

# **CYPQ Report**

Submitted to The Providence After-School Alliance  
Providence, Rhode Island  
July 2008

## **Results from the RIPQA Quality Improvement System Quality Standards Implementation in 19 After-School Programs**

Samantha Sugar  
ThomasDevaney  
Charles Smith

**The Center for Youth Program Quality (CYPQ)**

A Joint Venture of the High/Scope Educational Research Foundation and Forum for Youth Investment

# Table of Contents

## *Content*

Part I: Summary & Methods.....	Page 3
A. Summary.....	Page 3
B. Background & Methods.....	Page 3
Part II: Key Findings.....	Page 5
A. Point of Service Quality.....	Page 5
Appendix A: External vs. Self Assessment Benchmarks.....	Page 9
A. National 21 <sup>st</sup> CCLC Peer Comparisons (External Assessments) .....	Page 9
B. Michigan 21 <sup>st</sup> CCLC Peer Comparisons (Self-Assessments) .....	Page 10

## *Figures & Tables*

Figure 1: Pyramid of Youth Program Quality.....	Page 4
Table 1: 2007-08 YPQA Domain and Subscale Scores (Form A) .....	Page 6
Table 2: 20 Lowest Scoring Items (Form A) .....	Page 7
Figure 2: Inter-Network Comparison (Form A) .....	Page 8
Figure 3: Quality Scores by Content Area (Form A) .....	Page 8
Table 3: RI Scale Scores vs. 21st CCLC External Assessment Benchmarks (Form A) .....	Page 9
Figure 4: RI Domain Scores vs. 21st CCLC External Assessment Benchmarks (Form A) .....	Page 10
Table 4: RI Scale Scores vs. MI 21st CCLC Self Assessment Benchmarks (Form A) .....	Page 11
Figure 5: RI Domain Scores vs. MI 21st CCLC Self Assessment Benchmarks (Form A) .....	Page 11

# **Part I: Summary & Methods**

## **A. Summary**

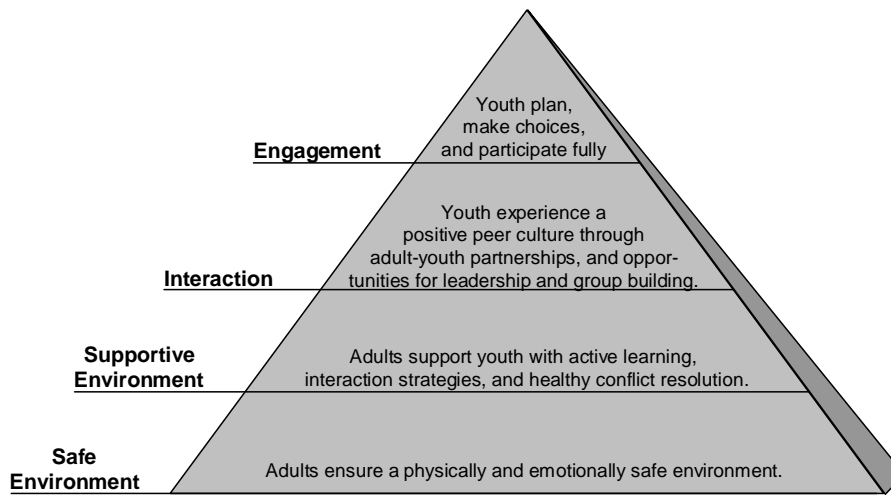
This report summarizes findings from the year 1 quality improvement system (QIS) in Rhode Island 21<sup>st</sup> Century Community Learning Centers as well as Providence After School programs. Key findings from this initial wave of data collection include:

- The RI QIS system successfully completed a pilot baseline data collection using an innovative self-assessment method, producing quality ratings for 77 activities in 19 organizations. Data was collected from the end of November 2007 to the end of June 2008. Two different networks were included in the data collection: Rhode Island 21<sup>st</sup> CCLC grantee sites and Providence After School Alliance (PASA) sites.
- Consistent with findings from other Youth Program Quality Assessment (Youth PQA) evaluations that we have conducted, sites assessed for this evaluation score highest in the safety and supportive environment domains and lowest in the interaction and engagement domains. In fact, 17 of the 20 lowest scoring items fall in the latter categories.
- Rhode Island's QIS baseline data was collected using a non-standard Youth PQA data collection methodology, pairing external and self-assessors to produce RIPQA quality scores. While this blended approach is highly recommended for lower stakes learning purposes, it is also likely to introduce some positive bias into scores.

## **B. Background & Methods**

Findings were produced using Form A of the Rhode Island Youth Program Quality Assessment (RIPQA). Form A of the RIPQA is a duplicate of High/Scope's Youth PQA, a research validated observational assessment tool (Smith & Hohmann, 2005) currently used in afterschool QIS models in several states. Figure 1 presents the contents of the PQA Form A. A rating for each of the four domains represented on the pyramid – (1) Safe Environment, (2) Supportive Environment, (3) Interaction and (4) Engagement -- was produced for each of 77 program offerings assessed in the 19 participating after-school sites. Data presented in the report are aggregated to the site-level unless otherwise noted.

**Figure 1: Pyramid of Youth Program Quality**



Form A of the RIPQA is an observational assessment tool designed to measure the quality of staff performances at the point of service in after-school programs. Specifically, the tool is intended to assess how frequently and how equitably key developmental and learning experiences are provided to youth in evaluated programs. In general, scores produced by the RIPQA may be interpreted as follows:

- A score of 1 means that the key experience was not delivered during the observation.
- A score of 3 means that the key experience was delivered informally or to only a portion of children/youth in the room.
- A score of 5 indicates that the key experience was delivered intentionally by the lead staff and that all children/youth in the space had access to it.

Prior research suggests that, in the aggregate, scores on Form A items reflect consistent characteristics of staff performance. That is, a score produced on an item during a single observation is more likely than not to capture how staff generally enact (or fail to enact) that practice across multiple sessions of the same offering.

The standard methodology used to generate Youth PQA ratings involves systematic observation by one or more trained observers during a selected program offering. Systematic observation requires observers to keep a written record of staff performances and student interactions during the observation period and then score Form A of the RIPQA (which includes 61 items or rubrics) using this written record. Note that a “program offering” is defined as one session in a multi-session sequence with the following characteristics: (1) A consistent learning purpose; and (2) the same or substantially similar group of staff and youth. Examples include traditional academic offerings such as math club, enrichment offerings such as “hip-hop dance” or computer graphics or blended academic and enrichment offerings such as “science in the kitchen.”

The specific methodology used to generate quality scores provided in this report relied on a unique blend of external- and self-assessment. Rhode Island’s process included the following steps:

- Trained external assessors with demonstrated reliability on the RIPQA Form A visited program offerings, collected observational data and scored rubrics using the methods outlined in the prior paragraph
- During each external rating, site staff shadowed the external assessors and collected a parallel set of performance data using Form A.
- The internal and external raters met to discuss the collected data and produce consensus quality scores.

Because this methodology blended the use of reliable external raters with site staff as self-assessors, we expect some positive bias in the scores. It is likely that this data collection design provided a powerful learning device for staff and facilitated institutionalization of ideas about program quality. *However, because systematic bias may be present due to the inclusion of self-assessors, this baseline quality profile is not appropriate for use in high stakes comparisons to quality scores across programs or over time.*

To provide some perspective on the efficacy of this blended methodology, we compared scores from the RIPQA to other 21<sup>st</sup> CCLC samples using either external assessment or self-assessment. As expected, these baseline scores fell between these two sets of scores. That is, they were slightly higher than external ratings and substantially lower than pure self-assessment (See Appendix A). We also examined the internal consistency of item scores within RIPQA Form A scales. These consistency statistics fall into an acceptable range (see Table 1), suggesting that the blended methodology produces reasonably reliable scores for the measured constructs.

With these caveats in mind, we offer that this blended methodology may: (1) provide strong support to staff learning; and (2) combine the accuracy of trained external assessors with the learning power of direct participation in the assessment process by site staff. However, because external assessors lead the process and because multiple observations were collected at each after-school site, it is also likely that these data are more accurate than those produced through a pure-self assessment process. *Ultimately, the process should be considered a very accurate form of self-assessment.*

## **Part II. Key Findings**

### **A. Point of Service Quality**

In this section we evaluate the RIPQA baseline quality data at both the aggregate, network and offering levels.

Figure 2 presents mean domain and subscale scores for Form A observational data collected from 19 sites (n=77 observations) in 2007 and 2008. Consistent with PQA evaluations conducted with similar after-school networks elsewhere in the country, mean quality scores are higher for the Safe and Supportive Environment domains than for the Interaction and Engagement domains. This ubiquitous performance gap between lower level (e.g., providing a safe and supportive environment for youth) and higher level professional practices (e.g., interacting with youth in ways that promote positive development and support deep engagement with content) draws into sharp relief several challenges facing efforts to improve the youth work field. Specifically:

- Many youth workers lack the training and skills required to interact with and engage youth in ways known to support positive youth development;

- After-school program designs and management practices limit practitioners' ability and/or willingness to enact practices aligned with youth development best practices
- After-school programs lack the measurement and planning infrastructure required to identify quality problems, intentionally adjust point of service practices and track the impact of enacted changes.

Although the quality gap between lower- and higher- level professional practices in Rhode Island's QIS may seem troubling, PQA time-series data collected elsewhere (e.g., point of service quality data collected at both the beginning and end of an improvement initiative) suggest that coordinated training, technical assistance and measurement based on the Youth PQA framework, coupled with a focused improvement initiative, can remediate this gap. In other words, with an effective set of integrated supports, continued use of the RIPQA is likely to re-orient management and point of service skill sets and practices in ways that increase access to key developmental experiences.

Note that, as stated above, the unique data collection methodology employed at the RIPQA baseline produced subscale scores with high levels of internal consistency (Cronbach's Alpha statistics of .596 and higher for all subscales). This suggests that Rhode Island's blend of external and self-assessors did not upset critical characteristics of consistency in the RIPQA data even though a standard external assessment method was not employed.

**Table 1: 2007-08 YPQA Domain and Subscale Scores (Form A)**

Domains and Scales	Mean Score (N=19 sites aggregated from 77 observations)	Cronbach's Alpha $\alpha^1$
<b>I. Safe Environment</b>	4.20	
IA. Psychological and emotional safety is promoted	4.52	.674
IB. The physical environment is safe and free of health hazards	4.74	.581
IC. Appropriate emergency procedures and supplies are present	4.47	.874
ID. Program space and furniture accommodate the activities offered	4.52	.761
IE. Healthy food and drinks are provided	3.75	.596
<b>II. Supportive Environment</b>	3.98	
IIF. Staff provide a welcoming atmosphere	4.41	.788
IIG. Session flow is planned, presented, and paced for youth	4.42	.707
IIH. Activities support active engagement	3.85	.761
III. Staff support youth in building new skills	3.79	.777
IIJ. Staff support youth with encouragement	3.49	.632
IIK. Staff use youth-centered approaches to reframe conflict	3.84	.865
<b>III. Interaction</b>	3.26	
IIIL. Youth have opportunities to develop a sense of belonging	3.65	.739
IIIM. Youth have opportunities to participate in small groups	3.02	.811
IIIN. Youth have opportunities to act as group facilitators and mentors	2.73	.631
IIIO. Youth have opportunities to partner with adults	3.63	.600
<b>IV. Engagement</b>	2.73	
IVP. Youth have opportunities to set goals and make plans	2.53	.894
IVQ. Youth have opportunities to make choices based on their interests	3.03	.813
IVR. Youth have opportunities to reflect	2.65	.802
<b>TOTAL</b>	<b>3.55</b>	

1: Cronbach's  $\alpha$  (alpha) is a statistic that measures the reliability of a psychometric instrument or construct.

Table 2 presents the 20 lowest scoring items (based on mean scores) from the PQA data collected in 2007-08. Consistent with the data summarized in Table 1, 17 of the 20 lowest scoring items come from the Interaction and Engagement domains. This reinforces our earlier contention that youth workers within the RI QIS, as with similar networks through the country, often lack the organizational supports or professional skills necessary to consistently deliver higher order learning experiences to their constituents.

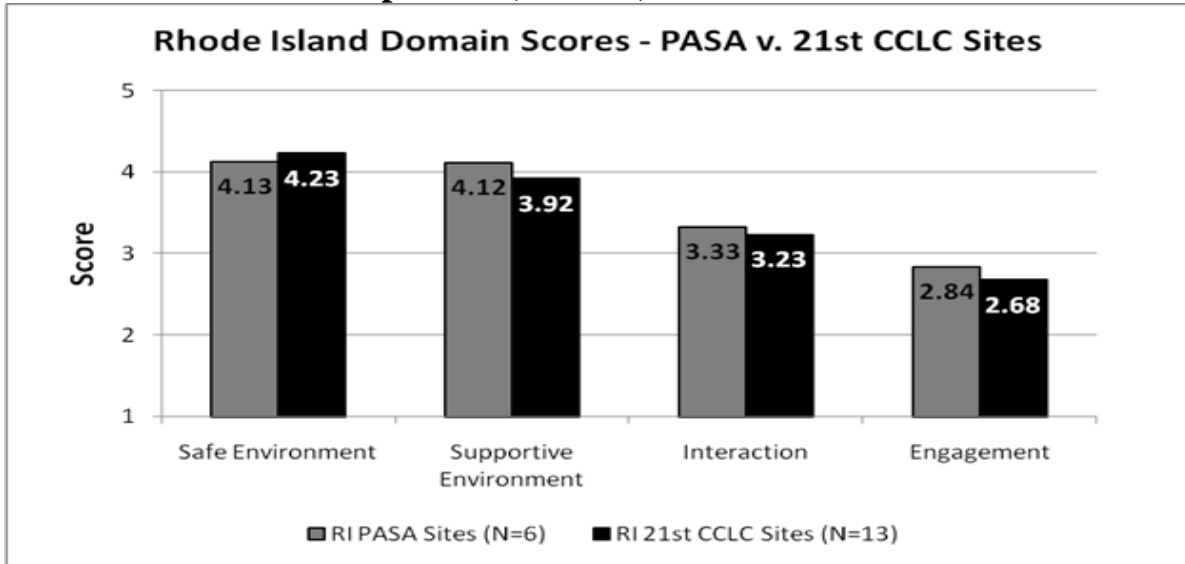
**Table 2: 20 Lowest Scoring Items (Form A)**

#	Item	Percent of offerings scoring a 1
1	RI.IIIN.3 – Youth have one or more opportunities to lead a group	66.7
2	RI.IVR.1 – Youth reflect on what they are doing	56.8
3	RI.IVR.3 – Youth make presentations	51.4
4	RI.IVP.1 – Youth have multiple opportunities to make plans	50.7
5	RI.IVP.2 – Two or more planning strategies are used	50.7
6	RI.IVR.2 – Youth reflect in multiple ways	50.0
7	RI.IIIN.2 – Youth have one or more opportunities to mentor an individual	42.7
8	RI.IIK.2 – Staff seek input from youth to determine cause and solution of conflicts	38.1
9	RI.IIJ.3 – Staff make frequent use of open-ended questions	37.7
9	RI.IIIM.2 – Staff use 2 or more ways to form small groups	37.7
11	RI.IIIM.1 – Activities carried out in at least 3 different groupings	36.0
11	RI.IVQ.1 – All youth make at least one open-ended content choice	36.0
13	RI.IVR.4 – Youth give feedback on activities	31.5
14	RI.IIIL.4 – Activities include structured opportunities to publicly acknowledge the achievements of youth	30.7
15	RI.IVQ.2 – All youth make at least one open-ended process choice	29.3
16	RI.IIK.3 – Staff encourage youth to examine the relationship between actions and consequences	28.6
17	RI.IIK.4 – Staff acknowledge conflicts and negative behavior and follow-up afterward	27.3
18	RI.IIIO.1 – Staff share control with youth	25.3
19	RI.IIIN.1 – Youth have multiple opportunities to practice group process skills	23.7
20	RI.IIIM.3 – Each small group has a purpose	23.4

In addition to aggregate level comparisons, it is possible to compare quality between the two networks involved in the Rhode Island QIS: the Rhode Island 21<sup>st</sup> Century Community Learning Centers as well as Providence After School programs. Note, however, that due to low sample sizes (n=13 CCLC sites and n=6 PASA sites) these comparisons must be interpreted cautiously.

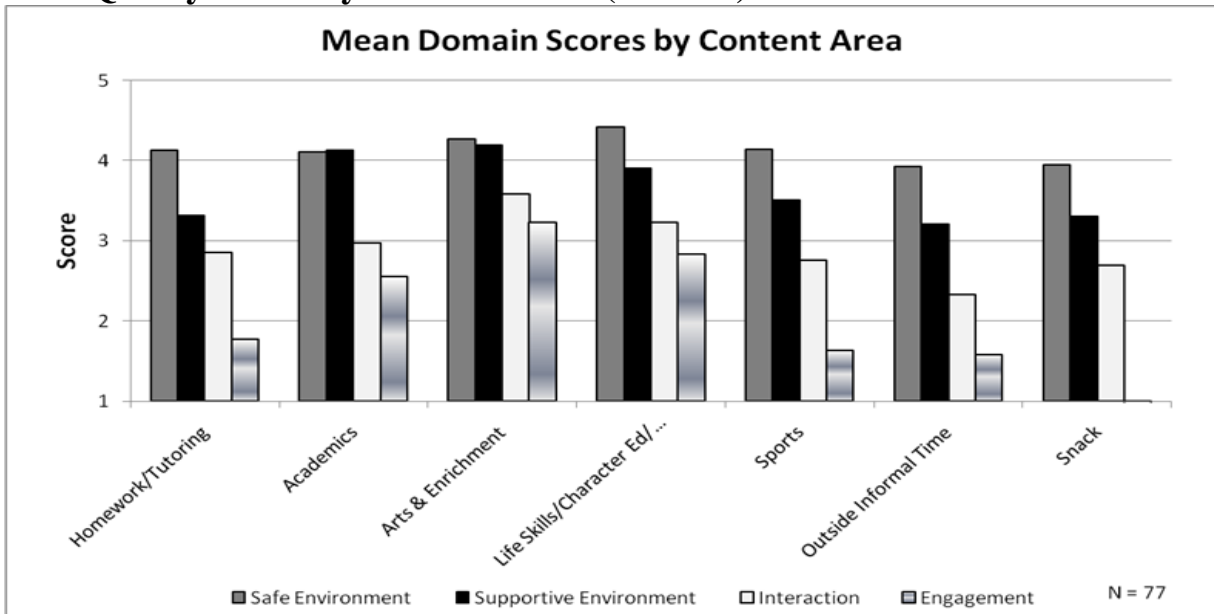
Figure 2 compares network level performance across the 4 domains of the RIPQA Form A. The PASA network has slightly higher scores than their 21<sup>st</sup> CCLC peers on 3 of the 4 RIPQA quality domains: Supportive Environment, Interaction and Engagement. The 21<sup>st</sup> CCLC network slightly outperforms PASA on the Safe Environment domain. However, in all of our prior work **age of youth** served is positively related to quality scores on Form A, leading to lower score profiles for 21<sup>st</sup> Century programs which are often located in elementary school settings.

**Figure 2: Inter-Network Comparison (Form A)**



Finally, we assess quality variation at the content offering level. That is, we analyze the RIPQA data to determine if there are systematic and meaningful quality differences between offering types. Figure 3 compares domain-level quality scores across 7 content areas: Homework/Tutoring (n=5), Academics (n=19), Arts & Enrichment (n=28), Life Skills/Character Educ/Health (n=8), Sports (n=13), Outside Informal Time (n=2) and Snack (n=2). Following patterns observed in other larger samples, the higher quality offerings fall into the arts enrichment and academic enrichment content areas while lowest scoring profiles are present in the content areas such as homework, outside/informal time and sport. Importantly, in these last content areas the quality construct represented by the RIPQA may not be an appropriate goal for staff performance.

**Figure 3: Quality Scores by Content Area\* (Form A)**



\* Data not aggregated to the site level.

# Appendix A. External vs. Self- Assessment Benchmark Data

## A. National 21<sup>st</sup> CCLC Peer Comparisons (External Assessments)

As discussed in the introduction to this report, the data collection methodology used for RI’s QIS baseline relied on a mixture of self- and external assessment. In this section we compare Rhode Island’s quality scores to normative data from two sources: (1) a national database of external Youth PQA assessment scores collected from 21<sup>st</sup> CCLC sites; and (2) self-assessment Youth PQA scores from 21<sup>st</sup> CCLC sites funded through the Michigan Department of Education. These data provide peer benchmarks against which Rhode Island’s 21<sup>st</sup> CCLC and PASA networks can:

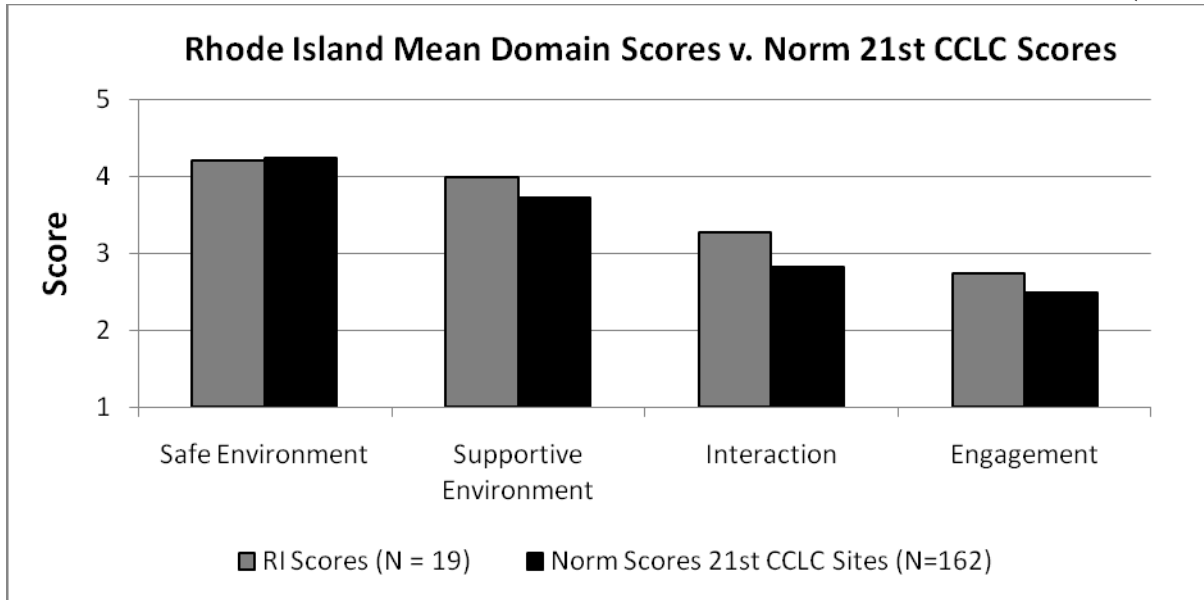
- Assess the overall quality of the programming they provide to youth
- Illuminate the impact of the modified data collection methodology on score levels.

Table 3 and Figure 4 compare mean quality scores for the 19 participating sites in Rhode Island to external Youth PQA scores from a national database of 21<sup>st</sup> CCLC programs. With only a couple of exceptions at the scale level and one exception at the domain level, Rhode Island’s scores *are higher than their peer norms*. We interpret this to mean that (1) the modified data collection methodology produced somewhat higher quality scores than could be expected through external observation only; but also that (2) on average, Rhode Island’s 21<sup>st</sup> CCLC and PASA programs are delivering point of service quality for most PQA domains and scales at least at the level of national benchmarks.

**Table 3: RI Scale Scores vs. 21<sup>st</sup> CCLC External Assessment Benchmarks (Form A)**

Youth PQA Scale Scores	RI Mean Score (N=19 sites aggregated from 77 observations)	Norm Score 21 <sup>st</sup> CCLC sites (N=162 observations)	Diff.
<b>I. Safe Environment</b>			
IA. Psychological and emotional safety is promoted	4.52	4.06	0.46
IB. The physical environment is safe and free of health hazards	4.74	4.65	0.09
IC. Appropriate emergency procedures and supplies are present	4.47	3.57	0.9
ID. Program space and furniture accommodate the activities offered	4.52	4.71	-0.19
IE. Healthy food and drinks are provided	3.75	4.15	-0.4
<b>II. Supportive Environment</b>			
IIF. Staff provide a welcoming atmosphere	4.41	4.02	0.39
IIG. Session flow is planned, presented, and paced for youth	4.42	4.34	0.08
IIH. Activities support active engagement	3.85	3.68	0.17
III. Staff support youth in building new skills	3.79	3.57	0.22
IIJ. Staff support youth with encouragement	3.49	3.43	0.06
IIK. Staff use youth-centered approaches to reframe conflict	3.84	2.96	0.88
<b>III. Interaction</b>			
IIIL. Youth have opportunities to develop a sense of belonging	3.65	3.24	0.41
IIIM. Youth have opportunities to participate in small groups	3.02	2.27	0.75
IIIN. Youth have opportunities to act as group facilitators and mentors	2.73	2.26	0.47
IIIO. Youth have opportunities to partner with adults	3.63	3.43	0.2
<b>IV. Engagement</b>			
IVP. Youth have opportunities to set goals and make plans	2.53	1.96	0.57
IVQ. Youth have opportunities to make choices based on their interests	3.03	2.91	0.12
IVR. Youth have opportunities to reflect	2.65	2.57	0.08

**Figure 4: RI Domain Scores vs. 21<sup>st</sup> CCLC External Assessment Benchmarks (Form A)**



**B. Michigan 21<sup>st</sup> CCLC Peer Comparisons (Self Assessments)**

Table 4 and Figure 5 compare mean quality scores for the 19 participating sites in Rhode Island to self-assessment Youth PQA scores produced by 21<sup>st</sup> CCLC programs in Michigan. Here we find a markedly different, but clearly complementary pattern of differences between the Rhode Island and Michigan data. With just a handful of exceptions at the scale level and no exceptions at the domain level, Rhode Island’s scores *are lower than the peer norms produced via a purely internal assessment process*. We interpret this to mean that the modified data collection methodology produced somewhat lower quality scores (*and probably more accurate scores*) than could be expected through self-assessment alone.

**Table 4: RI Scale Scores vs. MI 21<sup>st</sup> CCLC Self Assessment Benchmarks (Form A)**

Youth PQA Scale Scores	Mean Score (N=19 sites aggregated from 77 observations)	MDE 21 <sup>st</sup> CCLC Self- Assessment Scores (N=113)	Diff.
<b>I. Safe Environment</b>			
IA. Psychological and emotional safety is promoted	4.52	4.17	0.35
IB. The physical environment is safe and free of health hazards	4.74	4.52	0.22
IC. Appropriate emergency procedures and supplies are present	4.47	4.16	0.31
ID. Program space and furniture accommodate the activities offered	4.52	4.60	-0.08
IE. Healthy food and drinks are provided	3.75	4.49	-0.74
<b>II. Supportive Environment</b>			
IIF. Staff provide a welcoming atmosphere	4.41	4.34	0.07
IIG. Session flow is planned, presented, and paced for youth	4.42	4.40	0.02
IIH. Activities support active engagement	3.85	4.08	-0.23
III. Staff support youth in building new skills	3.79	4.45	-0.66
IIJ. Staff support youth with encouragement	3.49	4.02	-0.53
IIK. Staff use youth-centered approaches to reframe conflict	3.84	4.10	-0.26
<b>III. Interaction</b>			
IIIL. Youth have opportunities to develop a sense of belonging	3.65	3.72	-0.07
IIIM. Youth have opportunities to participate in small groups	3.02	3.73	-0.71
IIIN. Youth have opportunities to act as group facilitators and mentors	2.73	3.32	-0.59
IIIO. Youth have opportunities to partner with adults	3.63	3.81	-0.18
<b>IV. Engagement</b>			
IVP. Youth have opportunities to set goals and make plans	2.53	3.19	-0.66
IVQ. Youth have opportunities to make choices based on their interests	3.03	3.34	-0.31
IVR. Youth have opportunities to reflect	2.65	3.36	-0.71

**Figure 5: RI Domain Scores vs. MI 21<sup>st</sup> CCLC Self Assessment Benchmarks (Form A)**

