

Building Collective Capacity for Program Quality Improvement: Boston Beyond's Certified Observer Network

Emily Dodge

Boston Beyond

edodge@bostonbeyond.org

Abstract

This article describes Boston Beyond's effort to develop a network of out-of-school time program partner staff trained in implementing a program quality observation tool. Participant survey (n = 63) and interview (n = 4) feedback demonstrate that the network is meeting its goals of advancing participants' professional development, positively impacting organizations, and creating a system-level model of peer-to-peer program quality improvement. Areas of improvement are identified for each goal. Questions of sustainability, planned network improvements, and considerations for others seeking to establish similar networks are discussed.

Key words: observation, program quality, program improvement

Introduction

The out-of-school time environment has gained prominence in the last few decades due to its widely-acknowledged potential to positively impact crucial student academic and social-emotional outcomes (Durlak, Weissberg, & Pachan, 2010; Lauer et al., 2006). With the majority of United States youth attending some sort of out-of-school time program for at least 1 hour each week (Duffett & Johnson, 2004; Mahoney, Harris, & Eccles, 2006), there is a consensus that these programs must be measured for program quality and impact.

Program quality must be specifically defined in the out-of-school time setting because of context-specific factors such as high staff turnover and content changes (Shulman, 1986). Many



New articles in this journal are licensed under a Creative Commons Attribution 4.0 License. This journal is published by the [University Library System](#), [University of Pittsburgh](#) and is cosponsored by the [University of Pittsburgh Press](#). The Journal of Youth Development is the official peer-reviewed publication of the [National Association of Extension 4-H Youth Development Professionals](#) and the [National AfterSchool Association](#).

Building Collective Capacity

models have been developed to capture program quality. Common to each of these models is a primary focus on youth engagement and positive interactions between staff and youth (Eccles & Gootman, 2002; Larson, Rickman, Gibbons, & Walker, 2009; Pianta & Hamre, 2009).

Research has shown that improvements in out-of-school time program quality can be made when staff are trained in high-quality, developmentally-appropriate instructional practices; receive formative feedback via assessment tools; and are regularly coached by their more-experienced peers (Akiva & Yohalem, 2006; Hilberg, Waxman, & Tharp, 2003; Pianta, 2010; Smith & Akiva, 2008; Wilson-Ahlstrom & Yohalem, 2007). Boston Beyond is a nonprofit intermediary that promotes this sort of program quality improvement among its out-of-school time program partners through data collection, data analysis, and professional development. Programs choose among the suite of free measurement tools that capture program quality, youth experience, and youth skill development. The vast majority of the 275 programs in Boston Beyond's network receive a third-party observation using the Assessment of Program Practices Tool (APT).

In 2016, Boston Beyond developed the Certified Observer Network, a group of program partner staff trained to conduct APT observations of peer programs in Boston Beyond's larger network. The goal of this article is to detail lessons Boston Beyond learned over the last 3 years of the Certified Observer Network to provide guidance to other youth development organizations designing similar peer-driven, network-level program quality improvement initiatives.

The Assessment of Program Practices Tool

The National Institute on Out-of-School Time (NIOST) developed the APT specifically to assess out-of-school time program quality. This tool allows observers to evaluate programs on 19 dimensions of program quality under the umbrellas of structure, social emotional environment, and youth engagement and learning. These measures have research-demonstrated links to youth outcomes (Tracy, Charmaraman, Ceder, Richer, & Surr, 2016). Used nationwide in over 30 states by more than 800 out-of-school time programs for quality assessment, the APT has undergone extensive psychometric testing to establish its validity and reliability (Tracy et al., 2016; Tracy, Surr, & Richer, 2012).

Boston Beyond's Certified Observer Network

Boston Beyond has provided partner programs with annual APT observations since 2011. After an APT observation, programs receive quantitative ratings and qualitative notes from the

Building Collective Capacity

observer. A few years into implementation, it became clear to Boston Beyond that simply supplying ratings was insufficient; programs needed additional professional development to understand, interpret, and ultimately act upon the data. Furthermore, the growth of Boston Beyond's network of program partners from less than 30 in 2011 to 100 in 2015 meant an increasing number of costly APT observations were needed.

For these reasons, Boston Beyond created the Certified Observer Network, a group of youth development professionals from Boston Beyond partner programs who learn how to apply the APT to peer programs in Boston Beyond's wider network. These youth development workers, referred to as "certified observers," gain in-depth knowledge about the tool and develop a more nuanced understanding of the difference between low- and high-quality practices. They also have opportunities to glean novel best practices from peer sites during the training as well as during independent, external observations that they are then required to conduct for Boston Beyond. Certified observers are then better prepared to interpret the APT results their own site receives. They also earn a professional credential from Boston Beyond for completing this process. Every organization that participates in Boston Beyond's network must have at least one certified observer.

The three central goals of the Certified Observer Network are to advance individual participants' professional development, positively affect their organizations, and create a sustainable system for network-wide quality improvement. Firstly, drawing on research that links program quality improvement with individuals' awareness of high-quality practices (Akiva & Yohalem, 2006; Hilberg et al., 2003; Pianta, 2010; Smith & Akiva, 2008; Wilson-Ahlstrom & Yohalem, 2007), this professional development aims to educate participants on the APT's 19 research-backed, essential components of high-quality programming and recognize these practices in action. Secondly, it seeks to drive program quality improvement by teaching participants how to interpret APT results and develop action plans to address weaknesses. Lastly, it is envisioned as a system-level effort to sustain the financial viability of APT observations, encourage peer-to-peer observation exchange, and facilitate networking opportunities.

In its first iteration in 2016, the training consisted of a self-paced online APT tutorial followed by a practice field observation with a NIOST "master observer," an experienced APT observer trained and certified through NIOST. The training concluded with an in-person event to debrief the observation.

Building Collective Capacity

In May 2017, Boston Beyond adopted a two-phase training approach. First, participants attend a 2-day event where they familiarize themselves with the tool, an anchor chart that provides example program practices for each rating, and basic observation techniques. Participants then practice observing by watching short video clips of out-of-school time program activities, rating them on one APT section, and discussing their ratings. The 1st day concludes with a live observation of a local after-school program with a master observer. On the 2nd day, participants compare their ratings to their master observer's and review discrepancies. Phase two involves a test, called a "tandem visit," where participants complete an additional field observation with a master observer. Ninety percent of the participant's ratings must be within one point of the master observer's ratings to earn certification; if the percentage falls below this threshold, they must complete another tandem visit.

As of May 2019, Boston Beyond had executed 10 trainings with 183 participants across 79 programs, for a total cost of \$123,920 (see Table 1)¹. Of the 126 observers still in the network, 75 are certified and 51 are in the certification process. Certified observers conducted 180 observations in the 3 years since the network began, saving Boston Beyond a total of \$150,650. When training costs are accounted for, the net savings is \$26,730. Boston Beyond collected feedback via surveys and interviews with training participants to assess the Certified Observer Network's impact. This article chronicles that feedback, considers implications for network improvement, and offers guidelines that other regional networks aiming to develop similar initiatives should consider.

Table 1. Certified Observer Network Net Savings

Costs		Savings		Net savings
Number of trainings	Training cost ^a	Number of observations	Value of observations ^b	
10	\$123,920	180	\$150,650	\$26,730

^a Training cost involves the cost of trainings and tandem visits. ^b Value of observations is the number of observations conducted by certified observers multiplied by the price that Boston Beyond would have had to pay outside observers to conduct those observations.

¹ NIOST considers training and individual observation prices to be confidential information, so only aggregate data is reported.

Methods

Boston Beyond collected survey data in the spring of 2019. An anonymous survey was sent to the 119 training participants still in the Certified Observer Network, excluding seven participants employed at Boston Beyond and 57 participants no longer in the network. The survey consisted of questions related to the impact of the APT training on the participant and their organization as well as their thoughts on network improvement and a potential professional development opportunity. Participants indicated the extent to which they agreed with nine statements on a four-point scale ranging from "Strongly Disagree" to "Strongly Agree." Responses indicating disagreement were collapsed into a single "Disagree" category as there was very little strong disagreement. Participants also responded to five yes or no questions about whether they had applied their knowledge in other ways. Specifically, they indicated if they had applied the tool internally, reviewed results alone or with a colleague, created and implemented an improvement plan based on results, or trained a colleague on the tool. If participants reported reviewing results with a colleague, creating an improvement plan, or training a colleague on the tool, they were asked to describe these activities. Finally, participants answered several optional, open-ended questions about how the experience impacted them in other ways and their suggestions for how to improve the training and their confidence in applying the tool.

Of the 119 participants contacted, 63 (53%) responded to the survey, but five of these could not recall when they were trained and were removed from the sample for analyses involving cohorts. Of the remaining 58 respondents, four trained in 2016, 14 trained in 2017, 29 trained in 2018, and 11 trained in 2019. Due to the small number of 2016 cohort responses, these were combined with 2017 cohort responses for all analyses. Forty-seven of the respondents were fully certified and 38 had conducted at least one external, independent observation. These data are summarized in Table 2.

A series of chi-square goodness of fit tests were performed to compare the population and sample distributions by training year, certification status, and independent observation status. The sample did not significantly differ from the population in terms of training cohort. However, it did significantly differ in terms of certification status, $\chi^2(3, n = 63) = 5.95, p < .05$, and observation status, $\chi^2(3, n = 63) = 6.43, p < .05$. Network-wide, 60% of participants are certified, whereas 75% of survey respondents were certified. Similarly, while 44% of participants have conducted at least one independent observation overall, 60% of survey respondents had. Thus, respondents, while representative of the various cohort groups, were significantly more likely than non-respondents to be certified and to have independent

Building Collective Capacity

observation experience. Simple adjustments were made when necessary to estimate more accurate population percentages.

Table 2. Sample and Population by Training Cohort, Certification Status, and Observation Status

	Sample	Population, $N = 126$
Training Cohort ($n = 58$)		
2016-2017	31%	29%
2018	50%	52%
2019	19%	19%
Certified ($n = 63$)	75%*	60%*
Conducted an independent observation ($n = 63$)	60%*	44%*

* Significant at $p < .05$ level.

In order to get a deeper understanding of participants' experience, interviews were conducted with four certified observers who attended trainings in 2017 or 2018. Interviewees answered open-ended questions about their training experience, its impact on their professional development and their organizations, and their thoughts on a potential professional development opportunity. Interviewees represented programs that varied in terms of their program type, structure, and the age group with whom they worked. Interviewees all had independent observation experience; three had conducted multiple independent observations.

Results

The results will be discussed in terms of the training's impact on participants, its impact on their organizations, and participants' thoughts on a future professional development opportunity.

Impact on Participants

Respondents answered questions about the impact of the certified observer experience on their professional development (see Table 3).

Understanding and Identifying Best Practices

All respondents either somewhat or strongly agreed with the statement "Participating in the certified observer process has helped me better understand what makes an out-of-school time

Building Collective Capacity

program high-quality.” Nearly all (95%) respondents strongly agreed or somewhat agreed that their participation helped them “identify new best practices that I can apply to my own organization,” while 5% disagreed with that statement. Results did not significantly differ by training year, certification status, or observation status.

Peer Networking

Most respondents (73%) expressed some level of agreement with the statement that this experience had allowed them to “connect with peers in the youth development field”; 27% disagreed. Though responses did not significantly vary by certification or observation status, they did significantly vary by training cohort, $\chi^2(4, n = 58) = 14.49, p < .01$. Respondents in the 2016 or 2017 training cohorts were more likely to indicate disagreement (44%) compared to those in the 2018 cohort (24%) or the 2019 cohort (0%). The majority of 2018 and 2019 respondents at least somewhat agreed that they connected with peers. These results align with changes made in 2017 and 2018 to increase peer-to-peer learning. Boston Beyond introduced the 2-day in-person model in 2017 and modified the training to include more opportunities for discussion in 2018.

Additional Impact

Respondents were asked to describe if this experience impacted them in any other ways. Among the 49 respondents who commented, the most common response (27%) was that they learned from other programs. These respondents mentioned gaining a different perspective on programming or seeing high- or low-quality practices to mimic or avoid, respectively. As one respondent explained, “I really enjoyed seeing other programs directly within the city of Boston. I think this is something that doesn't happen nearly enough and there isn't nearly enough best practice sharing between the many youth serving organizations within the city.” Another 20% reported that the APT provided a framework for better understanding program quality. “Once you see through the lens of the APT, it can't be unseen,” observed one survey respondent, adding, “I find that I am constantly thinking about APT-related best practices when planning new programs.” Others noted how the tool facilitates staff communication by providing a “common language and lens through which to think about programming” or helping them “learn how to present the goals of programming in meaningful and blatant terms.”

Building Collective Capacity

Table 3. Actual and Adjusted Impact on Participants and Perceived Impact on Youth

	Sample, <i>n</i> = 63			Population estimate ^a , <i>N</i> = 126		
	Disagree	Somewhat agree	Strongly agree	Disagree	Somewhat agree	Strongly agree
Participating in the certified observer process has helped me better understand what makes an out-of-school time program high-quality.	0%	40%	60%			
Participating in the certified observer process has helped me identify new best practices that I can apply to my own organization.	5%	38%	57%			
Participating in the certified observer process has helped me connect with peers in the youth development field.	27%**	46%**	27%**			
I currently feel confident applying the Assessment of Program Practices Tool (APT) to other programs.	2%*,††	36%*,††	62%*,††	3%	42%	55%
I would recommend the certified observer experience to a colleague.	5%	32%	63%			
Ultimately, my participation in the certified observer process has positively impacted youth at my organization.	3%	54%	42%			

^a Population adjustments were needed only when item responses varied by certification status or observation status, which were over-represented groups. Therefore, when adjusted, strong agreement level decreased.

*Significant at $p < .05$ level for observation status.

**Significant at $p < .01$ for training cohort.

†† Significant at $p < .01$ level for certification status.

Building Collective Capacity

When asked how this experience had impacted their professional development, three of the four interviewees likewise indicated that it helped them identify new best practices to apply to their own programs. As one interviewee concluded, "Becoming a certified observer has served to make me more aware and mindful of gaps in my organization's programs." Two of the four mentioned that it provided a framework for program planning. One interviewee explained that the APT anchor chart is so useful for curriculum development that he keeps it in a nearby desk drawer for regular reference.

Participant Confidence

Nearly all respondents (98%) agreed or strongly agreed that they felt confident applying the APT to other programs. Responses did not significantly vary by training cohort. However, they did significantly vary by certification status, $\chi^2(2, n = 63) = 10.16, p < .01$, with 72% of certified observers strongly agreeing compared to 31% of those still in the certification process. Similarly, results differed by observation status, $\chi^2(2, n = 63) = 9.05, p < .05$. Among respondents who had conducted at least one independent observation, 76% strongly agreed that they felt confident applying the APT compared to only 40% of respondents without such experience. Taken together, these results suggest that confidence is linked to experience. Adjusted population percentages were calculated since this sample over-represented both certified observers and observers with independent observation experience. Specifically, results were multiplied by the ratio of the percentage of the population that represented each combination of certification status and observation experience to the percentage of the sample that represented that certification status and observation experience. For example, since 44% of the population is certified with observation experience and 60% of the sample fits that description, these respondents' answers were multiplied by 44/60 or 0.73. In the adjusted results, agreement was slightly lower but still 97%: 55% strongly agreed, 42% somewhat agreed, and 3% disagreed.

Respondents who did not strongly agree answered another question about what additional support they would need to feel confident. Common responses among the 24 respondents included needing more tool experience or practice (42%) or a refresher (13%). Conducting another practice observation with a master observer or certified observer was specifically suggested by 17%, and another 13% indicated that they would have a better sense of their confidence level when they received their tandem results. These suggestions provide further evidence that confidence is likely fostered through observation experience.

Building Collective Capacity

Training Experience

Most of the respondents (95%) expressed some agreement with the statement, "I would recommend the Certified Observer experience to a colleague." Training year, certification status, and observation status did not significantly affect responses. The four interviewees concurred that the 2-day training experience was valuable, with three mentioning the helpfulness of the practice videos. Two of the four interviewees said increasing the interactivity of the training would be helpful.

Survey respondents were also asked for suggestions for how to make the certified observer experience more meaningful. The most common suggestions (3 responses each) among the 43 respondents were for Boston Beyond to organize a refresher or convene a meeting to highlight observers' contributions and/or discuss and recalibrate scoring techniques. Three respondents mentioned observation debriefs. As one explained, "It might be nice to have the option to debrief with the observer after they have observed your site. . . . Sometimes having a conversation is more informative than reading data points."

Perceived Impact on Youth

Most respondents (97%) expressed some agreement with the statement that their participation had "positively impacted youth at my organization." Results did not significantly differ by training cohort, certification status, or observation status.

Impact on Organization

Respondents answered five questions relating to how they applied knowledge gained through the Certified Observer experience to their own organizations (see Table 4).

Internal Application and Results Debriefing

There were no significant differences by training year, certification status, or observation status with respect to respondents' applying the APT to their own organizations. Of the respondents who indicated that they had shared results with a colleague, 46 explained the process. Almost half of these respondents (44%) reported going over the results and identifying areas of improvement, with some mentioning that they also recognized areas of strength. Respondents shared results in a variety of settings, most frequently during staff meetings (26%) or with their leadership teams (11%).

Table 4. Actual and Adjusted Impact on Organization

	Sample, <i>n</i> = 63	Population estimate ^a , <i>N</i> = 126
Have you ever applied the APT internally (to your own organization)?	62%	
Have you ever reviewed your organization's APT results (either from an internal or external observation)?	83%	
Have you ever reviewed your organization's APT results (either from an internal or external observation) with any of your colleagues?	76%	
Have you ever created and implemented an improvement plan based on your APT results?	44%	
Have you ever trained any of your colleagues on how to use the APT?	16%* [†]	12%

^a Population adjustments were needed only when item responses varied by certification status or observation status, which were over-represented groups. Therefore, when adjusted, positive response level decreased.

*Significant at $p < .05$ level for observation status.

[†]Significant at $p < .05$ level for certification status.

Improvement

Of the 44% of respondents who affirmed that they had created and implemented an improvement plan based on their APT results, 85% mentioned adapting programming. Respondents most often (30%) noted that they based improvement plans on specific ratings such as youth autonomy, leadership, or reflection. Seven percent reported focusing on low ratings, either generally or those that aligned with their particular program objectives and 11% highlighted that they tailored staff trainings to APT results. As one respondent summarized, "We found pieces of feedback that stood out from APT, either on their own or because they tracked with other means of feedback, and made decisions about programmatic, training, or curriculum changes."

When asked about the impact of their Certified Observer experience on their staff, three of the four interviewees similarly indicated using APT knowledge for staff training. They reported reviewing results with staff, providing training on specific parts of the APT, and analyzing the prior year's results to identify areas of focus for trainings. One mentioned employing the APT anchor chart as a guide for appropriate staff responses to common student behavioral situations.

Building Collective Capacity

Training Colleagues on the APT

Among the 16% of respondents who reported training at least one colleague to use the APT, results did not significantly vary by training cohort, but they did significantly vary by certification status, $\chi^2(1, n = 63) = 4.05, p < .05$, with 21% of certified respondents indicating that they had trained a colleague compared to 0% of uncertified respondents. Likewise, observation status significantly affected responses, $\chi^2(1, n = 63) = 4.38, p < .05$: 24% of certified respondents versus 4% of uncertified respondents reported training a colleague on the tool. When adjusted for the over-representation of both certified observers and those with independent observation experience, the estimated percentage of the population who likely trained a colleague dropped from 16% to 12%.

Of the 10 respondents who described training a colleague, four mentioned formally training other staff members to apply the tool internally. "The Site Coordinators were given individual trainings on each [section]. Detailed explanations were given and then observations were made, then we returned to the APT and talked about what they saw and where they saw things that needed improvement and what areas worked well," explained one respondent. Another 40% either shared the APT materials with a colleague or explained the tool and/or observation process to them.

Participant Thoughts on Future Professional Development Opportunity

Interest in Professional Development Opportunity

Respondents answered three questions related to a professional development experience Boston Beyond was considering (see Table 5). This opportunity, called "quality advising," would involve learning how to coach other partner programs on their APT results. Quality advisors would conduct an observation and then debrief with the site by reviewing ratings and helping them identify areas of focus. After reading a description of quality advising, respondents reported the extent to which they agreed with the statement, "I would be interested in this professional development opportunity." Training year, certification status, and observation status did not significantly impact responses. All four interviewees indicated interest in the training.

Table 5. Thoughts on Professional Development Opportunity

	Disagree	Somewhat agree	Strongly agree
I would be interested in this professional development opportunity ($n = 63$)	29%	41%	30%
I would have the capacity to engage in this professional development opportunity ($n = 45$)	27%*,†	49%*,†	24%*,†
I would be interested in receiving a professional credential for this training ($n = 45$)	5%	42%	53%

*Significant at $p < .05$ level for observation status.

†Significant at $p < .05$ level for certification status.

Interested Participants' Capacity for Professional Development

Responses to statements about the quality advising opportunity did not significantly vary by observation status, but they did significantly vary by both training cohort, $\chi^2(6, n = 43) = 14.79, p < .05$, and observation status, $\chi^2(3, n = 45) = 8.02, p < .05$. Specifically, 100% of respondents in the 2016 and 2017 cohorts agreed to some extent compared to 70% of those in the 2018 cohort and 55% of those in the 2019 cohort. This may signal a bias in the sample in these earlier cohorts: perhaps the few participants who trained this long ago, are still in the network, and chose to respond to this survey have a particularly high capacity for professional development. Similarly, 88% of respondents with independent observation experience either somewhat or strongly agreed compared to 53% of those without such experience. To the extent that the sample over-represents both certified observers and those with independent observation experience, these results may be higher than in the population as a whole.

Interested Participants' Desire for Credential

Finally, nearly all (95%) of the 45 interested respondents expressed some level of desire to earn a professional credential for the additional training. Responses did not significantly vary by training cohort, certification status, or observation status.

Implications

The vast majority of respondents attest that the certified observer experience improved their understanding of program quality and exposed them to new program practices. To a lesser extent, they agreed that it allowed them to connect with peers. Nearly all feel confident

Building Collective Capacity

applying the tool, would recommend the experience to a colleague, and believe their students have benefitted from their involvement. Three-quarters or more review the results on their own or with colleagues, and slightly less than half act on these results by creating and implementing improvement plans. Over half have applied the tool internally and a small number train their colleagues on it. Most respondents indicated interest in a new professional development opportunity, and those interested generally believed that they had the capacity to engage in it.

In this section, we examine the extent to which Boston Beyond met its three original goals of the Certified Observer Network: promoting participants' professional development; positively impacting organizations; and creating a sustainable, participant-driven model of program improvement. Implications are considered for Boston Beyond as well as other regional networks that may want to develop similar models.

Participants Find Training Valuable; Provide More Training and Foster Peer Connections

Network involvement benefits participants in several ways for which it was initially designed: participants better understand high-quality program practices and, through observations, gain a fresh perspective on programming and see novel practices. They are confident applying the tool, particularly if they have observation experience through a successful tandem visit or an independent observation.

The link between confidence and experience has several implications for others aiming to create this sort of network. Firstly, it highlights the importance of ensuring that participants complete their certification and uphold their obligation to independently observe other programs. A year after they attended the 2-day training, 18 participants still employed at their organizations have not finished their certification. Around a dozen certified observers still in the network do not regularly conduct observations. These participants are therefore missing out on opportunities to further their professional development, and Boston Beyond is not receiving as much of a return on the investment it made in these participants. Regional networks looking to develop a beneficial, sustainable observer network would want to strongly encourage certification completion. Boston Beyond plans to strategically time the 2-day trainings such that there are ample tandem opportunities immediately afterwards, encourage participants to schedule their tandems at the training, and impose a certification deadline (e.g., within 3 months of the training). Like Boston Beyond, others may want to consider developing an accountability system to clearly communicate to program partners about requirement completion.

Furthermore, if participants feel more confident with experience, regional networks would want to build in ample opportunities for re-training. These could take several forms: a formal re-certification process, informal refreshers, or convening a training session. Boston Beyond will ask participants who trained over 3 years ago to get re-certified. This online process will entail applying the APT Activities section to a series of short videos. Ratings will then be compared to those established by a group of NIOST master observers to re-establish inter-rater reliability. Interested participants can first practice with other videos that include master observer ratings and explanations. Boston Beyond plans to open up this process to anyone seeking additional practice. Additionally, Boston Beyond will offer all certified observers another in-field tandem observation refresher, as requested by some respondents.

Boston Beyond will also convene a gathering where participants can review tool guidelines, recalibrate their rating approach, and exchange best practices to further refine their observation skills. This gathering would also address another area of improvement for the Certified Observer Network: peer learning and networking. Survey results demonstrate that respondents felt more connected to their peers when the training switched formats from online to in-person and when the training became more discussion-based. Others designing observer networks may want to structure trainings in ways that foster such peer-to-peer exchanges and connection.

Over half of respondents are interested in the quality advisor training, where trainees would learn how to coach observed programs on their results. Interestingly, several respondents separately advocated for their organizations to receive this sort of debrief after a site visit. Results indicate that not all respondents share results with staff or design action plans to address weaknesses. Therefore, this training will focus on teaching quality advisors how to coach sites on these two particular facets of program quality improvement. Since only about three-quarters of interested participants have the capacity for such a training, Boston Beyond will launch this pilot project with a small cohort of around a dozen trainees.

Participants Share Knowledge With Organizations, but Need Additional Support

As intended, participants are applying the knowledge gained through this training to better their organizations, though only to a certain extent. Most but not all participants share their APT results with colleagues. Participants should be explicitly encouraged to transfer this knowledge and supported in doing so. Participants could be taught how to analyze their results at the 2-day training and provided with a simple protocol for debriefing results. Regional networks in the

Building Collective Capacity

beginning stages of network development may want to consider fostering a sense of pride among early participants and suggesting that they have a responsibility to share their knowledge with colleagues. For example, they could refer to these trainees as “APT ambassadors.” If internal application of the APT is a network priority, this should be mentioned at the training and time should be dedicated to discussing this process.

System-Level Effort Working; Sustainability Is an Issue

The Certified Observer Network is meeting its final goal of harnessing the collective capacity of partner programs to create a system of program quality improvement. Certified observers are regularly and confidently evaluating each other and, in doing so, catalyzing program quality improvement amongst each other. However, the network’s sustainability is a persistent challenge. Consistent with the high level of turnover in the out-of-school time field, 31% percent ($n = 57$) of the 183 participants have either left their positions at partner programs or indicated that they no longer want to participate in the Certified Observer Network. Since Boston Beyond requires each organization to get only one staff member certified, this loss may mean the loss of the partner program’s APT knowledge altogether. Regional networks considering this sort of network should be very deliberate about recruitment. As Boston Beyond plans to do, others could specifically recruit mid-level staff. These staff are more stable than their direct-service counterparts and likely have a responsibility to train and manage staff, facilitating the sharing of APT knowledge throughout the organization. Furthermore, the network requirement should be explained to each potential participant, and they should be asked to commit to it.

The corresponding negative financial implications on sustainability are obvious. If participants do not conduct any independent observations, Boston Beyond does not get the expected return on the investment it made in training them. The 193 total participants have conducted only 180 observations, a lower output than would be expected from this input. Savings, while positive, are relatively modest compared to costs. To limit costs, Boston Beyond plans to decrease the number of annual trainings from four to three. Others thinking about developing this sort of network may want to also require additional independent observations to decrease the financial burden, particularly when first starting up.

Limitations

There are several limitations to this study. Firstly, only participants who are still in the Certified Observer Network were asked to complete the survey, so the 31% of participants ($n = 57$) who have left the network did not provide feedback. Furthermore, only 53% of participants ($n = 63$) completed the survey. Since completion was voluntary, observers who were more engaged with Boston Beyond and/or the Certified Observer Network were more likely to respond. Data were collected at one point in time, so participants who attended the two-day training more recently may not have yet had an opportunity to apply their knowledge. Certified observers and observers with independent observation experience were over-represented in this sample, though this was corrected for when necessary. Interviews were only conducted with four participants, a very small percentage of the overall population that may not be representative.

Conclusion

Boston Beyond's Certified Observer Network is meeting its three primary goals of cultivating participants' professional development, positively impacting program partner organizations, and creating a system-level model of peer-to-peer program quality improvement, though it is not yet fully sustainable. As Boston Beyond looks ahead to its 4th year, it aims to strengthen the network's efficacy by refining the recruitment, training, and post-training processes to better support, connect, and retain observers. In the years to come, Boston Beyond will harness interested certified observers' knowledge and experience to develop a group of quality advisors who can guide others through the improvement process to help them ultimately better serve their students.

References

- Akiva, T., & Yohalem, N. (2006). *Quality systems: Lessons from early efforts to disseminate the Youth PQA*. Washington, DC: Forum for Youth Investment. Retrieved from www.forumfyi.org
- Duffett, A., & Johnson, J. (2004). *All work and no play? Listening to what kids and parents really want from out-of-school time*. New York, NY: Public Agenda and the Wallace Foundation. <https://files.eric.ed.gov/fulltext/ED485306.pdf>
- Durlak, J. A., Weissberg, R. P., & Pachan, M. K. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology, 45*, 294-309. <https://doi.org/10.1007/s10464-010-9300-6>

Building Collective Capacity

- Eccles, J., & Gootman, J. (Eds.). (2002). *Community programs to promote youth development*. Washington, DC: National Academy Press. <https://doi.org/10.17226/10022>
- Hilberg, R. S., Waxman, H. C., & Tharp, R. G. (2003). Introduction: Purposes and perspectives on classroom observation research. In H. C. Waxman, R. G. Tharp, & R. S. Hilberg (Eds.), *Observational research in U.S. classrooms: New approaches for understanding cultural and linguistic diversity* (pp. 1-20). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511616419.001>
- Larson, R., Rickman, A. N., Gibbons, C. M., & Walker, K. C. (2009). Practitioner expertise: Creating quality within the daily tumble of events in youth settings. *New Directions for Youth Development*, 121, 71-88. <https://doi.org/10.1002/yd.297>
- 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, 275-313. <https://doi.org/10.3102/00346543076002275>
- Mahoney, J. L., Harris, A. L., & Eccles, J. S. (2006). Organized activity participation, positive youth development, and the over-scheduling hypothesis. *Social Policy Report: Giving Child and Youth Development Knowledge Away*, 20, 3-32. <https://srcd.onlinelibrary.wiley.com/doi/pdf/10.1002/j.2379-3988.2006.tb00049.x>
- Pianta, R. C. (2010). *Standardized classroom observations from pre-K to third grade: A mechanism for improving quality classroom experiences during the P-3 years*. <https://www.fcd-us.org/standardized-classroom-observations-from-pre-k-to-third-grade-a-mechanism-for-improving-quality-classroom-experiences-during-the-p-3-years/>
- Pianta, R., & Hamre, B. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38, 109-119. <https://doi.org/10.3102/0013189X09332374>
- Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15, 4-14. <https://doi.org/10.2307/1175860>
- 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. <https://doi.org/10.1093/acprof:oso/9780195327892.003.0011>
- Tracy, A., Charmaraman, L., Ceder, I., Richer, A., & Surr, W. (2016). Measuring program quality: Evidence of the scientific validity of the Assessment of Program Practices Tool. *Afterschool Matters*, 24, 3-11. <https://files.eric.ed.gov/fulltext/EJ1120616.pdf>

Building Collective Capacity

Tracy, A., Surr, W., & Richer, A. (2012). The Assessment of Afterschool Program Practices Tool (APT): Findings from the APT validation study. Wellesley, MA: National Institute on Out-of-School Time. <https://files.eric.ed.gov/fulltext/ED539180.pdf>

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. <https://www.issuelab.org/resources/1194/1194.pdf>