1C.4		Responds to Direction	45	0.76**
2A.1	Scene Analysis – Acting: Watching a filmed performance and providing an analysis of the actors in the scene	Analysis of Character	159	0.62**
2A.2		Analysis of Relationships	158	0.70**
2A.3		Writing Skills	161	0.63**
2B.1	Scene Analysis – Design: Watching a filmed performance and providing an analysis of the playwright’s intent	Analysis of Intent	13	0.11
2B.2		Justification of Artistic Interpretation	13	0.59*
2B.3		Writing Skills	12	0.82*

* $p < .05$, ** $p < .001$.

Table A9

Inter-rater Reliabilities of the Spring 2012 High School Benchmark Theater Assessments

Task Number	Task and Description	N Cases	Kappa
1.1	Collaboration	178	0.59**
1.2	Theater Performance – Acting: Vocal Skills	178	0.43**
1.3	Collaborating with a partner to complete Physicality	178	0.29**
1.4	the dialogue of a scripted scene and then Staging	178	0.41**
1.5	performing the scene Objectives	178	0.35**
1.6	Playmaking Structure	177	0.30**
2A.1	Directorial Point of View	135	0.65**
2A.2	Scene Analysis – Directing: Written Analysis of a scene in relation to the Relationship Between Characters	135	0.69**
2A.3	actors Scene Analysis	135	0.71**
2A.4	Writing Skills	135	0.68**
2B.1	Description of set and costume	40	0.69**
2B.2	Scene Analysis – Costume Design: Relationship between character and costume	40	0.79**
2B.3	Written Analysis of a scene in relation to the costume design Description of set and costumes for change in setting	40	0.51**
2B.4	Design Analysis	40	0.63**
2B.5	Writing Skills	40	0.49**

* $p < .05$, ** $p < .001$.

Table A10

Inter-rater Reliabilities of the Spring 2012 Fifth-Grade Benchmark Visual Arts Assessments

Task Number	Task and Description	N Cases	Kappa	
1.1	Elements of Art – Drawing Lines: Demonstrating knowledge of a variety of lines	–	–	
2.1	Elements of Art – Visual Textures: Demonstrating knowledge of visual textures	–	–	
2.2	Identifying Visual Textures Creating Visual Textures	–	–	
3.1	Drawing: Using lines and textures to create a drawing reflecting imaginative capacities	Achieves Expressive Quality	577	0.65**
3.2		Uses a Variety of Lines	577	0.45**
3.3		Uses a Variety of Texture	576	0.51**
3.4		Uses Space Appropriately	577	0.57**
4.1	Color Theory: Demonstrating knowledge of categories of color	Primary	–	–
4.2		Secondary	–	–
5.1	Tints and Shades: Creating tints and shades	Tints	–	–
5.2		Shades	–	–
6	Thinking Like an Artist – Elements of Art: Written response, Using elements of art in describing the artist’s process	573	0.62**	
7	In the Art Museum: Demonstrating knowledge of museum conventions	–	–	
8	Art History: Demonstrating an understanding of art history chronology	–	–	
9	The Artistic Process: Observing an artist at work and identifying the artistic process	–	–	
10	Artistic Expression: Written response on what inspires an artist	495	0.59**	
11.1	Elements of Visual Art: Looking at artwork and identifying elements in visual art	Geometric Shapes	–	–
11.2		Organic Shapes	–	–
11.2		Negative and Positive Space	–	–
12.1	Art Making: Creating a work of art	Uses a Variety of Colors	573	0.38**
12.2		Uses a Variety of Shapes	573	0.47**
12.3		Uses Space Appropriately	574	0.53**

Task Number	Task and Description	N Cases	Kappa	
12.4	Clearly Demonstrates Use of Negative and Positive Space	574	0.53**	
12.5	Handles Materials Competently	574	0.52**	
12.6	Demonstrates Imaginative Capacities	573	0.67**	
12.7	Depicts Figure and Setting	569	0.62**	
13.1	Reflective of Artwork: Reflection on student's art making process	Reflection of Artwork	564	0.64**
13.2		Visual Arts Vocabulary	563	0.70**
13.3		Writing Skills	562	0.65**

Note: Tasks 1.1 through 2.2, Tasks 4.1 through 5.2, Tasks 7 through 9, and Tasks 11.1 and 11.2 do not have kappa values, given that they had right or wrong answers.

* $p < .05$, ** $p < .001$.

Table A11

Inter-rater Reliabilities of the Spring 2012 Middle School Benchmark Visual Arts Assessments

Task Number	Task and Description	N Cases	Kappa	
1.1	Renders Shape	296	0.61**	
1.2	Draws Detail	296	0.58**	
1.3	Draws Object Large Enough to View	296	0.57**	
1.4	Drawing: Drawing from observation use of shape, texture, and lighting	Detail	296	0.59**
1.5	Uses Visual Texture	296	0.61**	
1.6	Uses Light and Shadow	296	0.61**	
1.6	Achieve Expressive Quality	295	0.62**	
2	Analysis of Artwork: Comparing and contrasting two works of art	291	0.77**	
3.1	Color Theory: Demonstrating knowledge of categories of color	Primary	—	—
3.2		Secondary	—	—
3.3		Complementary	—	—
3.4		Analogous	—	—
4.1	Qualities of Sculpture: Demonstrating knowledge of positive and negative space, symmetry, categories of shape, and visual arts styles	Positive and Negative Space	—	—
4.2		Symmetry	—	—
4.3		Shapes	—	—
4.4		Styles	—	—
5	Art History: Demonstrating an understanding of art history chronology	—	—	
6.1	Art Making: Creating a work of art	Addresses specific art assignment	299	0.68**
6.2		Uses Space Appropriately	299	0.72**
6.3		Demonstrates unity through color	299	0.71**
6.4		Demonstrates balances composition	299	0.63**
6.5		Handles materials competently	299	0.59**
6.6		Achieves expressive quality	298	0.66**
7.1	Reflection of Artwork	289	0.66**	

Task Number	Task and Description		N Cases	Kappa
7.2	Reflection of Artwork: Reflection on student's art making process	Expressive and Descriptive Written Response	290	0.71**
7.3		Writing Skills	290	0.73**

Note: Tasks 1.1 and 1.2 and Tasks 3.1 through 5 do not have kappa values, given that they had right or wrong answers.

* $p < .05$, ** $p < .001$.

Table A12

Inter-rater Reliabilities of the Spring 2012 High School Benchmark Visual Arts Assessments

Task Number	Task and Description	N Cases	Kappa	
1	Three Dimensionality: Looking at artwork and describing artist's technique	243	0.69**	
2.1	Elements of Art and Principles of Design: Writing about artwork in terms of Elements of Art and Principles of Design	Elements of Art	245	0.74**
2.2		Principles of Design	237	0.66**
3	Visual Art Vocabulary: Identifying vocabulary for art media	—	—	
4	Art History: Demonstrating an understanding of art history chronology	—	—	
5	Color Theory: Demonstrating knowledge of categories of color	—	—	
6	Art in the Community: Written response on the relationship between artwork and its environment	231	0.71**	
7.1	Curating an Art Exhibit: Choosing artwork by self-selected them and describing choices	Art Exhibit	237	0.61**
7.2		Writes for Intended Audience	225	0.64**
7.3		Expressive and Descriptive Written Response	225	0.73**
7.4		Writing Skills	225	0.73**
8.1	Art Making: Creating a work of art	Addresses Specific Art Assignment	240	0.67**
8.2		Uses Space Appropriately	240	0.66**
8.3		Demonstrates Three-Dimensionality and Volume	240	0.75**
8.4		Composition	240	0.66**
8.5		Use of Materials	239	0.70**
8.6		Demonstrates Realism	240	0.79**
8.7		Demonstrates Perspective and Scale	240	0.66**

Note: Tasks 3 through 5 do not have kappa values, given that they had right or wrong answers.
 $*p < .05$, $**p < .001$.