

*The Youth Program Quality Assessment is one of an emerging class of observational assessment tools that measure staff performances at the point of service and can help drive systemic approaches to workforce development and accountability.*

# 6

## Quality and accountability in the out-of-school-time sector

*Charles Smith, Thomas J. Devaney,  
Tom Akiva, Samantha A. Sugar*

YOUTH-SERVING ORGANIZATIONS that operate outside the traditional school day face increasing accountability pressures. According to a recent tally, fourteen states are implementing quality accountability systems for subsidized child care and another thirty are “exploring/designing or piloting” some type of quality improvement process.<sup>1</sup> In addition, several national philanthropic organizations and public agencies have invested heavily in after-school quality improvement projects in cities and counties across the country, and numerous place-based projects are moving forward with local resources.<sup>2</sup> Thus, more than ever before, theories of performance measurement and innovative quality accountability policies

This article was supported by the William T. Grant Foundation, the Picower Foundation, and the Robert Wood Johnson Foundation. Correspondence can be addressed to Charles Smith at [charles@cypq.org](mailto:charles@cypq.org) or David P. Weikart Center for Youth Program Quality, 124 Pearl Street, Suite 601, Ypsilanti, MI 48198.

are being brought to bear within and across out-of-school-time (OST) institutions.

Implicit in the calls to improve quality across the diverse universe of OST organizations is the notion that the quality of what teachers and youth workers do when they interact with youth influences how young people engage with content and, more broadly, supports developmental change on a range of cognitive, social-emotional, health, and other outcomes. Thus, the point of service—the microsetting where intentional interactions between adults and youth occur—is increasingly viewed as the place where the most important qualities of an OST program are enacted and where youth outcomes can be most directly influenced. Traditional accountability systems are not optimized to define, measure, or support high-quality practices in microsettings; rather, they are designed to measure and influence the inputs and outputs that bound these settings.

In the after-school field, discomfort with the applicability of outcomes-based policy has led to the development of alternative accountability and improvement models that draw on understandings of adult motivation, knowledge management, and performance change.<sup>3</sup> Rather than holding professional staff “accountable” for peak performances of children on standardized tests or for population-level outcomes such as teen pregnancy or recidivism rates, these models attempt to provide OST programs with a set of metrics, methods, and policies they can use to:

- Create learning communities focused on standards and professional learning
- Measure and track changes in the quality of staff performances at the point of service
- Align professional development resources across programs and policies

Although these approaches hold great promise for the OST field, several questions about their practicality and ultimate impact remain:

- How can point-of-service quality be defined in a way that makes sense within and across the many organizational types that make up the OST field?
- Which assessment technologies (defined as an assessment tool plus methods of use) can reliably measure and represent quantities of quality so defined?
- How can the system-level momentum and incentives necessary to pursue continuous quality improvement be generated and sustained in a highly fragmented policy environment?

In this article, we provide a definition of point-of-service quality that is relevant to and measurable across a broad range of OST settings; describe the Youth Program Quality Assessment and an associated continuum of methods as a set of technologies designed to serve different accountability purposes; and offer advice to OST organizations and decision makers interested in using quality metrics as a foundation for innovative accountability policies.

---

### *Defining quality*

In this opening section, we provide a detailed definition of point-of-service quality and argue that this construct is most profitably located and measured within, and compared across, OST program offerings in which adults and youth deliberately interact. That is, we suggest that a content-neutral quality construct, applied at the OST program offering level, can support (1) measurement of quality across the diverse universe of OST systems and sites and (2) systemic implementation of quality accountability and improvement policies.

### *What do we mean by quality?*

When we think about quality we start with a simple formula. Developmentally powerful environments provide (1) positive relationships as a context for (2) a learning task or content that (3) increases in complexity over time.<sup>4</sup> This rough formula is content neutral, that is, it can be applied across all academic and nonacademic settings

where adults and youth purposefully interact, and thus it represents a generic process for supporting positive developmental change.

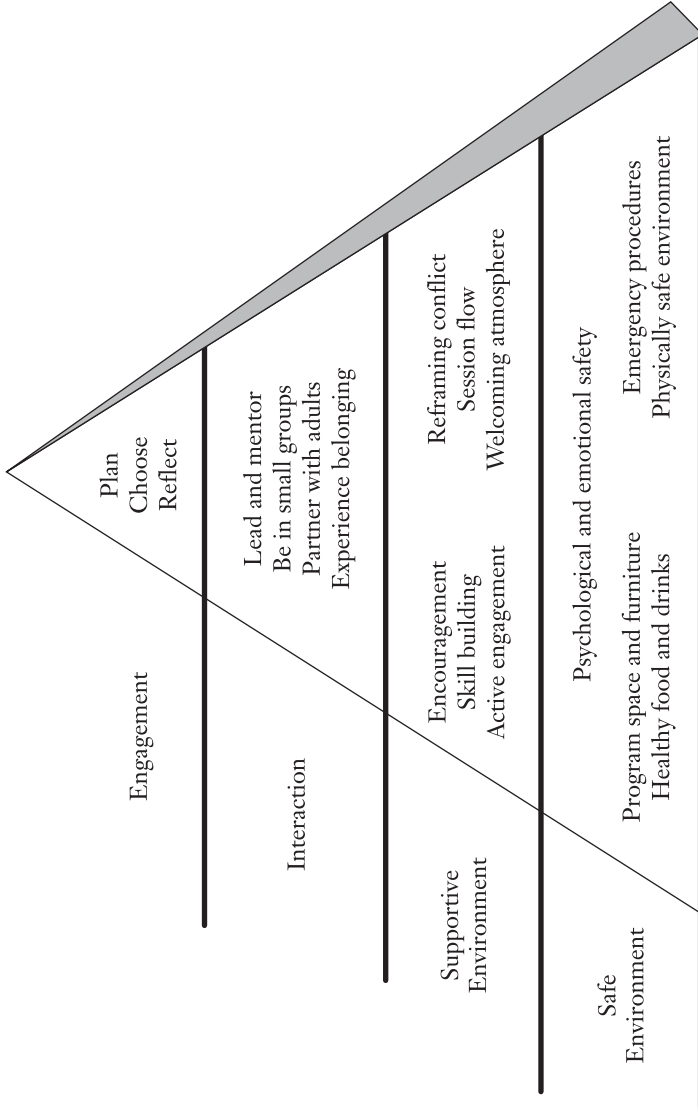
The specific content that an OST organization chooses to provide is of critical importance. However, because OST organizations seek to influence youth development trajectories through a variety of academic, enrichment, recreational, and other content-laden activities, it is virtually impossible to define quality related to particular content in a way that is relevant across the field. Consequently, we define quality in terms of factors that are ubiquitous across program types. Specifically, we define quality in terms of staff behaviors and the nature of their interactions with youth. For example, everyone agrees that relationships between staff and youth are a critical characteristic of all high-quality youth settings and that specific practices, such as welcoming participants, ensuring inclusion, and employing conflict resolution techniques, influence those relationships regardless of setting type. These kinds of practices represent the generic, content-neutral professional skills that are necessary, for example, to create and maintain a positive relational context and scaffold youth toward deep engagement with content.

Our construct for quality is thus focused on discrete staff behaviors that represent best practice in youth work methods and reflect important aspects of the research validated High/Scope active-participatory approach to learning.<sup>5</sup> The elements described in Figure 6.1 are conceptualized as a set of staff practices that provide key developmental experiences for youth and can be reliably observed. A recent compendium of assessment tools suggests wide agreement about specific elements of best practices in OST programs as defined in Figure 1.<sup>6</sup>

### *Where do we look for quality?*

Although there is general agreement about the basic elements of high-quality practice in youth settings, there is less agreement and even less discussion about where quality is located within those settings. The “where” question, however, is critical. Answering it reduces the scope of all possible things that can happen in an OST

**Figure 6.1. Elements of program quality**



program to the patterns of interaction and practice most likely to support positive youth development.

Program offerings are the key microsystem in OST contexts where access to powerful developmental and learning experiences is available and thus the place where technology is optimized to measure quality. In our definition of where, offerings consist of the same staff, same youth, and same learning purpose delivered over a sequence of sessions. For example, a mural painting workshop that meets at a residential treatment center every Tuesday and Friday is a program offering.

### *The language barrier*

Program offerings are the common context in which quality is enacted. However, the terminology used to define quality varies across places, networks, and programs. Many different social science languages include ideas related to positive youth development and are used to support policies and programs with parallel intentions but different names. Indeed, programs variously employ and mix the languages of positive youth development, resiliency theory, asset building, social-emotional learning, brain-based instruction, constructivist education, community service, and others. These approaches are all designed to optimize youth motivation and learning and include variations of several basic elements: emotional and active learning supports, cooperative learning methods, and higher-order cognitive engagement.

These competing languages about positive youth development reinforce policy fragmentation and can serve as barriers to cross-sector and cross-program improvement initiatives. The key to overcoming these barriers is to conceptualize quality in a way that can span these language divides. Defining quality at the point of service in terms of distinct elements of content neutral staff practices offers OST professionals a straightforward translation guide across different languages and program sectors. Table 6.1 describes a diverse profile of programs in twenty cities, counties, and states that planned to apply the quality construct described in Figure 6.1 to create place-based accountability and improvement systems as part of the Ready by

**Table 6.1. Diversity of program types in a place-based project using the Youth PQA**

	<i>Percentage (N = 124)</i>
Program’s primary purpose: What one or two of the following best represent your program’s primary approach to working with young people?	
Thriving	21.8
Learning	40.3
Connecting	74.2
Working	16.1
Leading	33.1
Age level served: Which age level does your program primarily serve?	
Elementary	64.5
Middle school	68.5
High school	62.9
Post-high school	24.2
Organization type: Which type best represents your organization?	
Local nonprofit	57.3
Nationally affiliated nonprofit	25.0
Unit of city or county government	14.5
Local school organization	10.5
Program model: Which type best describes your program model?	
After-school (mix)	75.8
After-school (sports)	4.8
School-day insert	21.0
Residential	4.0
Court referred/juvenile justices	8.9
Alternative school	2.4

*Note:* The data presented were generated using a survey tool developed by the High/Scope Educational Research Foundation and administered in October 2008 to 124 sites in twenty cities, counties, and states participating in a national project. Columns total to more than 100 percent because respondents could select all that applied.

21 Quality Counts Initiative.<sup>7</sup> A content-neutral definition of quality provides a common framework through which these diverse programs can conceptualize and create system-level quality improvement systems.

Taken together, our definitions of what quality is and where it can be found form the integral elements of a larger point-of-service theory.<sup>8</sup> This theory is grounded in the following assumptions:

- OST participation can provide youth with access to powerful development and learning experiences.
- Access to these key experiences occurs during content-driven offerings defined by continuity of staff, youth, and purpose.
- Performances of staff that lead offerings are the key units of quality in OST settings.
- High-quality performances represent consistent approaches to youth work and instruction over time within individual staff and across different staff within the same organization.
- Point-of-service performances cannot be sustainably improved without explicit attention to and support from the organizations and systems in which point-of-service interactions are embedded.

---

### *Assessing quality across diverse programs and purposes*

In this section, we discuss a widely used quality assessment tool—the Youth Program Quality Assessment (PQA)—focusing on several methodologies for its use and the distinct accountability purposes that these different methods serve. Note that although we focus on a specific instrument, we believe that the logic of alignment between methods and purposes applies across an entire class of observational tools that measure the quality of settings where children and youth spend their time.

#### *Youth Program Quality Assessment*

Point-of-service theory defines and locates quality within OST staff practices that occur within program offerings. As noted above, however, the terms used to define and manage positive youth development can differ substantially across places, networks, and programs. Although these distinct languages reflect substantially similar ideas, in practice they serve as barriers to cross-sector and cross-program improvement efforts. The key to effectively measuring quality in this fragmented environment is to define quality as a common, credible unit of study across programs and offerings.



The Youth Program Quality Assessment (PQA) is an observational instrument designed to assess the quality of staff performances during OST offerings.<sup>9</sup> The instrument requires an observer to document practices employed by staff and then use this written evidence to score multi-item rubrics, constituting the scales represented in Figure 6.1. Relative to other observational approaches, the PQA is a lower-inference observational metric since its items reflect counts of staff instructional behaviors. Youth PQA items are scored using a three-point scale: the lowest level indicates the absence of a practice, the midlevel indicates that a practice occurred but not all children had access to it, and the highest level for each item indicates that the practice was present and all youth participated. On average, the PQA requires an observer to collect objective, anecdotal evidence for one hour and one additional hour to score the instrument's multi-item rubrics.

Several samples of Youth PQA data produced by reliable external raters have been subjected to psychometric evaluation. The evidence from these evaluations indicates that data generated by the instrument are both reliable (interrater and internal consistency) and valid (concurrent and predictive validity) through association with multiple other sources of observational and youth survey data.<sup>10</sup> One of the key strengths of the Youth PQA is that it produces data that is clearly indexed to staff practice.<sup>11</sup> Another is the utility or consequential validity of the data it provides for end users. That is, the Youth PQA and its methods of use help youth workers, program managers, and other stakeholders develop common understandings of program quality and help build momentum for improvement initiatives. Indeed, Youth PQA users report that the process of using the instrument and reviewing the data it produces leads to program change.<sup>12</sup>

### *Continuum of methods and purposes*

The Youth PQA can be employed using a continuum of methodologies.<sup>13</sup> As depicted in Table 6.2, different methods of data collection serve different purposes and carry different strengths and

**Table 6.2. Continuum of methods, purposes, and resources for quality assessment**

<i>Lower-stakes program self-assessment</i>	<i>Hybrid approaches</i>	<i>Higher-stakes external assessment</i>
Site-based self-assessment teams.	Trained, reliable assessors recruit site-based self-assessment teams to coproduce quality scores.	Trained, reliable assessors not connected to the program.
Rough data to get staff thinking and discussing program quality in the context of best practice.	Rough and precise data comingled to produce more accurate scores than self-assessment alone.	Precise data for internal and external audiences for evaluation, monitoring, accountability, improvement, reporting.
Less time, less cost.	Process supports more specific planning and staff development but not appropriate for evaluation or accountability.	More time, more cost.
Impact internal audiences.	Most expensive, potentially highest learning impact.	Impact on internal and external audiences.
	Impact on internal audiences.	

limitations. It is worth reiterating that the Youth PQA's primary application has been in the context of quality accountability and improvement policies for OST systems. By definition, accountability and improvement policies are designed to measure performance on some set of behavioral norms and create incentives to improve performances on those norms. However, such policies do not always have the intended effect. In fact, some quality assessment technologies may paradoxically push staff into counterproductive accountability behaviors—resistance, avoidance, and minimum compliance—that deliver little return on investment.<sup>14</sup> In this article, our analysis of the various methods of deploying the Youth PQA technology is designed to surface the strengths and weaknesses of each method and highlight the impact of each method on improvement behaviors. Note that case studies describ-

ing the fit between external and self-assessment technologies and system purposes have been described in greater detail elsewhere.<sup>15</sup>

In the external assessment data collection method, a single Youth PQA is scored for a specific offering. Individual ratings can be aggregated to represent the quality of an overall program, usually using a sample drawn from all offerings at a site. External assessors are trained to acceptable levels of reliability to ensure precision, a process that takes up to two and a half days. Scores produced through the external assessment process can be a source of performance feedback to staff but can also serve higher-stakes purposes such as public release, program monitoring, and accountability for change over time or attainment of norms. In prior studies, we have sampled either two or three unique offerings with different staff and different learning purposes to produce a program-level quality score.

The primary limitation of the external assessment method is expense associated with data collection and scoring. It takes an average of two hours per observational rating plus logistics and travel. However, when collected by data collectors of known reliability, observational data can have powerful effects on site management perceptions of staff skills, suggesting that the additional cost of external observation may serve as a down payment toward more effective interventions to increase workforce skills, program quality, and ultimately youth development and learning.<sup>16</sup>

The program self-assessment data collection method is team based and designed to build learning community and acceptance for quality performance metrics at an OST site. To implement the self-assessment method, a site manager attends training (six hours of content, delivered live or online) and assembles a program assessment and improvement team, typically consisting of site staff but also frequently including volunteers, board members, the school principal, and parents. Individuals from this team take turns observing in fifteen- and twenty-minute increments during a selection of different program offerings (led by different staff) at a site. Then all anecdotal records are anonymously pooled and used to score one Youth PQA Form A for the entire site.

This method has several strengths. First, the program self-assessment process has been roundly endorsed by site managers as a way to build communities of learners around program operation and understanding of quality.<sup>17</sup> For example, in a recent survey of Michigan-based 21st Century Community Learning Center staff, 96.7 percent of respondents anonymously reported that using the Youth PQA as a state-mandated self-assessment “was a good use of time given the demands of the job” ( $N = 145$ ). Second, we believe that the program self-assessment process prepares teams to engage with and adopt the external assessment process in more productive ways.<sup>18</sup> This is so because program self-assessment builds familiarity among the staff with the metric that will eventually be used and begins a learning process that staff find valuable, increasing demand for more accurate data provided by external methods.<sup>19</sup>

From a research perspective, data produced in the program self-assessment process are limited in two primary ways:

Data are generally biased in a positive direction; that is, scores produced by self-reports are likely to be inflated relative to those that would be produced by external raters.

The flexibility of the process means that self-assessment does not produce a consistent, definable unit of study, that is, a single staff performance during a defined offering.

Consequently, it is not possible to make apples-to-apples comparisons between programs.

These limitations mean that program self-assessment data are not appropriate for comparison or aggregation. Use of such data is thus limited to immediate interpretation by the site team for purposes of building common understanding around the content of the rubrics.

Hybrid methods are also possible, as described in Table 6.2. This table represents one of the numerous ways to combine external and self-assessment to meet local objectives and conditions. Several years ago, the Maine 21st Century Community Learning Centers program introduced an innovative accountability model that produced both more rigorous baseline scores

than program self-assessment alone and built a learning community among management staff by training site managers to high levels of reliability and using them as external assessors in other 21st Century Community Learning Centers grantees.<sup>20</sup> More recently, the Rhode Island 21st Century Community Learning Centers program trained its quality improvement coaches as external raters. These coaches then recruited frontline staff to join in both the process of observational evidence gathering and scoring the Youth PQA.<sup>21</sup> Again, these methods were adopted to achieve specific purposes that balanced within-site learning and measurement precision.

### *System-level implications of quality assessment*

So far, we have argued that point-of-service theory, particularly as embodied in the Youth PQA, is a useful lens through which to measure and manage the quality of program offerings across the diverse range of missions, funding, and regulatory streams and client types that characterize the OST field. We have also suggested a rough alignment between quality assessment technologies and accountability purposes:

External assessment fits into an accountability context where scores are made available for various types of comparison across settings and over time and where external ratings are required through program rules.

Program self-assessment fits into an accountability context where the implementation of the team-based self-assessment process, whether voluntary or mandated, is enacted for purposes of building a learning community around the content of the quality metrics.

At the heart of this alignment argument is the idea that innovative accountability policies can be developed in ways that address both POS performances (staff accountability for practices) and organizational efforts to support workforce development and learning (system-level accountability for creating and sustaining professional learning communities).

Clearly the use of quality metrics at the point of service can have an impact on management practices and organizational policies. By finding moments of comparability across the great diversity of OST programs, it is possible to generate system-level outcomes such as greater coordination of resources across organizations and more system-like approaches to workforce development. The Iowa Collaboration for Youth Development (ICYD), for example, conducted a “quality snapshot” by collecting performance samples in thirty-eight organizations representing thirteen different youth funding streams in the state of Iowa. The snapshot was explicitly introduced in order to “launch a program quality improvement system across youth-serving programs funded by a variety of funding streams.” Based on findings from the report, ICYD recommended that “state agencies should continue to explore how best to meet the capacity needs of their grantees, and, when feasible, other youth-serving agencies not receiving state funding, as a way to improve *interactions* and *engagement* of youth [two low-scoring domains of quality from the snapshot report] . . .”<sup>22</sup> This suggests that the deployment of a uniform quality metric across the siloed education, human service, juvenile justice, and community-based program settings that are present in most cities and counties can initiate movement toward shared vision, planning, and resource allocation at higher levels of decision making. The Iowa example provides further suggestive evidence that accountability and improvement technologies like the Youth PQA can empower intermediary organizations to facilitate this kind of movement toward greater collaboration across agencies and across geography.

---

### ***Implementation advice***

There is ample evidence that a general definition of point-of-service quality can be relevant and meaningful across a range of OST settings and that observational assessment instruments like the Youth PQA represent performance assessment technologies that can provide a foundation for accountability and improvement

policies. We now offer some more general advice based on our experience as training and technical assistance providers to nearly thirty active accountability and improvement projects.

### *Purposes*

Purposes should drive decisions about how to use performance metrics like the Youth PQA. In fact, in our experience, mismatches between purpose and methods are likely to lead to implementation failure. Consequently it is imperative that systems and organizations that wish to deploy a quality improvement system clearly define their objectives (for example, high-stakes comparisons across time and settings versus low-stakes trust building and professional development) and then select a method or methods of use.

### *Combined methods*

Although there are relative strengths and weaknesses of the program self- and external assessment methods when deployed separately, our experience suggests that there are great benefits to deploying them sequentially as a means of building professional learning community and achieving buy-in for future higher-stakes accountability and improvement policies. In fact, the most successful implementations of the Youth PQA have used program self- and external observational assessments in ways that spur formation of a team culture and generate demand for more objective external assessment of staff performance against quality standards.<sup>23</sup>

This last point is critical. If systems and organizations do not provide proper incentives to managers to prioritize and pursue point-of-service quality initiatives, these efforts are likely to fade or never gain momentum. In OST settings (as in other education and human services organizations), managers set the priorities and make the decisions that influence the adoption and use of quality assessment and improvement practices. In Palm Beach County, Florida, for example, qualitative and quantitative data collected as part of a quality improvement system pilot suggest that low-stakes

deployment of program quality technology, coupled with substantial technical support from an intermediary organization such as Prime Time Palm Beach County, can produce measurable changes in point-of-service quality and management practices, as well as streamline adoption of accountability policies.<sup>24</sup>

### *Pilots and supports*

Sequenced deployment of assessment methods can build trust throughout stakeholders in a network and set the stage for higher-stakes accountability and improvement initiatives. Our experience suggests an additional strategy for ensuring the successful deployment of quality improvement initiatives using the Youth PQA or similar tools: begin with a small-scale pilot demonstration. Demonstration pilots offer systems the opportunity to adjust technology to context, demonstrate applicability across disparate program types, and align training and technical assistance that supports improvement. Furthermore, pilots allow word-of-mouth messages that the quality accountability and improvement policy works to spread, and they support the emergence of success stories and local champions for the work. Nearly all of our place-based and statewide projects have begun with pilot groups of ten to thirty programs implementing the Youth PQA and various packages of supporting training and technical assistance over six to eighteen months.

---

### *Conclusion*

The Youth Program Quality Assessment is one of an emerging class of observational assessment tools that measure staff performances at the point of service and, depending on the methodology of use, help create the conditions that managers and youth workers need to accept, adopt, and sustain quality accountability and improvement initiatives. Observational assessment tools are flexible enough to be used for program self-assessment (appropriate for lower-stakes, nonnormative learning purposes), external assessment



(appropriate for higher stakes, normative comparisons, and performance accountability), as well as various hybrids that combine elements from each. Innovative quality accountability and improvement policies are emerging in numerous places and statewide systems across the country, applying uniform standards of quality and professional performance at the point of service and increasing higher-level collaboration of funders, regulators, and site-level managers. These initiatives are often carried by intermediary organizations that can provide equal measures of disinterested objectivity and careful support for successful point-of-service change. We hope that the learning discussed in this article will contribute to the ongoing development and adoption of innovative quality accountability and improvement policies across the OST field.

### Notes

1. Data compiled by the National Child Care Information Center as of November 2006, cited in the 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. Smith, C., Akiva, T., Blazevski, J., Pelle, L., & Devaney, T. (2008). *Final report on the Palm Beach Quality Improvement System pilot: Model implementation and program quality improvement in 38 after-school programs*. Ypsilanti, MI: High/Scope Educational Research Foundation; 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*. New York: Oxford University Press.

4. Bronfenbrenner, U. (1999). Environments in developmental perspective: Theoretical and operational models. In S. L. Friedman & T. D. Wachs (Eds.), *Measuring environment across the life span: Emerging methods and concepts* (pp. 3–28). Bethesda, MD: National Institutes of Health, Child Health and Human Development Agency, Center for Research for Mothers and Children.

5. Ilfeld, E. M. (1996). *Learning comes to life: An active learning program for teens*. Ypsilanti, MI: High/Scope Press; Oden, S., Kelly, M., & Weikart, D. (1992). *Changing the potential: Programs for talented disadvantaged youth*. Ypsilanti, MI: High/Scope Press; Smith, C. (2005a). *Evidence of effectiveness for training in the High/Scope Participatory Learning approach*. Ypsilanti, MI: High/Scope Educational Research Foundation. For more information, see <http://www.highscope.org>.

6. Yohalem, N., & Wilson-Ahlstrom, A. (2007). *Measuring youth program quality: A guide to assessment tools*. Washington, DC: Forum for Youth Investment.

7. Data from this table were drawn from pilot sites participating in the Ready by 21, Quality Counts Initiative led by the Forum for Youth Investment and funded by the Robert Wood Johnson Foundation. For more information, see [www.forumforyouthinvestment.org](http://www.forumforyouthinvestment.org).

8. Smith, C., Peck, S., Denault, A., Blazeovski, J., & Akiva, T. (in review). *Quality at the point-of-service: Profiles of practice in afterschool settings*. Manuscript submitted for publication, January 8, 2009.

9. High/Scope Educational Research Foundation. (2005). *Youth PQA program quality assessment: Administration manual*. Ypsilanti, MI: High/Scope Press.; Smith, C., & Hohmann, C. (2005). *Full findings from the Youth PQA validation study*. Ypsilanti, MI, High/Scope Educational Research Foundation.

10. Smith & Hohmann. (2005); Blazeovski, J., & Smith, C. (2007a). *Inter-rater reliability on the youth program quality assessment*. Ypsilanti, MI: High/Scope Educational Research Foundation; Blazeovski, J., & Smith, C. (2007b). *After-school quality and school-day outcomes in Michigan's 21st CCLC program*. Ypsilanti, MI: High/Scope Educational Research Foundation; Smith et al. (2008).

11. A common challenge among observational measures is that all of the scales tend to be highly correlated, resulting in an inability to assess any but the most global dimensions of quality. We have tested theoretically relevant subsets of items on the Youth PQA to achieve dimensionality of measurement in replicated confirmatory analyses, suggesting that the measure can be used to identify different aspects of staff performance quality in OST offerings. For more information, see La Paro, K. M., Pianta, R. C., & Stuhlman, M. (2004). Classroom assessment scoring system (CLASS): Findings from the prekindergarten year. *Elementary School Journal*, 104(5), 409–425; Perlman, M., & Zellman, G. L. (2004). Examining the psychometric properties of the Early Childhood Environment Rating Scale-Revised (ECERS-R). *Early Childhood Research Quarterly*, 19, 398–412. Also see Smith et al. (in review).

12. Smith & Hohmann. (2005); Spielberg, J., & Lockaby, T. (2008). *Palm Beach County's Prime Time Initiative: Improving the quality of after-school programs*. Chicago: Chapin Hall Center for Children at the University of Chicago.

13. High/Scope Educational Research Foundation. (2005).

14. Smith & Akiva. (2008).

15. Smith, C., Akiva, T., Arrieux, D., & Jones, M. (2006). Improving quality at the point-of-service. In D. A. Blythe & J. A. Walker (Eds.), *Rethinking programs for youth in the middle years*. San Francisco: Jossey-Bass; Smith, C. (2005b). *Findings from the self-assessment pilot in Michigan 21st Century Learning Centers*. Ypsilanti, MI: High/Scope Educational Research Foundation.

16. Spielberg, J., & Lockaby, T. (2006). *The Prime Time Initiative of Palm Beach County, Florida—QIS development process evaluation: Year 2 report*. Chicago: Chapin Hall Center for Children at the University of Chicago; Spielberg & Lockaby. (2008).

17. Smith. (2005b).

18. Smith et al. (2008).

19. Spielberg & Lockaby. (2006). Spielberg & Lockaby. (2008).

20. Smith & Akiva. (2008).

21. Sugar, S., Devaney, T., & Smith, C. (2008). *Results from the RIPQA quality improvement system: Quality standards implementation in 19 after-school programs*. Ypsilanti, MI: Center for Youth Program Quality.

22. Sugar, S., Akiva, T., Wilson-Ahlstrom, A., Behrer, C., & Croll, A. (2008). *The Iowa Youth Program quality snapshot*. Ypsilanti, MI: Center for Youth Program Quality.

23. Smith et al. (2008); Spielberger & Lockaby. (2008).

24. Smith et al. (2008).

CHARLES SMITH *is the director of the David P. Weikart Center for Youth Program Quality.*

THOMAS J. DEVANEY *is the operations and knowledge manager at the David P. Weikart Center for Youth Program Quality.*

TOM AKIVA *is a doctoral student in the Combined Program in Education and Psychology at the University of Michigan and a consultant with the David P. Weikart Center for Youth Program Quality.*

SAMANTHA A. SUGAR *is a research assistant at the David P. Weikart Center for Youth Program Quality.*