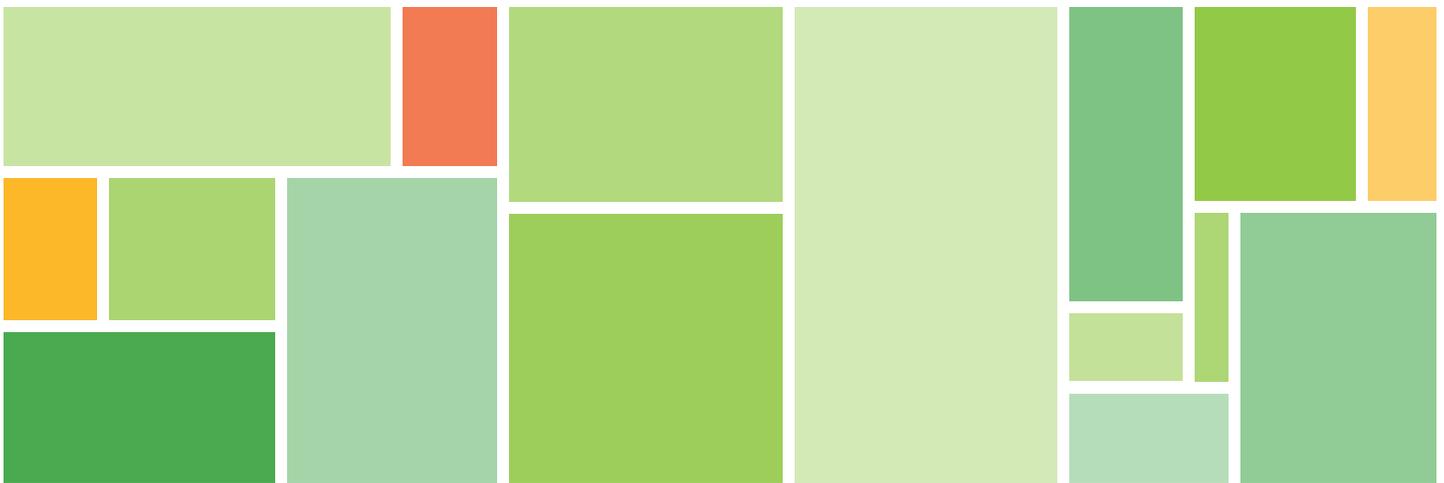


Final Report on the Palm Beach Quality Improvement System Pilot

*Model Implementation and Program Quality Improvement
in 38 After-school Programs*



Charles Smith, Tom Akiva, Juliane Blazeovski, and Thomas
Devaney
The David. P. Weikart Center for Youth Program Quality
HighScope Educational Research Foundation

Lisa Pelle
Independent Consultant, Prime Time Inc.

January 2008

Acknowledgements

We would like to acknowledge the following persons and organizations for their support of this work. At Prime Time Palm Beach County, Dominique Arrieux, Teal Chance, and Katherine Gopie provided collaboration and commitment to excellence in the design and delivery of the QIS. Elaine Mancini at Family Central has provided exceptional effort and rigorous quality control for both data collection and capacity building. At High/Scope, Linda Horne provided excellent and timely support in all phases of the project. At the Center for Youth Program Quality, our new joint venture between the High/Scope Foundation and the Forum for Youth Development, Samantha Sugar, Kynisha Johnson, Joe Bertolotti, and Erica Curry have all leant their skills and expertise to the project. A big thanks to the staff at the Forum for Youth Investment who offered support and advice throughout the project. Finally, perhaps most of all, we are indebted to the directors and staff of the 38 pilot programs in Palm Beach County who participated in the QIS. These folks demonstrated commitment to youth and the courage and strength to reflect on their own practice and to try new things.

Related Reports

The High/Scope Foundation prepared four reports following data collection for the QIS baseline: *Quality in the Palm Beach County QIS: Final report from the QIS Baseline Data Collection* (High/Scope, 2006); *Technical Report: Quality in the Palm Beach County QIS Baseline Data Collection* (High/Scope, 2006); *Training Satisfaction for High/Scope Workshops Delivered as Part of the Palm Beach QIS* (High/Scope, 2006); *Communities of Practice in the Palm Beach County QIS: A Preliminary Look at Findings from a Staff Survey* (High/Scope, 2006).

Three formal evaluation reports on the QIS development and implementation process have been prepared by an outside evaluation contractor, the Chapin Hall Center for Children at the University of Chicago. Two of these reports are publicly available at the Chapin Hall website and provide excellent detail on both the QIS background and the

process of capacity building at the intermediary and provider levels. These unique reports highlight key learnings about QIS policy development as well as the accountability concerns of individual program managers.

Copyright

Copyright © 2008 by David P. Weikart Center for Youth Program Quality at the Forum for Youth Investment. All rights reserved. FYI is a registered trademark and service mark of the Forum for Youth Investment.

Table of Contents

Acknowledgements	2
Related Reports	2
Copyright.....	2
Table of Contents	3
Summary	5
Part I. Overview of the QIS Pilot Study	7
Introduction	7
QIS Pilot History & Components	11
QIS Pilot Participants and Measures	15
Part II. QIS Pilot Findings	19
Change in Point-of-Service Quality During the QIS Pilot.....	19
Change in Organizational Practices and Policies During the QIS Pilot	23
QIS Model Fidelity	26
Part III. Formative Analysis	28
The QIS Model and Quality Change	28
Additional Formative Analyses	30
Part IV. Conclusions and Recommendations.....	33
Endnotes	34
References.....	35
Appendix A. Emerging QIS Theory of Change.....	38
Appendix B. Standards Crosswalk.....	39
Appendix C. Domains, Scales and Items for the PBC-PQA Instruments.....	42
Appendix D. Psychometric Performance of the PBC-PQA.....	47

Summary

This report on the Palm Beach County Quality Improvement System (QIS) pilot provides evaluative findings from a four-year effort to imagine and implement a powerful quality accountability and improvement policy in a countywide network of after-school programs. The Palm Beach QIS is an assessment-driven, multi-level intervention designed to raise quality in after-school programs, and thereby raise the level of access to key developmental and learning experiences for the youth who attend. At its core, the QIS asks providers to identify and address strengths and areas for improvement based on use of the Palm Beach County Program Quality Assessment (PBC-PQA)—a diagnostic and prescriptive quality assessment tool – and then to develop and enact quality improvement plans. Throughout this process training and technical assistance are provided by several local and national intermediary organizations.

We present baseline and post-pilot quality ratings for 38 after-school programs that volunteered to participate in the Palm Beach QIS pilot over a two-year cycle. This data is the routine output from the QIS system and is designed to support evaluative decisions by program staff and regional decision-makers. In addition to the typical QIS output, we also provide as much detail as possible about the depth of participation in the various elements of the improvement initiative and offer a few opinions about what worked.

Main findings include:

- ***Quality changed at both the point of service and management levels.*** During the QIS quality scores changed substantially at both the point of service and management levels, suggesting that the delivery of key developmental and learning experiences to children and youth increased between baseline and post-pilot rounds of data collection.
 - Point-of-service quality increased most substantially in areas related to environmental supports for learning and peer interaction, but positive and statistically significant gains were evidenced in all assessed domains of quality.
 - The incidence of organizational best practices and policies increased in all assessed management-level domains, especially staff expectations, family connections and organizational logistics.
- ***Planning strategies that targeted specific improvement areas were effective.*** Pilot sites registered larger quality gains on point of service metrics that were aligned with intentionally selected areas for improvement. This indicates that the quality improvement planning process effectively channels improvement energies.
- ***Site managers and front line staff participated in core elements of the QIS at high rates.*** Relative to other samples, participation by front line staff was especially high, suggesting that the core tools and practices of the QIS are reasonably easy for site managers to introduce into their organizations.
- ***The core tools and practices of the QIS were adopted at high rates.*** Thirty-five of 38 sites (92%) completed the self-assessment process and 28 sites (74%) completed all of the steps necessary to submit a quality improvement plan.

A number of exploratory and formative analyses related to policy were also conducted. These exploratory and formative analyses are presented in Part III of the report. All should be treated with caution since they are drawn from a small sample and, in some cases, less than perfect data sources. Findings from the exploratory and formative analyses include:

- ***The low stakes approach to accountability within the QIS model appears to have increased provider buy in.***

Through review of secondary documents and quantitative data, the QIS emphasis on partnership rather than external evaluation achieved buy-in from pilot group providers for the self-assessment and improvement planning process.

- *The self-assessment and improvement planning sequence was associated with change in quality scores.* Programs that participated in the self-assessment process were more likely than those that did not to experience improvement in their quality scores.
- *Structural characteristics such as organization type, licensing status, supervisor education and experience levels were not strongly related to point-of-service quality.* This suggests that the variables most often manipulated by reform initiatives are, at best, weak drivers of setting quality and thus less-than-ideal policy targets. Put another way, these several program “credentials”, while reasonably easy to measure, were poor proxies for quality.

Introduction

High quality after-school programs provide youth with access to key experiences that advance developmental and learning outcomes (National Research Council, 2002; Durlak & Weissberg, 2007; Lauer, Akiba, Wilkerson, Apthorp & Snow et al., 2006). However, many after-school settings miss opportunities to provide these key experiences for the children and youth who attend (Granger, Durlak, Yohalem & Reisner, 2007; Smith, Peck, Denault, Akiva & Blazevski, in submission). This report on the Palm Beach County Quality Improvement System (QIS) Pilot provides findings from a four-year effort to imagine and implement a powerful quality accountability and improvement policy in a countywide network of after-school programs.

The Palm Beach QIS is an assessment-driven, multi-level intervention designed to raise quality in after-school programs, and thereby raise the level of access to key developmental and learning experiences for the youth who attend. Unlike narrower interventions designed to produce specific effects with packaged curricula, the QIS supports a broad, developmentally focused intervention model targeting programs with varied content, structures, and missions. The QIS asks providers to identify and address strengths and areas for improvement based on use of the Palm Beach County Program Quality Assessment (PBC-PQA)—a diagnostic and prescriptive assessment that measures fidelity to the values and methods of positive youth development. The PBC-PQA and the QIS model meet a need for policy vehicles that simultaneously address the quality of proximal program experiences available to youth, organizational improvement systems, and place-based workforce development strategies (Wilson-Ahlstrom & Yohalem, 2007; Akiva & Yohalem, 2006).

Emerging Quality Measurement Systems

Numerous states and localities are focused on improving the return from existing investments in after-school programs. According to a recent tally, 14 states are implementing quality accountability systems for subsidized childcare and another 30

are “exploring/designing or piloting”¹ some type of quality improvement process. Several national funders have recently invested in after-school quality improvement projects in cities and counties across the country and numerous other place-based projects are moving forward with local resources.² Most of these system-level efforts entail some package of the core tools and practices present in QIS -- quality standards, observational metrics, improvement planning, and aligned coaching and training supports. However, this trend is relatively new and there are few experimental or descriptive studies that provide either evidence of effectiveness for specific models or that provide generic guidance about design.³ The Palm Beach QIS stands out as an exemplar for quality accountability and improvement policies in three areas: (1) an accountability approach that understands adult motivation; (2) a core focus on improving the developmental and learning experiences available to children and youth at the point-of-service; and (3) mobilization of authority, resources and partnerships in a place-based initiative.

In this era of high-stakes testing, the idea of accountability has taken on negative connotations for many educators and youth workers, often due to the assumed links between standardized test results, staff performance and school improvement (Halverson, 2005; Laitsch, 2006; Ryan & Brown, 2005; Wiggins, 1993). In contrast, emerging accountability and improvement models in the after-school field employ a different set of tools and assumptions, drawing upon our understanding of adult motivation (see discussion of self-determination theory in Ryan & Deci, 2000) and knowledge management systems (Mason, 2003). These emerging models are not premised on holding professional staff directly “accountable” for either peak performances of children on standardized tests administered during the school day, or for demonstrating change on population level indicators that individual programs are unlikely to influence in isolation (e.g., teen pregnancy rates for a community). Rather, the emerging models attempt to empower after-school managers and youth workers to improve the quality of their own performances – the instructional and therapeutic technologies

Part I.

Overview of the QIS Pilot Study

that they apply with children at the point of service in after-school programs - according to known standards and employ reliable metrics to mark progress toward goals. Because an individual's own performance is a program output (outcome) over which they are likely to have direct control, these models create and channel energies for improvement that flow from the desire to build skills and fulfill mission through self-improvement.

This approach can only be successful where agreement exists about what constitutes a high quality performance, and fortunately, youth development researchers have produced a number of influential models of positive youth development practice (Larson, 2000; Lerner, 2005), research-based setting features (National Research Council, 2002; Durlak & Weisburg, 2007) and community-based standards for supports and opportunities (Gambone, Klem, & Connel, 2002; Little, 2007). This research base (1) supports the powerful idea that the quality of youth experiences in after-school programs can influence positive youth development and learning; and (2) defines a core set of key experiences that high quality after-school environments should provide for their constituents.

Leveraging this research base, we have developed a three-level framework to address the hierarchy of components that can affect and sustain improvements in the quality of after-school settings (Smith, Akiva, & Henry, 2006; Smith & Akiva, 2008): Point-of-Service (POS) is where youth and adults spend time together; Professional Learning Community (PLC) is an organizational level where program managers guide staff through adoption of core tools and practices; and the System Accountability Environment (SAE) contains the values, incentives, and priorities of funders and intermediaries that seek to influence groups of after-

school providers through regulation and funding.

The POS directly represents the proximal experiences that occur where adults and youth meet in day-to-day after-school programming. Figure 1 presents the definition of quality at the point-of-service employed both in the High/Scope Foundation's ongoing research program on learning environments and setting change, and in the Palm Beach QIS.

Several studies of quality from a diverse range of after-school settings suggest that while moderate to strong levels of psychological safety and emotional support are typically available, the frequency of opportunities for interaction and engagement (see Figure 1) are substantially lower across a majority of programs, regardless of program type, content focus or age of children served (Smith, Akiva & Henry, 2006; Learning Points Associates & Berkley Policy Associates, 2006; Gramiak, Vanauken, Brugger & Young-Miller, 2006; INCRE & NIOST, 2006; Walker & Arbreton, 2004). These studies find that structured and purposeful experiences of peer interaction (cooperative learning, leadership) and deep engagement (reflection, decision-making) are less frequently available to youth in the programs than are experiences with safe environments and caring adults.

The critical counterpoint to these findings about point of service quality is that relatively few programs concentrate their improvement energy on the nature of staff performances. We recently examined early data from an ongoing study of 100 after-school programs in four states to discover that less than 16 percent of all intentional improvement efforts conducted in these programs during the last year focused on elements described in the top three levels of figure 1.⁴ In summary, then, point-of-service quality in after-school programs is often

low but few after-school managers are focused on improving it. Thus there is a clear need in the youth development field for accountability tools and performance management metrics that not only distinguish between high- and low- quality performances, but also provide regulators, funders, network leaders, program managers and front-line staff with the ability to (1) identify discrete POS quality improvement opportunities that are most likely to influence positive youth development; and (2) track incremental progress toward POS quality improvement goals. The Palm Beach QIS is an exemplar of such a system.

Figure 2 is a model of change which we believe represents the emerging quality accountability and improvement systems in the after-school field. The components of this model are detailed in Appendix A. This model serves to illustrate the strengths of the Palm Beach QIS. Much has been studied and written about instruction and youth work methods that constitute the POS level, and about professional development and other activities which occur at the PLC level of the organization and the Palm Beach QIS certainly represents best practice in these areas. However, it is the focus on using a high-capacity intermediary in partnership with regulators and funders – the key SAE-level actors – that makes the Palm Beach QIS unique and important. The QIS uses the three SAE-level inputs on the far left of figure 2—accountability messages, standards & training, and advising/coaching—with an explicit intent to affect the PLC/POS factors that lead professional staff to improve their own practice. Key among these factors are the ability to derive meaning from data as increased understanding about the level of alignment between expectations and performance, and creating energy for change at both the individual and system levels. As individual staff received targeted performance feedback they can strategize about where to focus their improvement efforts. As site level managers come together to engage the QIS process through T&TA, they become

aware their role in a county-wide movement.

Aims of Report

This report is written by a lead technical assistance contractor for the QIS and as such we do not claim, nor do we strive for, the perspective of an outside evaluator. The story that we tell is part of an ongoing effort to advance QIS policy in Palm Beach County. Our approach to the QIS is best described as a design experiment, where product development, field work and rigorous measurement are blended together in an iterative cycle of implementation, feedback, and revision (Blumenfeld, Marx, & Harris, 2006; Brown, 1992). An external evaluation of the QIS development process is available in two reports (Spielberger & Lockaby, 2006; 2008) which are excellent companion pieces to this report.

The primary purpose of this report, however, remains evaluative. We present QIS Pilot baseline and post-pilot quality ratings for 38 after-school programs that volunteered to participate in the project over a two-year cycle. This data is the routine output from the QIS system and is designed to support evaluative decisions by program staff and regional decision-makers. In addition to this routine QIS output, we also provide additional detail about the depth of participation in the various elements of the QIS as well as offer several hypotheses about which elements of the intervention worked.

These latter portions of the report reflect our growing realization that implementation and model fidelity are critical if complex intervention models

Figure 1. Quality at the Point of Service

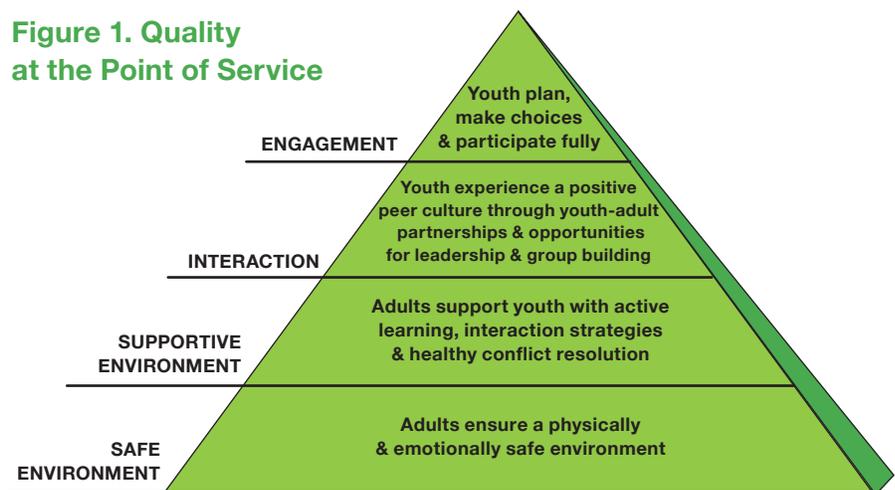
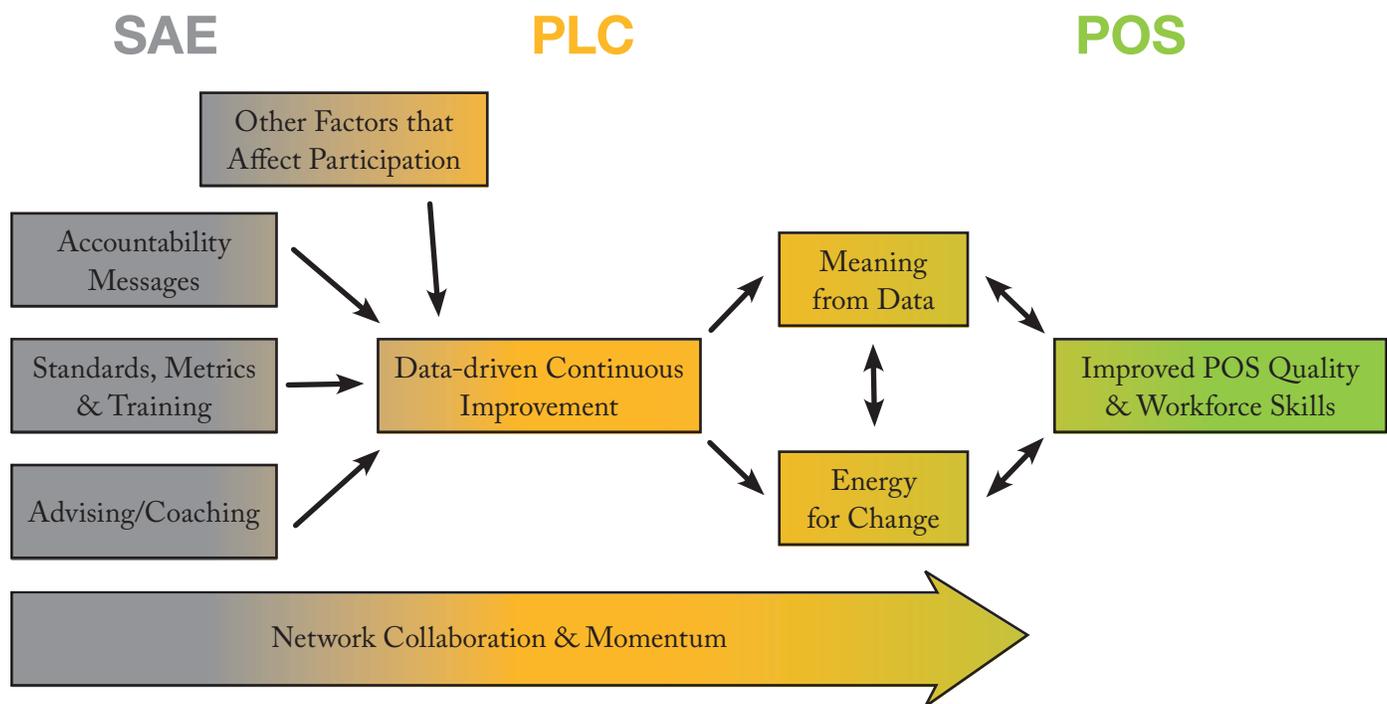


Figure 2. Model of Change Pathway from SAE Inputs to POS Improvement



like the QIS are ever to achieve scale in other places (Fixsen, Naoom & Blasé et al., 2005; Center for substance Abuse Prevention, 2002; National Research Council, 2002). Several research questions structure our assessment of QIS implementation fidelity:

At the program management level, we discuss fidelity to the QIS model as adoption of core tools and practices and use of QIS coaching and training resources. Specifically, we will address several questions: How much of the QIS did pilot programs actually participate in? How many of the pilot sites actually adopted the core tools and practices? How deeply did site managers involve direct staff in use of the core tools and practices?

Our second set of questions addresses actual change in organizational practices and the quality of youth worker performances with children and youth who attend after-school programs: Did positive change occur during the QIS intervention period? Were pilot sites able to undertake improvement plans that actually altered direct staff skill sets?

Although this study does not employ a research

design that supports causal interpretation, this is the first data set using the PBC-PQA measurement tools that will allow an examination of the magnitude and direction of score change during a formal quality improvement intervention. This is a issue for the success of all QIS-like systems and addresses the following question: Are quality metrics not only sensitive enough to differentiate between different programs of higher and lower quality at a single point in time, but also sensitive enough to capture real change that occurs over time within individual programs that are striving to improve?

The final questions we seek to address are formative: Which elements of the QIS were associated with positive change in program quality scores? Does quality differ systematically by characteristics of sites and staff? In effect, what organizational features and intervention components hold the most promise as quality drivers at both the point-of-service and professional learning community levels? Although our ability to make casual inferences about these questions is circumscribed by the lack of a true experimental design, we nevertheless feel that the body of evidence collected during and analyzed for

the QIS Pilot offers real guidance to policymakers, intermediaries and practitioners. Indeed, we ultimately make a strong circumstantial case for “what works” in the QIS to improve quality in after school settings.

QIS Pilot History and Components

Palm Beach County is the third largest county in Florida with an estimated 1.1 million residents and 237,459 children under the age of 18. More than 19 percent of those children live in poverty and nearly two-thirds (ages 6-17) live in single-parent households in which the parent works or two-parent households in which both parents work. The county has the 11th largest public school district in the nation with more than 160 schools and 166,000+ students. Fifty-two percent (52%) of those students are minorities, primarily African-American and Hispanic.

In 1996, key stakeholders in Palm Beach County, dedicated to developing quality after-school programs, formed the Out-of-School Consortium to share resources and enhance existing after-school and summer programs. From this activity, PRIME TIME Palm Beach County, Inc. was created in 2000 and incorporated as a 501(c)(3) in 2001. Since its creation, Prime Time has emerged as the county’s leading quality improvement intermediary, awarding professional development scholarships, providing training, facilitating Consortium meetings (provider networking), providing technical assistance, managing a resource lending library, developing activity modules for after-school programs, and offering other relevant services.

In 2004, Palm Beach county had an estimated 300 after-school providers operating more than 450 after-school programs including school-based, community-based and child care centers, operated by school staff, non-profit organizations (both with and without national affiliation), and for-profit organizations. Prime Time staff and other key stakeholders determined that the Palm Beach County after-school system lacked core features and frameworks to improve service delivery and build provider and workforce capacity. Specifically, Palm Beach County needed: (1) a set of common

quality standards; (2) metrics to assess compliance with those standards; (3) a framework for delivering training and technical assistance to program managers and front-line staff in support of quality improvement; and (4) countywide institutional relationships that could eventually grow into a system of credentialing for the after-school workforce. Since 2004, Prime Time has served as the catalyst for meeting these needs.

Development of Local Standards, Metrics and Outcomes Model

Quality standards establish a framework for common understanding and language about practices that positively affect program quality. Over 13 months (February 2004 – March 2005), a Standards Committee of diverse stakeholders was established and met to review the after-school literature and examples of existing quality standards from the National Afterschool Association as well as a variety of communities (Baltimore, Kansas City, Philadelphia, and others). The Standards Committee decided that while national standards were a good starting point, countywide standards needed to reflect local values and priorities. Key committee activities included: reviewing, sharing and discussing after-school research; discussing potential standards and indicators; and obtaining input from more than 1800 parents and 200 staff about their opinions on proposed quality standards and indicators. As a result of the committee’s work, five Quality After-school Standards for Palm Beach County were developed: (1) Administration, Program Organization, Procedures and Policies Provide Solid Framework for After-school Program; (2) Supportive Ongoing Relationships Between and Among Youth and Staff; (3) Positive and Inclusive Environment for Youth; (4) Youth Development and Challenging Learning Experiences; and (5) Outreach to and Activities for Families. Full detail for the Palm Beach County Standards for After-school Programs is provided in Appendix A.

To support adoption of the new standards, metrics for compliance were needed. Because it was not economically feasible to develop a tool and assessment process from the ground up, the committee used a competitive process to identify

the High/Scope Educational Research Foundation as a contractor to develop the Palm Beach County Program Quality Assessment (PBC-PQA). The High/Scope Youth PQA has a solid research base with established reliability and validity (Smith & Hohmann, 2005; Blazeovski & Smith, 2007) and provided the template for development of the PBC-PQA. The PBC-PQA is the core quality metric in the QIS and allows both program staff and external assessors to produce after-school program quality ratings which are aligned with the Palm Beach County standards. PBC-PQA Form A assesses quality at the point-of-service and Form B captures organizational practices and policies. Full detail for the PBC-PQA forms A and B is provided in Appendix C. Reliability and validity evidence for the PBC-PQA is provided in Appendix D.

A brief youth survey for older children and youth, grades 4 and up, was added to the QIS to provide youth voice as another source of evaluative evidence in the QIS system. The youth survey captures youth input on three dimensions: positive affect, sense of challenge, and program quality. Youth survey items and relationships between the youth survey and point of service quality ratings from PBC-PQA Form A are discussed in Appendix D.

A final task during QIS development was to design an outcomes model that linked program quality to youth outcomes. Figure 3 provides the model developed by the standards committee, summarizing the linkage between (1) key program inputs and standards and (2) intermediate and long-term outcomes. In general, the logic model suggests that when after-school programs deliver key developmental experiences to the youth that attend, desired intermediate and long terms outcomes will follow. Dashed lines indicate alignment between the outcomes model and QIS quality metrics.

The QIS model consists of supporting partnerships, model elements and sequence as well as a low stakes accountability approach. Each of these will be described separately in the main body of the report. It is important to note, however, that the overall QIS system relies upon the integrated implementation of the component pieces.

QIS Partners

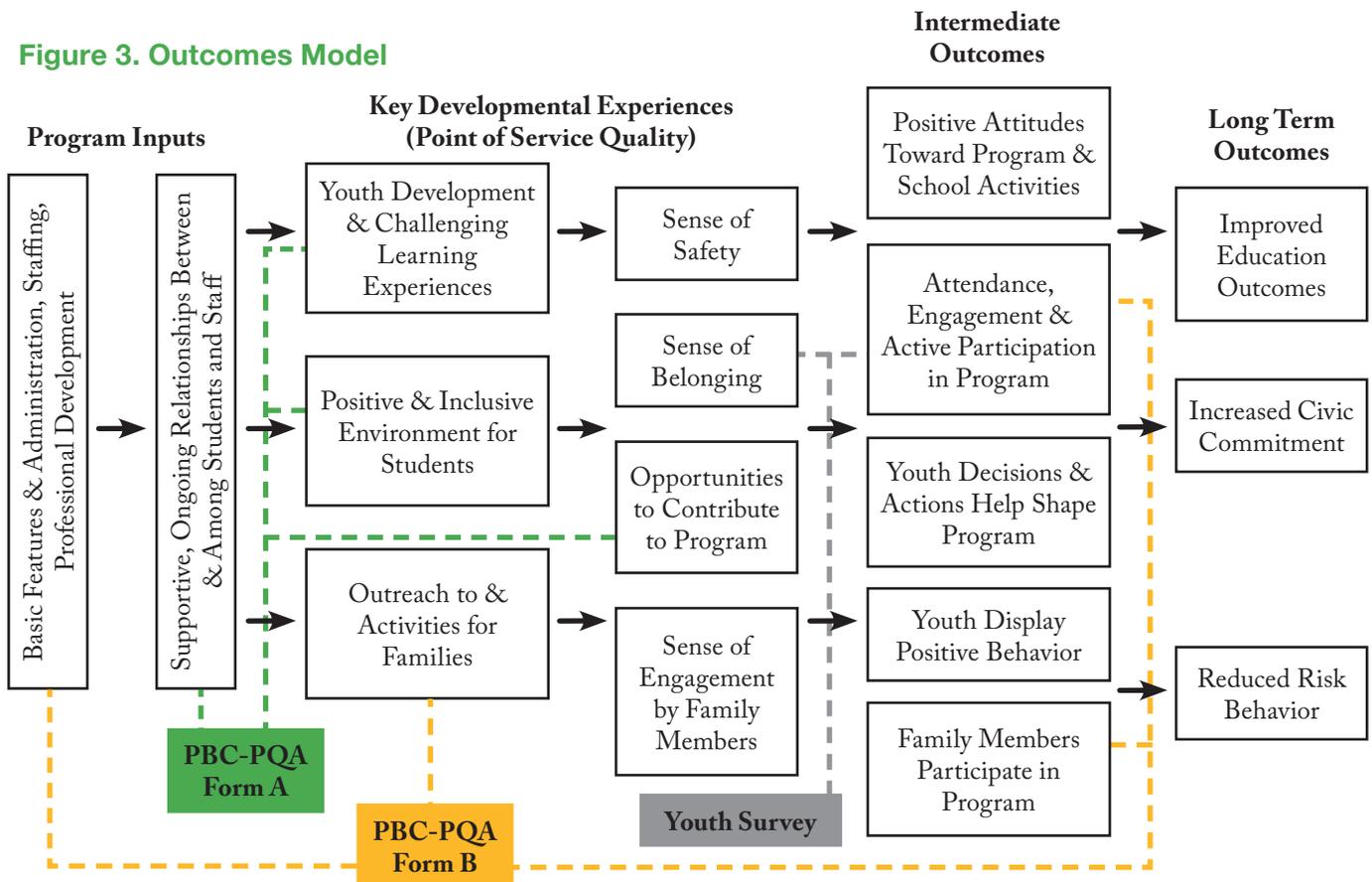
Several partners developed and implemented the QIS. Prime Time Palm Beach County Inc. is the county’s high capacity after-school intermediary, providing leadership, support and management of QIS activities:

- **Prime Time** contracts for assessment services; employs several quality advisors to provide technical assistance to directors and their staff on self-assessment, improvement plan development, and other areas; provides training linked to improvement opportunities; and analyzes impact of QIS on individual programs and across the afterschool system.
- **Children’s Services Council of Palm Beach County** (CSC) was created in 1986 as a special district of local government. The CSC serves as county’s primary investor of public funds for child and youth services, including the CSC’s investment in Prime Time and after-school programs. As of August 2007, CSC requires all of its contracted after-school programs to

participate in QIS.

- **Family Central** is an independent agency overseeing the external assessment process using the PBC-PQA. Family Central contracts with and trains after-school assessors, ensures assessor reliability, schedules assessment visits at after-school programs and provides individual program level reporting on assessment results. Assessment activities include observation of program offerings, interviews with each program manager and surveys of youth about their experiences in programs.
- Numerous **after-school providers** participated as partners to develop, pilot and implement the QIS. Provider representatives were involved in developing quality afterschool standards, selecting and testing PBC-PQA as an assessment tool, participating in pilot activities, and providing feedback on QIS activities.

Figure 3. Outcomes Model



- *The School District of Palm Beach County*, an active participant in QIS development, is the largest single provider of after-school services in the county through elementary and middle school programs operated by district employees.
- *Palm Beach Community College* supports a professional development pathway for youth workers in Palm Beach County, including school-age care certificates and the Advancing Youth Development curriculum.
- *High/Scope Foundation (Center for Youth Program Quality)* provided services including metric development, capacity building (in conjunction with Prime Time and Family Central), technology support for on-line learning modules, training for management and direct staff supporting implementation of the QIS components, and workshops emphasizing sustainability. Most training modules, particularly those designed to support managers' use of the PBC-PQA and introduce direct staff to new methods for working with youth, were developed and initially led by High/Scope staff, and then transferred to Prime Time through Training-of-Trainers workshops.
- The *Chapin Hall Center for Children at the University of Chicago* has provided external evaluation for the QIS Pilot. The Chapin Hall evaluations utilize a qualitative research methodology consisting of interviews, review of artifacts including original documents, and attendance at numerous on-site meetings and training workshops.

QIS Elements

The QIS model consists of the following core elements:

- *Program Quality Standards and Aligned Metrics.* The Palm Beach County standards were developed through a consensus-driven process to gain local expertise and build buy-in from various stakeholders for an

emerging countywide after-school system. While the standards established the “rules” for program delivery in the county, the PBC-PQA and youth survey were developed to assess compliance with the standards. The standards and aligned metrics are the core of the countywide QIS.

- *Self-assessment.* Site directors were trained to implement a self-assessment process, encouraging them to engage all levels of staff and then helping those staff understand afterschool quality by observing and interviewing each other using the PBC-PQA.
- *External assessment.* Local assessors were trained to produce reliable ratings using the PBC-PQA. Initial assessments were conducted to establish a program quality baseline. Programs were subsequently reassessed track progress towards improvement goals and overall quality improvement. This report draws primarily on PBC-PQA data.
- *Financial incentives.* Participating pilot programs were given monetary incentives for participation at various points throughout the QIS.
- *Program improvement plan development (PIP).* In addition to being trained on conducting self assessment, site directors and program staff were also trained on understanding and using their assessment data to create a Program Improvement Plan which identified strategies that were specific, measurable and achievable. Program directors attended the “Planning with Data” training to learn about the process of leading change, receive their external baseline assessment report and start developing program quality improvement plans. After receiving the aggregate baseline assessment scores and reviewing program improvement plans, Prime Time staff identified key opportunities for improvement and then translated those areas into technical assistance and training workshops.

- **Quality Advising.** Prime Time Quality Advisors helped program managers and staff to effectively engage in the improvement process by providing technical assistance for both the self-assessment process and improvement plan development, linking with community resources, and promoting training opportunities.
- **Training.** Prime Time and its partners delivered a variety of workshops for after-school staff based on areas of need identified during baseline assessments and the program improvement planning process. Training for youth workers was directly aligned with PBC-PQA scales included: (1) Youth Participation in Action, (2) Avoiding Conflict through Youth Participation, (3) Choices within Choices, (4) Effective Use of Small Groups, (5) Planning and Reflection, and (6) Developing and Sustaining a Youth Advisory Council.
- **Peer coaching.** Peer coaches were recruited, trained and deployed to work with directors and/or staff on implementing new strategies, modeling best practices and creating an environment that supports reflective practice. Peer coaches worked with programs to achieve specific goals as dictated in the PIP.

QIS Sequence and Timeline

Table 1 outlines sequence of QIS Pilot activities and the timeline for their delivery to the 38 Pilot sites.

Linkages between the “Low Stakes” Approach and Cross-level Cooperation

The QIS model evolved in important ways over time and this evolution, discussed at length in Spielberger and Lockaby (2006; 2008), may provide important lessons for the field. In this section we highlight two critical developments – the decision to implement the QIS as a low-stakes intervention and the impact of this decision on cross-level cooperation.

First, the QIS system is currently designed as a low-stakes system, meaning that while rigorous external assessment of quality at the point-of-service undergirds the system, programs do not

experience punitive consequences if they do not reach some absolute level of quality. For example, throughout the sequence of external assessment, self-assessment and planning, program directors were reminded that: (1) observation scores represent a snapshot that has both limitations and value; (2) all program scores are presented in aggregate only to avoid focusing on performances of the individual staff who were observed; (3) the overall story of the data is more important than the individual numbers; and finally, (4) Prime Time is primarily concerned with program-level use of the data, not on enforcing compliance or conducting cross-program comparisons.

Although program directors and front-line staff were not punished or rewarded based on performance against absolute quality norms, such standards were made available for pilot participants to form their own judgments. QIS Quality Advisors provided each program manager a “Summary and Interpretation of Program Quality Assessment” report that allowed Quality Advisors to reflect on the meaning of pilot site ratings. The reports noted that, in reference to the PBC-PQA’s five-level scales, scores above 4.0 are considered excellent and below 2.5 suggest need for attention.

The low-stakes approach to accountability coupled with QIS data transparency (that is, giving programs access to performance information and allowing them flexibility to self-manage quality) had important consequences for the intervention’s success. First, the design elements appear to have facilitated stakeholder buy-in. Second, these elements heightened the importance of on-site and ongoing contact between local experts – *Quality Advisors* from Prime Time or *Peer Coaches* from the ranks of other program directors in the county – and pilot sites. From the perspective of numerous stakeholders, the importance of the frequency and quality of these collaborative relationships appears to be a major factor in the success of the pilot (Spielberger & Lockaby, 2006, 2008; Prime Time, 2007a; Prime Time, 2007b).

Table 1. QIS Sequence and Timeline

<i>Activity</i>	<i>Timeline</i>
QIS Pilot Kick-Off Meeting	January 2006
External Assessors Complete Baseline Assessments	February - March 2006
Program Directors Attend Self-Assessment Training (provided by High/Scope)	March 2006
Quality Advisors Provide Follow-Up Training with Directors and/or Staff	March - April 2006
Quality Advisors Assist with Self-Assessment Scoring Process	March - April 2006
Program Directors Attend Planning with Data Training (provided by High/Scope)	May 2006
Program Directors and/or Staff Create Program Improvement Plans (PIP)	May - July 2006
Program Directors and/or Staff Attend Available Training	September 2006 - April 2007
Peer Coaches Work with Program Directors and/or Staff	September 2006 - April 2007
External Assessors Complete Reassessments	January - April 2007

QIS Pilot Participants and Measures

Participants

Participation in the QIS Pilot can be described at several levels: organization, site managers, direct staff, and youth.

Thirty-eight after-school programs, serving an estimated 4,100 children and youth annually, participated in the QIS pilot. While QIS participation was voluntary, the pilot group is representative of the wide variety of after-school settings in the county. Pilot sites served a mix of age groups with 11 of the sample sites serving elementary youth and another 18 serving a mix of elementary and middle school students. The remaining eight sites served only middle school or a mix of middle school and high school students. While all of the QIS programs are identified as after-school programs, 19 sites used a documented program or curriculum model. For example:

- **Beacon** sites are structured around the widely known Beacon model.
- **Champs** programs employed a locally developed after-school curriculum structured around activity embedded academic modules.

At baseline, 17 of these pilot programs were licensed by a state agency and another 15 were either exempt or in the process of acquiring licensed status.

Pilot after-school programs also differed on other dimensions such as setting management and geographic location in the county. Table 2 profiles participating programs on key dimensions.

We received completed surveys from 21 of the 38 program directors at the baseline data collection. On average this group had substantial experience in their role as program administrators with eight years of experience as program director and more than five years in their current position. Education levels at the time of survey completion were reported as follows: 5% high school certificate, 19% Associate's degree, 35% Bachelor's degree, 15% graduate coursework but no degree, and 25% graduate degree. Ten percent of these directors were certified teachers but none was trained as a social worker. Average monthly wages for the full-time program directors in this group were \$1,658.

A small number of direct staff (24 individuals from eight organizations) also responded to our survey requests. This group of front line youth workers averaged nearly four years of experience in the profession, with just over three years in their current position. Only 29% of this group had attained

Table 2. Dimensions of QIS Pilot Programs

<i>Program Dimension</i>	Pilot Program Participants
<i>Setting</i>	School- based: 15 programs Community-based: 23 programs
<i>Management</i>	Operated by school: 9 programs Operated by community-based organization: 26 programs Operated by parks and recreation department: 3 programs
<i>Geographic Location</i>	Represented variety of neighborhoods - Riviera Beach, Delray Beach, Greenacres, West Palm Beach, Boca Raton, Boynton Beach, Lake Worth, Pahokee, Belle Glade

education beyond the high school diploma and their average wage was \$8.52 per hour. In a recent survey of over 1,000 youth workers in eight cities, the percentage of youth workers with at least some post-secondary education is 85% (Yohalem, Pittman & Moore, 2006), suggesting the youth workers in Palm Beach have substantially lower levels of formal education than youth workers elsewhere in the country.

Youth surveys were collected for 592 youth in grades four and higher who attended an offering observed by an external PBC-PQA assessor. From this sample of 592 youth in 31 programs collected during the post-pilot data collection, we constructed the profile of youth participants in Table 3. Although the PBC-PQA was designed for use across after-school settings in Palm Beach County, the youth survey was only administered to children in grades four and five, even though the offerings for early elementary were part of the sample where PBC-PQA ratings were collected.

QIS Metrics

The PBC-PQA Form A, Form B, and the youth survey are the core quality metrics used in the Palm Beach QIS. These instruments are designed to support routine annual data flow from participating after-school providers in the county and to facilitate ongoing performance management processes. The PBC-PQA was designed for use both as a program self-assessment and a source of external review and evaluation. In addition to these core metrics, supplemental supervisor and staff surveys are also employed to further assess management practices, institutional culture and performance changes.

PBC-PQA self assessments were conducted by staff teams at each participating site. The self assessment process consisted of the following steps:

- Staff members observe one another’s program sessions and record anecdotal evidence about the quality of their peers’ interactions with youth.
- After sufficient observational data was collected, the team met, discussed the data and scored a single PBC-PQA Form A for the entire site.

This two-step self-assessment process was designed to (1) support staff learning about point of service quality and positive youth development; and (2) familiarize staff and increase their comfort with Form A in order to increase buy in for the external assessment process (which relies on the same assessment tool).

External assessments using the PBC-PQA Form A were conducted by external observers employed by the Family Central organization. External observations of randomly selected program offerings were collected at each pilot site. A program offering is defined as a component of an after-school program that involves the same staff and same youth meeting for the same purpose over multiple sessions, e.g., science club at Jones Middle School’s 21st Century after-school program meets every Tuesday and Thursday between 3:30 and 4:30 pm during fall semester 2007. In the QIS, the number of offerings selected for observation at each site was determined by a formula involving the number of children enrolled in its after-school program. Each

Table 3. Youth Characteristics (N = 592)

Average age	11 years
<i>Gender (8% missing)</i>	
Boys	43%
Girls	49%
<i>Frequency of Program Attendance</i>	
A few times each month or less	27%
Once per week	14%
A few times per week	49%
Required by parents to attend	69%
Participates in other after-school activities	63%

program received a minimum of three observations. External assessors were required to satisfy accurate scoring norms (accuracy at the level of 80% perfect agreement with “gold standard scores” during a series of video tests) prior to conducting observations and were trained to employ a data collection methodology designed to maximize score reliability. Appendix D provides a technical discussion of measurement properties PBC-PQA Forms A and B.

Surveys for program supervisors and direct staff were added as supplemental measures during the QIS pilot to provide evidence about staff practices and performance change. Supervisor surveys were only collected during the baseline data collection and are summarized in the report entitled *Communities of Practice in the Palm Beach County QIS (High/Scope 2006)*. Direct staff surveys were administered during both baseline and post-pilot data collections; however, response rates for the post-pilot data collection were very low. Approximately 20% of all direct staff, representing 39% of all organizations participating in the QIS Pilot, completed a survey.

Table 4 summarizes the content and usage of the five instruments. See Appendices C & D for detail.

Table 4. PBC QIS Pilot Metrics: Core and Supplemental

<i>Instrument</i>	<i>Type</i>	<i>Administration Date(s) & Sample Size</i>	<i>Description and Data Collection</i>
PBC-PQA Form A	Observation	Feb/Mar 2006 (N=139 observations) Feb/Mar 2007 (N=128 observations)	The PBC-PQA Form A is an assessment of best practices in after-school and community programs for youth. The assessment consists of four domains focused on quality at the “point-of-service”: Safe Environment, Supportive Environment, Interaction, and Engagement. Each of these domains is comprised of scales (measurement rubrics consisting of 2 - 6 items). Items are scored on a scale from 1 to 5 and then averaged up to scale and domain levels. This instrument was completed by external data collectors after observing a program session in an after school site. Multiple observations were completed at each site according to a formula that adjusts for program size during both of the administration dates listed above by the assessment contractor (Family Central).
PBC-PQA Form B	Interview	Feb/Mar 2006 (N=33) Feb/Mar 2007 (N=37)	The PBC-PQA Form B consists of four domains focused on organization/administration practices and policies: Youth Centered Policies, High Expectations for Youth and Staff, Organizational Logistics, and Family. Each domain is comprised of scales (measurement rubrics consisting of 2 - 6 items). Items are scored on a scale from 1 to 5 and then averaged up to scale and domain levels. This instrument was completed based on a phone interview with program directors (following a list of interview questions that accompanies the assessment). One interview was conducted per site during the administration dates listed above by the assessment contractor (Family Central).
Youth Survey	Self-Report Survey	Feb/Mar 2007 (N=592 surveys during 48 offerings at 31 programs)	The youth survey was administered by the external data collector (Family Central) at the end of his/her post-test PBC-PQA observation of a youth program session. Only children attending the observed sessions were asked to complete the survey (e.g., if Arts and Crafts was observed, then only students who participated in Arts and Crafts were surveyed). Survey directions and items were read out loud by the data collector. Surveys contained no individually identifying information but were linked to the program session (offering) in which they were administered.
Direct Staff Survey	Self-Report Survey	Feb/Mar 2006 (N=80) Apr/May 2007 (N=24)	The Staff Survey is a questionnaire designed to assess front-line staff’s professional background, use of best-practices (self-reported), beliefs about youth work, and perceptions of the professional learning community. With the exception of nominal descriptive data, survey items have response scales from 1 to 5. The survey was administered in paper format in 2006 and in online format (with follow-up via paper surveys) in 2007.
Supervisor Survey	Self-Report Survey	Feb/Mar 2006 (N=21)	The Supervisor Survey is a questionnaire designed to assess site supervisors’ professional background, use of best-practices, beliefs about youth work, and perceptions of the professional learning community. With the exception of nominal descriptive data, survey items have response scales from 1 to 5. The survey was administered in paper format.

Part II. QIS Pilot Findings

Change In Point-of-Service Quality During The QIS Pilot

One of the primary purposes of the Palm Beach QIS is to efficiently produce routine data on program quality so that sites can better manage performance. In this section, we present data regarding the quality of after-school environments and staff performances with youth in the 38 pilot sites. Our findings are based on a combined total of 264 observations using Form A of the PBC-PQA. Major findings presented in this section include:

- During the QIS, quality scores generally increased from the baseline to the post-pilot, especially in PBC-PQA domains concerning the quality of staff support and student interaction.
- Quality scores increased more in areas targeted by program directors for improvement, suggesting both that the intervention can focus improvement efforts and that skill sets of individual staff are malleable.

Although this evaluation of the QIS does not employ a research design to enable causal interpretation, several design elements improve our ability to interpret pre-to-post changes in the data. First, observational data was collected during the same time period in each of two program years. Pre- and post-test data was collected during the

same months of spring 2006 and spring 2007. This is important because it controls for naturalistic gains that occur over the course of a program year as staff-youth teams get better at working together and as students with behavioral problems exit the program for various reasons. Second, observational data was collected by an independent entity not tied to funding from Prime Time, High/Scope or any individual providers. This reduces the chance that scores will be biased by financial or other connections to the intervention. Third, we are able to construct a within-sample comparison group by contrasting baseline to post-pilot change scores across groups that did and did not select a given quality improvement area. For example, it is possible to compare all sites that selected “opportunities for reflection” as an improvement area with all of the sites that did not. The hypothesis driving the construction of such within-sample comparison groups is that pre-to-post change in a particular improvement area should be greater for the subsample of QIS sites that selected that improvement area than for sites that did not.. Finally, we are able to triangulate data from staff surveys, qualitative data sources and evaluation reports by Chapin Hall.

Quality scores are presented at three levels of aggregation. The four domain scores represent the most global level of measurement using the PBC-PQA Form A. Next, scale levels scores representing

Table 6. PBC PQA Form A Domains: QIS Pilot Scores & Comparisons

<i>Form A Domains I-IV</i>	<i>Baseline (N=38)</i>	<i>Post-pilot (N=37)</i>	<i>Comparison (Youth PQA Validation Study, N=71)</i>
Safe Environment	4.46	4.77**	4.4
Supportive Environment	3.86	4.29**	3.7
Interaction Opportunities	3.33	3.61**	3.0
Engaged Learning	2.61	2.85+	2.8
Statistical significance of differences established using a repeated measures <i>t-test</i> . Levels are: + = marginally sig at $p < .1$, * = sig at $p \leq .05$, ** = sig at $p \leq .01$.			

staff practice sets are assessed. Staff practice sets are specific youth development skill sets that staff purposefully employ as part of a specific youth work method or program philosophy. Evidence of change in *targeted* scale level scores is important because it suggests that the QIS intervention may be an effective strategy for workforce development. That is, when the intervention focuses improvement efforts on a specific practice area, consequently workforce skills in that area are improved. Finally, information is presented for selected items where high percentages of offerings received scores of “1” at the baseline. Items on the PBC-PQA are scored at levels 1, 3 or 5. In general these scores can be interpreted in the following way: 5 = quality element is available to all youth in the setting; 3 = quality element is available but not consistently or not for all youth in the setting; 1 = quality element is not available during the observation in the setting.

Global Quality Scores

Table 6 presents scores for the four primary domains of the PBC-PQA Form A with an additional score profile from an independent study presented for comparison. The first two columns in Table 6 present baseline and post-pilot scores, demonstrating that quality scores for all four domains increased over the course of the QIS intervention. Symbols in the post-pilot column denote differences that were statistically significant. Column three presents comparison scores from the Youth PQA Validation Study, an after-school sample similar to the QIS pilot (Smith & Hohmann, 2005).

Staff Practice-Sets

Table 7 presents baseline to post-pilot change scores for 37 of the QIS pilot sites at the scale level of the PBC-PQA Form A. The first column presents

Table 7. PBC PQA Form A Scales: Change Scores & Significance Tests

<i>20 Form A Scales</i>	<i>PBC Baseline (N=38)</i>	<i>Change Score (N=37)</i>
Psychological and Emotional Safety	4.51	0.31**
Physical Environment Safety	4.72	0.13*
Emergency Proc/Supplies	4.46	0.21+
Program Space/Furniture	4.57	0.26**
Healthy Food/Drink	4.04	0.68**
Welcoming Atmosphere	4.44	0.22
Session Flow	4.40	0.34**
Clear Limits	4.11	0.58**
Active Engagement	3.64	0.44**
Skill Building	3.63	0.71**
Encouragement	3.40	0.07
Conflict Approach	3.41	0.82**
Sense of Belonging	3.62	0.24**
Groups Strategies	2.36	0.30+
Shared Responsibility	2.97	0.27
Youth/Adult Partnering	3.37	0.31+
Positive Peer Relationships	4.37	0.36**
Setting Goals and Making Plans	2.61	0.06
Choices/Interests	2.62	0.38+
Reflection Opportunities	2.61	0.32*
Statistical significance of differences established using a repeated measures t-test. Levels are: + = marginally sig at p<.1, * = sig at p ≤ .05, ** = sig at p ≤ .01.		

Table 8. PBC PQA Form A: Percentage of Selected Items Scoring “1” at Baseline & Post-pilot

<i>Selected Form A Items</i>	<i>% Scoring 1 at Baseline</i>	<i>% Scoring 1 at Post-Pilot</i>
Staff make frequent use of open-ended questions (e.g., staff ask open-ended questions throughout the activity and questions are related to the context). II.K.3	49.6	43.8
Session consists of activities carried out in at least 3 groupings—full, small, or individual. III.N.1	50.7	39.8
Staff use 2 or more ways to form small groups (e.g., lining up by category and counting off, grouping by similarities, signing up). III.N.2	49.3	48.4
Each small group has a purpose (i.e., goals or tasks to accomplish), and all group members cooperate in accomplishing it. III.N.3	52.2	47.7
All youth have one or more opportunities to lead a group during program activities. III.O.2	43.5	43.0
In the course of the program offering, all youth are given a structured opportunity to set one or more long-term goals. IV.R.1	59.4	49.2
Time is regularly provided for young people to make (individual or group) plans for and/or to set goals for activities. IV.R.2	40.0	40.6
All youth have the opportunity to make at least one open-ended content choice within the content framework of the activities (e.g., youth decide topics within a given subject area, subtopics, or aspects of a given topic). IV.S.1	45.7	42.2
All youth have the opportunity to make at least one open-ended process choice (e.g., youth decide roles, order of activities, tools or materials, or how to present results). IV.S.2	45.7	34.4
All youth are engaged in an intentional process of reflecting on what they are doing or have done (e.g., writing in journals; reviewing minutes; sharing progress, accomplishments, or feelings about the experience). IV.T.1	66.7	59.4
All youth are given the opportunity to reflect on their activities in 2 or more ways (e.g., writing, role playing, using media or technology, drawing). IV.T.2	48.6	34.4
In the course of the program offering, all youth have structured opportunities to make presentations to the whole group. IV.T.3	44.9	37.5
This table is constructed from the total of all observations in the QIS Pilot: N=139 at baseline and N=126 at post-pilot.		

baseline quality scores for each scale while the second column presents the change scores calculated by subtracting the baseline score from the post-pilot score. Symbols in the change-score column denote differences that were statistically significant. All 20 of the scale scores demonstrate positive change and all but four of the positive differences are statistically significant.

In order to better understand the amount of change that these change scores represent, we need to think about the magnitude of the changes in light of the overall score variation across sites. We used

a variation on the “Cohen’s D” effect size formula (post-test score – pre-test score / standard deviation (SD) of pre-test score) to better understand the magnitude of the score changes. When these magnitude estimates were calculated, 45% of the scales had differences between the baseline and post-pilot scores that were nearly as large as the standard deviation for that scale score across programs. We interpret these changes to be substantial in magnitude.

Access to Selected Key Experiences

A final way to look at quality data from the QIS

is to examine items for which large proportions of offerings scored a level 1, i.e., external assessors found that staff did not deliver these experiences to youth during observed sessions. Our focus on these low-performing items is intended to: (1) estimate youths' access to key developmental experiences; (2) assess the magnitude of increased or decreased access to these experiences subsequent to the QIS intervention; and (3) provide insight in the skill base of and daily practices of youth workers.

In general, our analysis of low-scoring items reveals that large numbers of program staff fail to provide opportunities for youth to exercise voice and choice. The post-pilot changes scores, however, suggest that the QIS increases the incidence of these opportunities.

Table 8 provides a QIS perspective on low program quality at the level of individual staff practices by listing the PBC-PQA items for which 40% or more of the baseline offerings received a score of 1. Column one provides the baseline percentages and column two provides the post-pilot percentages. During the QIS intervention, the number of offerings where key developmental experiences were unavailable to youth declined in 11 of 12 areas. Substantial improvements occurred in the percentage of offerings that provided access to small

group work (III-N1), process choice (IV-S2), and reflection on that session's activities (IV-T2).

Success of Targeted Improvement

The prior discussion demonstrates that during the QIS, on average, point of service quality did improve across all pilot sites and across most of the staff practice sets assessed by the PBC-PQA. As a part of the PBC QIS, program managers (and in some cases their staff teams) selected specific practice sets in which to concentrate their improvement efforts. These areas (scales from the PBC-PQA) were selected after reflecting on scores from the external and self-assessment results.

Improvement plans generated by the QIS pilot providers identified a total of 96 improvement goals. As noted above, goals were selected after consideration of baseline PBC-PQA data and thus were closely aligned with scales on the point of service measurement tool. Targeted improvement areas were spread across the Form A construct, with 56% focused on scales in the interaction and engagement domains.

Table 9 presents change scores for pilot sites that selected a given improvement (column 1) in comparison to all other pilot sites that did not select the same improvement area (column 2). When

Table 9. Change Scores for Pilot Sites Selecting Improvement Areas versus Pilot Sites Not Selecting the Same Areas

<i>Improvement Area</i>	<i>Change Scores for Pilot Sites Selecting Improvement Areas</i>	<i>Change Scores for Pilot Sites NOT Selecting Improvement Areas</i>	<i>Difference of Change Scores</i>
Healthy Food/Drink (N=7)	1.32	0.53	0.79*
Encouragement (N=5)	0.64	-0.01	0.65*
Conflict Approach (N=11)	1.39	0.57	0.82*
Grouping Strategies (N=13)	0.43	0.23	0.20
Shared Responsibility (N=9)	0.89	0.08	0.81*
Youth/Adult Partnering (N=5)	0.38	0.30	0.08
Setting Goals & Making Plans (N=11)	0.74	-0.22	0.96*
Choices/Interests (N=11)	0.80	0.20	0.60
Reflection Opportunities (N=6)	0.50	0.25	0.25

Statistical significance of differences established using a repeated measures t-test.
Levels are: + = marginally sig at p<.1, * = sig at p ≤ .05, ** = sig at p ≤ .01.

looking at change from baseline to post-pilot, pilot sites that selected an improvement scale had, on average, substantially larger change scores for those scales compared to programs that did not target those areas. Column 3 provides the “difference of the difference scores” in the prior two columns. Symbols in column 3 denote differences that are statistically significant.

Additional Evidence of Change in Quality

As an attempt to triangulate evidence of change in program quality from different sources of data, surveys were administered to direct staff at baseline and post-pilot. Overall return rates for the post-pilot survey were low (N=24 staff in 15 programs). In addition, the number of sites with stable staffs (e.g., at least one front line staff consistent between pre- and post- measures) that returned baseline and post-pilot surveys was also very low (N=8 sites). Consequently, the data from the direct staff survey provides at best a limited window in the impacts of the QIS intervention. Given these caveats, we examined patterns of change in direct staff reports about their practices and beliefs about youth work. Statistically significant changes were found for direct staff reports about frequency of two practices: Youth use planning strategies and Youth reflect on their work. Statistically significant changes were found for direct staff reports about their own beliefs in one area: Emphasis on relationships.

Change in Organizational Practices and Policies during the QIS Pilot

In this section, we present evidence regarding change in organizational practices and policies in the 38 pilot sites based on a combined total of 70

interviews using Form B of the PBC-PQA. Major findings for this section include the following:

- Global scores for best practices at the organization level increased in all four of the PBC-PQA domains and the quality and usage of numerous management policies and practices improved during the QIS.
- Several specific management practices and policies demonstrated substantial increases in the percentage of sites adopting for the first time.
- Four management practices and policies related youth voice showed decreases in the percentage of sites employing these practices.

In this section quality scores are presented at three levels of aggregation. Domain scores represent the most global level of measurement using the PBC-PQA Form B. Scale levels scores representing dimensionality in the measures are also presented. The Form B scales define elements of quality that can be understood as specific kinds of practices and policies that program managers might put in place. Finally, information is presented for selected items where high percentages of pilot sites received scores of “1” at the baseline. Items on the PBC-PQA Form B are scored at levels 1, 3 or 5. In general these scores can be interpreted in the following way: 5 = the practice is part of intentional policies and procedures conducted at this organization; 3 = practice element is informally present or sporadically implemented; 1 = quality element is not present in the organization.

Table 10. PBC PQA Form B Domains: QIS Pilot Scores & Comparison

<i>Form B Domains V-VIII</i>	<i>Baseline (N=33)</i>	<i>Post-Pilot (N=37)</i>	<i>Comparison (Youth PQA Validation Study, N=71)</i>
Youth Centered	3.09	3.30+	3.81
High Expectations for Youth and Staff	4.12	4.60**	3.77
Organizational logistics	4.46	4.77**	NA
Family Connections	3.84	4.29**	NA
Statistical significance of differences established using a repeated measures t-test. Levels are: + = marginally sig at p<.1, * = sig at p ≤ .05, ** = sig at p ≤ .01.			

Table 11. PBC PQA Form B Scales: Change Scores & Significance Tests

<i>13 Form B Scales</i>	<i>QIS Baseline (N=33)</i>	<i>Change Score (N=33)</i>
Youth Interests/Build Skills	3.76	0.56**
Youth Influence on Activities	2.86	-0.27
Youth Influence on Policy	<i>missing data</i>	<i>missing data</i>
Staff Development	4.21	0.35*
Supportive Social Norms	4.39	0.54**
Support Academic Enrichment	3.83	0.40*
Commitment to Program Improvement	4.06	0.58**
Sound Business Practices	<i>missing data</i>	<i>missing data</i>
Organizational Logistics and Staffing	4.56	0.20+
Youth Tracking System	4.91	0.09
Staff Records/Policies	4.03	0.50
Communication with Families	3.75	0.69**
Support Family Involve	3.92	0.31*
Statistical significance of differences established using a repeated measures t-test. Levels are: + = marginally sig at $p < .1$, * = sig at $p \leq .05$, ** = sig at $p \leq .01$. Note: V-C, and VIIIH are not included due to missing data (unable to form scale for pre- or post-test).		

Global Scores for Organizational Practices and Policies

Table 10 presents scores for the four primary domains of the PBC-PQA Form B with an additional score profile from an independent study included for comparison. The first two columns in table 10 present baseline and post-pilot scores, demonstrating that quality scores for all four domains increased over the course of the QIS intervention. Symbols in the post-pilot column denote differences that were statistically significant. Column three presents comparative data from the Youth PQA Validation Study, an after-school sample with similar program characteristics (Smith & Hohmann, 2005).

Supervisor Practice Sets

Table 11 presents baseline to post-pilot change scores for management policies and practices at 33 of the QIS pilot sites. The first column presents baseline quality scores for each PBC-PQA Form B scale while the second column presents the change scores calculated by subtracting the baseline score from the post-pilot score. Symbols in the

Change Score column denote differences that were statistically significant.. Ten of the eleven scales for which data is available demonstrate positive change; of these score increases, 7 are statistically significant.

Selected Best Practices and Policies

A final way to look at quality data from the QIS is to examine items for which high percentages of offerings scored a level 1. This perspective most concretely describes best practices and policies that are not part of the culture of a network of organizations.

Table 12 describes specific management practices and policies for which 20% or more of the 33 administrative interviews received a score of “1” on the PBC-PQA Form B. Column one provides the baseline percentages and column two provides the post-pilot percentages. During the QIS intervention, these selected best practices and policies increased in 8 of the 12 areas that scored low at the baseline. Substantial improvements occurred in the areas of programmatic focus (V-A3), youth participation in program recruitment (V-C3), youth participation in governance (V-C5), staff participation in

Table 12. PBC PQA Form B Items: Change in Percentage of Selected Items Scoring “1”

<i>Selected Form B Items</i>	<i>Baseline % Scoring 1</i>	<i>Time 2 % Scoring 1</i>
Across all program offerings, the organization has a major and specific programmatic focus on 5-6 of the following areas: academic, cultural, service learning, life skills, career exploration, and recreation. V.A.3	41.2	8.1
Youth have influence on setting and activities in the organization. V.B.1	41.1	78.4
Youth and adults share decisions on programs and schedules. V.B.2	29.4	67.6
Youth take charge of and facilitate or lead program sessions or activities for peers or younger youth. V.B.3	38.2	21.6
Youth participate in program quality review and planning for improvement. V.C.1	36.4	63.0
Youth and staff share responsibilities for recruiting other youth to join organization or program offerings. V.C.3	24.2	3.8
Youth and staff share responsibilities for the character and nature of community outreach. V.C.4	42.4	51.9
Youth and staff share responsibilities for governing bodies. V.C.5	75.8	44.4
A majority of staff participate in at least one relevant professional development activity per year within the organization. VI.D.2	23.5	2.7
Planned activities have explicit objectives and/or learning goals.	29.4	8.1
Organization has established mechanisms for helping parents connect to their child’s school learning. VIII.L.3	20.6	13.5
Parents often participate in or have significant influence on organizational decision-making. VIII.M.4	32.4	27.0

professional development (VI-D2), and association of planned activities with explicit learning goals (VI-F1).

These improvements are encouraging and reflect the overall direction of the results. In fact, the central story of the QIS Pilot is that, on balance, the general quality of point-of-service performances and management policies and practices improves. It is important to note, however, that in 4 particular management areas performances deteriorated in the post-pilot measure. Specifically, substantial declines occurred in measures associated with organizational support for youth voice and choice (V.B.1, V.B.2, V.C.1, and V.C.4). In certain respects, it is no surprise that some performance measures improved while others declined; after all, the broad and low stakes nature of the intervention afforded programs much latitude in terms of selecting and pursuing quality improvements. That is, sites had the freedom to choose what to focus their improvement efforts on

and in many cases elected to prioritize issues other than youth voice and choice.

It is nonetheless disconcerting that the percentage of low scores on these particular items increased between the baseline and post-Pilot measures. Closer investigation of the specific program reveal that a high percentage of those sites that declined on two of these items also had much higher rates of supervisor turnover during the QIS pilot than other QIS programs. It is also true that as site managers are exposed to and buy into the concept of youth voice and choice, they may become more critical about both their organizations’ ability to provide systematic opportunities for youth agency and these more critical response patterns may produce a negative bias in the data, given the interview methodology employed to collect Form B data.

Additional Evidence of Change in

Organizational Practices and Policies

As described in Part IV, surveys were administered to direct staff in an effort to collect data that might corroborate our analysis of Form A and Form B data. As noted above, however, response rates for these surveys were very low for the post-pilot data collection. Our analysis of the collected data produced one notable finding: direct staff were asked a number of questions about supervisor practices and decision-making authority; however, none of these items evidenced statistically significant change.

QIS Model Fidelity

Implementation fidelity refers to the extent to which a program model actually gets implemented at a site. For the QIS Pilot, model fidelity is relevant at two levels, organization and point-of-service. In this section, we focus on QIS fidelity at the organizational level by considering (1) the extent to which program managers implemented the quality assessment and improvement planning elements of the QIS, and (2) how frequently program managers or their staff participated in training and technical assistance provided by Prime Time. We also test the depth of the QIS intervention by asking front-line staff about their degree of participation in the core elements of the QIS. Major findings for the section include the following:

- Organizational participation in training and TA components of the QIS was high.
- Adoption of key intervention practices of conducting self-assessment and improvement planning were high.
- Direct staff reported high levels of awareness of PBC standards and participation in PBC-PQA quality assessment process in comparison to other relevant samples.

Not all programs participated in each element of the QIS. For example, some program directors chose to implement self-assessment without assistance from a Prime Time quality advisor while others requested assistance from quality advisors for staff training or assessment scoring support. Table 5 identifies the

number of programs that participated in the several QIS elements.

In order to assess the depth at which the QIS was experienced in the pilot sites, the direct staff survey administered at the post-pilot data collection period asked about two key elements of the QIS: Were they familiar with the Palm Beach Standards? Had they participated in the PBC-PQA self-assessment process?

Responses from the post-pilot sample (N=24 direct staff at 15 of 38 pilot sites) indicate that 64% of direct staff had “seen or discussed” the Palm Beach Standards. For comparison, in a recent statewide sample of direct staff in 21st Century Community Learning Centers (N=154), only 4% reported familiarity with the states’ widely disseminated after-school program quality standards.

When direct staff in Palm Beach County were asked about their level of participation in the self-assessment and improvement planning processes, similarly large numbers appear to have been involved: 61% reported having “worked with your site supervisor to completed the PBC-PQA”; 43% had “conducted observation and made notes”; 46% had participated in scoring the PBC-PQA; and 57% had talked about program quality scores. Again for comparison, in a recent sample of 540 direct staff from 100 after-school programs in four states, only 30% reported ever having used any type of quality assessment tool. In the statewide 21st Century sample mentioned earlier, the number of direct staff that had used a formal quality assessment tool was only 8%.

Overall, we conclude that fidelity to the organizational level of the QIS intervention was quite high, with substantial rates of adoption of the core QIS elements and high rates of participation by site managers in training and technical assistance. Furthermore, it appears that the QIS intervention was successfully introduced to direct staff by site supervisors. However, direct staff reports must be treated with caution due to low response rates.

Table 5. Pilot Site Participation in QIS Elements

<i>Improvement Activity</i>	<i>Number of Pilot Sites</i>
PBC-PQA Baseline External Assessment (Family Central)	38
Training (High/Scope): PBC-PQA Self-Assessment	35
Quality Advisor TA: Self-Assessment Training	23
Quality Advisor TA: Self-Assessment Scoring	8
Self-Assessment Submitted to Prime Time	
Training (High/Scope): Planning with Data	31
Quality Advisor TA: Program Improvement Plan Development	26
Program Improvement Plan Submitted to Prime Time	28
Direct Staff Attended Training:	
• Avoiding Conflict through Youth Participation	8
• Bringing Yourself To Work Training	6
• Choice and Challenge	3
• Effective Use of Small Groups	8
• Youth Planning and Reflection	5
• Advancing Youth Development Training	7
Peer Coaching	7
External Time 2 Assessments	37

Part III. Formative Analysis

The final area of inquiry for this report is formative exploration of QIS elements and other factors that are related to positive changes in quality at the point-of-service. While our (lack of) research design does not permit causal inference, we nevertheless feel compelled to use available information to both suggest and evaluate some hypotheses about “what worked” in the QIS pilot. Taken together, this circumstantial evidence offers strong guidance to policymakers, intermediaries and practitioners in Palm Beach and elsewhere who are seeking effective ways to drive quality improvement in settings where adults and youth interact.

The following findings are described in this section:

- The preponderance of qualitative and quantitative evidence suggests Prime Time’s partnership-oriented approach and focus on supporting program directors to make meaning from data were key strengths of the QIS.
- Site supervisor and staff turnover did not affect quality improvement during the QIS pilot.
- Organizational attention to youth interests and commitment to program improvement are the management-level (Form B) practices most strongly related to point-of-service quality.
- Organizational characteristics such as management type, curriculum model and licensing status were not related to point-of-services quality at baseline or post-pilot.
- Supervisor education and experience levels were not related to point of service quality at baseline.

The QIS Model and Quality Change

This section makes use of primary source

documents, the Spielberger and Lockaby (2006) evaluation report, and quantitative information from the QIS. The primary source documents include meeting minutes from Prime Time’s QIS Steering Committee (2007a) and QIS Working Committee Minutes (2007b), notes on site progress collected by Quality Advisors, and notes compiled as part of each site’s Program Improvement Plan.

According to our review of information from these multiple primary and secondary sources, we suggest that Prime Time’s partnership orientation and the QIS focus on supporting program supervisors to make meaning from quality data are the critical success factors (CSFs) within the QIS model. Although we do not have an explicit research design in place to test these hypotheses, two sources of data are available, including data from training satisfaction surveys and QIS participation data, that provide insight into the perceptions of site managers toward the QIS improvement sequence. Data from these sources are cited where applicable to support our hypotheses regarding CSFs.

Prime Time Partnerships with Programs.

Both the QIS project and Prime Time as an organization are focused on building partnerships with programs in order to help them improve the quality of their after-school services. Although building partnerships with programs may have been part of Prime Time’s mission before embarking on the QIS, the partnership orientation appears to have become more deeply institutionalized during the QIS pilot. As the Chapin Hall Year 2 report asserted, “Prime Time’s new direction is far more focused and more tied to quality improvement” (Spielberger & Lockaby, 2006). Prime Time’s commitment to building relationships with programs is also evidenced by its current staffing structure which includes a QIS director, three quality advisors, and other personnel such as training coordinators who are increasingly involved in QIS activities.

In all of the primary source documents, this focus on building strong, non-adversarial relationships with programs is a core component of the change model employed by QIS staff. A participating program director summed this up at a Working Committee meeting (May 10, 2007):

What has made [the QIS pilot] work particularly for our sites... is the sense of partnership with Family Central and Prime Time. It has not been a critical intervention or oversight, [but] rather an open dialogue, and people have not been threatened by having observers and the feedback that was offered. And all the training has related to that feedback. The feedback is the provider's feedback—voluntary engagement in the system and the improvement plan is also voluntary. And how the providers implement the plan is also voluntary, based on the needs/capability of staff and program you have.

In our interactions with Prime Time staff, the Quality Advisors (QAs) frequently stressed the importance of (1) getting to know program directors and listening to their concerns; and (2) molding Prime Time services to best meet the directors' needs. Several QIS components, including Peer Coaching training, emerged directly from these priorities and Prime Time has established workflow structures that give these relationships prominence. In addition, the expansion of Quality Advising services (e.g., site visits, mini-trainings and goal check-ins) has allowed Prime Time to be in tighter contact with pilot sites and has provided another channel through which to support programs' improvement efforts. For example, the following entry was made by a Prime Time Quality Advisor in a program improvement plan:

Progress Made: Site visit; I spoke with [Director] about her goals.
Goal 1: Students have had several opportunities to make plans. At

their center they recently had a Mayan Presentation to celebrate Tecun Uman (a Mayan warrior) the children performed for their parents and had a feast of Guatemalan food. They are also planning another parent's night showcase for May 9th; the children were able to choose the type of performance/dance they will be performing.

Data from training satisfaction surveys certainly lends support to our contention that Prime Time successfully established partnerships with participating programs. Satisfaction surveys were administered at the end of each training day for the PBC-PQA self-assessment training as well as the Planning with Data workshops for QIS pilot supervisors. Several items on these surveys seek to understand the level of perceived "fit" between the training content and site managers' own organizational contexts. Two items from the survey are of particular interest. On a scale of 1 to 5 with 5 meaning "strong", the average rating on the item "level of administrative support at your program for implementing the content" was 4.68 (n = 33). For "Applicability of content to current job position", the mean rating was 4.63 (n = 30). These high mean scores indicate that the site supervisors were supportive of the self-assessment and improvement planning processes and that these QIS elements were aligned with their program environments.

Notably, Prime Time's ongoing efforts to expand its in-house training capacity—including the planned deployment of online training—are likely to further strengthen its partnership orientation by providing increased capacity to overcome geographic and time/resource barriers to QIS participation.

Helping Directors Make Meaning from Data.

Prime Time, in partnership with Family Central, now has capacity to generate external quality reports for programs. These reports are supported by the Planning with Data training, and other, less structured quality advising and coaching aids to help program directors make use of quality data. Together these components are designed to help the

PBC-PQA provide meaningful data for program directors to act on—data that program directors say fits, matters, and is useful (Spielberger & Lockably, 2007). Through the relationship building strategies discussed above, Prime Time is turning data into dynamic, valuable performance management information for that can be used to steer improvement efforts. Moreover, there is a important statistical relationship between participation in self-assessment and Planning with Data trainings and measured quality improvement.¹

The self-improvement sequence consists of three parts: self-assessment, improvement planning, and improvement itself. Requiring program directors to conduct self-assessment before receiving external reports and creating improvement plans seems to be particularly effective. Program self-assessment, which for program directors (and some staff) consists of attending a training (PQA Basics), collecting internal data, and self-scoring a PBC-PQA, has two main purposes: it gets program directors familiar with the PBC-PQA, and gets them familiar with reflecting on their own program within a best practices framework. Perhaps the most powerful outcome of the self-assessment is not the quality scores that programs generate, but the preparation to receive and work with external reports. That is, self-assessment prepares programs to create and carry out improvement plans. The message of improvement is unified throughout QIS. Indeed, Quality Advisors indicated that self-assessment and improvement planning helped program directors be more successful at accepting and interpreting quality data and more intentional about improvement plans that flowed from quality data.

1 In order to evaluate the hypotheses about the QIS focus on self-assessment and improvement planning, bi-variate correlations between QIS participation data (see table 5) and PBC-PQA Form A data were examined. The program quality score used in these analyses was an aggregate of PBC-PQA scales identified in table 7 as areas in which significant gains occurred. Only participation in the Self-Assessment Training and Planning with Data trainings were significantly correlated with this measure of post-test program quality ($r = .29$ and $r = .32$ respectively).

Additional Formative Analyses

Several additional analyses were conducted to explore relationships between important program characteristics and quality at the point-of-services.

Turnover and POS quality change.

Staff, and to a lesser extent supervisor, turnover is endemic to after school programs in Palm Beach and across the country, a fact that is often cited to explain the limited impact of reform efforts on the field. The QIS Pilot was explicitly designed counteract quality churn associated with staffing instability by: (a) providing continuous training, technical assistance and performance feedback through a strong intermediary; and (b) institutionalizing ideas about and creating a culture of quality at the program management level. In this section we ask: did the QIS Pilot's design drive positive quality change in spite of staff churn?

During the course of the QIS, seven of the 37 sites experienced supervisor turnover and 12 experienced more than 50% staff turnover. Bi-variate relationships between PBC-PQA (Form A) scores and supervisor and staff turnover were examined. Supervisor and staff turnover were not significantly correlated (Pearson-r coefficients) with post-pilot scores. Significant correlations were identified when examining the relationship between turnover and change scores (difference between post-pilot and baseline); specifically, 50% turnover in staff was correlated with greater positive change in I-E Healthy Food / Drink ($r = .42$), IV-R Goals/Plans ($r = .34$), and Domain I Safe Environment ($r = .33$). However, results of multi-variate analyses indicated that staff turnover was not a significant predictor of change in scores on these items when controlling for baseline scores. In other words, based on the available evidence, it appears that supervisor and staff turnover did not affect quality improvement during the QIS pilot.

Management practices and POS quality change.

Bi-variate relationships between PBC-PQA Form A and Form B were examined. The Form B practices most strongly/frequently related to point-of-service

quality scales were V-A *Attention to Youth Interests* and VI-G *Organizational Commitment to Program Improvement*. Specifically, V-A was significantly correlated with Form A scales II-I *Activities Support Active Engagement* ($r = .48$), III-O *Opportunities to Share Responsibilities / Mentor* ($r = .37$), and III-P *Opportunities to Partner with Adults* ($r = .33$), as well as with Domain III *Interaction Opportunities*; ($r = .39$) and Domain IV *Engaged Learning* ($r = .35$). VI-G was significantly correlated with Form A scales II-M *Sense of Belonging* ($r = .35$), III-O *Opportunities to Share Responsibilities / Mentor* ($r = .45$), III-P *Opportunities to Partner with Adults* ($r = .37$), III-Q *Positive Peer Relations* ($r = .36$), and IV-T *Reflection Opportunities* ($r = .46$), as well as with Domain III *Interaction Opportunities* ($r = .50$).

Significant bi-variate relationships between Form A and Form B domains were as follows: Domain V Youth Centered Policies and Practices was correlated with Domain III and IV Interaction Opportunities and Engaged Learning ($r = .38$ and $.36$ respectively). Domain VI High Expectations for Youth/Staff was correlated with Domain III Interaction Opportunities ($r = .43$). Strangely, Domain VII Organizational Logistics was negatively correlated with Domain I Safe Environment ($r = -.34$). Domain VII Family was not significantly correlated with any of the Form A domains. These correlations suggest that there is a relationship between what is emphasized / reported at an organizational level and what happens within the offerings at that site, with higher quality at the organizational level generally relating to higher quality at the point-of-service.

Program characteristics and POS quality change.

This section compares observed quality ratings across program characteristics, including program type and content focus. Table 13 compares programs on the four domains in the PBC-PQA (Safe Environment, Supportive Environment, Interaction, and Engagement). Overall there were very few significant differences related to program type or content focus and no clearly interpretable meaning to the pattern of significant differences that were found.⁵

A few caveats are in order when interpreting Table 13. First, these are not very sophisticated analyses and there may be program-level variables not included in our calculations that should be controlled for when comparing quality by program characteristics. Second, the analysis of the Champs sites is not executed at the optimal level of analysis. Table A compares overall quality at sites that use the Champs curriculum to overall quality at sites that do not. Due to incomplete data, we cannot make the more important comparison between Champs offerings and all other offerings. With these shortcomings noted, these findings do follow a pattern seen in other samples: offering level quality scores are not related to either the type or content focus of most programs.

Table 13. Comparison of Mean PBC-PQA Domains Across Program Types at Baseline & Post-Pilot

	<i>QIS Baseline</i>				<i>QIS Post-Pilot</i>			
	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>
Elementary (N=12)	4.63	4.07	3.42	2.49	4.85	4.36	3.65	2.76
Middle School (N=7)	4.56	3.99	3.49	3.24	4.87	4.29	3.61	2.65
Mixed Age (N=19)	4.32	3.69	3.21	2.46	4.69	4.28	3.59	2.99
Beacon (N=7)	4.49	3.72	3.28	2.50	4.88	4.42	3.91	3.01
Not Beacon (N=31)	4.45	3.89	3.34	2.64	4.75	4.28	3.54	2.82
Champs (N=12)	4.58	3.88	3.25	2.50	4.82	4.32	3.70	2.85
Not Champs (N=26)	4.41	3.86	3.37	2.66	4.75	4.30	3.57	2.86
Community Based (N=25)	4.38	3.83	3.25	2.49	4.69	4.23	3.59	2.83
School Based (N=11)	4.60	3.93	3.41	2.69	4.94	4.48	3.63	2.92
Not Licensed (N=6)	4.35	3.84	3.27	2.27	4.63	4.27	3.53	2.86
Licensed (N=17)	4.52	4.01	3.25	2.56	4.78	4.32	3.71	2.90
Exempt (N=13)	4.43	3.84	3.48	2.83	4.83	4.29	3.51	2.82
Licensing “In Process” (N=2)	4.46	3.89	3.14	2.67	4.78	4.33	3.60	2.59
Note: Domains: I. Safe Environment, II. Supportive Environment, III. Interaction, & IV. Engagement.								

Part IV.

Conclusions & Recommendations

The Palm Beach QIS Pilot was an assessment-driven, multi-level intervention designed to raise quality in after-school programs, and thereby raise the level of access to key developmental and learning experiences for the youth who attend. At its core, the QIS asked providers to identify and address strengths and areas for improvement based on use of the Palm Beach County Program Quality Assessment (PBC-PQA) – a diagnostic and prescriptive quality assessment tool – and then to develop and enact quality improvement plans. Throughout this process training and technical assistance are provided by several local and national intermediary organizations.

Overall, the impact of the Palm Beach QIS Pilot on program quality was quite positive. Both point of service and management level assessments made after the intervention was deployed are almost uniformly better than baseline quality measures. These results suggest that youth in QIS pilot programs received higher doses of quality programming as a result of the intervention. In other words, the QIS appears to be having the intended effect and thus should serve as a strategic exemplar to after school networks across the country.

Although this study does not have an experimental design and we cannot make strong claims about causality, there is substantial information available here and from other reports on the Palm Beach County QIS to offer informed opinions about “what worked.” Likely strengths of the QIS include:

- Multiple opportunities for site leaders and front-line staff to participate in elements of the QIS intervention (Intense Participation)
- The provision of performance data coupled with guided planning and quality coaching (Focused Improvement Planning)
- A “low stakes” orientation to quality measurement, including a “top down” commitment to providing support, not sanctions during improvement initiatives (Low Stakes Accountability)
- The intentional development by Prime Time Palm Beach County, the lead training and technical assistance intermediary, of a collaborative, quality-focused culture within and across participating sites (Collaboration between sites & Prime Time)

Taken together these findings have important implications for the design and deployment of quality improvement initiatives in the youth work field. Indeed, the Palm Beach QIS not only tells us that quality improvements are possible in systems marked by high staff turnover and other structural challenges, but also provides a detailed roadmap for realizing such gains in other places.

Endnotes

- 1 Data compiled by the National Child Care Information Center as of November 2006, cited in 2007 annual conference presentation for the National Association of Child Care Resource and Referral Agencies.
- 2 Both the Wallace Foundation and Robert Wood Johnson Foundation have recently funded large scale quality intervention efforts in cities, counties and states.
- 3 See discussion in Durlak, Taylor & Kawashima (2007) for a unique review of studies that employ system-level intervention models and assess effects at the child level. Some evidence exists that accreditation programs in early childhood can affect teacher performance (Hall & Cassidy, 2002; Bryant, Maxwell & Burchinal, 1999; Whitebook, Sakai, & Howes, 1997).
- 4 These unpublished findings come from the High/Scope Youth Program Quality Intervention study, a randomized field trial for a QIS-like intervention model currently underway in four states. For more information visit: YPQI.org.
- 5 Elementary programs scored significantly higher than mixed-age programs on Domain I-Safe Environment at Time 1; Elementary and mixed-age programs scored significantly lower than middle school programs on Domain IV- Engagement at Time 2; Beacon programs scored significantly higher than non-Beacon programs on Domain III-Interaction at Time 2; School-based programs scored significantly higher than community-based programs on Domain IV-Engagement at Time 1 and on I-Safe Environment at Time 2; There were NO significant differences by licensing status or Champs designation.

References

- Akiva, T., & Yohalem, N. (2006). *Quality Systems: Lessons from Early Efforts to Disseminate the Youth PQA*. Washington, DC: Forum for Youth Investment.
- Blazevski, J. and C. Smith (2007). *After-school quality and school-day outcomes in Michigan's 21st CCLC program*. Ypsilanti, MI, High/Scope Educational Research Foundation.
- Blumenfeld, P. C., Marx, R. W., & Harris, C. J. (2006). *Learning Environments*. In A. K.
- Renninger, I. E. Sigel, W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology*, 6th ed., (Vol. 4 *Child psychology in practice*, pp. 297-342). Hoboken, NJ: John Wiley & Sons.
- Brown, A. L. (1992). Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. *Journal of the Learning Sciences*, 2, 141-178.
- Bryant, D. M., Maxwell, K. L., & Burchinal, M. (1999). Effects of a Community Initiative on the Quality of Child Care. *Early Childhood Research Quarterly*, 14(4), 449-464.
- Center for Substance Abuse Prevention. (2002). *Finding the Balance: Program Fidelity and Adaptation in Substance Abuse Prevention* (Conference Edition). Washington, DC: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration.
- Durlak, J. A., Taylor, R. D., Kawashima, K., Pachan, M. K., DuPre, E. P., Celio, C. I., et al. (2007). Effects of positive youth development programs on school, family, and community systems. *American Journal of Community Psychology*, 39, 269-286.
- Durlak, J. A., & Weisberg, R. P. (2007). *The impact of after-school programs that promote personal and social skills*. Chicago, IL.: Collaborative for Academic, Social, and Emotional Learning.
- Fixsen, D. L., Naoom, S. F., Blasé, K. A., Friedman, R. M., & F. W. (2005). *Implementation research: A synthesis of the literature (No. 231)*. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.
- Gambone, M. A., Klem, A. M., & Connel, J. P. (2002). *Finding out what matters for youth: Testing key links in a community action framework for youth development*. Philadelphia: Youth Development Strategies Inc. & Institute for Research and Reform in Education.
- Gramiak, W., Lori Vanauken, Lauri Brugger, & Gillian M. Young-Miller. (2006). *Greater rochester after-school alliance (GRASA) assessment*. Rochester, NY: Children's Institute.
- Granger, R., Durlak, J.A., Yohalem, N., & Reisner, E. (2007). *Improving after-school program quality*. Unpublished manuscript, New York, NY.
- Hall, A. H., & Cassidy, D. J. (2002). An Assessment of the North Carolina School-age Child Care Accreditation Initiative. *Journal of Research in Childhood Education*, 17(1), 84-96.
- Halverson, R., J. Grigg, et al. (2005). *The new instructional leadership: Creating data-driven instructional systems in schools*. Madison, WI, Wisconsin Center for Education Research, University of Wisconsin-Madison.
- Quality in the Palm Beach County QIS: Final report from the QIS baseline data collection* (High/Scope, 2006)
- Technical Report: Quality in the Palm Beach County QIS baseline data collection* (High/Scope, 2006)
- Training satisfaction for High/Scope workshops*

- delivered as part of the Palm Beach QIS (High/Scope, 2006)
- Communities of Practice in the Palm Beach County QIS: A Preliminary Look at Findings from a Staff Survey* (High/Scope, 2006)
- Intercultural Center for Research in Education, & National Institute on Out-of-School Time. (2005). *Pathways to Success for Youth: What Counts in After-School*. Arlington, MA: United Way of Massachusetts Bay.
- Laitsch, D. (2006). *Assessment, High Stakes, and Alternative Visions: Appropriate Use of the Right Tools to Leverage Improvement*. Tempe, AZ, Arizonal State University College of Education: Education Policy Research Unit.
- Larson, R. (2000). Toward a psychology of positive youth development. *American Psychologist*, 55(5), 170-183.
- Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school time programs: A meta-analysis of effects for at-risk students. *Review of educational research*, 76(2), 275-313.
- Learning Point Associates & Berkely Policy Associates. (2006). *South Carolina extended learning time study: Final report*. Chicago, IL: Learning Point Associates.
- Lerner, R. M. (2005). Promoting positive youth development through community and after-school programs. In J. L. Mahoney, Larson, R.W., Eccles, J.S. (Ed.), *Organized activities as contexts of development: extracurricular activities, after-school and community programs* (pp. 4). Mahwah, NJ: Lawrence Erlbaum Associates.
- Little, P. (2007). *The quality of school-age child care in after-school settings*: Harvard Family Research Institution.
- Mason, S. A. (2003). *Learning from data: The roles of professional learning communities*. American Educational Research Association conference, Madison, WI..
- National Research Council and Institute of Medicine, Eccles, J., & A.Gootman, J. (Eds.). (2002). *Community programs to promote youth development*. Washington, DC: National Academy Press.
- Prime Time.(2007a)
- Prime Time.(2007b)
- Ryan, R. M. and K. W. Brown (2005). Legislating competence: The motivational impacts of high stakes testing as an educational reform. *Handbook of Competence*. C. Dweck and A. E. Elliot. New York, Guilford Press.
- 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., & Henry, B. (2006). Quality in the out-of-school time sector: Insights from the Youth PQA Validation Study, *Paper presented at the Society for Research on Adolescence biennial meeting*. San Francisco, CA.
- Smith, C., Blazeovski, J., Akiva, T., & Peck, S. J. (in submission). An empirical profile of after-school practices and pedagogies. *American Journal of Community Psychology*.
- Smith, C. and C. Hohmann (2005). Full findings from the Youth PQA validation study. *High/Scope Youth PQA Technical Report*. Ypsilanti, MI, High/Scope Educational Research Foundation.
- Spielberger, J., & Lockaby, T. (2006). *The prime time initiative of palm beach county, Florida: QIS development process evaluation: Year 2 report*. Chicago, IL: University of Chicago.
- Spielberger, J., & Lockaby, T. (2008). *The prime*

time initiative of palm beach county, Florida: QIS development process evaluation: Year 2 report.
Chicago, IL: University of Chicago.

Walker, K., & Arbreton, Amy J.A. (2004). *After-school pursuits: an examination of outcomes in the San Francisco Beacon Initiative.* San Francisco, CA: Public/Private Ventures.

Whitebook, M., Sakai, L., & Howes, C. (1997). *NAEYC accreditation as a strategy for improving child care quality: An assessment.* Washington DC: National Center for the Early Childhood Workforce.

Wiggins, G. P. (1993). *Assessing student performance: Exploring the purpose and limits of testing.* San Francisco, Jossey-Bass.

Wilson-Ahlstrom, A., & Yohalem, N. (2007). *Building Quality Improvement Systems in the Youth-Serving Sector: Lessons from Three Emerging Efforts.* Washington, DC: Forum for Youth Investment.

Yohalem, N., Pittman, K., & Moore, D. (2006). *Growing the next generation of youth professionals: Opportunities and Challenges.* Cornerstones for Kids and Forum for Youth Investment, Washington D.C.

Appendices

Appendix A. Emerging QIS Theory of Change

Figure A provides a generic theory of change for systems such as the QIS from the context of molding the SAE, in this case through intermediary inputs, to achieve improvement.

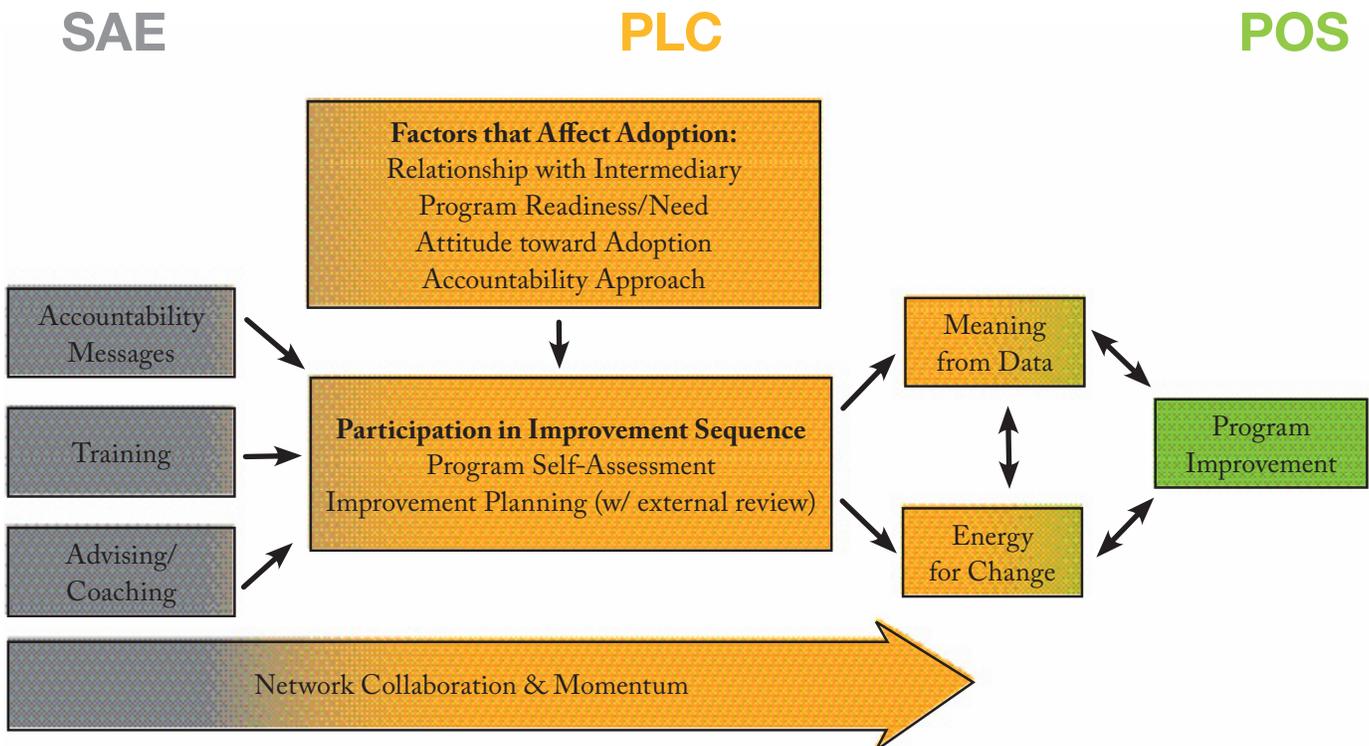
Figure A: Generic Quality Improvement System Theory of Change

Prime Time contributes to QIS the inputs of advising program directors, offering training for directors and staff, and communicating messages about accountability, incentives, and the nature of QIS. Advising and coaching are used synonymously, though in QIS advising is done by Quality Advisors and coaching by Peer Coaches trained and contracted by Prime Time. Advising consists of technical assistance for program directors as they participate in all the QIS components. Training workshops involve the PBC-PQA tool, and employing the best practice methods identified in the tool. Throughout the QIS, Prime Time

consistently communicates the importance of the initiative to program directors and staff. The consistency and face validity of these messages has an impact on how QIS is perceived and on level of participation and adoption.

The improvement sequence, though it contains several points of contact, is conceptually simple: program directors are taught how to use the PBC-PQA for self-assessment and how to use their external review to improve their program. The model assumes that the factors that affect adoption listed in the box above act as mediators—they influence the degree to which participation in the intervention components lead to site-based improvements. The degree to which meaning is made from data and energy for change exists determine improvements at sites. Throughout the QIS process, intermediary staff, program directors, and program staff interact frequently. This leads to networking and collaboration that would otherwise not occur. Though it is not a direct content focus of the QIS, this networking has a real impact. As this is successful, network momentum builds and positively affects program adoption.

Figure A: Generic Quality Improvement System Theory of Change



Appendix B. Standards Crosswalk

Table B1. Crosswalk between the PBC-PQA (1.41) and QIS Standards

<i>PBC-PQA</i>	<i>PBC Standards</i>	
	<i>I. STANDARD ONE</i>	
	<i>IA. PROGRAM ORGANIZATION INDICATORS:</i>	
VII-H	1	The administration utilizes sound business practices.
PB-B	2	Staff/ child ratios and group sizes permit the staff to meet the needs of children and youth.
VII-I.2	3	Program manager on-site.
VII-I.3	4	Daily transportation meets needs of children.
I-E(1-3)	5	The program serves healthy foods and drinks that meet the needs of children and youth.
	6	Suitable space and materials are available for program.
I-D.5	-	Adequate outdoor space for a variety of activities
I-D.1	-	Ample comfortable indoor space used for a variety of activities simultaneously
VII-I.4	-	Adequate office space for staff
I-D.3	-	Comfortable furniture in sufficient quantities
II-G.3	-	Sufficient materials for multiple activities
	7	Polices and procedures are in place to protect the safety of children and youth.
I-B(1-2)	-	No observable safety or health hazards in program space
I-C.2	-	Participants checked in and out
I-B(3-4)	-	Ventilation (heat and ac) is adequate and in good working order
I-C.1	-	Emergency procedures
VI-G.3	8	Ongoing assessment of program
VII-J.1	-	Up-to-date participant records
VII-J.2	-	Reliable and valid attendance tracking
VI-G.1	-	Regularly assess participant outcomes
	<i>IB. STAFFING AND PROFESSIONAL DEVELOPMENT INDICATORS:</i>	
VII-K	1	Staff recruitment – written job description; criminal backgrounds, drug screening and driving check; diversity
PB-C(1-3)	2	Staff qualifications
	3	Training
VI-D.1	-	Youth development orientation training
PB-D(1-2)	-	Annual in-service training
VI-D(2-3)	-	Professional development plans
VI-G.2	-	Annual evaluation
PB-E	4	Staff retention
VII-I.1	5	Process for managing staff absences
VII.K	6	Benefits for full-time staff
VI-D(4-5)	7	Staff meet regularly to plan curriculum and activities.
VI-G.4	8	Staff engaged in annual quality improvement process.
	<i>II. STANDARD TWO</i>	
	<i>IIA. STAFF TO YOUTH RELATIONSHIP INDICATORS:</i>	

<i>PBC-PQA</i>	<i>PBC Standards</i>	
III-P(1-2)	1	Staff engage youth as partners in program activities.
II-F.3	2	Youth experience positive gestures and words from staff.
II-I.3	3	Activities provide youth opportunities to communicate their thoughts, opinions and evaluation of experience to others.
I-A.2	4	Staff address stereotypic comments or slurs.
II-J.1	5	Staff encourage youth to take on challenging tasks.
II-J.2	6	Staff give verbal and non-verbal cues to all youth that suggest that they can succeed.
III-P.1	7	Staff share responsibility for program activities with youth.
IIIB. YOUTH TO YOUTH RELATIONSHIP INDICATORS:		
III-N(1-3)	1	Small group activities are available, have purpose and all group members are cooperating in accomplishing it.
III-M(1,2), II-H(1,2)	2	Youth identify with each other and with organization expectations for personal behavior.
III-Q(1-2)	3	Youth frequently experience positive gestures and words from other youth.
III. STANDARD THREE		
IIIA. ENVIRONMENT INDICATORS:		
II-F.1	1	A staff member greets each child by name daily.
II-F.2	2	Staff use a warm tone of voice and respectful language.
I-A.2	3	Staff understand, celebrate and reflect diversity.
V-A(1-2)	4	Special needs are identified and followed-up.
	5	Youth have opportunity to develop sense of belonging.
III-M.1	v	Youth have structured opportunities to get to know each other;
III-M.3	v	Youth have ownership and like the program; and
III-M.4	v	Youth are acknowledged for their achievement, works and contributions.
IIIB. BEHAVIOR MANAGEMENT INDICATORS:		
II-L(1-4)	1	Staff use youth-centered approaches to resolving conflicts.
II-L.1	2	Staff approach conflicts in a non-threatening manner.
II-L.2	3	Staff seek input from participants to determine cause and solution of conflicts/practice problem solving strategies.
II-H(1-2)	4	Staff communicate and reinforce clear limits and rules.
VI-E(2,3)	5	Staff receive regular training about promoting positive behavior.
II-H.3	6	Staff deal effectively around bullying issues and intimidation.
IIIC. ACTIVITIES INDICATORS:		
V-A.1	1	Activities are based on student interests and needs.
V-A.2	2	Activities are developmentally appropriate.
IV-S(1,2)	3	Youth are able to make choices about their activities.
II-I(1-4)	4	Program activities actively engage youth and are hands-on.
V-A.3	5	Balance of academic enhancement, cultural, service learning, life skills,, career exploration and recreational opportunities is provided.
III-P.(1-2)	6	Staff are constantly engaged with youth in activities.

<i>PBC-PQA</i>	<i>PBC Standards</i>	
	<i>IV. STANDARD FOUR</i>	
	<i>IVA. YOUTH DEVELOPMENT INDICATORS:</i>	
VI-D1	1	Staff have been trained in positive youth development principles and practices.
VI-D2	2	Staff have been trained in developmental stages.
IV-R(1-2)	3	Youth have time to reflect about activities.
IV-R.4	4	Youth have structured opportunities to provide feedback about activities.
III-O(1-3)	5	Youth have opportunity to assume leadership roles during group activities.
V-C.5	6	Youth have an active advisory committee to provide input on operations.
V-C.3	7	Youth are recognized for recruiting their peers to the program.
V-C.4	8	Youth have responsibilities for community outreach/service learning activities.
	<i>IVB. LEARNING APPROACH INDICATORS:</i>	
VI-F.2	1	Activities connect with school curriculum or learning standards.
VI-F.1	2	Planned activities have explicit objectives/learning goals.
VI-F.3	3	Program/staff communicate with regular school teachers to better understand and meet individual needs of youth.
	<i>V. STANDARD FIVE</i>	
	<i>FAMILY INVOLVEMENT INDICATORS:</i>	
VII-L.2	1	Staff and families interact with each other in positive ways.
VIII-L.3	2	Program helps parents connect with school and child's education.
VIII-L.1	3	Several mechanisms are used to regularly communicate with family.
VIII-M.1	4	Family have several opportunities to visit program to see child perform or be recognized for their accomplishments.
VIII-M(2-4)	5	Staff support families' involvement in the program.

Appendix C. Domains, Scales and Items for the PBC-PQA Forms A & B and Staff Survey

Table C1. Form A

<i>I. SAFE ENVIRONMENT</i>
A. Psychological and emotional safety is promoted.
(I-A1) The emotional climate of the session is predominantly positive (e.g., mutually respectful, relaxed, supportive; characterized by teamwork, camaraderie, inclusiveness, and an absence of negative behaviors). Any playful negative behaviors (not considered offensive by parties involved) are mediated (countered, curtailed, defused) by staff or youth.
(I-A2) There is no evidence of bias but rather there is mutual respect for and inclusion of others of a different religion, ethnicity, class, gender, ability, appearances, or sexual orientation.
B. The physical environment is safe and free of health hazards.
(I-B1) The program space is free of health and safety hazards.
(I-B2) The program space is clean and sanitary.
(I-B3) Ventilation and lighting are adequate in the program space.
(I-B4) The temperature is comfortable for all activities in the program space.
C. Policies and procedures protect children and youth.
(I-C1) Written emergency procedures are posted in plain view.
(I-C2) All young people are checked in and out of the program.
(I-C3) Access to outdoor program space is supervised during program hours.
D. Program space and furniture accommodate the activities offered.
(I-D1) Program space is ample for youth and adults to move freely while carrying out activities (e.g., accommodates all participants without youth blocking doorways, bumping into one another, crowding around).
(I-D2) Program space is suitable for all presented activities (e.g., furniture and room support small and large groups; if athletic activity is offered, then program space supports this).
(I-D3) Furniture is comfortable and of sufficient quantity for all youth participating across program offering.
(I-D4) Physical environment can be modified to meet the needs of the program offering (e.g., furniture and/or supplies can be moved).
(I-D5) Outdoor program space is ample for youth to move freely while carrying out various activities (e.g., accommodates all participants, plentiful room for group physical activities such as team sports.)
E. Healthy foods and drinks are provided.
(I-E1) Drinking water is available and easily accessible to all youth.
(I-E2) Plentiful food and drinks are available at appropriate times for all youth during program session.
(I-E3) Available food and drink is healthy (e.g., fresh fruit, vegetables, real juice, homemade dishes).
<i>II. SUPPORTIVE ENVIRONMENT</i>
F. Staff provide a welcoming atmosphere.
(II-F1) All youth are greeted by (a) staff within the first 15 minutes of the program session.
(II-F2) Staff mainly use a warm tone of voice and use respectful language during program activities.
(II-F3) Staff mainly wear a smile, use friendly gestures, and make eye contact during program activities.
G. Session flow is planned, presented, and paced for youth.
(II-G1) Staff start and end session within 10 minutes of scheduled time.
(II-G2) Staff have all materials and supplies ready to begin all activities (e.g., materials are gathered, set up).
(II-G3) There are enough materials and supplies prepared for all youth to begin activities.

(II-G4) Staff explain all activities clearly (e.g., youth appear to understand directions; sequence of events and purpose are clear).
(II-G5) There is an appropriate amount of time for all of the activities. (e.g., youth do not appear rushed, frustrated, bored, or distracted; most youth finish activities).
H. Staff effectively maintain clear limits.
(II-H1) Staff communicate clear limits and rules.
(II-H2) Staff consistently reinforce stated limits and rules.
(II-H3) Staff effectively deal with direct and indirect incidents of bullying and intimidation.
I. Activities support active engagement.
(II-I1) The bulk of the activities involve youth in transforming (creating, combining, reforming) materials or ideas OR improving a skill through guided practice.
(II-I2) The program activities lead (or will lead in future sessions) to tangible product(s) or performance(s) that reflect youth ideas or designs.
(II-I3) The activities provide all youth one or more opportunities to talk about (or otherwise communicate) what they are doing and what they are thinking about to others.
(II-I4) The activities balance concrete experiences involving materials, people, and projects (e.g., field trips, experiments, interviews, service trips, creative writing) with abstract concepts (e.g., lectures, diagrams, formulas).
J. Staff support youth in building new skills.
(II-J1) All youth are encouraged to try out new skills or attempt higher levels of performance.
(II-J2) All youth who try out new skills receive support from staff despite imperfect results, errors, or failure; staff allow youth to learn from and correct their own mistakes and encourage youth to keep trying to improve their skills.
K. Staff support youth with encouragement.
(II-K1) During activities, staff are almost always actively involved with youth (e.g., they provide directions, answer questions, work as partners or team members, check in with individuals or small groups).
(II-K2) Staff support at least some contributions or accomplishments of youth by acknowledging what they've said or done with specific, nonevaluative language (e.g., "Yes, the cleanup project you suggested is a way to give back to the community." "I can tell from the audience response that you put a lot of thought into the flow of your video.").
(II-K3) Staff make frequent use of open-ended questions (e.g., staff ask open-ended questions throughout the activity and questions are related to the context).
L. Staff use youth-centered approaches to reframe conflict.
(II-L1) Staff predominantly approach conflicts and negative behavior in a nonthreatening manner (i.e., approach calmly, stop any hurtful actions, and acknowledge youth's feelings).
(II-L2) Staff seek input from youth in order to determine both the cause and solution of conflicts and negative behavior (e.g., youth generate possible solutions and choose one).
(II-L3) Staff encourage youth to examine the relationship between actions and consequences in helping youth to understand and resolve conflicts and negative behaviors.
(II-L4) Staff acknowledge conflicts and negative behavior and follow up with those involved afterward.
III. INTERACTION
M. Youth have opportunities to develop a sense of belonging.
(III-M1) Youth have structured opportunities to get to know each other (e.g., there are team-building activities, introductions, personal updates, welcomes of new group members, icebreakers, and a variety of groupings for activities).
(III-M2) Youth exhibit predominately inclusive relationships with all in the program offering, including newcomers.
(III-M3) Youth strongly identify with the program offering (e.g., hold one another to established guidelines, use ownership language, such as "our program," engage in shared traditions such as shared jokes, songs, gestures).

(III-M4) The activities include structured opportunities (e.g., group presentations, sharing times, recognition celebrations, exhibitions, performances) to publicly acknowledge the achievements, work, or contributions of at least some youth.
N. Youth have opportunities to participate in small groups.
(III-N1) Session consists of activities carried out in at least 3 groupings—full, small, or individual.
(III-N2) Staff use 2 or more ways to form small groups (e.g., lining up by category and counting off, grouping by similarities, signing up).
(III-N3) Each small group has a purpose (i.e., goals or tasks to accomplish), and all group members cooperate in accomplishing it.
O. Youth have opportunities to share responsibilities.
(III-O1) All youth have multiple opportunities to practice group process skills (e.g., actively listening, contributing ideas or action to the group, doing a task with others, taking responsibility for a part).
(III-O2) All youth have one or more opportunities to lead a group during program activities.
(III-O3) All younger (K-6) youth have one or more opportunities to help another youth with a task during program activities; all older (6+) youth have one or more opportunities to mentor an individual during program activities.
P. Youth have opportunities to partner with adults.
(III-P1) Staff share control of most program activities with youth, providing guidance and facilitation while retaining overall responsibility.
(III-P2) Staff always provides an explanation for expectations, guidelines, or directions given to youth.
Q. Youth have opportunities to develop positive peer relationships.
(III-Q1) Youth mainly use a warm tone of voice and use respectful language with each other.
(III-Q2) Youth mainly smile, use friendly gestures, and make eye contact with each other.
IV. ENGAGEMENT
R. Youth have opportunities to set goals and make plans.
(IV-R1) In the course of the program offering, all youth are given a structured opportunity to set one or more long-term goals.
(IV-R2) Time is regularly provided for young people to make (individual or group) plans for and/or to set goals for activities.
(IV-R3) Young people are encouraged to share their plans and represent their plans in a tangible way using words, writing, diagram, etc. (e.g. a small group draws a diagram before building; staff helps full group make a large idea web to plan an event, etc.)
S. Youth have opportunities to make choices based on their interests.
(IV-S1) All youth have the opportunity to make at least one open-ended content choice within the content framework of the activities (e.g., youth decide topics within a given subject area, subtopics, or aspects of a given topic).
(IV-S2) All youth have the opportunity to make at least one open-ended process choice (e.g., youth decide roles, order of activities, tools or materials, or how to present results).
T. Youth have opportunities to reflect.
(IV-T1) All youth are engaged in an intentional process of reflecting on what they are doing or have done (e.g., writing in journals; reviewing minutes; sharing progress, accomplishments, or feelings about the experience).
(IV-T2) All youth are given the opportunity to reflect on their activities in 2 or more ways (e.g., writing, role playing, using media or technology, drawing).
(IV-T3) In the course of the program offering, all youth have structured opportunities to make presentations to the whole group.
(IV-T4) Staff initiate structured opportunities for youth to give feedback on the activities (e.g., staff ask feedback questions, provide session evaluations).

Table C3. Staff Survey

SECTION I. PROFESSIONAL LEARNING COMMUNITY
Supportive Staff / Shared Norms $\alpha = .88$
1. My beliefs and values about the mission of the program are shared by most of my co-workers
2. I feel that everyone in our program is working together toward common goals
3. I am supported by other staff to try out new ideas
4. I can get good advice from other staff if I have problems with the youth
Staff Empowerment $\alpha = .82$
1. I am regularly involved in making decisions that affect our program
2. I regularly have an active role in planning about our program
3. I have a significant role in shaping the program's norms, values, and practices
Supervisor Quality Emphasis $\alpha = .89$
1. My supervisor emphasizes sharing control with youth as a core program value
2. My supervisor emphasizes active learning with youth as a core program value
3. My supervisor emphasizes a strong sense of belonging as a core program value
Supervisor Support $\alpha = .75$
1. My supervisor gives good feedback about how I work with youth
2. My supervisor challenges me to innovate and try new ideas
3. My supervisor knows what I am trying to accomplish with youth
4. My supervisor makes sure that program goals and priorities are clear to me
Decisional Capacity – POS $\alpha = .76$
How much control do you have over decisions about...
1. When and how daily activities take place in activities that you lead for youth
2. The types of daily activities that occur in sessions that you lead for youth
3. The availability of supplies that you need
Decisional Capacity – ORG $\alpha = .68$
How much control do you have over decisions about...
1. How much you are paid
2. How often you work late
3. Using paid time to plan for your program offerings
4. Cutting back on the number of hours that you work
SECTION II. BELIEFS ABOUT YOUTH WORK
Professional Self-Efficacy $\alpha = .86$
1. I am successful in providing the experiences that I want to provide for youth
2. I can build a positive relationship with even the most difficult or unmotivated youth
3. With patience and goodwill, I can help any youth to learn
4. I am adequately trained and prepared to work with the youth at my current job
Adult Control $\alpha = .71$
How important is it...
1. For youth to work on homework quietly and by themselves
2. To solve problems for youth so conflict does not arise

3. For staff to occasionally demonstrate authority using punishment or reprimand
4. For adults to step in and make decisions when youth are talking
5. For youth to be quiet and respectful so adults will respect them
6. To limit choices for youth that have too many (unhealthy) choices already
Shared Control $\alpha = .65$
How important is it...
1. For youth to be involved in establishing rules for the activity or session
2. To provide opportunities for unstructured or informal time
3. For youth to be involved in hiring new staff
4. For youth to be involved in how the organization's budget is spent
5. For youth to learn routines so they take responsibility for their own program
Emphasis on Relationships $\alpha = .78$
How important is it...
1. To make strong relationships between staff and youth the highest priority
2. To make strong relationships among the program youth the highest priority
Modeling $\alpha = .63$
How important is it...
1. For staff to use the same behaviors they want from youth
2. For supervisors to interact with staff like they want staff to interact with youth
SECTION III. SELF-REPORTED PRACTICES
Meaningful Learning Experiences $\alpha = .82$
1. The youth activity ends with a product or performance
2. The youth take planned assignments or activities in new directions
3. Youth are asked to talk about what they are doing and thinking
4. Youth are encouraged to try new skills (Ex. writing a poetry, using a saw)
5. Planning strategies are used (Ex. brainstorming, idea webbing)
6. Youth review or reflect on their work (writing in journals, reviewing minutes)
Shared Control – POS $\alpha = .79$
1. Youth have opportunities to teach or coach others
2. Youth have opportunities to lead a group
3. Youth have an assignment or project that they decide how to complete
4. Youth have input on what activities are offered

Appendix D. Psychometric Performance of the PBC-PQA

The QIS baseline and post-pilot data were used to test the psychometric performance of the PBC-PQA. Tables D1, D2 and D3 present findings for analyses that use the total sample of 139 offerings at baseline and 128 offerings at post-pilot.

Findings

- PBC-PQA domains demonstrate acceptable levels of internal consistency at all three data collection timepoints (including the YPQA Validation Study)
- Bi-variate correlation and factor analyses from QIS baseline and post-pilot suggest that the PBC-PQA domains are related but distinguishable constructs, but that the structure of the domains might be improved changing the position of a few scales
- Preliminary evidence suggests that it may not be necessary to sample more than three offerings per site, regardless of program size
- Rater reliability on the PBC-PQA was 70% perfect agreement overall at QIS Time 1 (N=13 rater pairs)

This appendix the following sections: internal consistency, rater reliability, correlation between domains, factor analyses, predictive and concurrent validity, and evaluating the data collection formula.

Internal Consistency (Scale Reliability)

Table D1 provides internal consistencies (alphas) for the four PBC-PQA observational domains. Alphas are not reported for domain I (Safe Environment) because the items on this scale operate more like a dichotomous checklist rather than a one-dimensional construct and therefore do not conform to the underlying assumptions for reliability testing. At post-pilot and for the YPQA Validation study, the internal consistency coefficients meet or exceed the general rule (>.7) for acceptable scale performance.

Rater Reliability

Table D2 presents rater reliability information for the QIS baseline data collection. Raters were trained through a two-day training and “anchored observation” where they collected data with an expert rater and compared scores and spent substantial time analyzing differences. The results from the table below are drawn from the anchored observation for 13 raters. Because each item on the PBC-PQA is based on a three-point scale, random rater agreement (the trainee guesses on every item) would achieve 33% perfect agreement. Our goal for “anchored” raters is 80% perfect agreement at the item level. Our goal for the anchored observation during the QIS baseline (since it was the first “check” on the raters’ accuracy) was 70% overall.

Table D2 presents the item-level perfect agreement, averaged to the scale level, for 13 rater pairs. The rate of perfect agreement across all raters at the

Table D1. Internal Consistency for the PBC-PQA Domains at Baseline, Post-Pilot, and Youth PQA Validation Study

<i>Youth PQA Observation Domains</i>	QIS Baseline Internal consistency (alphas) <i>N=139 ratings</i>	QIS Post-Pilot Internal consistency (alphas) <i>N=128 ratings</i>	YPQA Validation Internal consistency (alphas) <i>N=199 ratings</i>
I. Safe Environment (5 scales)	No alpha reported. The scales in this section do not meet reliability model assumptions.		
II. Supportive Environment (6 scales)	.85	.86	.85
III. Interaction Opportunities (4 scales)	.66	.74	.70
IV. Engaged Learning (3 scales)	.68	.72	.81

Table D2. PBC-PQA Reliability by Scale

<i>Scale</i>	<i>Percent Perfect Item Agreement Averaged to Scale</i>
I-A. Psychological and emotional safety are promoted.	78%
I-B. The physical environment is safe and healthy for youth.	79%
I-C. Policies and Procedures protect children and youth.	84%
I-D. Rooms and furniture accommodate activities.	79%
I-E. Healthy foods and drinks are provided.	77%
II-F. Staff provides a welcoming atmosphere.	84%
II-G. Session flow is planned, presented, and paced for youth.	76%
II-H. Staff effectively maintain clear limits.	65%
II-I. Activities support active engagement.	63%
II-J. Staff support youth to build new skills.	63%
II-K. Staff support youth with encouragement.	72%
II-L. Staff use youth-centered approaches to reframe conflict.	70%
III-M. Youth have opportunities to develop a sense of belonging.	60%
III-N. Youth have opportunities to participate in small groups.	69%
III-O. Youth have opportunities to share responsibilities.	56%
III-P. Youth have opportunities for adult-youth partnerships.	68%
III-Q. Youth have opportunities to develop positive peer relationships.	81%
IV-R. Youth have opportunities to set goals and make plans.	68%
IV-S. Youth have opportunities to make choices based on interests.	85%
IV-T. Youth have opportunities to reflect.	72%

“anchored observation” was 72%. It should be emphasized that after these scores were generated, the trainer and new data collector spent time analyzing difference – a process that we know raises reliability on items where errors occur when data collectors are learning the tool.

Correlations between Domains

Table D3 presents bi-variate correlation coefficients for each domain, suggesting that the domains are related, but not redundant.

Factor Analyses

Table D4 presents results from exploratory factor analysis for 15 PBC-PQA scales in domains II, III, and IV. In Table C all loadings below .30 were suppressed unless they fell within the designated domain (see boxed areas). The boxes in

each column present the theoretically established domains. Results from baseline and post-pilot provide moderate support for the current instrument structure. In general, scales loaded on the predicted domain (factor). Factor loadings were generally acceptable (<.30), although some scales which were problematic at both baseline and post-pilot (particularly II-I, II-K, III-M, III-Q, and IV-T) may be usefully repositioned in a slightly different factor structure – at least for research purposes. Domain I is not included in this analysis because it represents a different scalar level (counting things that are present or not present and do not vary over time e.g., presence of a fire extinguisher) than the behavioral items in domains II-IV.

Predictive and Concurrent Validity

Evidence for the predictive and concurrent validity

Table D3. Correlation Coefficients for Domains I-IV at Baseline and Post-Pilot

	QIS Baseline			QIS Post-Pilot		
	I	II	III	I	II	III
I. Safe Environment (5 scales)						
II. Supportive Environment (6 scales)	0.47**			0.36**		
III. Interaction Opportunities (4 scales)	0.46**	0.61**		0.31**	0.51**	
IV. Engaged Learning (3 scales)	0.23**	0.45**	0.47**	0.02	0.40**	0.50**

**p>.01

of the PBC-PQA domains is provided in Table D5. In support of predictive validity (i.e., Do the variables predict or correlate with theoretically expected outcomes), we found that scores that a program receives for the Interaction Opportunities domain are related to student reports of positive affect (e.g., “I felt a sense of pride in what I was able to accomplish.”). Specifically, Interaction Opportunities was significantly correlated with Positive Affect ($r = .31$) and was identified as a significant predictor of this variable in multivariate

and multi-level analyses which controlled for frequency of attendance, parent requirement to attend, age and gender (Table D6). We also found that scores that a program receives for the Engaged Learning domain are related to student reports of Challenge (e.g., “I really had to concentrate to concentrate to do the activities.”, $r = .29$). This relationship remained even when controlling for frequency of attendance, parent requirement to attend, age and gender in multivariate and multi-level analyses (Table D6).

Table D4. Factor Analysis for PBC-PQA scales F-T at Baseline and Post-Pilot

	QIS Baseline		
	Factor 1	Factor 2	Factor 3
Variance Explained	21%	20%	15%
Score II-F	0.64		
Score II-G	0.74		
Score II-H	0.80		
Score II-I	0.29	0.66	
Score II-J	0.42	0.51	
Score II-K	0.46	0.65	
Score II-L	0.59		
Score III-M	0.42		0.34
Score III-N			0.76
Score III-O			0.65
Score III-P			0.70
Score III-Q	0.71		-0.01
Score IV-R		0.81	
Score IV-S		0.21	0.57
Score IV-T		0.78	

Extraction Method: Principal Component Analysis.
(Forced 3-factor model at post-pilot)
Rotation Method: Varimax with Kaiser Normalization.

In support of concurrent validity (i.e., Do the variables correlate positively with other measures of similar constructs), we found that the Youth Perception of Program Quality scale score was significantly correlated with a program’s score on Interaction Opportunities ($r = .33$), which was largely driven by the correlation between the survey item, “Kids worked together to solve problems” and this domain ($r = .31$). The relationship between Interaction Opportunities and Youth Perception of Program Quality remained even when controlling for frequency of attendance, parent requirement to attend, age and gender in multivariate and multi-level analyses (Table D6).

Evaluating the Data Collection Formula

The PBC-PQA (Form A) was used to observe and assess 38 sites in Palm Beach County. Because these sites varied in size, a larger amount of observational data was collected from larger sites. The number of observations per site ranged from two to eight, with the majority having three. In order to determine the necessity of collecting varying amounts of data, sites with four or more observations were analyzed (N=13 at baseline; N=12 at post-pilot). Two data

Table D5. Correlation between Youth Survey Scales & Form A Domains

		<i>Safety</i>	<i>Support</i>	<i>Interaction</i>	<i>Engaged Learning</i>
<i>Predictive Validity</i>	<i>Positive Affect</i> (= .74 M = 3.09) <ul style="list-style-type: none"> • I was interested in what we did • The activities were important to me • I got better at things I care about • I felt a sense of pride about what I had accomplished 	-0.04	.22	.31*	.21
	<i>Challenge</i> (= .70 M = 3.06) <ul style="list-style-type: none"> • I was challenged in a good way • I tried to do things I have never done before • I really had to concentrate to complete the activities • I was using my skills 	-0.13	.18	.22	.29*
<i>Concurrent Validity</i>	Youth Perception of Program Quality (= .68 M = 2.97) <ul style="list-style-type: none"> • Staff and students treated each other with respect • Staff explained things in another way if I was confused • Kids worked together to solve problems • I had a lot of choice about what we did 	-0.02	.23	.33*	.22

Table D6. Hierarchical Linear Models (Level 2) of Youth Challenge, Positive, Affect, and Perception of Quality (N = 48 offerings; 487 youth)

<i>Estimated Effects</i>	<i>Challenge</i>	<i>Positive Affect</i>	<i>Perception of Quality</i>
Intercept (<i>B0</i>)			
Base	2.780**	2.737**	2.608**
Domain III Score	Not in model	.162*	.181*
Domain IV Score	.116*	Not in model	Not in model
Youth Attendance Rate (<i>B1</i>)			
Required to Attend (<i>B2</i>)	.117**	.098**	.098**
Age (<i>B3</i>)	.056	.046	.074
Gender (<i>B4</i>)	-.012	.004	-.001
ICC			
Proportion of individual level variance explained by Level 1 variables	.032	-.040	.021
Proportion of between-offering variance explained by addition of Level 2 variables	.066	.022	.151
Notes. Coefficients reported in their original metric (not standardized). * p<.05, ** p<.01. Youth Attendance Rate is coded 1=almost none, 2=a few times each month, 3=once each week, 4=a few times each week. Required to Attend is coded 1=yes, 0=no. Age is coded in original metric (e.g., 7=7 years). Gender is coded 1=male, 0=female.			

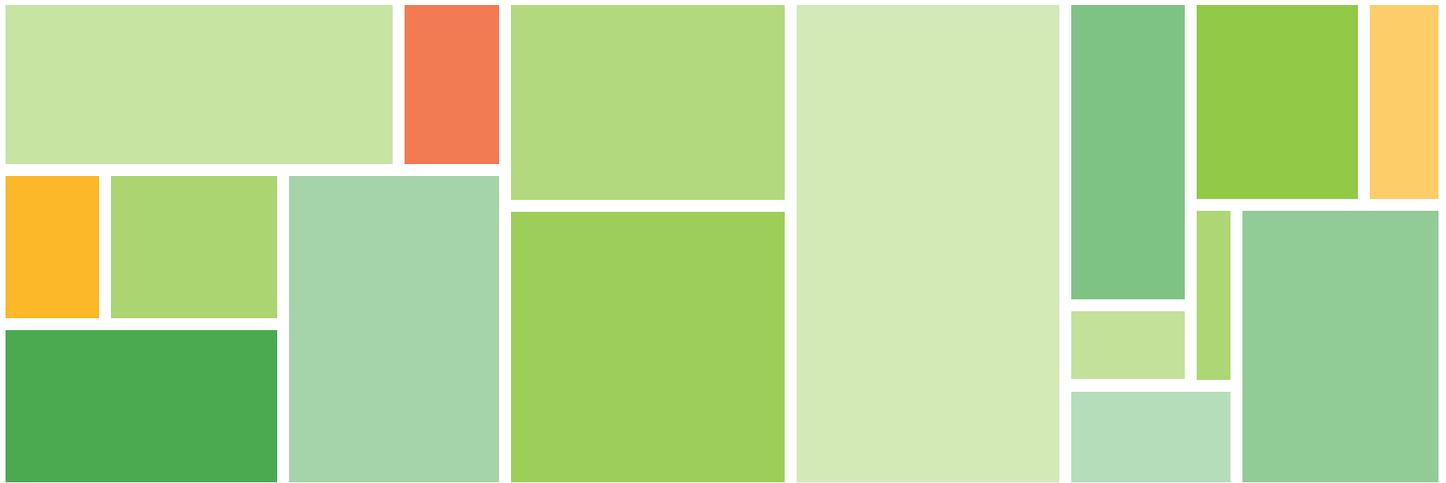
sets were created; the first composed of the mean scores of each scale and domain using all of the collected Form A scores. The second set was created by randomly selecting three offerings for each site and using the mean of scores for each scale and domain using data associated with just that subset of offerings.

Table D7 presents mean differences and correlation coefficients for organizational quality scores calculated using a random selection of three offerings versus the formula-driven score. There

were no statistically significant differences between the mean scores for any of the PBC-PQA scales or domains. Based on these results, it appears that the additional observations (beyond three offerings) do not significantly alter the mean site-level quality scores. However, additional, more sophisticated analyses are underway to investigate optimum number of observations per site. Accordingly, alteration of the data collection formula is not recommended at this time.

Table D7. Comparing Form A Score Aggregated to the Org Level: Mean Difference Between 3 Randomly Selected Offerings and Formula-Driven Aggregate

	<i>QIS Baseline</i>		<i>QIS Post-Pilot</i>	
	<i>Mean Difference</i>	<i>Correlation (Pearson-r)</i>	<i>Mean Difference</i>	<i>Correlation (Pearson-r)</i>
I-A. Psychological and emotional safety are promoted	.13	.87	.00	.92
I-B. The physical environment is safe and healthy for youth	.00	.95	.00	.97
I-C. Policies and Procedures protect children and youth	.01	.96	.04	.97
I-D. Rooms and furniture accommodate activities	.74	.06	.00	.98
I-E. Healthy foods and drinks are provided	-.07	.97	.00	.97
II-F. Staff provides a welcoming atmosphere	.10	.94	.12	.56
II-G. Session flow is planned, presented, and paced for youth	-.01	.82	.06	.85
II-H. Staff effectively maintain clear limits	.00	.91	.10	.83
II-I. Activities support active engagement	.01	.91	-.05	.93
II-J. Staff support youth to build new skills	.01	.92	.02	.82
II-K. Staff support youth with encouragement	-.02	.83	.11	.95
II-L. Staff use youth-centered approach to reframe conflict	-.10	.85	-.12	.89
III-M. Youth have opportunities develop sense of belonging	-.00	.90	-.09	.88
III-N. Youth have opportunities to participate in small groups	.12	.86	-.16	.91
III-O. Youth have opportunities to share responsibilities	.04	.86	.06	.96
III-P. Youth have opportunities to partner with adults	.06	.85	-.02	.94
III-Q. Youth have opps to develop positive peer relationships	-.01	.92	-.01	.95
IV-R. Youth have opportunities to set goals and make plans	.06	.88	.04	.93
IV-S. Youth have opps to make choices based on interests	.05	.88	.01	.90
IV-T. Youth have opportunities to reflect	-.07	.82	.10	.93
<i>Domain Scores</i>				
I. Safe Environment	-.11	.91	.00	.97
II. Supportive Environment	.07	.87	.03	.84
III. Interaction Opportunities	-.04	.83	-.04	.96
IV. Engaged Learning	.02	.96	.04	.97



David P. Weikart Center for Youth Program Quality

124 Pearl Street, Suite 601

Ypsilanti, Michigan 48197

Phone: 734.961.6900

Web: www.cypq.org