David P. Weikart
Center for Youth
Program Quality

Quality-Outcomes Study for Seattle Public Schools Summer Programs

Summer 2015 Program Cycle

This project was funded by The Raikes Foundation.
Forum for Youth Investment; Grant ID 346;
Submitted to the Raikes Foundation, December 22, 2015

Charles Smith, Ph.D.
Katharine Helegda
Ravi Ramaswamy
Barbara Hillaker, Ph.D.
Gina McGovern
Leanne Roy

Contents

Summary

In the summer of 2015 the Seattle Public Schools (SPS), the Raikes Foundation, School's Out Washington (SOWA), and David P. Weikart Center for Youth Program Quality (Weikart Center), collaborated on a quality-outcomes study. The study was conducted in 30 summer learning program offerings at 11 school sites in Seattle, Washington, during the summer 2015 program cycle. Key findings include:

- On average, SPS summer program offerings demonstrate high levels of instructional quality compared to summer programs in other cities. Programs were also well attended.
- Embedded assessments for math and literacy suggest that most students improved academic skills during the summer program cycle.
- Summer programs can be divided into three performance subgroups with distinct quality profiles: Very high, moderate, and lower quality.
- Students in offerings with very high quality instruction had more positive change on math and literacy assessments when compared to students receiving moderate and lower quality instructional experiences.

These findings suggest that benchmarks for high quality instruction can be identified and that high quality instructional experiences are associated with gains in academic skills. These findings align with a skill development theory suggesting that only high quality out-of-school time (OST) experiences produce patterns of skill growth that are observable during short program cycles. Further, these findings suggest that investments to improve the quality of summer offerings may produce returns in academic skills development.

As the evaluation design of this study did not include a rigorously matched comparison group or directly address questions about the effects of summer program participation on school success outcomes during a subsequent school year, these findings should be interpreted cautiously. This quality-outcomes study is part of a larger sequence of evaluations focused on the design and scaled implementation of a quality improvement intervention for summer learning programs.

Recommendations include (1) description of the SPS summer curriculum and best practices as implemented by the high quality subgroup and (2) replication of the study in a future year with better skill measures and a more rigorous matched control group design.

Introduction

Since 2012, a collaborative of funders, public and private summer learning service providers, and several technical services organizations has partnered to design, implement, and evaluate an evidence-based quality improvement system (QIS) intervention for summer learning networks and summer school programs. This collaboration has produced a series of implementation-design studies seeking to improve intervention fidelity at scale. (Ramaswamy, Gersh, Sniegowski, McGovern, & Smith, 2014; C. Smith, Ramaswamy, Gersh, & McGovern, 2015; C. Smith, Ramaswamy, Hillaker, Helegda, & McGovern, 2015). The collaborative has studied design and implementation of the QIS intervention in summer learning networks in Denver, CO, Grand Rapids, MI, Oakland, CA, Seattle, WA, and St. Paul, MN.

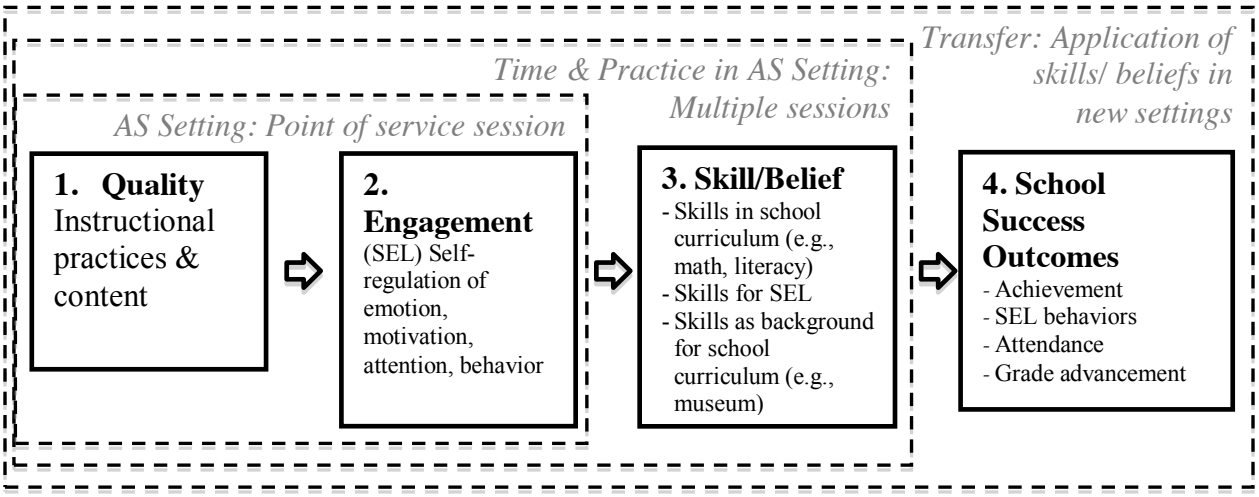
In the summer of 2015, the Weikart Center, SPS, SOWA, and the Raikes Foundation collaborated to extend this work with a quality-outcomes study intended to describe the effectiveness of summer programs in terms of instructional quality and academic skill growth. Primary goals of the study were to (1) describe the effectiveness of SPS summer programs in terms of instructional quality and youth skills and (2) to accumulate further validity evidence for the Summer Learning Program Quality Assessment (Summer Learning PQA).

Study Design

Figure 1 presents an OST skill development theory for use in the OST field. The Quality, Engagement, Skills, Transfer (QuEST) model (C. Smith, S. Hallman, et al., 2012) describes the quality of youth learning opportunities, first in terms of instructional practices and the given subject matter content. In turn, high-quality learning opportunities should stimulate interest and motivation to engage students. Repeated high quality sessions with high student engagement should result in mastery experiences for specific skills. Mastery of specific skills should promote transfer of these skills to other academic contexts, such as school day classrooms.

The QuEST model draws from a broad evidence base to suggest that (a) setting qualities influence student skill development, (b) motivation is an important correlate of learning, (c) skill building requires intentional adult supports (coaching, modeling, scaffolding, facilitating) and time to practice those skills, and (d) skills learned in one setting do not automatically transfer to a different setting. A practical theory template like QuEST allows local actors to fill in details about their specific program designs (e.g., how they define quality) and the specific skills they are trying to build. In this case, students attending high quality summer offerings will develop targeted skills to a greater extent than students attending lower quality summer offerings because students in higher quality offerings will be more engaged with the content and receive more opportunities to practice skills.

Figure 1: QuEST: A logic model template for skill development and transfer theories [Excerpted from Smith et al. (2009), Figure D1]



Prior research has consistently demonstrated the relationship between higher quality instruction (box 1) and both higher levels of youth engagement with OST content (box 2) and higher levels of social and academic skill during the school day (box 4) in afterschool programs (Akiva, Cortina, Eccles, & Smith, 2013; Naftzger, 2014; Naftzger, Devaney, & Foley, 2014; Naftzger et al., 2013).[1] This quality-outcomes study represents one of only a few evaluations linking high quality data for specific components of instructional quality (box 1) to student skill growth demonstrated in the OST context (box 3).

Method

The study addresses the following primary questions: What is the quality profile for SPS summer learning programs? Can meaningful quality subgroups be identified? Are academic skills enhancements for math and literacy related to the quality of the summer learning programs?[2]

Intervention Design

The summer program intervention design was delivered during 30 summer learning offerings, where an offering is defined as the same group of staff serving the same group of students for the same learning purpose over multiple sessions. In each offering, a staff team served the same group of

[1] Several evaluations using quality-outcomes evaluation designs and employing same/similar measures have been conducted in the Texas 21st Century Community Learning Communities program. These evaluations are available at: http://tea.texas.gov/Reports_and_Data/Program_Evaluations/Out-of-School_Learning_Opportunities/Program_Evaluation__Out-of-School_Learning_Opportunities/

[2] Two additional research questions will be addressed as time and additional resources allow: Do students in the lowest performance quartile at baseline gain more in higher quality programs than in lower quality programs? How are two new the Summer Learning PQA scales for Math and Literacy practices related to change in academic skills?

approximately 20 students as they rotated between two academically oriented activities each day, one math and one literacy. Literacy activities for different grade levels were drawn from three online literacy curricula from Houghton Mifflin Harcourt: *iRead, System 44*, and *Read 180*[3]. The math curriculum, *Summer Staircase*, was locally developed by SPS staff.

Each offering operated five days per week over a six-week period between 8:30 a.m. and 12:30 p.m. (with a greeting circle from 8:30 a.m. to 9:00 a.m.) for a total 29 program days and 90 academic contact hours. This structure allowed us to tie a unique quality rating (based on two observations of the staff team) to a unique group of students who were taught by the same team of teachers each day.

Participants

The study staff consisted of 40 individual teachers grouped into 30 teacher teams across 11 summer program sites. Offerings included 20 grade 3-4 offerings and 10 offerings for grade 1-2 or grade 2-3. Offerings were almost evenly distributed among the 11 school sites. Table 1 provides detail regarding the study sample and the numbers of students for whom the academic skills data were available.

Table 1 – Sample Characteristics

	Seattle Public Schools
Number of school sites	11
Number of offerings	30
Number of instructors	40
Grades served	1-4
Total number of students served	500
Students with math assessment data	224
Students with literacy assessment data	404
Total number of students with assessment data	421
Total number of students with complete assessment data	158

Source: Seattle Public Schools Summer Learning Program Quality Assessment (2015) and Seattle Public Schools Math and Literacy Scores (2015)

Table 2 indicates that quality data were collected for nearly all offerings (two offerings had only a single quality rating) and that some academic data were collected in 27 of 30 offerings. A test was conducted to determine if lower quality offerings might have more missing data for students, and they did not.

Table 2 - Sites, Offerings, and Completeness of Data by Site

	Number of	Offerings with	Offerings with student

[3] For more information, visit these Houghton Mifflin Harcourt product websites: System 44: http://www.hmhco.com/products/system-44/; Read 180: http://www.hmhco.com/products/read-180/; or iRead: http://www.hmhco.com/products/iread/

	offerings	complete quality assessment data	assessment data
Viewlands Elementary	3	3	3
John Rogers Elementary	2	2	0
B.F. Day Elementary	3	3	3
West Seattle Elementary	3	3	3
Highland Park Elementary	3	3	3
Emerson Elementary	3	2	2
Van Asselt Elementary	2	2	2
MLK Jr. Elementary	2	2	2
Hawthorne Elementary	3	2	3
Leschi Elementary	3	3	3
Sand Point Elementary	3	3	3

Source: Seattle Public Schools Summer Learning Program Quality Assessment (2015) and Seattle Public Schools Math and Literacy Scores (2015)

Measures

Measures for the study included the following:

Summer Learning PQA - Form A. Form A is an observation-based measure designed to rate the quality of instructional practices in six domains: Safe Environment, Supportive Environment, Interaction, Engagement, Math, and Literacy. To complete Form A, assessors collected systematic, anecdotal notes and a detailed running record of staff behavior and youth responses during an offering session. Assessors then used the anecdotal records to score 58 rubrics, typically requiring about 60 minutes of time. Two Form As were completed during two separate offerings and then averaged together to produce a single rating for each offering.[4]

Summer Learning PQA – Form B. Form B is an interview-based assessment of management practices. To complete Form B, the assessor interviews the program manager and records written responses. Later this written record is used to score 11 rubrics, typically requiring about 30 minutes. The Form B interview with the program manager assesses management practices in four domains: Planning, Staff Training, Family Connection, and Individualization. One Form B was completed for each school site. Because the Form B data were not of central concern in this report, all further results are presented in Appendix A.

Math Scores. Math assessments were constructed from a bank of approved items developed by SPS staff and aligned with the *Summer Staircase* math curriculum. Offering leads were allowed to select different sets of test items for the assessment but were required to administer the same assessment as the pre- and post-assessment. We calculated a percent change score for each student by subtracting the

[4] A substantial validity argument exists for the PQA assessments and for the construction and use of composite ratings (C. Smith, T. Akiva, et al., 2012; Charles Smith, Akiva, Sugar, & Hallman, 2012) and an emerging body of evidence specifically about the Summer Learning PQA (see reliability and validity appendices in Smith et al., (2015) and Smith et al., (2015)).

number correct at pre-assessment from the number correct at post-assessment and then divided by the total number of items at pre-assessment.

Literacy Scores. Literacy assessments were counts of completed lessons/units recorded online as students completed one of three SPS online literacy curriculum. Because all three curricula were developed by the same publisher, we were able to request normative cut points to produce a four-level scale which described the number of lessons/units completed for each curriculum as a categorical proficiency level: low, mid-low, mid-high, and high.

Attendance. Attendance was measured as the total number of offering sessions attended for each individual student out of a total possible 29 sessions.

Data Collection

All data collection was conducted by SOWA in conformance with data collection and data management protocols approved by the Weikart Center and SPS staff. All raters were required to have a current reliability endorsement for the Summer Learning PQA. Raters observed for one entire 8:30 a.m. to 12:30 p.m. session on each of two days at least 1.5 weeks apart and produced one complete Summer Learning PQA Form A rating for each observation day. First observations were conducted between June 29 and July 21, 2015. Second observations were conducted between July 16 and July 27, 2015. Form B interviews were completed during the second observation date. All programs ended on July 31, 2015.

SPS staff coordinated student data collection and supplied the Weikart Center with a complete, de-identified data file for analyses.

Analytic Approach

The primary purpose of the quality-outcomes evaluation design is to first differentiate intervention (e.g., summer learning offerings) subgroups by quality of instruction, and then to compare rates of individual student growth (e.g., pre-to-post change) across the quality subgroups. This "skill growth by levels of quality" design has been used with some frequency in early childhood evaluations (e.g., Karoly, 2014; Thornburg, Mayfield, Hawks, & Fuger, 2009). While this design does not achieve the high certainty of inference entailed by randomized or some quasi-experimental designs that seek to equate groups at baseline, it does (1) make cost-effective use of data already produced by the QIS and (2) transparently aligns with the theory that quality of the service is integral for student skill change.[5]

Analyses and reporting were conducted by the Weikart Center and an analytics subcontractor during the months of October and November 2015. Findings are summarized in this report. A supplementary technical discussion of methodology, analyses, and findings are provided in Albright &

[5] The quality-outcomes design can be improved by using propensity score methods to equate students in high quality offerings with students in lower quality offerings at baseline. See Recommendations section.

Guyon-Harris (2015). The pattern of findings described here reflects our best effort to fit the available data into the models that make optimal sense given the resources available.[6]

Results

In this section, we summarize findings from analyses conducted by Weikart Center staff and by the Weikart Center's subcontractor, Methods Consultants of Ann Arbor (see Albright & Guyon-Harris, 2015).

Instructional Quality and Quality Subgroups

This study produced detailed information about the quality of instructional practices overall, in relation to other summer learning samples, and as a profile of quality subgroups.

Instructional Quality in 30 Offerings and in Comparison to other Summer Learning Samples. Figure 2 shows average quality ratings for the six Form A domains. Overall, the 30 summer learning offerings demonstrated levels of quality that are high in comparison to the Weikart Center's normative databases for the Safety, Support, and Interaction domains, which are nearly identical to the widely used Youth Program Quality Assessment (PQA). The Instructional Total Score is the average of the Support, Interaction, and Engagement domain scores.

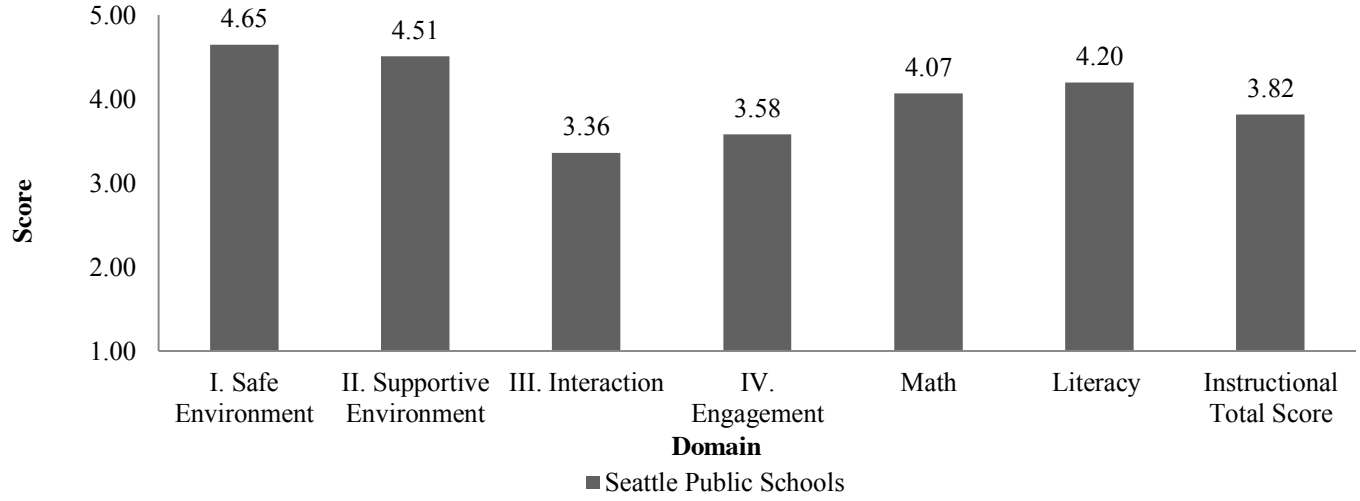
Instructional Quality in Comparison to Other Samples. A better comparison for SPS summer learning programs is with data from other summer learning network samples using the same Summer Learning PQA measures. Figure 3 presents SPS average domain scores in comparison with two additional samples. Comparison sample 2015 includes 31 summer program offerings at 20 sites in two cities (C. Smith, Ramaswamy, Hillaker, et al., 2015). Comparison sample 2014 includes 32 summer program offerings at 21 sites in four cities (Ramaswamy et al., 2014). The sample of 30 summer offerings described in this study were of substantially higher quality than both of the comparison samples.

Instructional Quality Subgroups. Figure 4 describes three performance subgroups identified by subjecting the quality ratings for the 30 SPS offerings to a latent profile analysis. The higher quality subgroup included eight offerings and 131 students. The middle quality subgroup included 18 offerings and 306 students. The lower quality subgroup included three offerings and 44 students. Detailed discussion of the latent profile analysis method used to identify the quality subgroups is available in Albright and Guyon-Harris (2015).

---

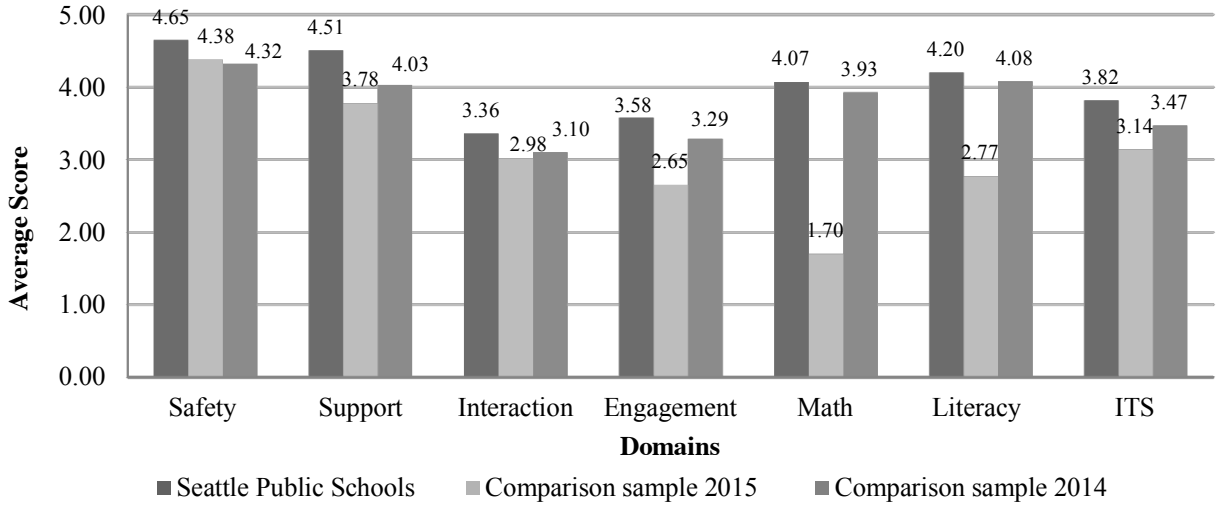
[6] We also executed three level hierarchical models to test for the effects of school site, student attendance, student gender, and the interaction of the school site indicator with offering quality. The basic pattern of results reflected in this report was maintained, although each of these variables explained a unique portion of the variance in the outcomes in some of the models. These results are not reported due primarily to our skepticism about using linear models to detect quality-outcome relationships. We provide recommendations to strengthen the evaluation design, and hence models that help us understand the data, in the final section of the report.

Figure 2 – Quality of Summer Program Instruction: Six Domains



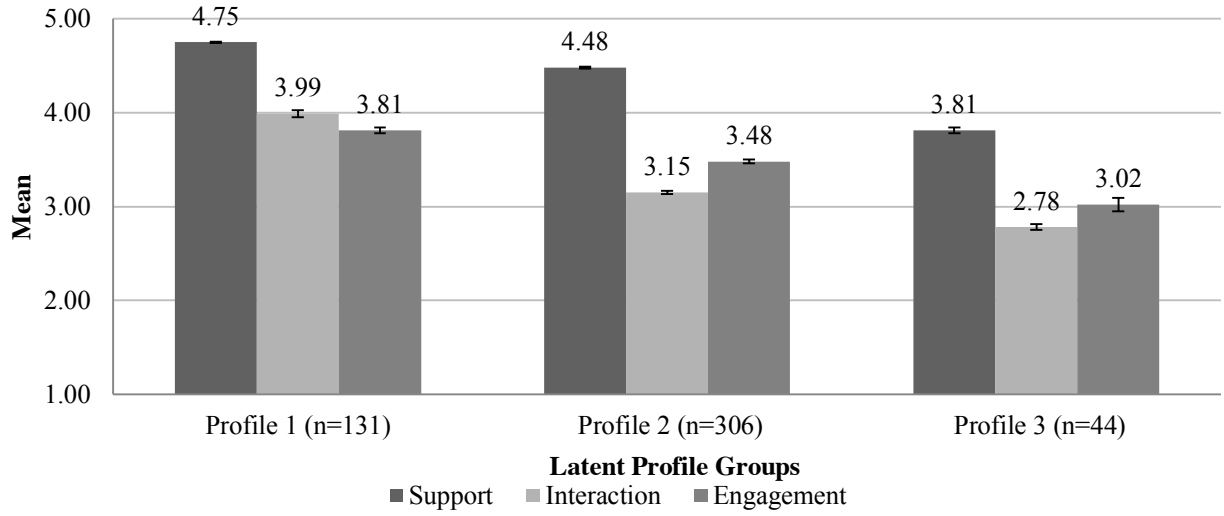
Source: Seattle Public Schools Summer Learning PQA Scores Reporter (2015), N=58

Figure 3 – Summer Program Domains 2015 and Two Comparison Samples



Source: Seattle Public Schools Summer Learning PQA Scores Reporter (2015), N=58, Summer Learning Program Quality Intervention Phase III Interim Report (2015), N=31, Summer Learning Program Quality Intervention (SLPQI): Phase Two Feasibility Study (2014), N=32

Figure 4 - Latent Profile Analysis Classifying Three Quality Subgroups



Source: Source: Albright and Guyon-Harris (2015).

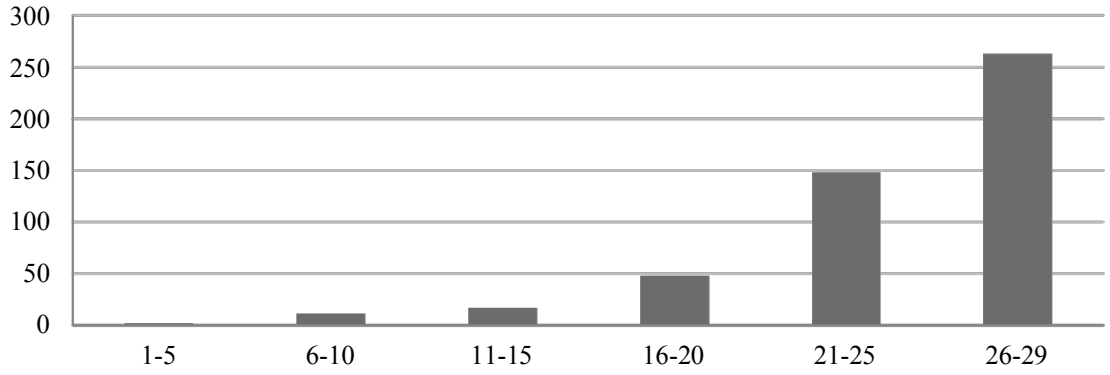
Student Attendance and Academic Skills

Attendance. In general, attendance was very high at SPS summer programs with mean attendance across all offerings of 24.40 days (SD=4.83, range 3-29). The frequency of days attended is shown in Figure 5.

Academic Skills. Students in SPS summer learning programs were tested on the same math items at the beginning and end of the summer program cycle and a percent change math score was calculated for each student. The distribution of percent change in student math scores is provided in Figure 6, indicating that most SPS students made gains in math skills during the summer cycle.

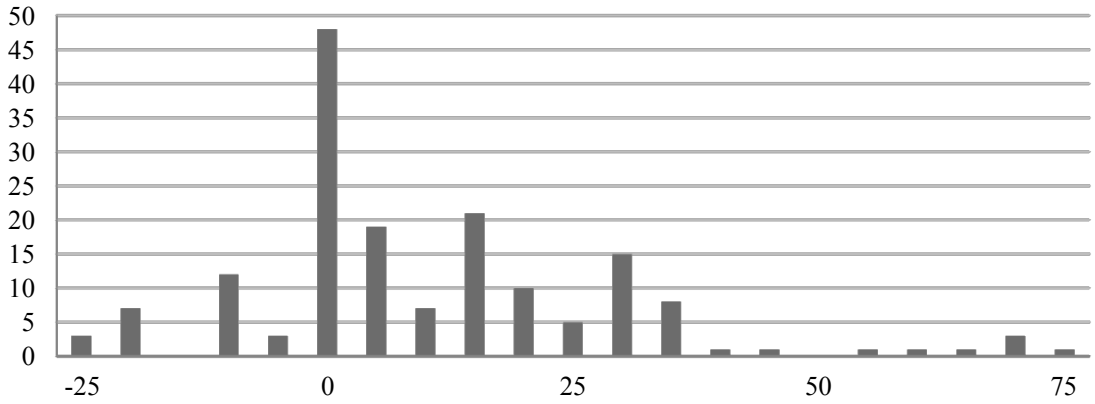
During the summer cycle, all students also completed a leveled language arts curriculum with multiple lessons/units. Four normative proficiency categories were created for the number of lessons/units completed. The frequency of SPS summer students falling within in each proficiency category is provided in Figure 7. Most students (68.8 percent) fell within the high performance category.

Figure 5 – Frequency of Student Days of Attendance



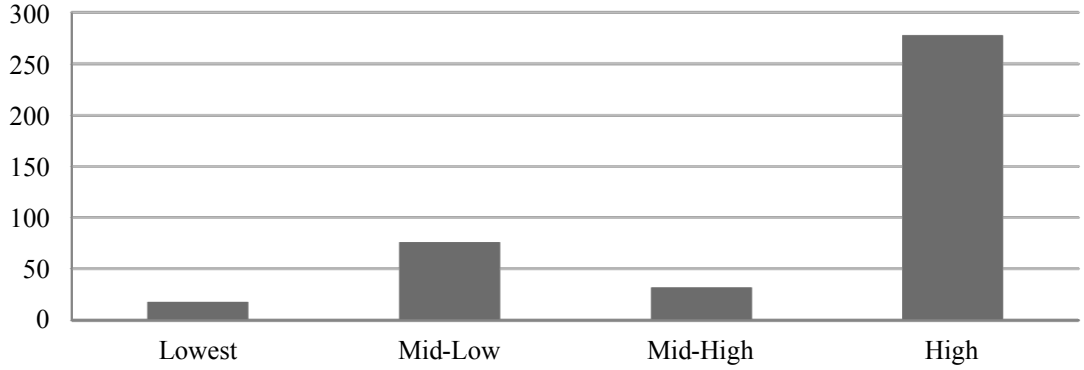
Source: Seattle Public Schools Math and Literacy Scores (2015), N=500

Figure 6 – Frequency of Student Math Change Scores



Source: Seattle Public Schools Math and Literacy Scores (2015), N=500, n=167

Figure 7 – Frequency of Student Literacy Proficiency Levels



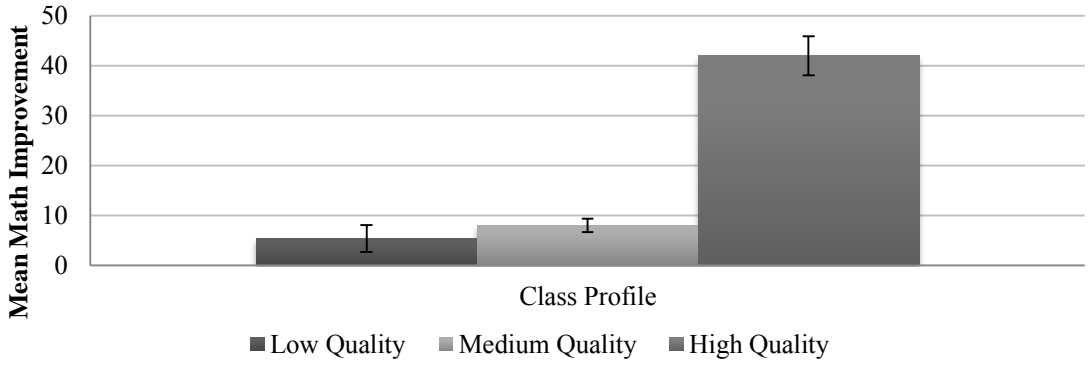
Source: Seattle Public Schools Math and Literacy Scores (2015), N=500, n=404

Skill Growth by Quality Subgroup

Math Skill Growth. Figure 8 presents math skill change scores by each of three quality subgroups. The difference between average student performance in the high quality subgroup and student performance in the other two subgroups was large and statistically significant.

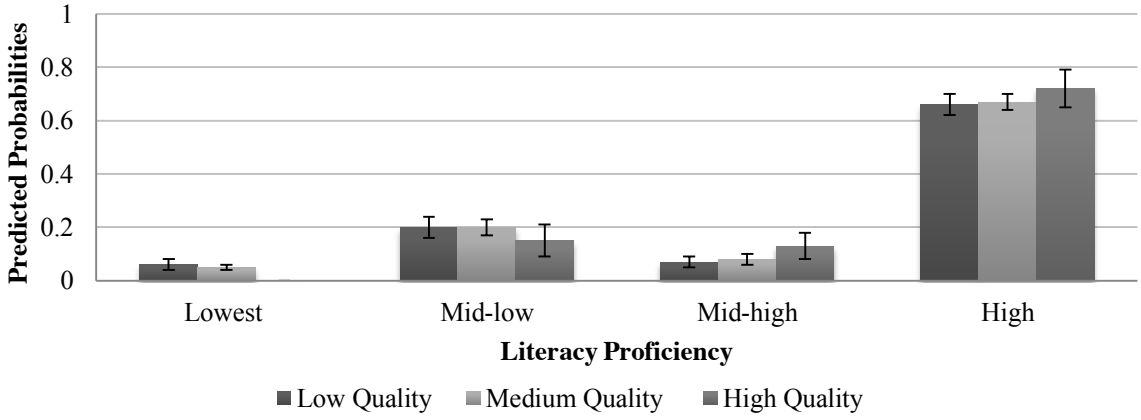
Literacy Skill Growth. Figure 9 presents the probability (ranging between 0 and 1) that a student is in one of the four levels of the literacy outcome. The average probability for all students in the high, medium, and lower quality subgroups is then compared. Here the evidence of differences by subgroup is smaller but still statistically significant. In particular, students in the high quality subgroup were more likely to be in the highest literacy proficiency level while none of the students in the high quality subgroup were in the lowest literacy proficiency level.

Figure 8 - Math Improvement by Latent Profile for Each Groups



Source: Albright and Guyon-Harris, 2015.

Figure 9 - Predicted Literacy Proficiency Probabilities by Latent Profile



Source: Albright and Guyon-Harris, 2015.

Discussion of Findings and Recommendations

This quality-outcomes study is the most recent in a series of evaluations focused on the design and scaled implementation of a QIS intervention for networks of summer learning programs. The primary goals of this study were to extend that work by (1) describing the effectiveness of SPS summer programs in terms of instructional quality and youth skills and (2) to accumulate further validity evidence for the Summer Learning PQA as a quality standard for summer learning programs focused on school success outcomes.

Context for Evidence and Policy. Summer learning programs are positioned to play an important role in reducing summer learning losses that disproportionately affect disadvantaged students (Alexander, Entwisle, & Olson, 2007; Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996), and summer learning programs with an explicit focus on improving academic skills are an important part of the OST landscape (Boss & Railsback, 2002; Newhouse, Neely, Freese, Lo, & Saili, n.d.).

While a growing literature suggests that summer learning programs can impact academic and other school-related skills (Borman & Dowling, 2006; Chaplin & Capizzano, 2006; McCombs, Augustine, & Schwartz, 2011; McCombs et al., 2014; Roderick, Engel, & Nagaoka, 2003), few rigorous studies have closely examined the specific features and practices that mediate or moderate relationships between summer program participation and school success outcomes (Arbreton et al., 2008; Spielberger & Halpern, 2002).

This relatively oblique understanding about the specific practices that support skill development in young learners limits the potential of summer learning programs. In particular, without measures of practice that are both sufficiently precise and feasible to implement, it is difficult to provide either validated standards that drive planning for high quality services or to produce performance feedback necessary for accountability and improvement. Because explicit off-the-shelf classroom interventions and curricula have proven difficult to implement with fidelity or at scale, the identification of best practices related to student learning is a subject of growing interest across educational fields (Jones & Bouffard, 2012).

The Summer Learning PQA is a measure of best practices for academically focused summer instruction that, when implemented as part of an effective QIS intervention[7], is designed to advance best practices at scale. This study presents the first direct evidence that the Summer Learning PQA standard

[7] Increasing evidence suggests that the continuous improvement approach may prove to be an effective way to bring best practices to scale (C. Smith & Akiva, 2008). The *Summer Learning Program Quality Intervention* (SLPQI) is a quality improvement intervention for summer learning systems that includes four core components: (1) a standard and measures for quality of management and instructional practices—the Summer Learning PQA used in this study, (2) training and technical assistance supports, (3) performance data products, and (4) a continuous improvement cycle that aligns the prior three elements with local circumstances and resources.

for instructional quality is related to positive change in academic skills demonstrated during the OST program (linking box 1 to box 3 in Figure 1). By identifying quality subgroups, this study further suggests that standards for high quality summer programs can be benchmarked using an existing performance measure.

Findings

The findings from this quality outcomes study suggest that the key findings for the study include:

1) On average, SPS summer program offerings demonstrate high levels of instructional quality compared to summer programs in other cities. Programs were also well attended.
2) Embedded assessments for math and literacy suggest that most students improved academic skills during the summer program cycle.
3) Summer programs can be divided into three performance subgroups with distinct quality profiles: Very high, moderate, and lower quality.
4) Students in offerings with very high quality instruction had more positive change on math and literacy assessments when compared to students receiving moderate and lower quality instructional experiences.

The evaluation design of this study did not include a rigorously matched comparison group or directly address questions about the effect of summer program participation on school success outcomes during a subsequent school year. Results should be interpreted cautiously.

Recommendations

We offer three primary recommendations that follow from the discussion of the policy context and study findings:

First, it is clear that the quality of instructional practices in SPS summer programs is high, particularly in the high quality subgroup. These exemplary offerings should be the subject of a curriculum and best practices study that would manualize the sequence of content and activities that teachers plan for the summer session and the responsive practices that they use to keep youth engaged as learning or interpersonal challenges occur. This documentation could extend to performance benchmarks for instructional quality so that the Summer Learning PQA can support high fidelity implementation.

Second, because summer school is both a huge public investment and because much summer school is apparently not sufficiently high quality, it is critical to test the effects of high quality with a more rigorous evaluation design. If the high quality subgroup is a threshold for effects—as it was for both math and literacy skills in this study—then moderate quality may not be worth the investment. A more

rigorous design that would produce an extension of the results while maintaining very low cost (perhaps double, depending on the number of sites) should include:

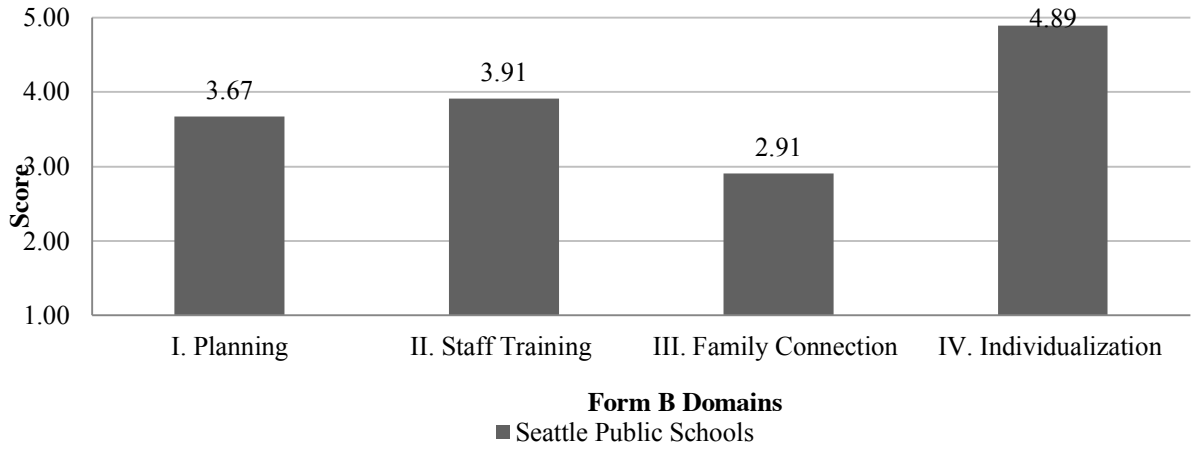
Improvement of Measures. Measurement of academic skills demonstrated during the summer offering could be improved by having greater control over the skill measures and data collection and adding a middle time point.

Improve Rigor of Impact Estimates. Several methods could be employed to increase the certainty of inference about effects of offering quality on academic skills. In order to replicate the test for a relationship between high quality and academic skill growth, a more rigorous, matched control group design that employs propensity score methodology should be used to match students in the high quality offerings to students in the lower quality offerings who were very similar at baseline. This design would also allow the two groups of students to be compared on subsequent school year performance as well, creating an impact estimate for high quality summer school on school day achievement. This method could be further extended to include matching of students who were similar to those in high quality at baseline but who did not participate in summer programs (the no-program group). This design, and the several variations mentioned here, is post hoc analysis in that matching of students occurs after the program has taken place, dramatically reducing need for control over assignment of students and cost.

Appendix A – Results for Management Practices

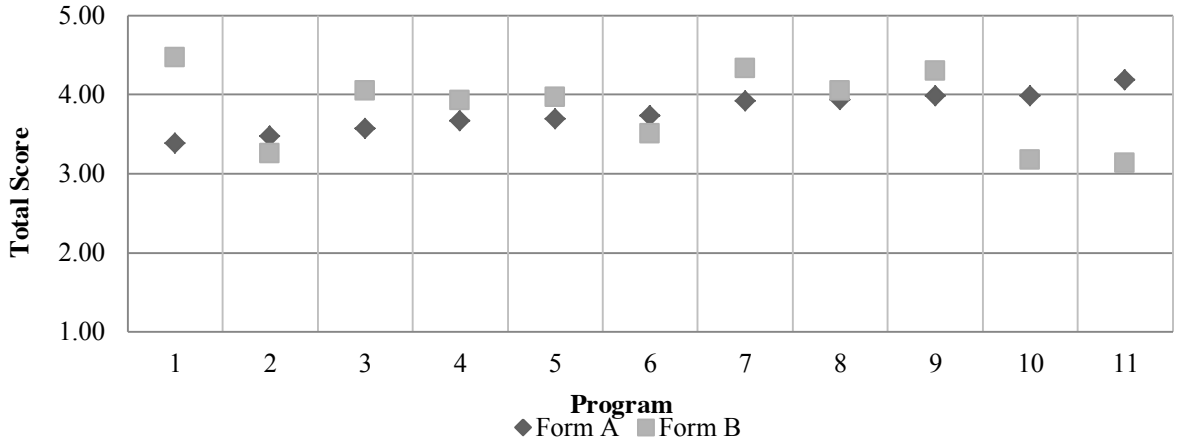
The Summer Learning PQA Form B includes four domains: Planning, Staff Training, Family Connection and Individualization. Figure 5 provides domain averages for all 11 program sites in the study sample. Figure 6 provides the Form B total score for management practices, a mean score across the four domains, by the Form A Instructional Total Score to present a profile of program site quality in terms of management practices and instructional practices. For the 11 school sites, these two composite scores have a Pearson-*r* correlation coefficient of *r*= -0.29.

Figure A.1 – Quality of Management Practices



Source: Seattle Public Schools Summer Learning PQA Scores Reporter (2015), N=58

Figure A.2 - Management and Instructional Quality by Site



Source: Seattle Public Schools Summer Learning PQA Scores Reporter (2015), N=58

References

Akiva, T., Cortina, K. S., Eccles, J. S., & Smith, C. (2013). Youth belonging and cognitive engagement in organized activities: A large-scale field study. *Journal of Applied Developmental Psychology, 34*(5), 208-218. doi:http://dx.doi.org/10.1016/j.appdev.2013.05.001

Albright, J. J., & Guyon-Harris, K. (2015). *Results of SLPQA Data Analysis*. Retrieved from Ypsilanti, MI:

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007). Lasting Consequences of the Summer Learning Gap. *American Sociological Review, 72*, 167-180.

Arbreton, A., Sheldon, J., Bradshaw, M., Goldsmith, J., Jucovy, L., & Pepper, S. (2008). *Advancing achievement: Findings from an independent evaluation of a major after-school initiative*. Retrieved from San Francisco:

Borman, G. D., & Dowling, N. M. (2006). Longitudinal Achievement Effects of Multiyear Summer School: Evidence From the Teach Baltimore Randomized Field Trial. *Educational Evaluation and Policy Analysis, 28*(1), 25-48.

Boss, S., & Railsback, J. (2002). *Summer School Programs: A Look at the Research, Implications for Practice, and Program Sampler*. Retrieved from Washington, D.C.:

Chaplin, D., & Capizzano, J. (2006). *Impacts of a Summer Learning Program: A Random Assignment Study of Building Educated Leaders for Life*. Retrieved from

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The Effects of Summer Vacation on Achievement Test Scores: A Narrative and Meta-Analytic Review. *Review of educational research, 66*(3), 227-268.

Jones, S., & Bouffard, S. (2012). *Social and emotional learning in schools: From programs to strategies*. Retrieved from

Karoly, L. A. (2014). *Validation Studies for Early Learning and Care Quality Rating and Improvement Systems: A Review of the Literature*. Retrieved from

McCombs, J. S., Augustine, C. H., & Schwartz, H. L. (2011). *Making summer count: How summer programs can boost children's learning*. Santa Monica, CA: RAND Corporation.

McCombs, J. S., Pane, J. F., Augustine, C. H., Schwartz, H. L., Martorell, P., & Zakaras, L. (2014). *Ready for Fall*. Retrieved from Santa Monica, CA:

Naftzger, N. (2014). *A Summary of Three Studies Exploring the Relationship between Afterschool Program Quality and Youth Outcomes*. Paper presented at the Ready by 21 National Meeting, Covington, KY.

Naftzger, N., Devaney, E., & Foley, K. (2014). *Summary of Analyses Related to Nashville After Zone Alliance Program Outcomes*. Retrieved from Chicago, IL:

Naftzger, N., Manzeske, D., Nistler, M., Swanlund, A., Rapaport, A., Shields, J., . . . Sugar, S. (2013). *Texas 21st Century Community Learning Centers: Final Evaluation Report*. Retrieved from Naperville, IL:

Newhouse, C., Neely, P., Freese, J., Lo, J., & Saili, W. (n.d.). *Summer Matters: How Summer Learning Strengthens Student's Success*. Retrieved from

Ramaswamy, R., Gersh, A., Sniegowski, S., McGovern, G., & Smith, C. (2014). *Summer Learning Program Quality Assessment 2013 Phase I Pilot Report*. Retrieved from Ypsilanti, MI:

Roderick, M., Engel, M., & Nagaoka, J. (2003). *Ending Social Promotion: Results from Summer Bridge*. Chicago, IL: Consortium on Chicago School Research.

Smith, C., & Akiva, T. (2008). Quality accountability: Improving fidelity of broad developmentally focused interventions. In H. Yoshikawa & B. Shinn (Eds.), *Transforming Social Settings: Towards Positive Youth Development*: Oxford University Press.

Smith, C., Akiva, T., Devaney, T., & Sugar, S. (2009). Quality and accountablility in the out-of-school time sector. In R. Granger, K. Pittman, & N. Yohalem (Eds.), *New Directions for Youth Development: Defining and measuring quality in youth programs and classrooms* (Vol. 121). San Francisco: Jossey-Bass.

Smith, C., Akiva, T., Sugar, S., Lo, Y. J., Frank, K. A., Peck, S. C., & Cortina, K. S. (2012). *Continuous quality improvement in afterschool settings: Impact findings from the Youth Program Quality Intervention study*. Retrieved from Ypsilanti, MI:

Smith, C., Akiva, T., Sugar, S. A., & Hallman, S. (2012). *Leading indicators measurement system: Analysis of Oklahoma data - Technical appendix to the Oklahoma 21st Century Community Learning Centers statewide evaluation*. Retrieved from Ypsilanti, MI:

Smith, C., Hallman, S., Hillaker, B., Sugar, S., McGovern, G., & Devaney, E. (2012). *Development and early validation evidence for an observational measure of high quality instructional practice for science, technology, engineering and mathematics in out-of-school time settings: The STEM supplement to the Youth Program Quality Assessment*. Retrieved from Ypsilanti, MI:

Smith, C., Hillaker, B., & McGovern, G. (2015). *Academic Skill-Building Program Quality Assessment: Measure Development Study*. Retrieved from Ypsilanti, MI:

Smith, C., Ramaswamy, R., Gersh, A., & McGovern, G. (2015). *Summer Learning Program Quality Intervention (SLPQI): Phase Two Feasibility Study*. Retrieved from Ypsilanti, MI:

Smith, C., Ramaswamy, R., Hillaker, B., Helegda, K., & McGovern, G. (2015). *Summer Learning Program Quality Intervention Phase III Interim Report*. Retrieved from Ypsilanti, MI:

Spielberger, J., & Halpern, R. (2002). *The Role of After-School Programs in Children's Literacy Development*. Retrieved from Chicago, IL:

Thornburg, K. R., Mayfield, W. A., Hawks, J. S., & Fuger, K. L. (2009). The Missouri quality rating system school readiness study. Columbia, MO: Center for Family Policy & Research.