District 75, New York City Department of Education
Everyday Arts for Special Education
Impact Evaluation

i3 Development Grant – US Department of Education

Dr. Rob Horowitz, Evaluator
ArtsResearch

December 2016

The research reported here was supported by the Office of Innovation and Improvement (OII), U.S. Department of Education, through Grant # U396C100275 to District 75, New York City Department of Education.

rob.horowitz@artsresearch.net www.artsresearch.net 212-781-3730

Contents

District 75, New York City Department of Education
Everyday Arts for Special Education
Impact Evaluation

1. Summary of Intervention

Everyday Arts for Special Education (EASE) was funded through an i3 Development Grant to District 75, New York City's special education district. When the grant began in 2010, District 75 employed 4,000 teachers in 56 schools within each of New York City's five boroughs. Of the District's 23,000 students, 71% were eligible for Title I support, 86% were from minority populations, 15% were English Language Learners, and 60% were assessed on New York State alternate academic achievement standards. Students' disabilities encompassed a wide range, including autism spectrum disorders, cognitive disabilities, emotional disturbance, severe learning disabilities, and multiple handicaps (physical and cognitive). The students required highly specialized educational programs and support systems.

The participating EASE students were all in elementary grades (K-5) within 10 District 75 schools. However, the meaning of "district," "school" and "grade" were different than a conventional setting. District 75 is not a geographical district, but is instead an organizational structure across the entire city that encompasses students who require the special services available through the district. The EASE "schools" were not buildings, but were also organizational structures, comprised of multiple sites, within various school buildings which might also have other programs. In total, the EASE program operated in 37 sites throughout NYC. While all students were in elementary grades, these grades did not reflect a level of academic achievement, but instead were based upon each student's age.

EASE students had 4 categories of disabilities: autism spectrum, intellectual disabilities, emotional disturbance and multiple disabilities. Students varied in the range of severity of disabilities within each category. EASE instruction took place within each EASE teacher's classroom. The classes were sometimes grouped according to disability, but often contained various disabilities. EASE class sizes varied, as well, with up to 12 students per class. Most EASE classes had 6 students. The EASE program provided extensive professional development (PD) for the teachers that included a repertoire of arts-based strategies that they could employ to address each students' Individual Education Plan (IEP) needs, particularly in communication, socialization and related academic areas.

The Program and Implementation

EASE teachers learned strategies across multiple arts disciplines (music, dance, visual arts, and theater) through a series of professional development workshops and extensive in-school support. There were four key program components, as shown in Table 1

Table 1: Program Components

Intervention Activity	Key Indicators for Each Component
Professional Development Workshops	Participating teachers and administrators meet with teaching artists for full-day workshops.
Collaborative Classroom Modeling	Teaching artists collaborate with teachers in the classroom to implement curricula learned in the professional development workshops, differentiating instruction for all age and ability levels.
On-Site Professional Development	Teaching artists conduct 45-minute on-site professional development sessions with participating teachers, focused on differentiation and documentation of best practices.
Classroom Instruction	Teachers address an IEP goal through an EASE instructional activity.

The PD was led by Manhattan New Music Project (MNMP) and Urban Arts Partnership (UAP) teaching artists.[1] Teachers learned differentiated arts-based strategies designed to meet the communication, socialization, academic, and arts goals of each student's IEP across multiple arts. The strategies were developed through many years of experience by a core group of teaching artists, who had worked in other District 75 and special education projects. The strategies were codified at the outset of the program by a curriculum developer/program designer. The PD sessions were experiential, with opportunities for practice and reflection. Artist visits to each classroom provided modeling and on-site coaching.

A new cohort of 60 teachers and their students were added to the EASE program each program year, while ongoing PD and support was provided for previous years' participants. The level of direct teaching artist support was reduced in each subsequent year of teachers' participation as they became more capable of implementing the program without assistance. Upon reaching their fourth year, the most

[1] District75's arts partner was MNMP for the first 2 years of the grant. MNMP was absorbed by UAP, which then became District 75's partner for the final 3 years.

effective teachers were identified and trained to become mentor teachers, in order to help disseminate the EASE program to others.

Table 2 shows the levels of support for each year's cohort of teachers. Table 3 shows activities for each level.

Table 2: EASE Teacher Cohorts – Levels of Program Support Over Five Years

	Cohort 1	Cohort 2	Cohort 3	Cohort 4	Cohort 5
YEAR 1	Beginner Level support				
YEAR 2	Intermediate Level support	Beginner Level support			
YEAR 3	Advanced Level support	Intermediate Level support	Beginner Level support		
YEAR 4	Selected teachers become mentors	Advanced Level support	Intermediate Level support	Beginner Level support	
YEAR 5	Selected teachers become mentors	Selected teachers become mentors	Advanced Level support	Intermediate Level support	Beginner Level support

Table 3: Activities for Teachers at Different Program Levels

	General Description	PD Workshops	Collaborative Classroom Modeling	On-Site PD
Beginner Level	Teachers learn and implement EASE with extensive support from teaching artists	4 full-day workshops per teacher	20 in-class teaching artist visits per class	20 forty-five minute sessions per teacher
Intermediate Level	Teachers implement EASE with new students, taking more responsibility from teaching artists	2 full-day workshops per teacher	16 in-class teaching artist visits per class	16 forty-five minute sessions per teacher
Advanced Level	Teachers implement EASE with new students, requiring minimal support from teaching artists	1 full-day workshop per teacher	8 in-class teaching artist visits per class	8 forty-five minute sessions per teacher

1.1. Research Questions

The impact evaluation used a cluster (site) quasi-experimental design to examine the impact of EASE on special education (SPED) students' academic achievement and social-emotional behavior. Academic achievement was measured using the New York State Alternative Assessment (NYSAA): a teacher-selected performance task rated on a 4-point scale designed to be administered to students with severe cognitive disabilities. Social-emotional behavior was measured using the Student Annual Needs Determination Inventory (SANDI): an 88-item questionnaire completed by teachers.

Research questions for the impact studies included:

- What is the impact of two years of exposure to EASE on 4th grade special education (SPED) students' reading achievement as measured by NYSAA?
- What is the impact of two years of exposure to EASE on 4th grade SPED students' mathematics achievement as measured by NYSAA?
- What is the impact of two years of exposure to EASE on 2nd-5th grade SPED students' social-emotional behavior as measured by SANDI?

2. Impact Evaluation Design

The impact studies employed a cluster (site) quasi-experimental design to examine the impact of EASE on special education students' academic achievement and social-emotional behavior. The studies were conducted in multiple sites within the 10 EASE schools.

2.1. Sample Identification, Selection and Assignment

2.1.1. Identification/Selection of Study Schools and "Sites"

At the outset of the program, District 75 schools were invited to apply for participation via a survey, with the following questions.

1. State your current elementary-age population across all sites. (Pre-K through 6th grade)
2. State the number of elementary-aged classes you currently have across all sites. (Pre-K through 6th grade)
3. How many classes are served by these arts specialists, at which sites, and how frequently?
4. List the cultural organizations which provided arts programs/residencies (if any) at your school for the past three years.
5. How did you hear about EASE and why do you think your school will be a good candidate for participation in EASE?

6. What changes to you expect to see in your students, teachers, and administration as a result of the EASE program?
7. EASE may require schools to adjust their schedules from time to time in order to facilitate teaching artist visits, teacher meetings, in-school professional development, and interviews by the evaluation team. What do you see as the challenges to this?
8. How will you address these challenges?
9. EASE will require that participating teachers be released up to 4 times per year for mandatory all-day workshops. Although funds are available for hiring substitutes, schools sometimes find it difficult to release multiple teachers on the same day. How challenging is it for you to find substitutes?
10. How many teachers do you feel you could reasonably expect to release on any given day?
11. As part of the project's research component, participating teachers will be required to submit weekly online reports tracking students' progress in EASE activities. What strategies will you employ to ensure your teachers are compliant with this requirement?
12. Although EASE training is provided to teachers, the program will function more smoothly with other school staff on board. How will you turnkey the knowledge and skills gained by teachers to administrators, related service providers, and paraprofessionals?
13. The EASE program will require a school point person for EASE coordination. What are the qualities that would make an individual effective in this position?
14. Who would be the EASE point person at your school?
15. In addition to any grant-imposed teacher selection criteria, indicate the criteria you will use to select teachers for participation in EASE.

Applications were reviewed by District 75 and MNMP staff. The 10 schools were selected based upon their capacity, leadership and ability to sustain commitment to the program. The program was implemented in 37 sites within the 10 schools.

The treatment groups for the QED impact studies were drawn from these 37 sites, based upon students matching the selection criteria described in the subject selection section, below. The comparison groups were from the same 10 schools, but in sites where there was no EASE participation. The academic achievement (reading and math) analysis included 28 EASE sites and 18 comparison non-EASE sites (within 8 of the 10 schools implementing EASE). The analysis of social-emotional behavior included 23 EASE sites and 22 comparison non-EASE sites (in 10 of 10 schools implementing EASE). The treatment sample for academic achievement was 4th grade students after 2 years of exposure, and the treatment sample for SEL was 2nd, 3rd and 4th grade students after 2 years of exposure.

The counterfactual conditions were "business as usual" in the comparison sites. Students in the business-as-usual condition were not exposed to any EASE activities.

2.1.2. Identification/Selection of Students

The sample used to examine academic achievement included two cohorts of 4th grade students (2011-12 and 2012-13). To be included in this analysis, students were required to meet these criteria:

- Student identified as an alternative assessment student (i.e., eligible to take the NYSAA)
- Student had a NYSAA third grade score (in order to have a pre-intervention measure of academic achievement).

The sample used to examine social-emotional behavior included two cohorts of students in 2nd-4th grade (2012-13, and 2013-14) that had fall pretest (2012, 2013, respectively) scores and spring posttest scores (2014, 2015, respectively) on the SANDI. Table 4 below shows the treatment and comparison samples for each measure.

Table 4: EASE Impact Study Samples

	Schools	Sites		Students		
		EASE	Comparison	Grade(s)	EASE	Comparison
Academic Achievement Sample	8	28	18	4th	83	74
Social-Emotional Behavior Sample	10	23	22	2nd, 3rd, 4th	190	569

Multi-Year Intervention

The sample used to examine academic achievement included two cohorts of 4th grade students (2011-12 and 2012-13). Academic outcomes (i.e., NYSAA scores) were analyzed at the end of fifth grade (2013 and 2014 respectively). The sample used to examine social-emotional behavior included two cohorts of students in 2nd-4th grade (2012-13, and 2013-14). Outcomes (i.e., SANDI scores) were analyzed at the end of 3rd, 4th and 5th grades, respectively (in spring 2014 and 2015).

2.1.3. Treatment of Missing Data

The analysis sample was defined as being comprised of students with non-missing pre-tests and non-missing outcomes. Therefore, the analysis used casewise deletion for students that are missing pre-tests or outcomes. Estimates of the intervention's effects were based on analysis of data without any imputed outcome values.

2.1.4. Baseline Equivalence Testing

For each contrast baseline equivalence was established in the analytic sample (those students with pretest and posttest scores, using: (a) 3rd grade NYSAA scores for the academic achievement contrasts and (b) the SANDI administered to students in the fall prior to the school year of EASE exposure.

Baseline equivalence was examined using a modified version of the impact model (shown in the Appendix) with the pretest as the dependent variable and the inclusion of two independent variables: (1) a treatment indicator (2) a cohort (block) indicator. The treatment-comparison difference in Table 5 is the parameter estimate on the treatment indicator taken from this model.

3. Data Collection

3.1. Measures

New York State Alternative Assessment (NYSAA). The study sample of students took the New York State Alternative Assessment (NYSAA) in reading and math each year, in the fall and spring. Only students with severe cognitive disabilities are eligible for NYSAA, based upon these criteria: (1) the student has a severe cognitive disability and significant deficits in communication/ language and significant deficits in adaptive behavior; and (2) the student requires a highly specialized educational program that facilitates the acquisition, application, and transfer of skills across natural environments (home, school, community, and/or workplace); and (3) the student requires educational support systems, such as assistive technology, personal care services, health/medical services, or behavioral intervention (New York State Education Department [NYSED], 2010).

When administering the NYSAA, teachers select indicators for each content area that are appropriate for each student. The indicators range "across a spectrum of complexity from least to most complex." The indicators are selected based upon the students' grade and age, the appropriateness of the challenge, the students' academic ability, and the school year's curriculum content. Teachers select an assessment task that matches the indicator, that can be performed in one day, and that is observable and measurable. Teachers gather two pieces of verifying evidence for each indicator, and score the resulting "dataportfolios" on two dimensions: Connection to Grade Level Content and Performance. Level of Accuracy and Level of Independence are considered when scoring Performance. They are calculated as a percentage from 0% to 100% and then assigned to a 4-point scale (1 to 4). Accuracy is determined by comparing the student's number of correct responses with the total number of expected responses. Independence is determined by dividing the number of steps or items not requiring prompts or cues by the total number of steps in the task (NYSED, 2010).

Student Annual Needs Determination Inventory (SANDI). SANDI was developed by the Riverside County, California, Office of Education, Special Education Unit. SANDI is designed to be used as an assessment tool to determine students' functional skills. In addition, SANDI aids instructional decisions based upon a students' performance level, progress on IEP goals and identification of educational needs (Riverside County Office of Education, 2008).

Students are asked to perform tasks that reflect IEP goals. SANDI items are scored according to 4 categories (Not Engaged, Engaged, Supported, Independent) which reflect the percentage of correct responses, considering additional scoring criteria such as the number of prompts needed to complete

the task. The scores are used by the teacher to identify areas of need and to refine IEP goals. Based upon the scores, the students are assigned to one of four categories for each IEP assessed (including social/emotional learning goals, the domain of interest in the SEL impact study): Proficient, Basic, Below Basic, Far Below Basic (Cahill, Silva & Chappell, 2012).

4. Analysis

The impact evaluation used a cluster (site) quasi-experimental design to examine the impact of EASE on special education (SPED) students' academic achievement and social-emotional behavior. The sample used to examine academic achievement included two cohorts of 4th grade students (2011-12 and 2012-13). Academic outcomes (i.e., NYSAA scores) were analyzed at the end of fifth grade (2013 and 2014 respectively). The sample used to examine social-emotional behavior included two cohorts of students in 2nd-4th grade (2012-13, and 2013-14). Outcomes (i.e., SANDI scores) were analyzed at the end of 3rd, 4th and 5th grades, respectively (in spring 2014 and 2015). A list of contrasts (Table 6) and the analytic model are in the Appendix.

4.1. Baseline Equivalence Testing

For each contrast, baseline equivalence was established in the analytic sample (those students with pretest and posttest scores), using third grade NYSAA scores for the academic achievement contrasts, and the SANDI administered to students in the fall prior to the first school year of EASE exposure.

Baseline equivalence was examined using a modified version of the impact model with the pretest as the dependent variable and the inclusion of two independent variables: (1) a treatment indicator and (2) a cohort (block) indicator. The treatment-comparison difference in Table 5 below, is the parameter estimate on the treatment indicator taken from the model. Baseline equivalence was established for each contrast.

Table 5: Baseline Equivalence of Treatment and Control Groups

Contrast and Measure	Treatment Group N	Comparison Group N	Unadjusted Treatment Group SD	Unadjusted Comparison Group SD	Unadjusted Comparison Group Mean	Treatment – Comparison Difference	Standardized T-C Difference
Reading Proficiency (NYSAA)	83	74	.896	.846	3.68	-0.026	-0.03
Math Proficiency (NYSAA)	83	74	.480	.404	3.88	-0.048	-0.11
Social-Emotional Learning (SANDI)	190	569	77.89	73.63	141.68	12.34	0.16

4.2. Impact Findings

Analysis of the impact studies indicated a program effect on reading achievement and social-emotional learning. There was a substantively meaningful effect of the intervention on students' reading skills (Effect Size = .42) as measured by NYSAA. There was a modest but significantly significant effect on students' SEL (Effect Size = .18) as measured by SANDI. Analysis of NYSAA math proficiency indicated indeterminate effects with no treatment-comparison difference (p = .97).

Table 6: Impact Estimates

Contrast	Treatment Group SD	Comparison Group SD	Impact Estimate	Standardized Effect Size	Impact Standard Error	p value	Degrees of Freedom
Reading	.549	.771	.278	.42	.139	.054	33.34
Math	.610	.652	.033	.05	.088	.971	37.86
SEL	73.65	66.09	12.21	.18	4.11	.005	36.48

4.3. Student Assessment: Communication and Socialization Skills

As part of the EASE program evaluation, every participating teacher rated each of their students weekly on a 3-point scale on these indicators: (1) communication skills, (2) socialization skills, (3) following directions, (4) time on task, (5) self-esteem, (6) engagement, and (7) arts proficiency. Teachers also identified an IEP goal for each student that they addressed through EASE program teaching strategies, and indicated the degree of improvement towards meeting those goals. The criteria for progress in each area were determined by the teachers, who had the best understanding of the students' disabilities and IEP. Teachers used their professional judgment and knowledge of each student's disability (and degree of disability) to determine if the student was making progress. Teachers received professional development in developing criteria and rating students. The teacher ratings were submitted weekly for 23 weeks during the school year. Teachers were also asked to provide qualitative examples of behaviors that indicated progress in each indicator. The assessments were conducted online using ArtsResearch software, which was developed to expedite timely responses by teachers.

Analysis demonstrated significant gains in each domain. Table 7 shows significant increases (p < .001) from November to May in the fourth year of the program, when most surveys were gathered and the program was up to a full complement of 180 teachers (n = 19,689 assessment surveys were gathered that year).[2]

[2] Because students were rated on a 3-point scale, mean scores could range from 1 to 3.

Table 7: Student Assessment Surveys

Assessment Domains	November Mean	May Mean
Communication Skills	1.64	2.05
Socialization Skills	1.66	2.08
Follows Directions	1.69	2.11
Time on Task	1.67	2.09
Engagement	1.78	2.17

In the assessment surveys, teachers described behaviors that supported their high ratings in communication, socialization, and engagement. These included: taking turns, leading activities, successfully interacting with peers and teachers, responding to modeling, sharing materials, appropriately interacting, making eye contact, taking turns, verbalizing or physically expressing ideas, demonstrating self-control, focusing, staying on task, and using imagination and creativity when responding to a set of criteria in a sustained task. Teachers observed that students were better able to participate longer on challenging tasks.

5. Summary

The program was determined to have an effect on reading scores and social-emotional learning. There was a substantively meaningful effect of the intervention on students' reading skills (Effect Size = .42) as measured by NYSAA. There was a modest but significantly significant effect on students' SEL (Effect Size = .18) as measured by SANDI. Analysis of NYSAA math proficiency indicated indeterminate effects with no treatment-comparison difference (p = .97).

These effects were consistent with analysis of survey and observational data. These additional data will be published in spring 2017 in a book chapter describing the EASE evaluation in the context of evaluating special education and the arts.

The improvement in reading skills and social-emotional learning may be due to the increased engagement of students with disabilities. The EASE activities were inherently interactive, involving peer-to-peer and teacher-student communication in verbal, artistic and kinesthetic domains. These were the kinds of activities that children would otherwise not be consistently exposed to in their usual special education classrooms. The degree of support provided by District 75 and its partners, MNMP and UAP, through the grant was likely instrumental in obtaining the improvements described here. The instructional model we studied is based upon the interlocking program components of professional development, in-school support and classroom instruction. The expertise of seasoned teaching artists was a major factor as well. The core pool of visiting teaching artists had extensive practical experience in working in District 75 classrooms. Additional research may determine if similar results could be obtained without this degree of prior expertise and systemic support.

6. References

New York State Education Department, Office of State Assessment. (2010). New York State Alternative Assessment: Administration manual 2010-11 (pp. 1-31). Albany, NY.

Riverside County Office of Education, Special Education Unit. (2008). Student annual needs determination inventory: (3rd ed.). Riverside, CA.

7. Appendix

Table 6: List of Contrasts

Contrast Name	Design	Treatment Group			Comparison Group	Outcome				Baseline		
		Description	Age/grade during intervention	Exposure	Description	Domain	Measure	Unit of Observation	Timing of Measurement	Measure [Scale]	Unit of Observation	Timing of Measurement
1. Reading Proficiency	QED with treatment and matched comparison students from same schools	Everyday Arts for Special Education: special education students in grade 5 with 2 years of intervention	a. grade 4-5 (2011-12 – 2012-13) b. grade 4-5 (2012-13 – 2013-14)	2 years	Business as usual	Reading Proficiency	New York State Alternate Assessment: Reading	Student	a: Between Oct 2012 and Feb 2013, during grade 5 b. Between Oct 2013 and Feb 2014, during grade 5	New York State Alternate Assessment: Reading	Student	a: Between Oct 2010 and Feb 2011, during grade 3 b: Between Oct 2011 and Feb 2012, during grade 3
2. Math Proficiency	QED with treatment and matched comparison students from same schools	Everyday Arts for Special Education: special education students in grade 5 with 2 years of intervention	a. grade 4-5 (2011-12 – 2012-13) b. grade 4-5 (2012-13 – 2013-14)	2 years	Business as usual	Math Proficiency	New York State Alternate Assessment: Math	Student	a: Between Oct 2012 and Feb 2013, during grade 5 b. Between Oct 2013 and Feb 2014, during grade 5	New York State Alternate Assessment: Math	Student	a: Between Oct 2010 and Feb 2011, during grade 3 b: Between Oct 2011 and Feb 2012, during grade 3
3.Social/Emotional Learning	QED with treatment and matched comparison students from same schools	Everyday Arts for Special Education: special education students in grade 5 with 2 years of intervention	a. grade 4-5 (2012-13 – 2013-14) b. grade 4-5 (2013-14 – 2014-15)	2 years	Business as usual	Social/Emotional Learning	Student Annual Needs Determination Inventory	Student	a. Spring 2014 b. Spring 2015	Student Annual Needs Determination Inventory	Student	a. Fall 2012 b. Fall 2013

Two-level model used to estimate impacts for each contrast

Level-1 Model: Student Level

$$Y_{ij} = \beta_{0j} + \beta_{1j}(Y^*_{ij}) + \beta_{2j}(\text{Disability}_{ij}) + \beta 3_j\ (\text{Cohort}_{ij}) + \varepsilon_{ij}$$

where

Y_{ij} is the spring posttest score for student i in site j;

β_{0j} is the conditional mean posttest score for control students in site j,

Y^*_{ij} is the pretest score for student i in site j;

Disability is 1 if autistic 0 if not for student i in site j.

Cohort is 1 if in cohort 2, 0 if in cohort 1 for student i in site j

ε_{ij} is the student-level random error representing the difference between student ij's score and the conditional predicted mean score for block j. These residual effects are assumed normally distributed with mean 0 and variance σ^2.

Level-2 Model: Site Level

$$\beta_{0j} = \gamma_{00} + \gamma_{01}T_j + \sum_{s=1}^{n} \gamma_{0(1+s)}(Site_s) + \mu_{0j}$$

$$\beta_{mj} = \gamma_m$$

γ_{00} is the mean for control classes;

γ_{01} is the treatment effect (difference between treatment and control site means);

T_j= 1 if site j is an intervention class, and 0 if control;

$Site_s$=1 of student is in Site s, = 0 else.

μ_{0j} is the deviation of site j's mean from the grand mean