

Assessing the Evaluation Capacity of Large Nonprofit Organizations

A Detailed Account of the Methods,
Findings, and Lessons Learned from the
YMCA Evaluation Capacity Assessment

December 2012

Assessing the Evaluation Capacity of Large Nonprofit Organizations

A Detailed Account of the
Methods, Findings, and Lessons Learned from the
YMCA Evaluation Capacity Assessment

December 2012

Table of Contents

Introduction	3
YMCA Evaluation Capacity Assessment Purpose and Goals	3
Overview of Evaluation Capacity Assessment Process.....	4
Purpose and Organization of This Report	5
Step 1: Becoming Acquainted with Y Culture, Structure, and Evaluation Needs.....	6
Organizational Structure.....	7
Organizational Culture	9
Evaluation Capacity.....	10
Evaluation Capacity Assessment Process.....	12
Lessons Learned	17
Step 2: Evaluation Capacity Assessment Survey	18
Methods	18
Results	20
Lessons Learned	25
Step 3: Virtual Focus Group.....	27
Methods	27
Results	34
Lessons Learned	35
Step 4: Key Informant Interviews	37
Methods	37
Results	40
Lessons Learned	43
Step 5: Site Visits.....	45
Methods	45
Results	49
Lessons Learned	52
Summary of Lessons Learned	54
Conclusions from the Evaluation Capacity Assessment Process	57
Y Specific Conclusions	57

Conclusions About the Process Overall	58
References	59
Appendix A. Evaluation Capacity Assessment Survey	60
Appendix B. Virtual Focus Group Protocol.....	69
Appendix C. Key Informant Interview Guide	75
Appendix D. Site Visit Protocol.....	85

Introduction

Nonprofit organizations are increasingly challenged to provide a growing number of community services, influence public policy, and be innovation leaders in their fields, while at the same time facing resource constraints and increased pressure from funders to be accountable for achieving tangible results. Many nonprofits are turning to organizational capacity building as a strategy to run more effective and efficient operations, enhance sustainability, and produce better results from their programs (Connolly and York, 2002). Specifically, some organizations are undertaking evaluation capacity building. By building evaluation capacity, these organizations seek to improve their accountability, competitiveness for funding, and effectiveness by using results from high-quality evaluations to enhance program development and operations. The trend in building evaluation capacity is increasingly being supported by social sector funders to improve nonprofit performance and maximize the impact of philanthropic investments (Welsh and Morariu, 2011).

As activity in this area expands, a challenge for nonprofit organizations and funders alike is the lack of clarity and comprehension around what it means for an organization to possess evaluation capacity and what the best approaches are for assessing such capacity. There is a dearth of research in the emerging field of evaluation capacity, and what research exists uncovers no singular or widespread definition of evaluation capacity or best practices for its assessment (Nielson, Lemire, and Skov, 2011). For example, more simplistic definitions such as the ability to perform high-quality evaluations and use evaluation findings are available, but other definitions suggest that evaluation capacity is significantly more complex, intrinsically linked to an organization's structures, processes, culture, human capital, technology, and other attributes that traditionally comprise an organization's capacity more broadly. Additional research into how organizations define evaluation capacity and approach evaluation capacity assessment can contribute valuable knowledge to the field and expand understanding of how nonprofits can approach this work.

YMCA Evaluation Capacity Assessment Purpose and Goals

The purpose of this report is to describe the methods, findings, and lessons learned through an evaluation capacity assessment process completed with the YMCA (the Y) in 2012. The National Council of the Young Men's Christian Association of the United States of America (YMCA of the USA or Y-USA) is one of the top ten largest nonprofit organizations in the United States, and the nation's leading nonprofit committed to strengthening communities through youth development, healthy living, and social responsibility. For more than 165 years, the Y has worked to strengthen communities through programs and services that build healthy spirit, mind, and body for all. There are approximately 2,700 local Ys across the country (collectively, the Y movement) serving more than 10,000 communities and 21 million individuals, including 12 million adults and 9 million youth under the age of 18. As a mission-driven organization, the Y must use evaluation and measurement to assess whether they are reaching their goals and advancing their cause. The Y must also be accountable for producing results that demonstrate the impact the Y has on the lives of individuals and within communities. To this end, Y-USA sought to learn more about how local Ys were approaching evaluations and what their level of capacity was to conduct program evaluations that produce meaningful results.

Recent organizational initiatives within the Y had placed a strong emphasis on measuring the impact of Y programs and services on individuals, and the Y-USA Research and Evaluation Department anticipated the need to plan to provide additional evaluation support to local Ys. However, the Y lacked an organization-wide view of what evaluation skills, systems, and tools might be needed by Ys to respond to the need for more robust evaluations of their work. Evaluation capacity assessment was a process identified by Y-USA as a method for gaining a stronger understanding of current Y

evaluation activities, facilitators and barriers to evaluation, and what capacity building efforts may be needed. This process also had the potential to surface best practices among Ys that are worthy of nationwide replication.

In spring 2012, Y-USA submitted a successful proposal to the Bruner Foundation seeking funding to conduct an evaluation capacity assessment of the Y movement. The purpose of the evaluation capacity assessment was to expand Y-USA's understanding of the Y movement's capacity for evaluation, with an emphasis on program outcome evaluation. Once equipped with this information, Y-USA sought to determine how the Y national office might work to support the evaluation needs of local Ys.

Overview of Evaluation Capacity Assessment Process

In April 2012, Y-USA hired Altarum Institute (Altarum), a non-profit health systems research and consulting organization, to oversee and facilitate the evaluation capacity assessment process during May-November 2012. Altarum proposed to use a participatory, iterative approach that would build on existing models of nonprofit and evaluation capacity assessment. These existing models would provide a framework and foundation for the evaluation capacity assessment process, whose elements could then be tailored to the unique cultural and structural attributes of the Y. This approach was accomplished through the following process:

1. **Become acquainted with Y-USA and Y organizational culture, structure and evaluation-related needs.**

Gain a broader understanding of current Y capacity and processes in order to inform the development of assessment process that best suits the Y's organizational culture and goals for evaluation capacity development, built on a strong foundation of understanding around defining evaluation capacity, organizational attributes, and organizational goals.

2. **Conduct site visits and other informational sessions with Ys to understand the depth and breadth of the Y movement's capacity to evaluate program outcomes.**

Conduct a thorough needs assessment of Y evaluation capacity using three data collection strategies:

- ▲ *Evaluation Capacity Assessment Survey*—an online instrument designed to establish a broad baseline measure of Y status on various attributes of evaluation capacity, and to assist in categorizing the capacity of Ys for further investigation.
- ▲ *Virtual Focus Group*—an online, facilitated group discussion with a small group of lower capacity Ys designed to gather insights, perceptions, and beliefs around evaluation.
- ▲ *Key Informant Interviews*—one-on-one telephone interviews with a cohort of medium capacity Ys to learn more in-depth information around Y evaluation practices, data collection strategies, and evaluation supports.
- ▲ *Site Visits*—onsite interviews conducted with multiple staff at high capacity Y sites designed to learn about best practices in evaluation conduct and management and to surface potential models that may be replicable by other Ys.

3. Complete a final report including but not limited to recommendations for how Y-USA can best support Y evaluation practices.

Following the capacity assessment process, analyze all results and deliver a final report that compiles a summary of all learnings, along with conclusions and recommendations for moving forward with evaluation capacity building. In addition, Altarum would provide a final report suitable for external use that describes the evaluation capacity assessment process. This report is the latter final report.

Purpose and Organization of This Report

This report is intended for audiences who are interested in conducting or already embarking on an evaluation capacity assessment of a large nonprofit organization. It provides in-depth detail about how Altarum approached the evaluation capacity assessment process, which methods were used and why they were selected, what types of data were obtained, lessons learned about what worked well and not as well, and our overall conclusions about the process. It is important to note that this is not a report of all of the results of our evaluation capacity assessment process; rather, it is an account of our experiences conducting this process with results from the Y presented as illustrative examples.

The report is organized into chapters which articulate each step of the evaluation capacity assessment process. In each chapter, or step, we walk through the why, hows, and whats of our process, presenting specific examples from our work, including agendas and presentation slides, data collection protocols, and results from our efforts. We also discuss our thinking behind particular strategies used in our process and conclude each chapter with a lessons learned section that includes facilitators and barriers we faced within each step of the process. Finally, we conclude with a summary of lessons learned and conclusions about the approach we took to assessing the evaluation capacity of the Y.

Step 1: Becoming Acquainted with Y Culture, Structure, and Evaluation Needs

The first step in our capacity assessment process was to develop an understanding of Y-USA and local Y organizational culture, structure, and evaluation-related needs, including defining what evaluation capacity meant within the context of the Y. Although we entered the evaluation capacity assessment process with a general sense of the approach we would take, we recognized the need to assess organizational culture and goals to further refine and tailor the approach as well as to confirm the intended approach would meet the organization's goals and answer their evaluation questions. To accomplish this step, we facilitated a one day, in-person meeting with Y-USA at their headquarters in Chicago, IL. The objectives of this meeting were to:

- ▲ Understand the factors that have led Y-USA to engage in an evaluation capacity assessment process and what they hope to achieve through this process.
- ▲ Develop a deeper understanding of Y-USA and Y organizational structures and cultures, and how these may affect future evaluation capacity assessment and building efforts.
- ▲ Identify key strengths and weaknesses of current Y-USA and Y evaluative capacity.
- ▲ Come to consensus on the appropriate framework and methods for conducting the evaluation capacity assessment process.

To accomplish these objectives, Altarum staff facilitated a discussion with Y-USA Research and Evaluation staff on topics summarized in Table 1. These topics were selected based on a review of the literature around evaluation capacity assessment. Altarum utilized a PowerPoint presentation during the meeting. This presentation included slides of key topic discussion questions and diagrams, as well as section quotes. These quotes were selected from the literature and designed to direct the group's thinking to each new topic as well as present interesting findings related to the topic and organizational evaluation capacity. In addition to projecting the presentation, we projected the real-time notes we were recording as the discussion unfolded. By projecting this information, all participants could refer to questions, quotes, and diagrams used in the slides as well as the transcription of their responses as we took notes. This allowed Y-USA staff to make clarifications to our understanding if we misinterpreted any of the discussion points.

To further build our understanding of the Y, we requested organizational documents as well as documents that staff referred to that we felt may have relevance to the evaluation capacity assessment process. Materials gathered at the meeting included:

- ▲ Organizational charts
- ▲ Presentations on Y-USA evaluation activities, data, and approaches
- ▲ Reports on Y operations and evaluations conducted by Ys
- ▲ Resource materials Y-USA had developed related to evaluation
- ▲ Program surveys
- ▲ Evaluation plans and logic models for Y programs

Table 1. Y-USA and Altarum kickoff meeting agenda topics.

Organizational Structure	Organizational Culture	Evaluation Capacity Goals	Evaluation Capacity Assessment Process
Governance and Leadership	Evaluation Approaches	Defining Evaluation Capacity	Tools
Human Capital	Y Evaluation Culture	Goals for Assessing Evaluation Capacity	Frameworks
Technology and Systems	Prior Capacity Assessment or Building Efforts	Key Stakeholder Identification	
Strategic and Evaluation Planning Processes		SWOT Analysis	

Below, we present how we facilitated the agenda, including questions from the presentation slides we used during the meeting.

Organizational Structure

“In the State of Evaluation 2010, Innovation Network reports the results of its national survey of over 1,000 nonprofits regarding evaluation practice and capacity. The survey reveals that many nonprofits have significant barriers to conducting effective evaluation. In fact, 71 percent said that limited staff time is a significant barrier to evaluation, followed by lack of financial resources (57 percent), lack of sufficient in-house evaluation expertise (43 percent) and lack of leadership support for evaluation (29 percent).”

-Major and Brennan, 2011

Governance and Leadership

Leadership structure and support is integral to evaluation capacity. We began the discussion by understanding how the organization is structured, who is responsible for decision-making, and where support for evaluation exists at the local and national levels of the organization. We engaged Y-USA staff in this topic through a variety of discussion questions (see Figure 1) and a review of organizational charts. The information we obtained through this discussion helped to build our knowledge of how Y-USA and Ys are organized and operate, and where decision-making power that would affect evaluation capacity is housed within each part of the organization.

Figure 1. Governance and leadership discussion questions.

- ▲ How are Y-USA and Ys structured organizationally?
- ▲ How are decisions made at the local Y level? National level?
- ▲ What are Ys required to provide to Y-USA in terms of data collection and evaluation?
- ▲ Is there support for evaluation work at the board and executive levels?
- ▲ Are there clear expectations for evaluation roles at Y-USA? For Ys?
- ▲ How does Y-USA support Y evaluation efforts?
- ▲ Are there dedicated financial resources for evaluation?

Human Capital

Human capital is the staff available to conduct evaluations as well as staff evaluation experience, technical skill, and training. We sought to understand who was conducting evaluations at Ys, whether these staff had formal education or training in evaluation, and how these factors might differ across sites. In this discussion, we learned that Y staff conducting evaluations may serve in a number of primary roles, often not solely dedicated to evaluation, including marketing, membership services, financial development, and program implementation. Because of the vast array of staff with evaluation responsibilities and no central repository with information on staff backgrounds (e.g., formal education), some of these questions were difficult to answer for Y-USA staff. In the absence of organization-wide data, we asked Y-USA about their own experiences with Y staff on evaluation—anecdotes from their experiences and beliefs about what human capital existed.

Figure 2. Human capital presentation slides.

- ▲ What types of evaluation staff are available?
 - At Y-USA, at Ys
 - Types of formal education
 - Practical knowledge and evaluation experience
- ▲ Technical skill
 - To design evaluations
 - To conduct evaluations
 - To analyze results
 - To select quality indicators
- ▲ How does evaluation staff differ among Ys?
 - Describe examples of typical high, medium, low capacity sites

Technology and Systems

Technology and systems, such as software, evaluation models and data collection techniques, and systems for providing training and technical assistance and managing data were discussed at the Y-USA and local Y levels. Gathering this information helped us to understand what systems were

available to collect, analyze and store data, and how institutionalized processes were for providing training and technical assistance on evaluation. Y-USA had formed a partnership with SEER Analytics to create BaseMetrics, an integrated suite of tools designed provide a common, movement-wide framework to evaluate how Ys engage and strengthen the community in three performance areas: public perception, program impact, and operational effectiveness. Ys can utilize BaseMetrics for a fee. We learned that Y-USA routinely gathers data from Ys using online survey software, and Ys also frequently use similar systems to collect data on their members. Data management systems were available, but they were designed to track membership data, not data from programs being run at the Y. Tracking Y member data and program satisfaction were believed to be very common among Ys, but Y-USA staff believed few Ys were conducting the type of individual-level outcome evaluations—such as assessing changes in knowledge, attitudes, and behaviors—that Y-USA was most interested in.

Strategic and Evaluation Planning Processes

To understand how embedded evaluation was within the organizational structure, we inquired about the use of strategic plans and evaluation plans. We wanted to know what types of information were in these plans, as well as who participated in the design, implementation, and measurement of progress against such plans. Were Ys using evaluation plans? Did they include processes for monitoring quality of training for staff, program fidelity, and individual outcomes? Who met regularly within the organization to discuss evaluation? How was success measured? We learned that Ys often do have strategic plans, but they are not required to be shared with Y-USA, so they were unavailable to us. Evaluation plans were usually only created for large-scale (e.g., large national grants) programs, and there was no specific process or structure involved in their creation. It was unknown whether Ys routinely used evaluation plans, or whether evaluation was integrated into strategic plans.

Organizational Culture

“What funders seek to learn is often not what nonprofits seek to learn, or are equipped to measure.”

-Peter York in Major & Brennan, 2011

Evaluation Approaches

For this topic, we first sought to learn about what approaches Ys take in conducting evaluations. Do they typically use participatory approaches, directive approaches, collaborative approaches, or something else? We also wanted to know the extent to which these processes usually include formative, process, outcome, and impact evaluation, and whether the organization had developed a formal process to identify and disseminate the most effective programs. We learned that Ys typically do evaluations for reasons of accountability, at the behest of agencies that fund their work. These agencies are the primary beneficiaries of evaluation results, which are not necessarily utilized to improve program processes and outcomes. Evaluations that are driven by Ys tend to focus exclusively on customer satisfaction versus individual program outcomes.

Evaluation Culture

Understanding the culture of evaluation is important in ascertaining information about staff and organizational attitudes towards evaluation and assessing the staff and organizational readiness to participate in evaluation processes. It is not uncommon for evaluation to be viewed as a complex and complicated process. This could make staff feel anxious and worried about evaluation, which makes them less inclined to participate in evaluative processes and could have influenced their participation in the evaluation capacity assessment. We found that the culture of evaluation within Ys varies considerably, with some staff being very confident, while others not being aware of the need to use evaluation to drive improvement. Y-USA staff were concerned that some Ys may have reservations

about participating in the evaluation capacity assessment process, and may feel they are being “tested”. They also felt there was a threat of social desirability bias in assessing results; Ys may tell us what they think we want to hear instead of what they are actually feeling and doing with respect to evaluation.

Past Evaluation Capacity Assessment or Building Efforts

We sought to learn whether Y-USA had ever undertaken an evaluation capacity assessment process in the past. If they had, we wanted to learn what worked well about this process, and what could be improved in the current process based on those findings. Similarly, we wanted to know if they had engaged in evaluation capacity building efforts, and if so, what the results of these efforts have been. We found that the organization had not participated in a comprehensive evaluation capacity assessment or building process in the past, but had hired a third party to perform segmentation analysis of Ys with respect to performance measurement. Their goal in that process had been to learn about the current state of performance measurement among Ys of different sizes and characteristics, and how enthusiastic Ys were to adopt a system for performance measurement. We reviewed the report compiled from this analysis to gain insights into how Ys might feel about evaluation, and how different Y characteristics, such as budget (a proxy for size) might influence evaluation activities and attitudes.

Evaluation Capacity

“A review of the existing literature uncovers four discernible trends: (a) widespread conceptual pluralism; (b) differing opinions regarding the purpose of ECB [evaluation capacity building]; (c) the lack of a comprehensive empirical base for the various models; and (d) a significant focus on approaches and methods for tackling capacity building.”

-Nielsen, Lemire, and Skov, 2011

Defining Evaluation Capacity

As noted previously, evaluation capacity and evaluation capacity assessment are emerging topics in the field of evaluation and often have differing definitions in literature. This makes it challenging to determine precisely what elements are needed to have high capacity, or to assess whether organizations possess this capacity or some variation of it. When embarking on an evaluation capacity assessment process with an organization, it is imperative to define evaluation capacity in terms that are relevant and meaningful to ensure results are valid. To determine which definition of evaluation capacity most resonated with the Y, we discussed various ways to define evaluation capacity summarized by Nielson and colleagues (2011), and provided examples of each from literature (Table 2). A combination of the first and second definitions resonated most with Y-USA. Their definition of evaluation capacity was ability to perform high-quality evaluations *and* make use of evaluation results to ensure the best use of Y resources.

Table 2. Definitions of evaluation capacity and their use in the literature.

Definition	Examples
Ability to perform high-quality evaluations	“A context-dependent intentional action system of guided processes and practices for bring about and sustaining a state of affairs in which high-quality program evaluation and its appropriate uses are ordinary and ongoing practices within and/or between one or more organizations/programs/sites.” (Stockdill, Baizerman, and Compton, 2002)
Ability to make use of evaluation results	“Sustainable evaluation practice—where members continuously ask questions that matter, collect, analyze, and interpret data, and use evaluation findings for decision marking and action.” (Preskill and Boyle, 2008)
Ability to successfully manage one’s organization	“An organization’s ability to bring about, align, and sustain its objectives, structure, processes, culture, human capital, and technology to produce evaluative knowledge that informs ongoing practices and decision-making in order to improve organizational effectiveness.” (Mackay, 2002)
Ability to be accountable	“Systematic evaluation helps management assistance providers increase their accountability, articulate the value of their work, and compare the effectiveness of different capacity-building activities and it also allows funders to improve their capacity-building grant-making strategies” (Connolly and York, 2002)

Goals for Assessing Evaluation Capacity

Y-USA and Altarum staff then discussed both short- and long-term goals for the project. Short-term goals were defined in terms of desired results from the evaluation capacity assessment process, and included:

- ▲ Increased knowledge of breadth and depth of current Y-USA and Y evaluation capacity and evaluation strengths and weaknesses across the Y movement.
- ▲ Development and execution of an evaluation capacity building plan.

Long-term goals following an anticipated future capacity building effort were also articulated. Although these outcomes would not be measured during the project due to the brief timeline of the evaluation capacity assessment process, it was important for the group to define their ultimate desired outcomes as a result of capacity assessment and building work. Long-term goals included:

- ▲ Institutionalization of ongoing evaluation capacity development and improvement processes across the Y movement.
- ▲ Ys have improved ability to evaluate/apply evaluative thinking to initiatives locally and nationally.
- ▲ Y movement can collectively demonstrate positive outcomes in key areas based on valid and reliable methods.
- ▲ Data-based decision making is the norm across the Y movement.

By understanding these goals, we could ensure that we designed an assessment process that would inform Y-USA in priority areas. We also used these goals to design a logic model, which is further discussed on page 15.

Key Stakeholders

Key stakeholders were defined as those who would be involved in or affected by the development, implementation, or findings from the evaluation capacity assessment process. It is important to engage stakeholders at various points of the assessment process to promote ownership of the process and results and to ensure the results will be meaningful to those that have the ability to use the information to improve evaluation capacity. After identifying key stakeholders, we discuss how and when these stakeholders should be involved in the evaluation capacity assessment process. For example, Y-USA identified staff in one of their signature programs as stakeholders for the process. We met with these stakeholders during the kickoff meeting, as well as incorporated them into the process by having them vet some of our data collection materials prior to their use.

Strengths, Weaknesses, Opportunities, and Threats (SWOT) Analysis

A SWOT analysis was then performed to identify Y-USA’s current beliefs around where Ys were excelling and struggling with evaluation, as well as where the opportunities and threats were to assessing and building evaluation capacity. This activity was used to break up the interview dialogue, and provided an opportunity to reflect on all of the past evaluation capacity attributes discussed in a summative manner. The results from this process are summarized in Table 3.

Table 3. SWOT Analysis findings.

Strengths	Weaknesses
<ul style="list-style-type: none"> ▲ Invested Y-USA Research and Evaluation Department/national organization interest ▲ Y interest, openness, and awareness of the importance of evaluation ▲ Existing momentum around impact measurement 	<ul style="list-style-type: none"> ▲ Evaluation practices are not integrated or institutionalized ▲ Lack of understanding of what resources would be required for evaluation capacity building/unsure of what can be accomplished ▲ Y-USA lacks the capacity to assist all Ys with evaluation –there needs to be other alternatives or approaches to assisting Ys beyond assistance from Y-USA
Opportunities	Threats
<ul style="list-style-type: none"> ▲ Opportune timing—Y-USA is in the beginning phases of providing evaluation resources to Ys ▲ Large nonprofit with infrastructure to reach many staff in a variety of ways ▲ Ability to communicate about work in new terms ▲ Potential for fundraising 	<ul style="list-style-type: none"> ▲ Competing priorities for time and resources ▲ Ys might not appreciate the benefits of evaluated programs ▲ Ys might feel that they already know what they need to know and are doing well with evaluation (i.e., staff might have difficulty assessing their Y’s evaluation capacity).

Evaluation Capacity Assessment Process

Having learned a significant amount about Y-USA and Y structure and culture, the last part of the meeting was dedicated to planning for the evaluation capacity assessment process itself. This included a discussion of potential data collection tools and conceptual frameworks, which are presented below.

Tools

There are a relatively small number of tools available to measure an organization's evaluation capacity and what tools are available have not undergone validity and reliability testing. We identified several tools specific to this use, including the Bruner Foundation's Modified Evaluative Thinking Assessment Tool (2011), BTW Consultants, Inc.'s Evaluation Capacity Diagnostic Tool (n.d.) and Capacity 4 Health's Monitoring and Evaluation Capacity Assessment Tool (2011). A brief description of each of these instruments was given to Y-USA and is provided below.

- ▲ The Modified Evaluative Thinking Assessment Tool was developed by the Bruner Foundation to assess the extent to which evaluative thinking is present in various areas within an organization. The tool includes 90 questions across 15 categories of organizational capacity, including mission, strategic planning, governance, finance, leadership, fundraising, evaluation, program development, client relationships, communications and marketing, technology acquisition and planning, staff development, human resources, business venture development, and alliances and collaboration. Respondents score each item as present or not present and whether the item is a priority for the organization (high, low, or not a priority). The tool is designed to be completed by the leader of an organization.
- ▲ The Evaluation Capacity Diagnostic Tool was created by BTW Consultants, Inc. as an instrument to assist organizations in assessing their readiness to take on various types of evaluation activities. The tool includes 32 questions around six topic areas: organizational culture and practices around evaluation, organizational commitment and support for evaluation, using data to inform ongoing work, existing evaluation knowledge and experience, developing a conceptual model for designing outcome evaluations, and benchmarks and indicators. Organizations score themselves on a four-point scale (strongly agree to strongly disagree). Responses are then used calculate subsection scores and an overall score that translates to one of four evaluation capacity levels, ranging from need for increased capacity to adequate level of capacity in place. The tool is designed to be completed by a single person within an organization, either through self administration or with the assistance of an external evaluation consultant.
- ▲ The Monitoring and Evaluation Capacity Assessment Tool is an instrument developed by Capacity 4 Health based on five existing organizational capacity assessment tools. This 27-item instrument assesses evaluation capacity according to five key areas: organizational infrastructure, measurement and evaluation (M&E) planning, M&E content and data capacity, M&E systems, and use of M&E. Organizations score themselves on a four-point scale ranging from poor to excellent. The tool is designed to be completed in a group setting to gather multiple perspectives and to create a forum for dialogue around evaluation capacity attributes. After completing the tool, the group is instructed to complete a "Next Steps" capacity building worksheet to identify action steps, barriers, and ways to address barriers in areas where the organization is weak.

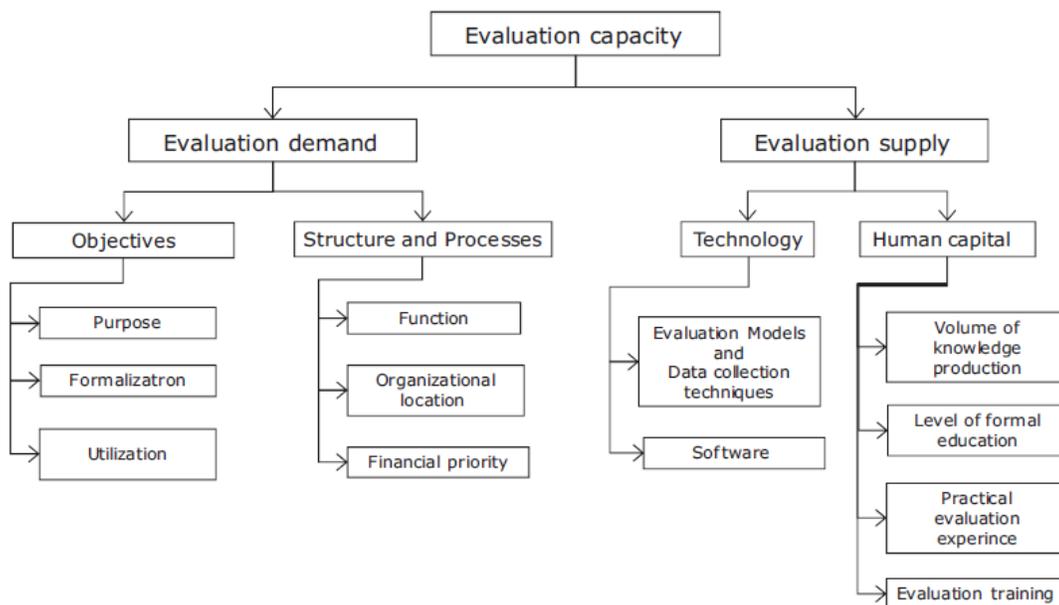
Content areas across all surveys are similar and commonly addressed the same types of organizational attributes (e.g., staff knowledge, funding for evaluation). The primary differences among these tools are the level of complexity and specificity of the questions and the length of the instrument. Y-USA was interested in conducting a broad assessment of the Y movement, so consideration needed to be made for how the final instrument could be completed by numerous Y staff across multiple geographic areas with varying degrees of evaluation experience and expertise. Ultimately, the BTW tool was selected because of its moderate length, but specific questions around key areas of inquiry desired by Y-USA to answer their evaluation questions for the project. The Bruner Foundation tool

was determined to be too long given these parameters. The Capacity 4 Health Tool was determined to not address enough specifics about the evaluation process and organizational structures to achieve Y-USA’s goals, and was not feasible to implement in a group setting as intended.

Frameworks

We sought to identify an approach that would serve as the foundation and guiding framework for the Y evaluation capacity assessment process. Several potential models were identified by Altarum and presented for consideration to Y-USA at the kickoff meeting for the project, including approaches developed by Bruner Foundation (2011), McKinsey (2011), Nielson and colleagues (2011), and the Urban Institute and The Center for What Works (2006). After discussing the Y’s organizational structure and culture, current evaluation activities, and their goals for the project, the group quickly coalesced around a conceptual model articulated by Nielson and colleagues as the framework for the Y evaluation capacity assessment process (Nielson, Lemire, and Skov, 2011). Drawing from literature in organizational theory, as well as the growing body of knowledge in evaluation capacity and evaluation capacity building, the authors constructed an evaluation capacity model that depicts evaluation capacity in terms of evaluation supply and demand (Figure 3). This model was selected because the elements and organization of the model resonated with Y-USA and their view of evaluation capacity.

Figure 3. Conceptual model for measuring evaluation capacity.



Source: Nielson, Lemire, and Skov, 2011

The evaluation demand side of the model reflects the authors’ positing that evaluation capacity of an organization is intrinsically linked to the general functions of an organization, such as objectives and structures and processes. Objectives include the organization’s reasons for conducting evaluations, formality of evaluation practices, and use of evaluation findings. Structures and processes are related to the infusion of evaluation in the organization’s activities, including the establishment of an evaluation function, location, and finances within an organization. The evaluation supply side of the model is the foundation for developing and implementing evaluations, and includes technology and human capital. Technology comprises evaluation models and data collection techniques, as well as software systems available for entering, managing, analyzing, and storing data. Human capital

includes knowledge gained through conducting evaluation activities, formal education of staff, and staff experience in conducting evaluations. Definitions of all elements in the Nielson conceptual model are presented in Table 4.

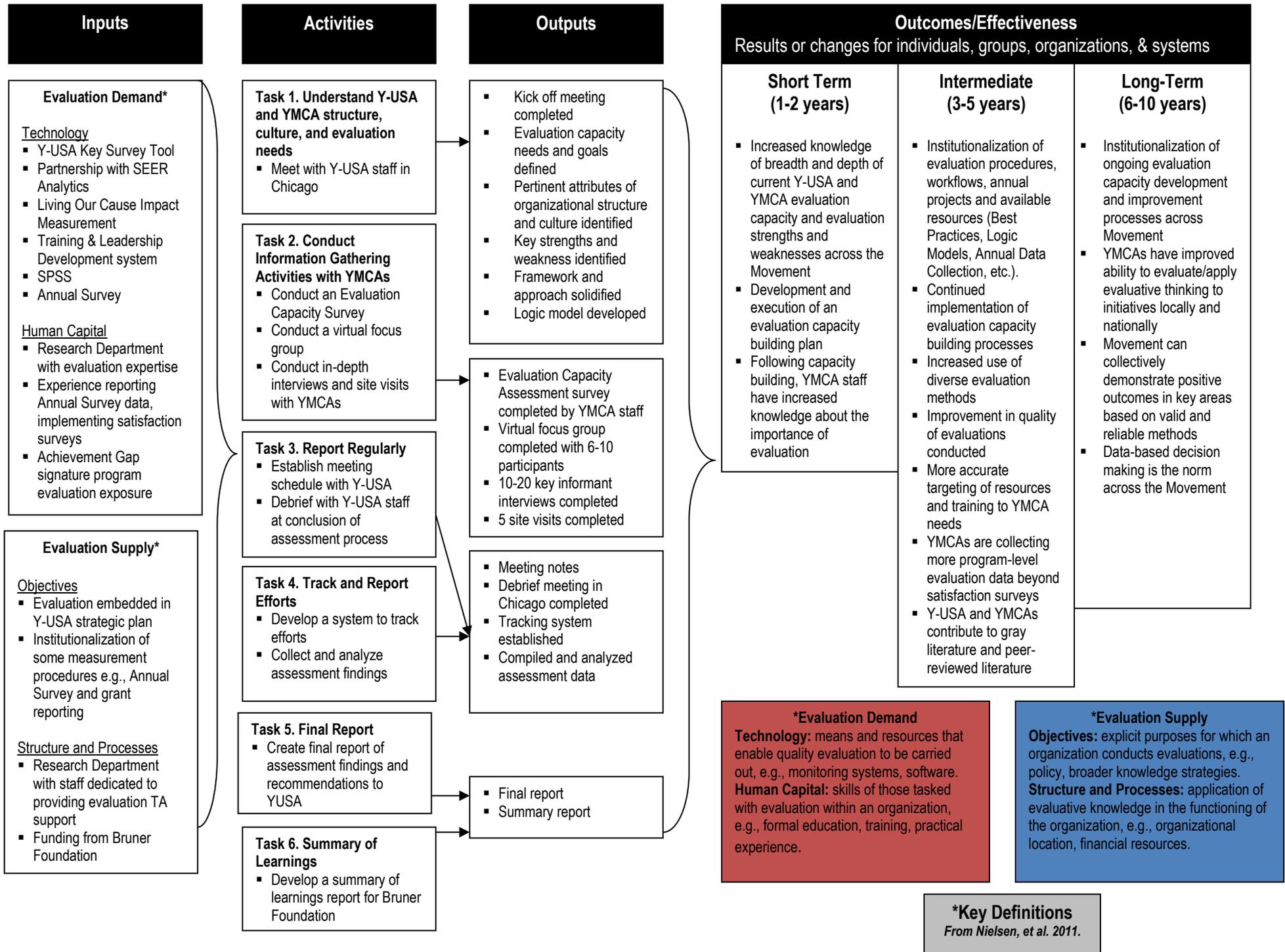
Table 4. Operational definitions for conceptual model (Nielson, Lemire, and Skov, 2011).

Conceptual Model Element	Definition
Evaluation Demand	
Objectives	
Purpose	Purpose for conducting evaluation
Formalization	Formalized evaluation policy, guidelines, and annual plans
Utilization	Utilization of evaluation practices and findings in decision-making
Structures and Processes	
Function	Whether a specialized evaluation function is established
Organizational location	Location of evaluation specialization
Financial priority	Financial resources dedicated to evaluation
Evaluation Supply	
Technology	
Evaluation models and data collection techniques	Codification of evaluation models and techniques in evaluation manuals
Software	Monitoring systems or software for quantitative and qualitative data collection and analysis
Human Capital	
Volume of knowledge production	Spillover, volume of other knowledge production as data collection and analytic techniques are similar
Level of formal education	Educational degrees and formal training in evaluation
Practical evaluation experience	Experience managing and conducting evaluations

Logic Model Development

After concluding the meeting, we used the information gathered from the meeting to develop a logic model (Figure 4) depicting our process and Y-USA's goals. The inputs reflect the organization's existing capacity in the evaluation demand and supply terms from our conceptual framework. These inputs were filled in using data gathered from the kickoff meeting and review of materials provided by Y-USA. The activities articulate the evaluation capacity assessment process. Outputs are the measureable results from our activities. Finally, we articulated short, intermediate, and long-term outcomes of the evaluation capacity assessment process, based on our conversations at the kickoff meeting about goals for the project. This logic model then served as a roadmap for the evaluation capacity assessment project, articulating our process and end goals.

Figure 4. YMCA Evaluation Capacity Assessment Logic Model



Lessons Learned

Our one-day kick off meeting with Y-USA was extremely useful in gathering information about how Y-USA and Ys operate, feel about evaluation, and use evaluation. We learned about potential pitfalls to consider when implementing our assessment process, such as Y concerns about feeling tested and large variance in Y staff experience and use of evaluation. We also learned about strengths of the organization that we could use to benefit the assessment process, such as survey software that was commonly used to gather information across the organization and a previously performed segmentation analysis of Ys that suggested many were interested in performance measurement and improvement.

In addition, it was critical that we defined evaluation capacity for the Y and identified a conceptual framework that embodied the attributes identified to be most important in possessing this capacity. This was important for ensuring Y-USA and Altarum shared a common understanding about what evaluation capacity entails for the Y movement, and what aspects of evaluation capacity are considered most relevant to the Y. To facilitate this process, it was useful to present varying definitions of evaluation capacity, along with a variety of tools and approaches that could illustrate different ways to view this work, and select the combination of these that aligned with the goals for this project. Creating a logic model helped to ensure that Y-USA and Altarum staff had a common understanding about how the project would operate and what the common goals were for the project.

What made this process challenging was the need to cover so many topics in a short period of time. Due to the time constraints of the project (7 months duration), the assessment process had to begin almost immediately in order to collect, analyze, and report on the data. Also, because Altarum and Y-USA are located in different states, we were not able to conduct the meeting across multiple days which may have lessened the burden of covering a vast amount of topics in one sitting. Instead, we established a regular weekly conference call schedule that offered us more time for discussion as the project unfolded.

Step 2: Evaluation Capacity Assessment Survey

The next step in our evaluation capacity assessment process was implementing the YMCA Evaluation Capacity Assessment Survey. We selected a survey tool as our first data collection method because surveys are a low-cost strategy for gathering substantial data from a large population. When assessing the evaluation capacity of an organization as large and vast as the Y, we determined that a survey was the only viable option to obtain a representative view of the entire organization’s evaluative capacity within our time and resource constraints.

Methods

Data Collection Tool

The YMCA Evaluation Capacity Assessment Survey (see Appendix A) was based off of the Evaluation Capacity Diagnostic Tool (BTW Consultants, Inc, n.d.). Slight revisions were made to the tool to align with our intent and audience. Instructions were added to describe the purpose of the survey and to provide definitions for evaluation and outcome evaluation (Figure 5). These definitions were added to clarify the use of these words in the instrument. To ensure that the survey would resonate with Y staff, we worked in collaboration with Y-USA to make minor changes to the survey language. For example, instead of questions asking about “our organization”, the survey asked about “our Y”. We were careful not to change the intent of any of the questions in order to maintain the integrity of the tool. Along those lines, we also chose to include all of the original survey questions; however, we added questions to gather additional information that was important to Y-USA and that was not addressed adequately in the tool. Four of these added questions were close-ended survey questions with likert-scale response options, asked in a manner consistent with the rest of the tool. Other questions (both closed and open-ended) were added at the end of the tool to gather specific information related to the use of particular evaluation tools, existence of data management systems, use of logic models, and conduct of outcome evaluations beyond participant satisfaction. For analysis purposes, these additional questions were excluded to maintain the scoring rubric outlined by the original tool. Once adaptations were completed, the survey was pre-tested by a hand-selected sample of Ys; their feedback was discussed via conference call. Minor revisions to survey language were made following this call.

Figure 5. Evaluation definitions provided at the beginning of the survey.

Evaluation concepts sometimes mean different things to different people. So that we are all on the same page and you can answer consistently with your fellow YMCA peers, please use the following definition:

Evaluation is the process of examining and rating something based on important features. You can evaluate a program, process, policy or initiative.

Program Evaluation typically includes outcome and process evaluation.

- *OUTCOME evaluation* assesses changes in knowledge, attitudes, and behaviors as a result of participating in the program.
- *PROCESS evaluation* monitors program environment, as well as if activities were completed and how well they were implemented.

Sampling

Our sampling strategy involved selecting all staff at Ys who had a role in the evaluation process, either in a leadership, implementation, analysis, or reporting role. Y-USA staff familiar with the job roles of Y staff selected a group of job codes thought to participate in these roles, producing a sample of 10,539 staff. Because sampling was by job code instead of by Y, there was potential for multiple individuals from the same Y to participate in the survey. The survey was administered electronically via Y-USA’s survey software. All potential participants received an email and survey link from Y-USA asking them to complete the survey. Our goal response rate for the survey was 10%. Although this response rate is generally considered to be low in survey research, Y-USA had significant experience conducting surveys organization-wide and reported that this was considered an adequate response rate for their organization. To achieve this response rate, Y-USA sent follow up messages to participants to remind them to complete the survey, and extended the open date of the survey several days beyond the initial time period to allow additional respondents to complete it. The survey was open from June 18-July 16, 2012 and was completed by 1,378 individuals, a 13% response rate.

Analysis

Survey results were exported into Microsoft Excel where data were cleaned. Respondents who skipped one or more questions from the original BTW tool (but not those who skipped questions added by Y-USA), were excluded from analysis. This resulted in the removal of 265 (19%) respondents, leaving 1,113 surveys for analysis. Using the BTW scoring rubric, data were analyzed and evaluation capacity scores calculated for each respondent. Respondents were then grouped into four capacity levels—need for increased capacity; emerging level of capacity in place; moderate level of capacity in place; adequate level of capacity in place (Table 5).

Table 5. BTW Consultants, Inc. scoring rubric used for categorizing Y evaluation capacity.

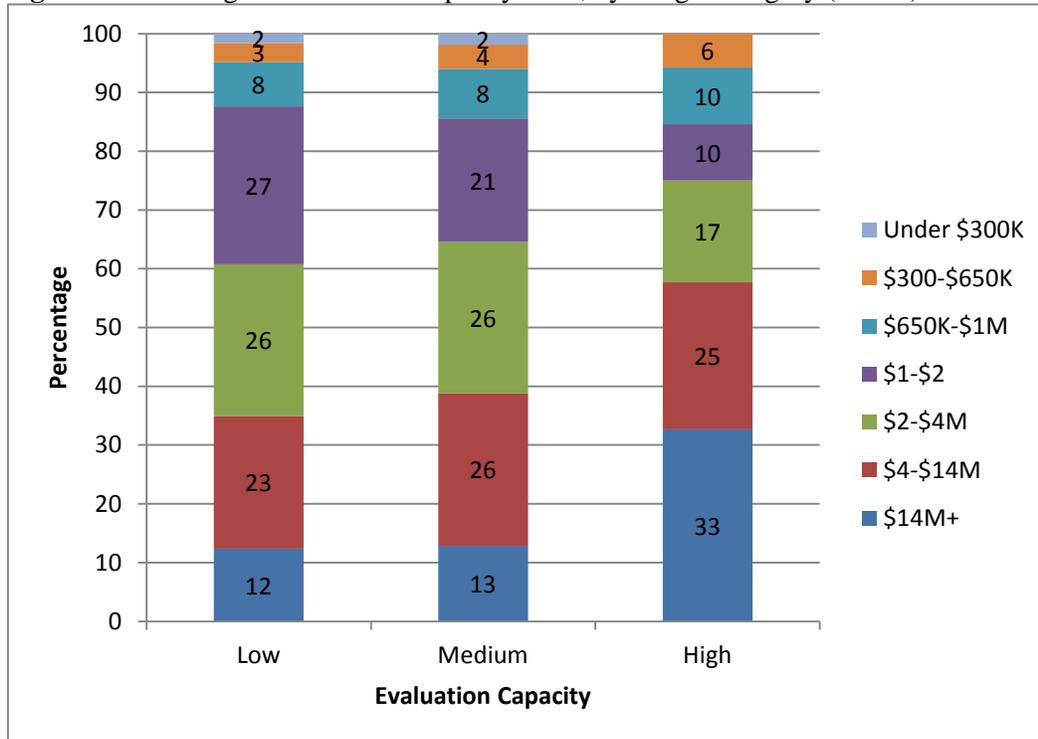
Score	1.00-1.51	1.52-2.49	2.50-3.48	3.49-4.00
Capacity Level	Need for increased capacity	Emerging level of capacity in place	Moderate level of capacity in place	Adequate level of capacity in place

After analysis, the two lowest capacity levels—need for increased capacity and emerging level of capacity in place— were combined into a single category of “low capacity” to better align with our assessment process (i.e., the use of an additional three data collection methods). Respondents were distributed into the following categories:

- ▲ Low Capacity: 29.6% (n=329)
- ▲ Medium Capacity: 61.4% (n=684)
- ▲ High Capacity: 9.0% (n=100)

This categorization was used to determine appropriate Ys to be included in the subsequent methods—a virtual focus group (low capacity), key informant interviews (medium capacity), and site visits (high capacity). Figure 6 shows a breakdown of these groups by budget category. Branch Ys are excluded from this group because budget categories are only assigned at the association level.

Figure 6. Percentage of Ys at each capacity level, by budget category (n=557)



Results

We reported our results in tabular and graphic form. We often grouped those who agreed with particular statements (survey responses agree or strongly agree) versus those who disagreed (survey responses disagree or strongly disagree) in order to more clearly see trends in capacities that Ys had or did not have. Below is a sample of the results from the survey.

Across all groups, respondents were in strongest agreement regarding organizational practices being supportive of evaluation. The overwhelming majority of respondents (89% agreed or strongly agreed) felt their organizational leaders support and value program evaluation and evaluative thinking. The majority (78% agreed) indicated their Y uses evaluation findings to modify its course of action when making decisions that are organization-wide and program specific. Questions on whether Ys prioritize evaluation by providing financial support (50% disagreed or strongly disagreed) or included a budget line item to ensure ongoing evaluation activities (65% disagreed strongly agreed) elicited some of the highest rates of disagreement.

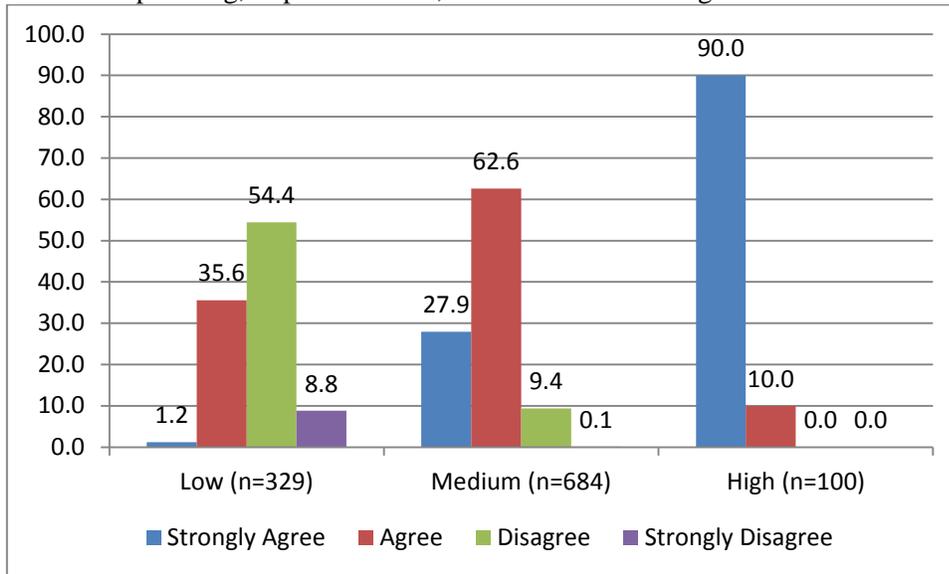
In terms of staff knowledge and experience, the majority of respondents were confident their Y staff have a basic understanding of evaluation (77% agreed or strongly agreed), know how to analyze and interpret data (67% agreed or strongly agreed), and can develop recommendations based on evaluation findings (68% agreed or strongly agreed). While most (78% agreed or strongly agreed) regularly assess member satisfaction; far fewer (57% agreed or strongly agreed) regularly assess member outcomes like changes in knowledge, skills, attitudes, and behaviors. Almost half of respondents (46% disagreed or strongly disagreed) did not feel their Ys identified indicators for measuring the impact of their work. When asked if their Ys have the appropriate capacity, systems, and expertise in place to ensure that high-quality, purpose-driven evaluation and measurement are practiced, half of respondents disagreed.

The following sections provide examples of results for each part of the survey.

Organizational Culture and Practices around Evaluation

The high and medium capacity respondents were in overwhelming agreement on questions relating to their organizational practices and culture. All of the respondents rating their Y as high capacity agreed or strongly agreed and 99% medium capacity respondents agreed or strongly agreed that evaluation is seen as integral to their work. In contrast, far fewer of the low capacity respondents agreed with this statement (62% low capacity respondents agreed or strongly agreed). When asked if their Y involves program staff and leaders in evaluation planning, implementation, and discussion of findings, there was strong consensus across the high and medium capacity Ys (100% high capacity respondents; 91% medium capacity respondents agreed or strongly agreed) whereas only one third of low capacity respondents agreed (37% low capacity respondents agreed or strongly agreed) (Figure 7).

Figure 7. Percent of low, medium, and high capacity Ys agreeing or disagreeing with the statement, "The organization involves program staff and organizational leaders (as appropriate) in meaningful ways in evaluation planning, implementation, and discussion findings."

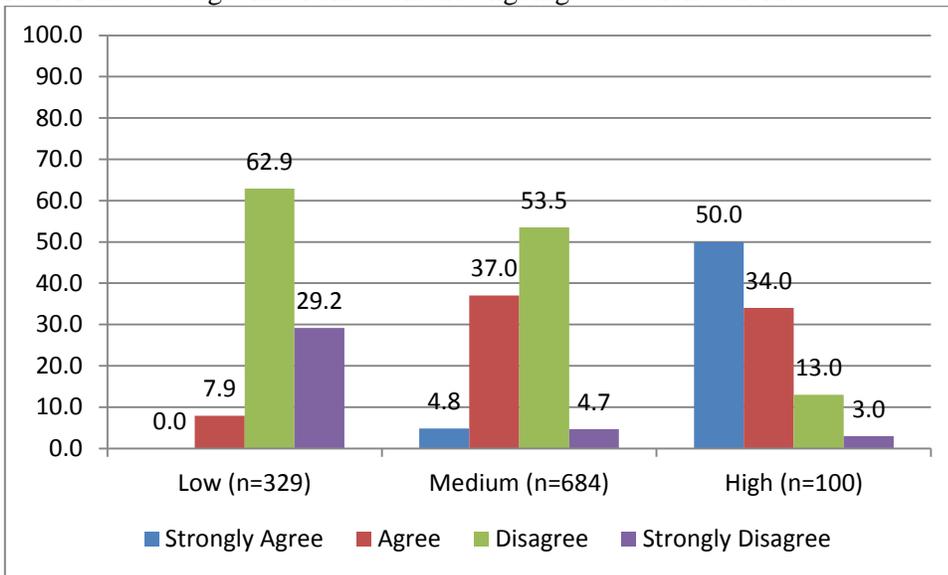


Organizational Commitment and Support for Evaluation

Although the majority of survey respondents felt their organizational leaders support and value program evaluation and evaluative thinking, fewer believed that their Y established clear expectations for staff roles related to evaluation. In the practices of establishing clear expectations for the evaluation roles of different staff and ensuring staff have the necessary information and skills to support evaluation efforts, high capacity respondents reported much higher rates of agreement (99% agreed 100% agreed or strongly agreed, respectively) compared to the medium (74% and 79% agreed or strongly agreed, respectively) and low capacity (13% and 16% agreed or strongly agreed, respectively).

Rates of agreement fell across the board in response to questions about financial support for evaluation, which reflected one of the biggest struggles for local Ys. When asked whether their Y has a budget line item to support ongoing evaluation activities, the majority of high capacity respondents (84%) agreed or strongly agreed but among medium capacity respondents, less than half (42%) agreed or strongly agreed, and even fewer low capacity respondents (8%) agreed or strongly agreed with this statement (Figure 8).

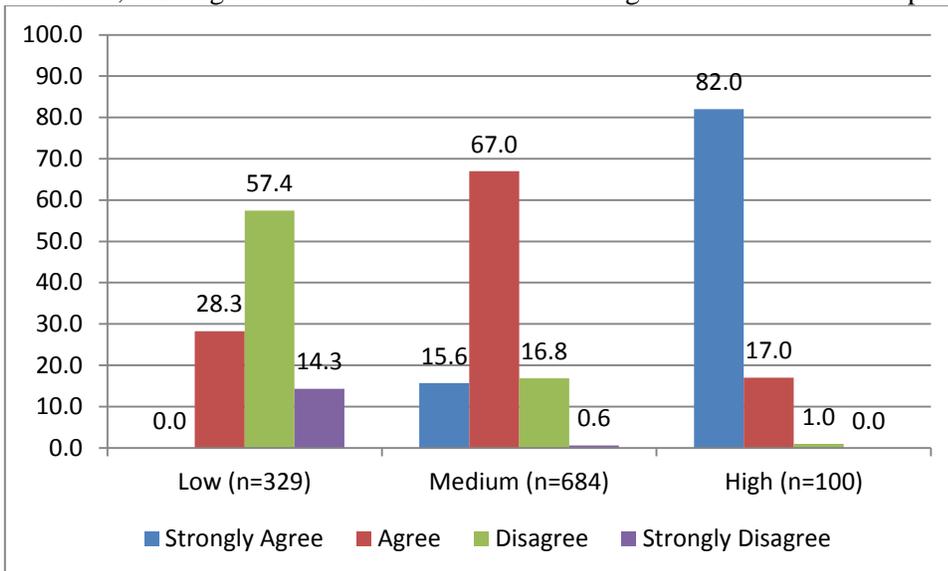
Figure 8. Percent of low, medium, and high capacity Ys agreeing or disagreeing with the statement, "Our local Y has a budget line item to ensure ongoing evaluation activities."



Using Data to Inform Ongoing Work

A few survey questions inquired about the extent to which Ys use data to inform ongoing work. With respect to using evaluation data to set staff goals and evaluate staff performance, rates of agreement dropped. Among high capacity respondents, almost all (99% agreed or strongly agreed) reported they incorporated this practice; the majority of medium capacity respondents (83% agreed or strongly agreed) did so compared to only one-third of low capacity respondents (28% agreed or strongly agreed) (Figure 9).

Figure 9. Percent of low, medium, and high capacity organizations agreeing or disagreeing with the statement, "Managers use evaluation data to set staff goals and evaluate staff performance."



Existing Evaluation Knowledge and Experience

Both high and medium capacity respondents were confident in their staff's knowledge and experience. They believe (99% of high capacity respondents agreed or strongly agreed; 91% of medium capacity respondents

agreed or strongly agreed) their staff have a basic understanding of evaluation; less than half of low capacity respondents (44% agreed or strongly agreed) felt the same (Figure 10). When asked specifically about their staff's experience developing data collection tools and collecting data utilizing a variety of strategies, rates of agreement decreased. High capacity respondent rates remained high (97% agreed or strongly agreed) but medium capacity (67% agreed or strongly agreed) and low capacity respondents (20% agreed or strongly agreed) reported less staff experience related to data collection (Figure 11).

Figure 10. Percent of low, medium, and high capacity Ys agreeing or disagreeing with the statement, "Our local Y has staff that have a basic understanding of evaluation (e.g., key evaluation terms, concepts, theories, assumptions)."

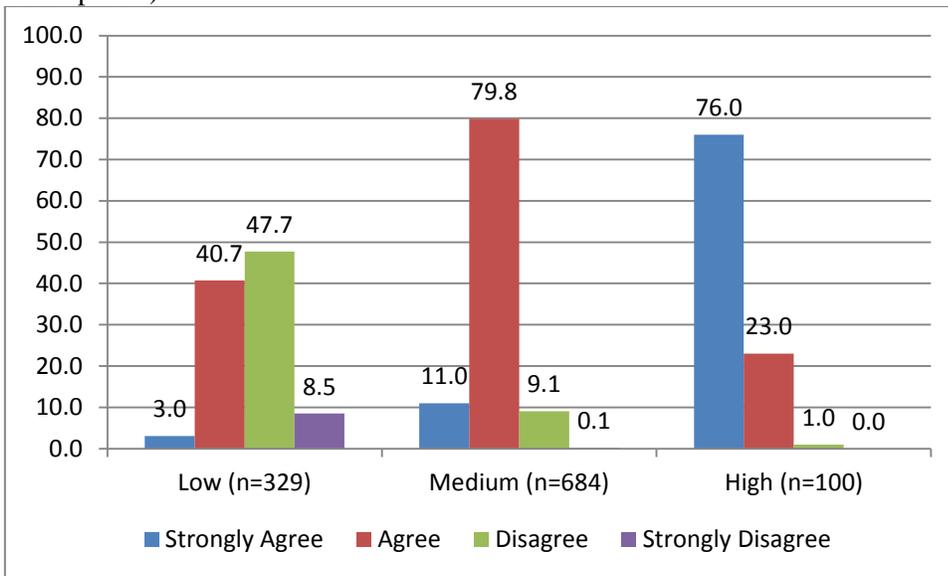
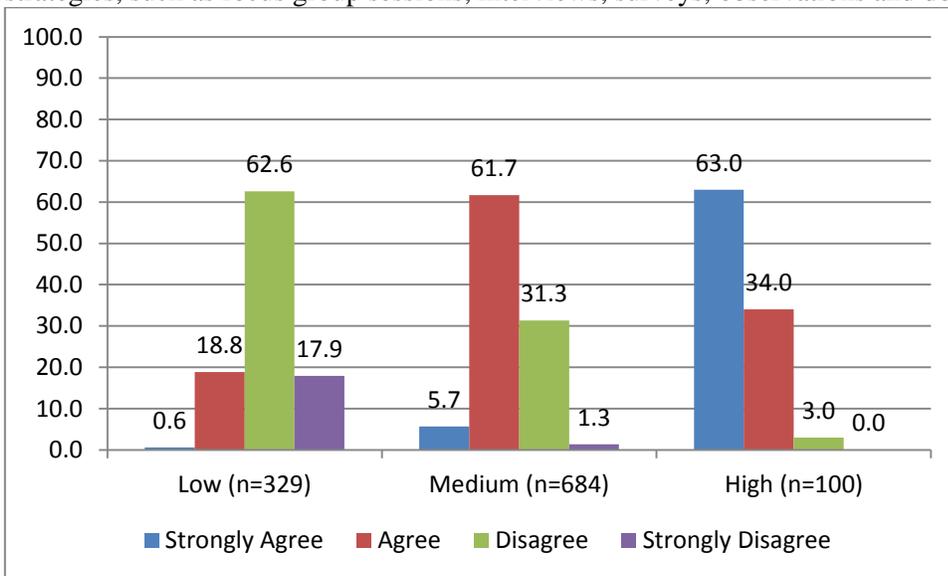


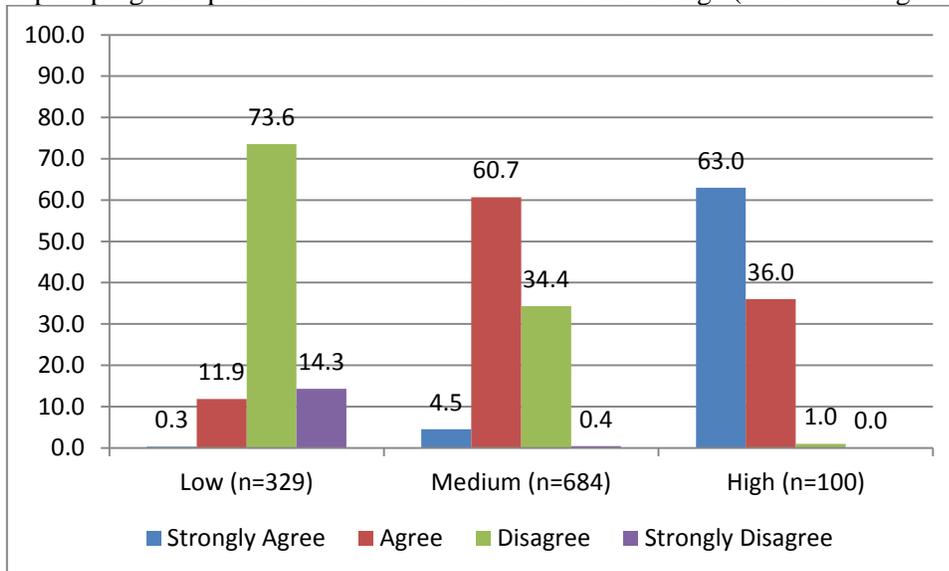
Figure 11. Percent of low, medium, and high capacity Ys agreeing or disagreeing with the statement, "Our local Y has staff with experience developing data collection tools and collecting data utilizing a variety of strategies, such as focus group sessions, interviews, surveys, observations and document reviews."



Developing a Conceptual Model for Designing Outcome Evaluations

Several survey questions inquired about how Ys clarify and define their program outcomes. In terms of being able to articulate how change is expected to occur among program participants (such as through a logic model) almost all (99%); of high capacity respondents agreed or strongly agreed, the majority (65%) of medium capacity respondents agreed or strongly agreed, and few (12%) low capacity respondents agreed or strongly agreed (Figure 12). A greater percentage of respondents were more confident in their ability to clarify what outcomes they want to accomplish in the short term (e.g., one to three years) and what success will look like. All the high capacity respondents agreed or strongly agreed; the majority (80%) of medium capacity respondents agreed, and about one-third (29%) of low capacity respondents agreed. A comparable percentage of respondents felt they also had the tools and methods for evaluating outcomes (99% high capacity respondents; 82% medium capacity respondents; and 22% low capacity respondents agreed or strongly agreed).

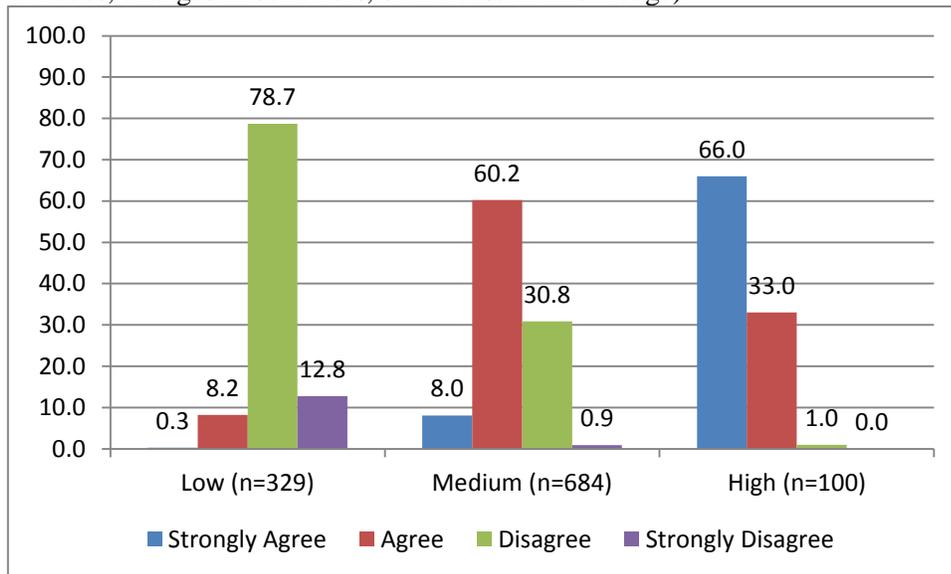
Figure 12. Percent of low, medium, and high capacity Ys agreeing or disagreeing with the statement, "Our local Y has articulated how we expect change to occur among individuals in our programs, and how we expect program specific activities to contribute to this change (such as through a logic model)."



Defining Benchmarks and Indicators

The majority of all respondents across capacity groups regularly assessed member satisfaction (98% high capacity respondents agreed; 87% medium capacity respondents agreed; 52% low capacity respondents agreed). Fewer reported assessing member outcomes such as changes in knowledge, skills, attitudes, behaviors (95% high capacity respondents agreed; 70% medium capacity respondents agreed; 17% low capacity respondents agreed). When asked about whether respondents have identified relevant outcome indicators, all high capacity respondents agreed (100%); the majority of medium capacity respondents agreed (80%) and about a quarter of low capacity respondents agreed (23%) (Figure 13). On a related question on the ability to identify appropriate indicators for measuring impact, rates of agreement decreased (99% high capacity respondents agreed; 68% medium capacity respondents agreed; 9% low capacity respondents agreed).

Figure 13. Percent of low, medium, and high capacity Ys agreeing or disagreeing with the statement, "Our local Y has identified what indicators are appropriate for measuring the impact of our work (e.g., changes in attitudes, changes in behaviors, or increase in knowledge)."



Additional Questions

As noted above, Y-USA added several questions to the survey to gather specific information of interest. Use of logic models was low (32%), with no significant differences across levels of evaluation capacity. Similarly, evaluating individual outcomes beyond satisfaction was low (29%) with no marked differences across evaluation capacity levels.

Lessons Learned

The YMCA Evaluation Capacity Assessment Survey was a good tool for getting a broad view of the organization's capacity. We were able to gather information on evaluation capacity from a variety of staff at Ys across the country and in Ys of various sizes in a relatively short period of time. Further, the scoring of this survey helped us to segment the audience into groups of Ys with similar evaluation capacity. This broke down a large population into more manageable groupings upon which we could conduct further assessment. The low, medium, and high capacity designations assigned through this survey were the basis of our sampling for all other data collection methods.

In terms of barriers, we had concerns about the validity of the survey results. Knowing that evaluation and evaluation capacity may not be well understood by many staff, we were concerned about respondents' ability to self-assess their organization on a variety of evaluative capacity attributes. Also, knowing respondents' may select the most socially desirable answer, we were somewhat concerned that Ys would rate their evaluation capacity higher than it actually was. Y-USA also expressed concerns that, conversely, higher capacity Ys might rate themselves lower because they understood evaluation better and would more readily recognize their Y's weakness in evaluation. To address these barriers and concerns, our evaluation capacity assessment process utilized a mixed methods approach that involved numerous data collection efforts designed to gather more in-depth information that would help us to assess whether these concerns were warranted.

We also faced challenges in analyzing open-ended questions. Some of the questions were not well understood by many respondents, who entered inappropriate responses in some cases (e.g., when asked for

the name of the data collection tool, a respondent wrote, “multiple choice”). This may have resulted from a lack of understanding of evaluation in general or poorly worded/defined questions. Because of the large number of respondents, these types of questions yielded large volumes of data that took significant time to code, sort, and analyze, but did not offer many insights in terms of evaluation capacity. Based on this experience, we would not recommend including open-ended response questions in surveys distributed to large audiences, or limiting such questions to one or two very specific questions that are very clearly worded/defined.

Step 3: Virtual Focus Group

A virtual focus group was conducted with staff at Ys categorized as low evaluation capacity based on the YMCA Evaluation Capacity Assessment Survey results. The focus group was intended to help us understand the factors that influence evaluation practices at lower capacity. Much like traditional focus groups, virtual focus groups are an excellent method for gathering insight on perceptions and beliefs. The group dynamic of this technique helps to facilitate lively discussions as individuals are inspired by or build off one another's experiences, ideas, and comments. A virtual group also permits gathering rich qualitative data from multiple staff in a short time frame at minimal cost compared to other qualitative data collection methods. This data collection approach was specifically selected to be used for collecting information on Ys categorized as low evaluation capacity because we believed that a group discussion format would encourage participants to build off one another's responses if they are less familiar and have less experience with evaluation topics. We were purposeful in populating the group exclusively with Ys rated as lower capacity because we posited that Ys of similar capacity would be more comfortable in candidly sharing their evaluation experiences, both positive and negative, among peers who may have similar experiences and perceptions about evaluation.

Methods

We first chose a date to hold the virtual focus group which would allow approximately three weeks time to recruit participants. We elected to schedule the group for 2:00 pm ET, which is a standard time that Y-USA uses for their webinars because it is convenient for staff in all time zones.

Sampling

Our recruitment goal for the virtual focus group was 8-10 individuals. Based on our experience, we have found this to be an ideal group size that allows sufficient time for each participant to contribute to the discussion and is also a manageable size for the facilitator. After categorizing Y respondents according to their score on the Evaluation Capacity Assessment Survey, a list of all respondents that categorized their Y as low evaluation capacity (survey score 1-2.49, n=329) was obtained. To minimize bias in the selection process, the list of respondents was uploaded into a computerized random list generator to produce a random permutation of respondents. We then selected participants from the randomized list until we were confident we could achieve our target participation rate. We purposefully over-recruited participants, assuming not everyone would be available on the given day and time. We also factored in Y-USA feedback as to how many individuals we would need to recruit to reach our target number of participants, based on past history hosting meetings.

If more than one participant was selected from the same Y, they were excluded so that no Y could have more than one participant in the focus group. We believed that having more than one representative from a given Y may skew results towards a single Y's experience with evaluation, and that Ys may not feel comfortable openly expressing their viewpoints and on experiences with evaluation in the company of their colleagues and supervisors. A total of 49 respondents that were selected, 14 were excluded due to having another representative from the same Y already selected. The remaining 35 respondents were recruited to participate via email invitation from Y-USA. We chose to have Y-USA send out the email invitation because Y staff are accustomed to receiving invitations from Y-USA staff, whereas they may not recognize an invitation from Altarum staff. On the day of the focus group, 9 individuals (26% response rate) logged into the system but only 8 individuals were active participants.

Facilitation

The focus group lasted 90 minutes. Once participants logged into Adobe Connect, they were able to view content on their computer screen and complete real-time poll questions. We felt that presenting content for participants to view would be more effective in keeping them engaged and less likely to multi-task.

We started the group by introducing Altarum and providing some background on the project as well as clarifying the purpose of the focus group. We reviewed some ground rules for the discussion and features of the web platform, Adobe Connect. To help manage the discussion, we asked participants to click on the icon for ‘hand raising’ when they wanted to contribute to the discussion. Next we had everyone introduce themselves and reveal a little about their experience with evaluation through some ice breaker questions (Figure 14).

Figure 14. Ice breaker questions.

Let's get started. I'd like to start by having each of you share a little about yourself.

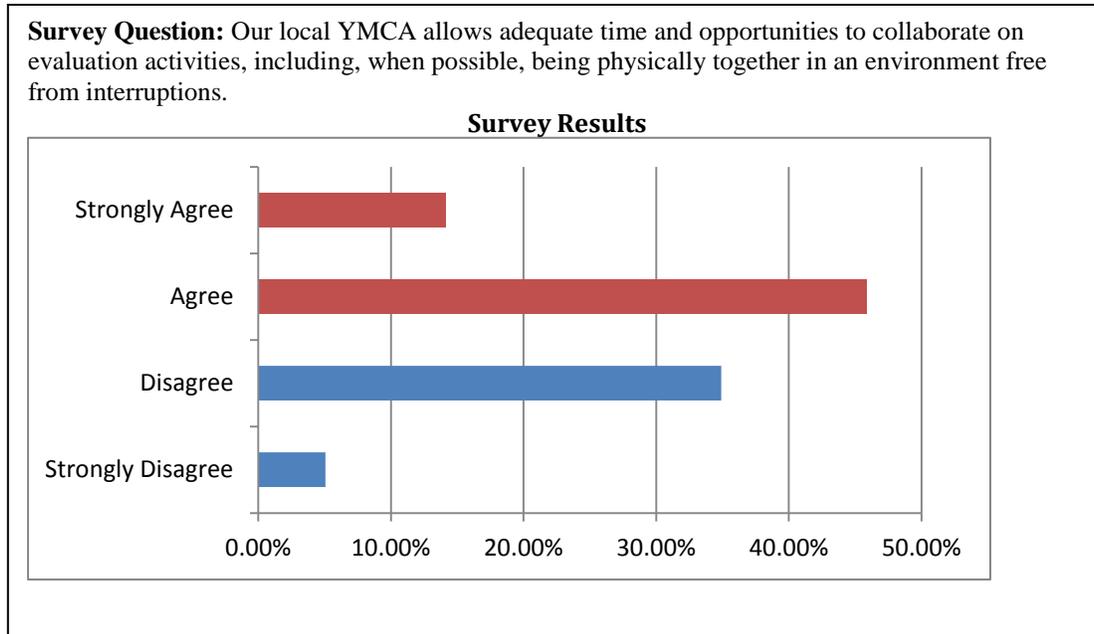
1. My name is _____.
2. I work for the _____ YMCA located in _____.
3. My experience with evaluation has been _____.

After introductions, we reviewed the working definitions related to the topic at hand—evaluation capacity, evaluation, program evaluation. The intent was to help everyone be on the same page. Next, we moved onto a series of guided questions contained in a discussion guide (see Appendix B) that was used to facilitate the group. This guide was designed to build on information collected through the YMCA Evaluation Capacity Assessment Survey.

The focus group covered the following topic areas: organizational culture and support; capacity; developing conceptual framework; and defining benchmarks and indicators. Survey results were presented to the group at various points as a basis for discussion, with the facilitator following up with targeted questions to gather additional information and insights. Because we did not have time to discuss all the survey questions, we selected seven survey questions to feature during the focus group. We selected questions in each of the topic areas with the highest rates of agreement or disagreement.

For each question that was discussed, we presented the survey results and then the facilitator directed some follow up questions to the group. A few examples follow. Figure 15 is an example of a slide that was presented early in the focus group showing the overall results for the survey question related to collaboration.

Figure 15. Focus Group Slide: Collaboration around Evaluation

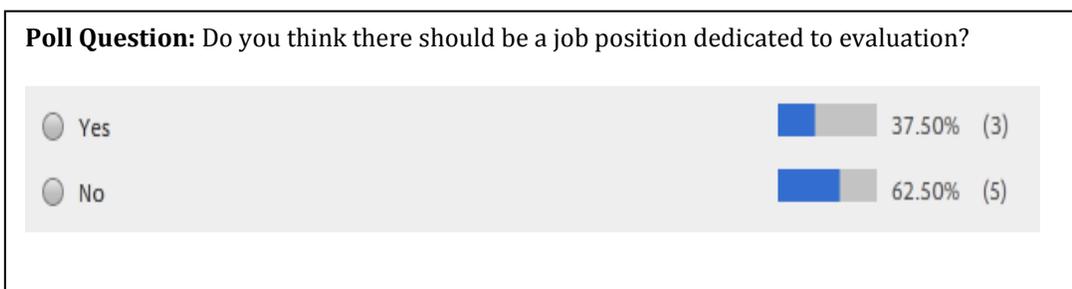


Once this slide was shared with the group, the facilitator followed with several questions related to the topic of collaboration.

- ▲ How are evaluation activities carried out? Is there one person who does this in isolation? Or, as a group?
- ▲ How well does this process work?
- ▲ What are the barriers to collaborating on evaluation?

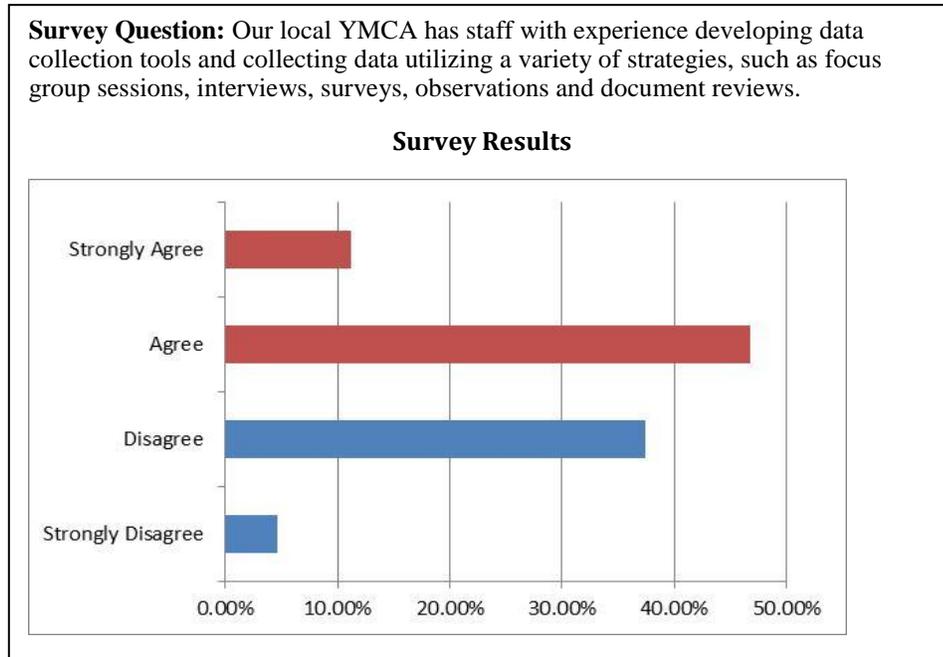
Polls were also used to gather additional information from all respondents. A poll was used while discussing the topic of expectations and collaboration to see if participants thought there should be a staff position dedicated to evaluation. Figure 16 shows the results of that poll.

Figure 16. Focus Group Poll



Following the discussion on organizational support, the group transitioned to the topic of capacity. We discussed the kinds of knowledge and experience necessary to do this work and to make appropriate evaluation decisions, such as deciding what data collection methods to use. Figure 17 is the slide showing the results for the survey question related to whether staff have experience developing various data collection tools.

Figure 17. Focus Group Slide: Staff Experience

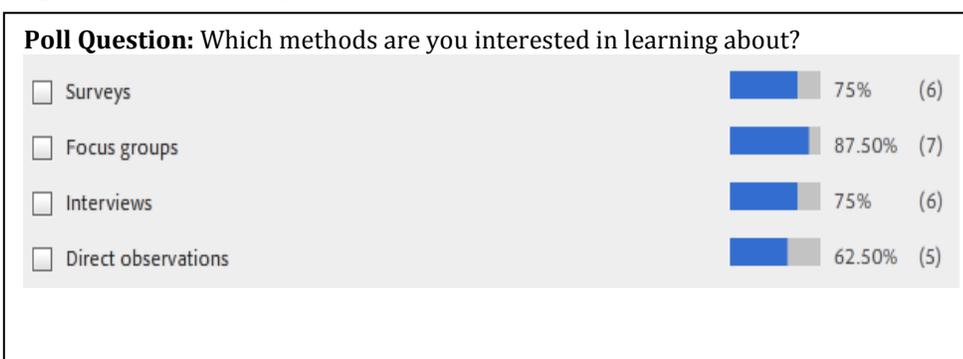


After presenting this slide, the facilitator asked the following questions of the group.

- ▲ In thinking about your local Y, what might be some of the reasons behind this?
- ▲ Do you think this is an issue of training?
- ▲ Are there other ways to improve knowledge about data collection methods?

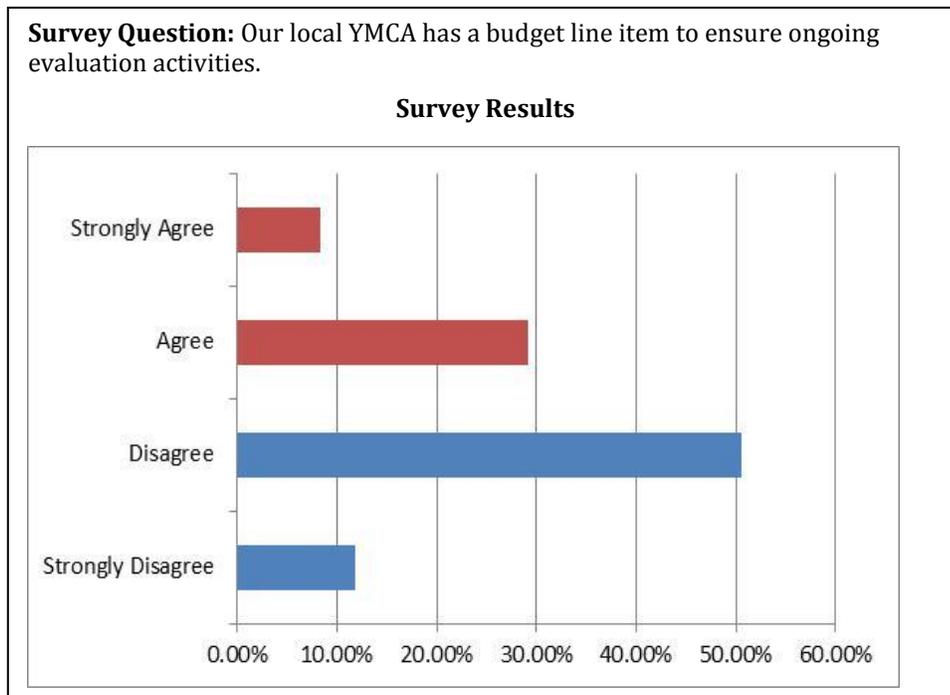
After discussing these questions, the facilitator used another poll to get more information about the types of training in data collection methods that would be useful. Figure 18 shows those results.

Figure 18. Focus Group Poll



Beyond staff, we were also interested in learning about other aspects of capacity, like the ability to support evaluation financially. We presented the survey results regarding whether Y's have a budget line item to support evaluation activities. This slide is shown in Figure 19.

Figure 19. Focus Group Slide: Financial Support



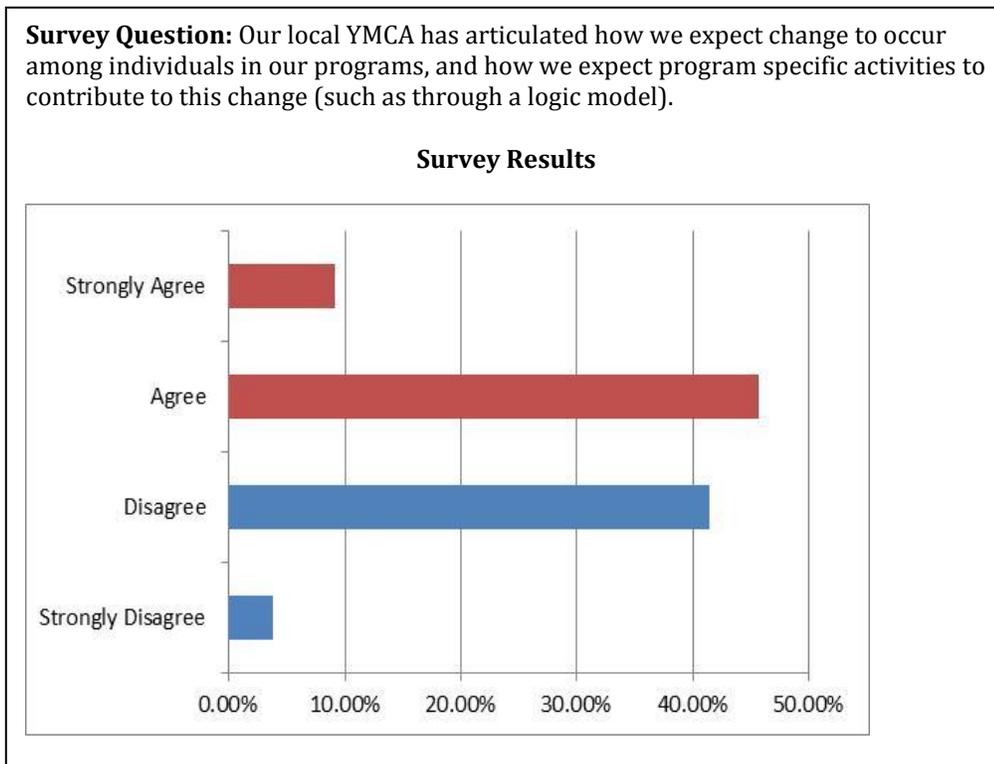
While participants were viewing this slide, the facilitator asked the following questions:

- ▲ Do you typically have funds allocated to evaluate your programs?
- ▲ If no, why do you think sufficient funds are not allocated?
- ▲ Why do you think evaluation is not a funding priority?
- ▲ What is the effect on your work of not having sufficient evaluation funding?

The next topic for discussion was developing conceptual models. We were interested in learning whether Y's use conceptual models when designing their evaluations. We were particularly interested in learning if Y's use logic models as well. Figure 20 showing the survey results to this question was on screen while the facilitator asked the following questions:

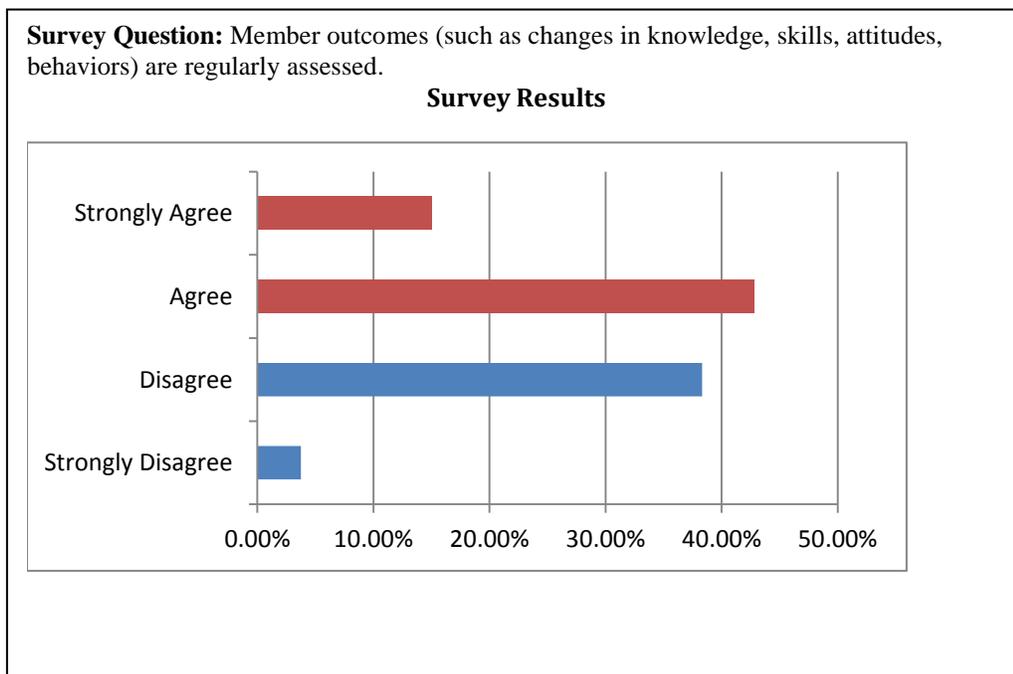
- ▲ When we talk about making a connection between activities and changes, what does this mean to you?
- ▲ In thinking about the program we just saw, does your Y use logic models or something else, describing the changing you're expecting to see? What has worked well?
- ▲ How are they useful?

Figure 20. Focus Group Slide: Connecting Activities with Outcomes



Related to this topic was the kind of outcomes that Y’s regularly assess. We were interested in whether the participants were assessing changes in knowledge, skills, attitudes and behaviors. Figure 21 shows the slide that was shared with the focus group participants.

Figure 21. Focus Group Slide: Member Outcomes



During this discussion of connecting activities to outcomes, the facilitator pulled up a logic model (Figure 22) on screen and talked it through with participants. The logic model was on provided to us by Y-USA for an outdoor leadership program. On the one hand, we think this was useful for those participants that were less familiar with logic models and did not have much experience using them. On the other hand, in hindsight, we should have presented a simpler logic model with fewer components.

Figure 22. Focus Group Slide; YMCA of Greater Seattle G.O.L.D/B.O.L.D Outdoor Leadership Program Logic Model

RESOURCES	ACTIVITIES	OUTPUTS	OUTCOMES	GOALS
<ul style="list-style-type: none"> - Staff/Volunteers - 1-3 Seasonal staff instructors, paid and volunteer 1:3 Staff/volunteer to student ratio - Participant recruitment, systems and networks for outreach. - Participants (30 youth 11 – 18 years, 65% from traditionally under-served populations; mix of city and suburban youth). - Program setting (Wilderness areas) - Program curriculum (5-16 day leadership curriculum on communication, decision making, and personal confidence/capacity) - YMCA BOLD/GOLD Pilot Site Expansion Grant, Participant fees, Local Y in-kind and contributed funds. - Outdoor Equipment borrowed from Seattle YMCA Base Camp. - Northwest environment, geography, and culture of outdoor recreation. - YMCA Seattle BOLD/GOLD program model, staff leadership & partnership. 	<p>Level One Expeditions: (Ages 11-14)</p> <ul style="list-style-type: none"> - Wilderness based experiences. - Mentoring by staff and volunteers. - Physical challenge and adversity used intentionally. - Daily rituals such as evening council, chow circle, etc. - Group games and initiatives. - Guided exploration of natural settings. - Opportunity lessons on environmental science. - Lessons on basic communication and multicultural leadership skills. 	<p>Wilderness Experience: 30 youth for 16 hours a day for 7 days. 210 total program days. 3360 total student contact hours.</p> <ul style="list-style-type: none"> - Physical activities: 4 miles hiked a day, plus fishing, yoga, and other enrichment. - Evening circles with guided reflections. - 30 youth gain 16 hours of group leadership as designated leader of the day. - At least 2 daily workshops each day on leadership, communication, decision-making, cooking, packing, Leave No Trace. 	<p>Short Term Outcomes</p> <ul style="list-style-type: none"> - Increased self-confidence & self-esteem. - Enhanced self-perception as capable leader - Increased sense of identity, including positive view of gender & body image - Increased ability to be team player and get along with peers from diverse backgrounds - Authentic relationships with peers and adult leaders - Increased communication and problem-solving skills - Increased knowledge about making positive, healthy choices in their lives. - Increased view of the outdoors as a fun, adventurous place to explore. <p>Longer Term Outcomes</p> <ul style="list-style-type: none"> - Individuals more likely to take positive action in their lives - Individuals more likely to contribute positively to their home, schools & communities - Individuals experience a stronger connection to nature/the natural world - Individuals return to the YMCA as alumni on longer expeditions and/or apprentices and staff 	<p>Young people have the courage and compassion to make positive change in their lives.</p> <p>Young people have the emotional intelligence and practical skills to be positive multi-cultural leaders.</p> <p>Increased diversity of those pursuing outdoor education careers.</p> <p>Inspire the next generation of conservationists and outdoor leaders.</p>

Copyright © 2013 YMCA of the USA

While this was on screen, the group discussed the following questions:

- ▲ For those that answered yes, can you give us some examples of the changes you are tracking? How do you use this information?
- ▲ For those that answered no, why do you think your Y is not tracking these changes? What makes it difficult?

We ended the discussion by asking if there were any issues or questions and then thanked everyone for their participation.

Analysis

The focus group was recorded, poll results were saved, and a transcript was prepared for analysis. Altarum used the transcript to conduct a content analysis by question, looking for themes across participants as well as examples to highlight in the final report. We also used the transcript to pull out direct quotes as illustrative examples on the specific discussion topics. We find that direct quotes are more effective at communicating

important points that can sometimes get lost in a general summary. The following section includes a high level summary of the focus group results.

Results

This opening question to the group revealed that most participants had negative experiences with evaluation and used words such as “painful”, “confusing”, “disappointing”, and “frustrating” to describe past experiences.

Organizational Culture and Support

For this series of questions, participants were asked about collaborating on evaluation and whether this was standard practice in their Ys. One individual described collaborative practices involving staff, management and board members but most individuals did not feel their Ys were collaborative when it came to evaluation. Two individuals described environments where evaluation decisions are made the management level without involving other staff. One person felt that as a program director, her input should be solicited by management. Most agreed that the biggest barrier to collaboration is a lack of time.

“I very much agree with those two, when we do it and how we do it is up to the program director to form their own surveys. I feel that we never have any time to debrief with the other staff members or leaders in the organization to interpret the information correctly and thoroughly. I don’t see the follow-through or continuity with it.”

Participants described feeling so overwhelmed with existing duties that evaluation ‘falls through the cracks’ without the team sitting periodically to discuss results and what is working.

“I don’t know if any of us are proud of this, but at some point we’re doing swimming lessons because we have a pool or you’re doing a basketball program because you have a gym. In a situation where the people are so stretched and busy to get things done, taking a moment to step back and talk about the Y doesn’t get to happen very often. Maybe at a higher level, but it doesn’t happen very often for us.”

Capacity

One of the largest barriers was related to financial support for evaluation activities. One individual comment that if funds were allocated for evaluations, then she would want to ensure that funds are being well spent. Several described how this would be a financial challenge for them.

“I would agree with [name redacted] that it’s a challenge financially to meet this. Unfortunately we’re faced with a two-fold issue: part of the issue is I don’t have the money to support this because I am paying minimum wage to people leading the exercise classes who have certifications under their belt. However, now more than ever we need to be doing evaluations to see what’s working and what’s not so we can adequately spend our funds properly. How do we decide between the two? We’re in a situation now to figure out what do we have to cut to make this happen? I’m not ready to make the sacrifice.”

“I know when we applied to several grants at the state level; they tell us that 7% of the total program budget must go to evaluations. If we did that to our programs, what would that do to us? We have to struggle shrinking dollars and budgets, revenue shrinking all over the place and reduced memberships”

While financing evaluation seemed challenging to participants, most acknowledged that their organizations could benefit from undertaking more rigorous evaluations. The information from these evaluations could be useful for internal purposes such as improving program effectiveness but can also make them more appealing to potential funders who are increasingly emphasizing the need for evaluation.

There were a few questions about whether local staff have the knowledge and experience to make appropriate evaluation decisions. A few individuals felt their staff lack the skills to conduct meaningful evaluations and do not have the opportunity to receive the necessary training. A few individuals were interested in going beyond member satisfaction to measure outcomes, like the extent to which the programs support their core values and behavior changes as a result of their programs. Some noted the challenge of identifying outcomes and developing the measure and questions that will capture the relevant information. The interest in improving their outcomes focused evaluation seems to be driven by both internal interested as well as external pressure from funders.

“I think it comes down to that we’re being asked to do evaluations and we want to, but we just haven’t been given the tools or the training to do it. It’s important to keep up with the tools and figure what the trends are. Also the needs of our specific community—for us gangs are a big problem. We’re telling people to get involved in the YMCA because that’s a way to prevent kids from being involved in the grants. We are being asked through a federal grant (gang prevention grant) and the outcomes we’re looking at are: what will change the mind of a child not to go into gangs—how am I supposed to do that? How do I show that? I can’t. I don’t think that should be our outcomes, but how to come up with something to answer that question, or figure out what questions I should ask. I need help in learning how to do that and get that information to our program directors. “

We conducted a real-time poll to see what the data collection methods were of interest to participants. Currently, all participants use surveys as their primary method but are interested in learning more about multiple methods, with the greatest interest in focus groups. Beyond learning about collecting data, participants also expressed a need to improve knowledge of interpretation and analysis, which some viewed as the most difficult component of evaluation.

Developing a Conceptual Framework

Focus group participants were asked about whether their local Y really made the connection between program activities and the long-term changes, and particularly, whether they use logic models for this purpose. Not surprisingly, this was an area of weakness among the focus group participants. Only a few individuals reported using them and in both cases; they were requirements for United Way funded programs.

Defining Benchmarks and Indicators

In this section we discussed the kinds of changes that participants might be tracking. Most of participants (6) did indicate they are measuring these types of outcomes and when asked for examples, mentioned: changes in literacy levels; changes in cardiovascular capacity and strength; and changes in weight. When asked about whether programs are evaluation other aspects of implementation, such as program fidelity, a few mentioned conducting program audits.

Lessons Learned

We felt the virtual focus group was effective and worked well. We had a high level of participation throughout the group with only one individual that had minimal participation. The Adobe Connect ‘raise hand’ feature allowed the facilitator to control the discussion and ensure that no one person dominated the conversation. Overall it was relatively easy to organize with the greatest effort related to developing the discussion guide and slides. We felt that presenting content on-screen for participants to view was useful in keeping them focused on the discussion at hand. We found that participants were interested in learning about the survey results but felt that the logic model was too complex to review, especially with Y’s considered to be low capacity. In hindsight, we should have either not presented the logic model or selected one that was simpler and easier to digest. We also overlooked conducting an evaluation and getting direct feedback on the virtual focus group.

Overall, this approach was valuable in getting more in-depth information on the specific perceptions of those individuals associated with Y's that were categorized as low capacity. It was a good complement to the quantitative data that were collected through the YMCA Evaluation Capacity Assessment Survey and provided us with insight into the survey results. Participants' comments reinforced some of the issues identified in the survey and contributed to interpretation of survey results. For example, from the survey we knew that most Y's did not designate funds to support evaluation, but in the focus group we learned about the barriers that made this prohibitive.

Step 4: Key Informant Interviews

One-on-one key informant interviews were conducted with staff at Ys categorized as medium evaluation capacity based on YMCA Evaluation Capacity Assessment Survey results. This method provided Ys with a confidential forum to discuss evaluation successes, as well as areas for improvement. Interviews were selected as a method to gather in-depth information from individual Ys on key questions of greatest interest to Y-USA, probe deeper into responses that key informants had selected in the survey, and gauge level of comprehension of evaluation terminology that was used in the survey. This data collection approach was specifically chosen for medium capacity Ys because it offers significant opportunity to gather rich qualitative information on various evaluation capacity attributes for a single Y. Also, we felt that medium capacity Ys, presumably with higher evaluation capacity than low capacity Ys, might have more information to share with respect to evaluation strategies, systems, and experiences that could serve as models for evaluation capacity building.

Methods

Sampling and Recruitment

Our recruitment goal for key informant interviews was 10-20 individuals. This goal was selected primarily based on resources available for data collection and analysis. Using the results from the YMCA Evaluation Capacity Assessment Survey, a list of all respondents from Ys rated as medium evaluation capacity (survey core 2.5-3.47, n=684) was uploaded into a computerized random list generated to produce a random permutation of respondents. We then selected the first 20 respondents from the list. Selected respondents were contacted up to four times by Altarum staff, through email and phone calls, to schedule an interview before being declared unresponsive. Of those contacted for interviews, two were unresponsive, two respondents were no longer at that Y, and one respondent refused to participate in the interview because they were unfamiliar with Altarum's involvement on this project. Two respondents requested confirmation of Altarum's relationship with Y-USA before participating in the interview. In those instances, requests were forwarded to staff at Y-USA, who then contacted the appropriate individuals at the Y and the interview was rescheduled. In total, fifteen respondents (75% response rate) completed a key informant interview. Table 6 includes information on the key informant characteristics. Each interview was scheduled around the availability of the respondent and Altarum set aside a 4-6 week period to conduct interviews.

The majority of respondents (n=12) were part of a Y Member Association or branch that offered traditional programming; however, there were a few exceptions. Two respondents' positions dealt solely with camp programs, one respondent was housed at a Y conference center that does not offer traditional programs, and one respondent was housed within the clinical services branch of a Y.

Table 6. Key Informant Interview Characteristics

Job Title	Time in Job Role	Survey Score	Budget Size
Director of Programming	14 months	2.84	Branch Y
Regional Director [title abbreviated for confidentiality]	3 ½ years	3.03	4 - small
Senior Program Director- Membership and Wellness	2 ½ years	3.00	Branch Y
Gymnastics Director	11 years	2.94	3 – medium/small
Outdoor Recreation Director	2 years	2.63	3 – medium/small
Program Development Director	7 years	3.19	2 – medium
Vice President of Finance	6 months	2.88	Branch Y

Program Director	2 years	3.00	Branch Y
Executive Director	5 years	2.88	Branch Y
Vice President of YMCA	3 years	3.00	Branch Y
Health and Wellness Director	4 ½ years	3.00	Branch Y
Associate Program Director	3 years	2.84	2 - medium
COO, Interim CEO	5 years	2.63	2 - medium
Financial Development Director	5 ½ years	2.50	Branch Y
Director, Youth and Family Services [<i>title abbreviated for confidentiality</i>]	1 ½ years	3.03	Branch Y

Facilitation

Each interview was conducted via telephone and lasted approximately 45 minutes. A protocol with interview questions and probes was used to facilitate the interview (see Appendix C). The call began with a reminder of the purpose of the interview and informed consent to participate in and record the interview. All but one respondent granted permission for recording. The interview topics were similar to those addressed in the Evaluation Capacity Assessment Survey and respondents were often asked to expand upon their survey responses. The topics that were covered during the interviews included: organizational commitment to and support for evaluation, existing evaluation knowledge and experience, developing a conceptual model for designing outcome evaluations, defining benchmarks and indicators, and overall evaluation capacity strengths and weaknesses. These topic areas were selected because they addressed key questions that Y-USA had about Y evaluation capacity, and had large inter-group variance in survey results and we wanted to probe further into why this was the case.

The first questions in the interview were simple questions designed to provide us with background information on the interviewee and help to establish a rapport with the interviewer. These questions included information on staffs' role at their Y, how long they had served in that role and evaluation responsibilities. Next, we delved into specific evaluation capacity topics. In some cases, we asked the exact same question the respondent had already answered in the survey. The reason for doing this was twofold: (1) to gain insight in the reliability of the survey and (2) to set the stage for further inquiry into a particular topic area. Figure 23 below provides an example of this type of questioning.

Figure 23. Example of a probing interview question including a survey question.

<p>Are there financial resources allocated to evaluate program activities?</p> <p>a. [<i>If yes</i>] Where does this funding come from? Is there a line item in the budget for evaluation? How is the amount of this funding determined? Is this amount of funding sufficient to evaluate program activities?</p> <p>b. [<i>If no</i>] Why do you think there is no funding for evaluation? How does this affect your Y's efforts to evaluate programs?</p>

For other questions, we provided the interviewee with their survey response and asked them to elaborate on why they selected their response. This allowed us to dig deeper into why interviewees felt a particular way about their Y's evaluation capacity in relation to particular topics beyond what we could learn through the survey response alone (Figure 24).

Figure 24. Example of probing interview question based on a survey response.

On the survey, we asked whether your Y has integrated evaluation processes purposefully into ongoing organizational practices and you [*insert survey response*] with this statement. [Question 13 on survey]

- a. [*If agree/strongly agree from survey*] Can you provide examples of how your Y has done this? (Probe: do you have examples for whether or not you have documented strategies and policies in place for applying evaluation?)
- b. [*If disagree/strongly disagree from survey*] Do you have thoughts on why your Y has not done this? (Probes: lack of time, staff training, funding)

We also asked unique questions, not used in previous data collection efforts, to gather additional qualitative information. These included comprehension questions and closing questions. Comprehension questions were used to ascertain interviewee knowledge of key evaluation terms. We knew that some evaluation concepts, such as indicators and outcome measures, may be less familiar to Ys and felt that some Ys may not have understood all of the questions in the Evaluation Capacity Assessment Survey. The interviews provided an opportunity to have Ys define these terms in their own words so we could understand what they knew about particular concepts. Figure 25 below shows an example of this type of question. Closing questions were used at the end of the interview to provide an opportunity to add additional comments on topics that we might not have covered and ask for recommendations of how Y-USA might build evaluation capacity. We also asked the interviewee to reflect on their organization's capacity overall. This was a complex question that required the interviewee to consider their organization across all attributes, as discussed earlier in the interview, and provide their final overarching assessment of their capacity in their own words. An example of this type of question is provided in Figure 26.

Figure 25. Example of a comprehension and probing interview question.

20. Are you familiar with the term 'indicators' used in evaluation? What does this mean to you? (If unknown or for clarification: An indicator is a metric that provides information to measure performance, that is, it is a sign that something exists or is true.) Is your Y able to identify outcome indicators that are important/relevant to your work?

- i. [*If yes*] What is the process your Y uses to identify the indicators that are most relevant to your work? Has your Y already identified the indicators that are appropriate for measuring the impact of your work (e.g., changes in knowledge, attitudes, or behaviors) for key programs? Which programs has this been done for?
- ii. [*If no*] If you knew the indicators that were most relevant to your programs, would this improve your ability to evaluate your programs? Why or why not?

Figure 26. Example of a closing question

22. When you think about your Y's evaluation capacity as a whole, where do you feel your Y is strongest? Weakest?

Analysis

Fourteen interviews were recorded and a transcript of each was prepared for analysis. For the individual who denied permission to record the interview, the interviewer took extensive notes which were used for analysis. All interview transcripts were uploaded in NVivo8 qualitative analysis software. A coding scheme was developed based on key topics and themes identified in the transcripts. The codes were used to label sections of interview text (words, phrases, sentences, paragraphs, and quotes) according to theme. Codes and accompanying text were then synthesized and organized according to the main topic areas of the interview guide. Particular attention was paid to recurring ideas and thoughts, as well as opposing viewpoints. Specific quotes that embodied thoughts, ideas, and viewpoints were selected from the text and integrated in the key informant interview findings chapter of a final report for Y-USA. The report also identified the number of respondents that reported specific answers to interview questions. Reporting qualitative data in quantitative terms can sometimes be misleading if all participants are not asked identical sets of questions, and therefore, the denominator could be different for each question making a response seem more or less common than it actually was. However, because our interview was highly structured and because organizations often wish to know how common or unique particular results are, we chose to present this data.

Results

Below is a sample of results obtained through key informant interviews with Ys.

Organizational Commitment and Support for Evaluation

In this section, key informants were asked to describe how evaluation is supported within their Ys. About a quarter of informants stated that their staff have access to tools to support evaluation, such as program evaluation, which are available through their internal network or have been provided by Y-USA or other organizations. Most interviewees (n=10) felt their staff had the time to collaborate on evaluation activities, such as developing surveys and reviewing responses.

“We’ve revised our survey process a number of times over the past few years. Every time we do that we bring in the key staff for that area to give their input on what kinds of questions we should be asking, wording, and what they would like to see in a survey tool.”

Four of the interviewees felt that their staff did not have the time to collaborate and work on evaluation activities primarily because evaluation work was not viewed as a staff priority.

When asked about financial resources allocated for evaluation, a few interviewees stated that funds were allocated for these activities at the branch or association level. Most respondents stated that there was no line item for evaluation, but that cost is built into program budgets and includes primarily staff time. A few stated they are trying to obtain grants to fund evaluation activities in particular. Specifically, one Y stated:

“We’ve had some financial problems in the past few years which have prevented us from dedicating money to evaluation.”

When asked whether their Ys integrate evaluation processes into ongoing organizational practices, almost all key informants (n=13) had previously reported they do. The following is an example provided during the interview:

“For the annual membership satisfaction survey, we do that every year, but also based on the results each year the association sets goals for us the next year. So, each year as the results come back, we see how we did and if we reached our goals. When we get responses and ask for suggestions, we will then make changes to our programming accordingly.”

The few respondents who stated that their Ys do not integrate evaluation processes into organizational practices felt it was due to a lack of leadership support or resources.

Interviewees were asked to describe evaluation expectations communicated to staff. Almost all respondents (n=13) stated that the evaluation roles for staff were clearly established at their Y and provided the following as examples:

- ▲ Six interviewees stated that staff are expected to distribute surveys to all program participants during specific time frames.
- ▲ One interviewee established specific targets for survey response rates.
- ▲ Three respondents expected staff to use evaluation results to implement program improvements.

One Y stated that there is no explanation to staff as to why or how they are to evaluate programs, just the expectation that something will be done. Evaluation roles are most commonly explained to staff during annual performance reviews (n=8) or staff meetings (n=2). A few interviewees stated there were no evaluation expectations communicated to staff, although one is currently working on implementing evaluation policies.

This included questions about expectations for evaluation roles of staff at their Y. Of the individuals who stated that there are clear expectations for the evaluation roles of staff, five respondents stated that the expectation was to “get them done.” In addition, three respondents mentioned an action plan that comes from the results of these evaluations.

“[The expectation is] that we implement some learning from it, we’re not just going to survey out members and not do anything with the information.”

Interviewees mentioned several methods to insuring high quality and purpose driven evaluations. Staff training, quality tools, utilizing third party resources and organizational commitment and knowledge were all cited as ways that high-quality evaluations could be insured. However, those same areas of strength for some respondents were a weakness for other Ys and the reason that they felt their Y could not insure high quality evaluations.

“I believe that they’re purpose driven because they all come back to the same goals and strategies that we’re trying to move towards as an association.”

“I don’t think I have a clear message on what the expectations are.”

When asked if their Y had the capacity, systems and expertise in place to ensure high-quality and purpose-driven evaluations, respondents both agreed (n=9) disagreed (n=6). Of those who agreed with the statement, six referenced an organizational commitment and understanding of evaluation importance. However, for the individuals who disagreed with the statement, four individuals cited a lack of leadership or organizational support as to why they disagreed. From one respondent’s interview:

“At our branch it was our executive director who was in charge of sending out the annual member satisfaction survey on top of all his other responsibilities. There were times that we would ask to see results, but he wasn’t able to explain very well what the data was trying to explain. He didn’t know what the goals of the survey questions were.”

Existing Evaluation Knowledge and Experience

Five questions were taken directly from the Evaluation Capacity Assessment Survey and respondents were asked supplemental questions to provide context for their answer. Although respondents were reminded of their response to the survey question, the response given in interview was not always the same. There was one question in particular that difference between survey responses and interview response (n=6). Under the survey section on existing evaluation knowledge and experience, respondents were asked if they agreed with the statement: “Our local Y can identify which data collection methods are most appropriate for different outcome areas (e.g., changes in norms require determining what people think about particular issues, so surveys, focus groups and interviews are appropriate).”

Of those who agreed with this statement (n=10), six individuals changed their response during the key informant interview when asked to qualify their answer. These respondents stated that there was no specific process for identifying the appropriate method of data collection because only one or two types were used (n=5), or they felt that there was no specific strategy in place (n=1). Other survey discrepancies were not common in any other question, and there were no more than 3 differences between survey response and interview response for any individual.

Within the survey section on existing evaluation knowledge and experience, respondents were asked if they or anyone at their Y had experience in developing data collection tools for a variety of methods (focus groups, surveys, interviews, etc.) Most commonly, those with experience developing data collection tools were individuals housed at the association level (n=5) or were program directors (n=5). In addition, respondents were asked if there were any specific people at their Y whose job, or components of their job, was to work on evaluations. Of those who agreed with the question (n=6), five individuals were housed within an association-level job, and one individual was a branch director. At the association level, jobs with evaluation components included Training and Development VP, Membership Executive, regional office manager, Information Systems (IS) and communication departments, and a branch-level outcome coordinator.

Developing a Conceptual Model for Designing Outcome Evaluations

When participants were asked if they articulate how they expect a program to create change in a participant, the majority of individuals (n=8) stated that their Y does not do this. Of these individuals, the reason that they believe articulating member change is not essential is because their evaluations focus on member satisfaction versus program outcome.

“They’re [CEOs, senior staff] looking for information so they can put it in their published materials. For example, something like “90% of our participants agree that it’s important to help other through participating in our programs”. Surveys are not a tool that they’re using to see where they can improve”

Logic models were not used by any of the key informants in their programming. One individual stated that they’ve used one in the past for a whole summer-program experience, but not for a particular program. When asked their opinion on logic models, a key informant shared:

“Logic models are fine if all staff working with them can get good training. If not, it becomes a rote process and the value is limited. The staff don’t understand it, so it just becomes something to do and fill out and not something that drives the process of what they’re doing”

For some of the individuals who had heard of logic models, but have not used them:

“I think they’re interesting, but I don’t know if I would use them. I think it depends on what it would be for, honestly. Our association wants us to do more large-scale association-wide programming and I think on a grander scale a logic model would be beneficial. But for small, everyday-type things it would probably be a burden and get redundant. But for big events I think that it would definitely be beneficial”

“Not for our programs here but I have seen them before. I think that they’re great and very useful because it helps you to understand how you need to measure results. They can get beyond just measuring the numbers involved in a program”

However, the most common answer to this question (n=7) regarding logic models was that the respondent was not familiar with logic models, and had never heard of them before.

Lessons Learned

Overall, we felt that the key informant interviews were very effective and provided quality information on current evaluation activities and environments. We achieved a high response rate, and each interview was insightful and respondents were forthcoming with answers, both positive and negative. The key informant interviews were valuable for understanding what types of evaluation activities are currently in place in local Ys. The key informant interviews also provided valuable context and explanation for the Y Evaluation Capacity Assessment Survey responses. The respondents were able to qualify their survey responses and in some circumstances, it was noted that their response in the interview did not wholly align with their survey response. For example, many Ys stated in the survey that their Y could distinguish the appropriate type of data collection method, but when they were asked about the process they used to determine this, they explained that they did not distinguish between data collection methods or that only use one type of data collection method.

Although the interviews were successful overall, there were several areas that could have been improved. One challenge associated with the interview protocol was the length and complexity of some of the questions. The survey qualification questions were difficult for respondents to listen to, consider, and respond to. There were several instances where respondents requested the question be repeated, broken down into smaller sentences or further explained. One respondent specifically suggested providing the protocol prior to the interview to aid those who are visual learners, and act as a guide during the interview process.

Another challenge we encountered was that many respondents were unfamiliar with evaluation terminology and this made probing for more information challenging. For example, only six respondents reported having seen a logic model before and only one respondent had actually used a logic model. We did not anticipate that logic models would be new to respondents and could have provided more definition and perhaps an example in advance for respondents to refer to during the interview. Although we offered definitions of terminology at some points during the interview, we had specifically chosen to not further define all terms so that we could better determine interviewees’ level of understanding of evaluation terms. Although it was beneficial to get a better understanding of their familiarity with such terminology, to probe more in-depth into key areas, definitions for some items could have been provided during the interview for respondents to view. To be most helpful, these definitions also could have included relatable Y examples of what our intent with the question was as well as examples of what it was not. For example, Y-USA very specifically wanted to understand outcome evaluation beyond participant satisfaction. We could have provided a definition of

outcome evaluation and included examples of what we mean (in this case, changes in knowledge, attitudes, and behaviors) and what we do not mean (participant satisfaction with a program).

Finally, because a variety of respondents were invited to participate in the YMCA Evaluation Capacity Assessment Survey, our key informant interview cohort reflected a diversity of Y and staff roles. While useful for gathering a broad view of the organization, some respondents in less traditional roles did not always have applicable experiences to respond to the questions, and as such were unable to answer or provided atypical responses. For example, the clinical services respondent in particular felt it was difficult to respond to the interview questions because her work was specific to the patient and their family, and evaluation and assessment was done on an individual clinical basis versus a program basis. Again, this was a positive and a negative to our approach. We could have possibly gathered more in-depth data on program outcome evaluation had we selected specific positions to include as key informants although our approach may have yielded a more comprehensive picture of the variety of Ys and staff experiences.

Step 5: Site Visits

Site visits were conducted with five high capacity Ys, which were identified based on the YMCA Evaluation Capacity Assessment Survey results. The site visits were used to gather detailed information about evaluation activities at a single Y from multiple staff in a variety of roles. We chose to conduct site visits, the most intensive of our data collection techniques, with high capacity Ys because we felt these Ys had the most potential to surface models and approaches of evaluation that could be replicated by other Ys. Meeting with multiple individuals at each Y was also helpful in validating the survey results and understanding the layered perceptions of evaluation capacity that may exist within a Y.

Methods

Sampling and Recruitment

The results of the YMCA Evaluation Capacity Assessment Survey were used to identify a sample of Ys eligible for the site visits (n=100, survey score 3.50 or higher). To ensure we visited a diverse group of Ys, we purposely selected sites that were heterogeneous with respect to geography and budget category (a proxy for size). We made the decision to focus on sites located in the Northeast, South, and Midwest and eliminated Y's located in the West due to project budget and time constraints. Y-USA staff further narrowed the list based on knowledge of Y evaluation activities, yielding a list of 8 potential sites. This list was further narrowed by cross referencing the list for other YMCA Evaluation Capacity Survey scores. We searched the entire list of survey respondents and if we identified other results from the same Y, scores were averaged across respondents and sites were eliminated if their average score was less than 3.50. The final list of potential sites was shared with the Y-USA project officer who vetted the sites with regional Y Resource Directors. Resource Directors support local Ys in activities such as strategic planning and have detailed knowledge of a Y's capacity. This cumulative input helped guide our final selection, which resulted in the identification of 6 sites located in 5 states.

Selected sites received an email invitation from Altarum staff to participate in a site visit. This message included a description of the project and purpose of the site visit. In some cases, this approach worked well but was challenging in other situations. Some sites were reluctant to agree to the site visit without having contact with Y-USA. In these cases, the Y-USA project officer either phoned or emailed the site directly to ease their concerns and provide additional information when needed. In one situation, the CEO was not emailed directly and felt that he should have been the recipient of this invitation, rather than someone on his staff that had completed the survey.

Based on the project schedule, we identified an 8-week time frame for the site visits. Each site was asked to complete a site visit intake form to identify potential dates and to help us determine appropriate individuals that would participate in the site visit interviews. We were interested in meeting with leaders and staff members with key evaluation roles. One Y declined to participate because staff were busy with other obligations during the potential visit dates. Another Y was reluctant to participate because the CEO did not respond to the survey and disagreed with the staff person that completed the survey and rated their Y as high capacity across the evaluation measures. The Y-USA project officer did speak with him and provided the rationale supporting his participation and in the end, he agreed to the site visit. Site visits were conducted at a total of five Ys (83% participation rate). Table 7 includes information on the site visit characteristics.

Table 7. Site Visit Characteristics

Region	BTW Score	Budget Size
Mid-Atlantic	3.66	4 - small
Mid-Atlantic	3.50	2 - medium
Mid-Atlantic	3.96*	1 - large
Midwest	3.92+	Branch Y
New England	3.75	2 - medium

*Average score from 4 respondents

+Average score from 2 respondents

Facilitation

Altarum staff scheduled pre-site visit phone calls with Ys in advance of the visit. These calls lasted about an hour and were used to discuss the purpose of the visit, confirm a date, develop a draft agenda, and respond to any questions. Typically, there were several individuals from each site on these calls. One individual at each site took the lead in working with us to organize the visit, which involved coordinating the interviews with the identified staff and scheduling a room for the interviews.

Altarum conducted site visits in October and November, 2012 and each lasted one day and approximately 4-8 hours, depending on the number of interviews scheduled. Typically, there was a mix of both individual and small group interviews and staff with similar roles or in the same department, were often interviewed together. The following staff were usually interviewed as part of the site visit:

- ▲ Chief Executive Officer, Chief Operating Officer, and/or Vice Presidents
- ▲ Marketing and Membership Directors
- ▲ Program Directors
- ▲ Site Coordinators
- ▲ Other staff leading evaluation activities

The following (Table 8) is a sample agenda used with one of the larger Ys and included staff from all major program areas as well as the COO and directors of marketing and membership.

Table 8. Sample site visit agenda.

10:00 am	School age Childcare Director Pre-school Childcare Director All day Childcare Director School age Childcare Director
11:30 am	Camp Director Asst. Camp Director
12:30 pm	Lunch Break
1:00 pm	Aquatics Director Asst. Aquatics Director Competitive Aquatics Director
2:00 pm	Physical Director Asst. Physical Director Health and Wellness Director

	Asst. Health and Wellness Director
3:30 pm	Senior Director, Member Services Marketing and Membership Director Asst. Health and Wellness Director
4:30 pm	Senior Director, Member Services Senior Director, Childcare, youth and teens Chief Operating Officer
5:30 pm	Conclude site visit

A protocol was used to guide interviews with Y staff (see Appendix D). Interview topics were selected from the YMCA Evaluation Capacity Assessment Survey and covered the following topic areas: organizational culture and support; capacity; developing conceptual framework; and defining benchmarks and indicators. In particular, interview questions were included for survey questions that had yielded the largest variations between high, medium, and low evaluation capacity sites. This was done to assess how high capacity Ys achieve capacity in areas where other Ys struggle.

Altarum started the interviews by providing background information on the project and the purpose of the site visit. We found that individuals asked to participate in the interviews were often not briefed on the discussion topics or clear on the purpose of the visit. Next, we asked some introductory questions to get a better sense of the interviewees’ background and their evaluation experience. We asked each person to describe his/her role at the Y, responsibilities related to evaluation, and the number of years in this role.

Next, we moved into the topic areas mentioned above. The interviewer followed the protocol most of the time, beginning with the questions on organizational commitment and ending with the ones related to evaluation design. This allowed us to begin interviews by discussing broader organizational approaches and practices around evaluation and then move into the more specific and detailed information related to the specific outcomes that were assessed and how. This approach worked well for us although there were instances where the interviewee responses did veer off topic and touched on other issues to be addressed later in the interview.

The interview protocol included two sets of questions—one for those in leadership positions and another for staff. Some questions were included in both sets. The following (Figure 27) are some questions that were posed to all interviewees related to organizational commitment and culture.

Figure 27. Example of general questions posed to both leadership and staff

- | |
|--|
| <ul style="list-style-type: none"> ▪ Do you think evaluation is seen as integral to your organizational mission? Please explain. ▪ What prompted you to undertake evaluation? ▪ Are staff involved in evaluation-focused discussions, in terms of planning or implementing? |
|--|

In other cases, questions were tailored. This approach was used to learn about the different perspectives on a given topic. We were interested to see in what ways the viewpoints of leadership aligned or diverged from those of staff. This was possible to ascertain because leadership was interviewed separately from program staff. Figure 28 shows an example of these tailored questions on the topic of evaluation expectations, also discussed in the organizational commitment and culture section.

Figure 28. Example of tailored leadership and staff questions

Leadership Questions: What are the expectations for staff in terms of evaluating programs? How do you ensure staff expectations are clearly communicated?

Staff Questions: Do you feel that leadership has established clear expectations for staff roles? Please explain.

Tailoring questions was also helpful to examine consistencies across practices and policies and how they are communicated throughout the organization. For example, this approach was important in the section on evaluation knowledge and experience, since staff capacity is such an important component of overall evaluation capacity. It was important for us to hear from both staff and leaders at these high capacity Ys how they successfully address building evaluation knowledge through either experiential learning or formal trainings. Figure 29 includes example questions on this topic.

Figure 29. Example of tailored leadership and staff question

Leadership Questions: Is training on evaluation provided to staff? What types of training and how often? Are staff encouraged to seek evaluation training? If yes, how?

Staff Questions: Are you provided with training on evaluation? What types of training and how often? Is this training sufficient? Are you encouraged to attend evaluation training?

Even though we had both leadership and staff questions for each topic area, we assumed that interviewees would not be able to address all the interview topic areas. We expected that those in leadership would be most comfortable discussing areas such as organizational commitment and culture and that program staff would have the most detailed knowledge of evaluation design and implementation. In other words, not all interviewees were expected to answer all questions. In the end, because we had a good sample of staff at each Y, this approach yielded sufficient information in each topic area.

As mentioned before, the section on evaluation design and implementation included very specific questions on different aspects of evaluation design, including the use of a conceptual framework with the assumption that these high capacity Ys were actively conducting program evaluation. We also asked to see examples of logic models and data collection instruments when possible. Figure 30 includes a sample of the questions posed in this section to both leadership and staff.

Figure 30. Example of detailed questions posed to both leadership and staff

- ▲ Have you described how activities will lead to short-term and long-term outcomes, and eventually achieve your program goals? Have you done this through logic models? If so, how important are logic models in conducting evaluation work? Can you share some examples?
- ▲ Are you assessing outcomes in any of your programs? If so, what types of outcomes are you assessing—changes in knowledge, skills, attitudes, behaviors? Please describe by program. How did you determine these?
- ▲ Are you assessing fidelity or how programs are being implemented (i.e. process evaluation)? If so, please describe examples (and share relevant documents) of programs and how processes are being tracked.
- ▲ What data collection methods does your Y use? What factors influence the selection of these methods? What staff preparation is necessary to use these methods?
- ▲ What data collection instruments did you use? Who developed these instruments? Can you please share these? Do you feel these instruments captured the necessary data?

We concluded the site visit interviews by asking interviewees to summarize their overall strengths and weaknesses related to evaluation; and to identify any resources or trainings that would be useful in supporting their evaluation activities.

Analysis

Interviews were recorded and transcribed. The Altarum staff member who conducted the site visit was responsible for preparing a written summary for each visit. The summaries were organized by topic area and included a synthesis of interviewee responses and supporting quotes. The cases where interviewees had divergent opinions or perspectives were noted. These summaries were considered internal Altarum documents that were used for analysis. A content analysis was conducted of the five site visit summaries, looking for themes across the sites or for examples to highlight. Findings were presented for each site under each topic area with supporting quotes that were illustrative of key points.

Results

The following is a summary of site visit findings by interview topic.

Organizational Culture and Practices around Evaluation

All of the sites expressed a willingness to conduct evaluation and an interest in implementing organizational practices that support evaluation and continuous improvement. This viewpoint was particularly apparent at two sites that viewed data as an important tool that can validate strategic action.

“Evaluation means measuring things so that you can strategically take action. The evaluation tool is an important aspect of it, and methodology is important, but ultimately it’s about assessing how well you’re doing and always striving to be better. Evaluation isn’t measuring for the sake of measuring, you need to understand and be able to isolate variables.” (large Y in the Mid-Atlantic)

Part of creating an organizational culture that is evaluative in nature is supporting a learning environment. The Midwest site successfully created an environment of critical thinking where program directors meet with their site staff and ask key questions while reviewing their site data. The staff work together and use various data sources (e.g., school grades) to depict what is happening at the site level. Of all the sites, the concept of a learning environment was most strongly observed at a small Mid-Atlantic area Y where the leadership models and reinforces the practice of ongoing reflection with staff. All staff members commented how the CEO uses reflection as a way to help staff be very intentional about actions taken and why. The concept of intentionality was also mentioned during the site visit at the medium Y in the New England area where leadership sees evaluation as a vital tool, a philosophy promoted among staff. Evaluation is linked explicitly to decision-making and sharing lessons learned from one program with another.

An important component of the organizational culture is the extent to which evaluation results are communicated, both internally and externally. All five sites described regular practices they have adopted to ensure that staff are kept informed and these often include information about evaluation results. The following are some strategies used by sites:

- ▲ regular team meetings (weekly, monthly) between program directors and staff
- ▲ all staff meeting (monthly)
- ▲ preparing program reports for leadership
- ▲ staff newsletter
- ▲ staff retreats (annual)

Organizational Commitment and Support for Evaluation

To some extent, all the sites have integrated evaluation into organizational processes. Leadership across all five sites expressed a high level of commitment to evaluation, which was supported by staff comments during the site visits. Staff at all sites expressed that there is strong leadership support for evaluation both at the program level as well as the management level.

In terms of financial support for evaluation, most of the sites did not have a specific budget for evaluation activities, aside for funds allocated for the SEER surveys. For the most part, evaluation is integrated into all staff duties as opposed to having dedicated evaluation staff. Program staff at all the sites are responsible for overseeing periodic member surveys.

One of the highest levels of organizational commitment is hiring staff dedicated to evaluation. The large Mid-Atlantic Y is the only site that hired staff to specifically support evaluation efforts. The CEO felt they needed to improve measurement and data quality and made the decision to designate a department—Quality Assurance and Risk Management—at the association level to focus specifically on improving their evaluation. This department takes the lead in assessing quality through site audits. Quality measures are defined and assessed at the association level as opposed to the local level. This arrangement is unique because evaluation activities are primarily centralized at the association level.

At the medium Y in New England, even though management believed that staff expectations were clearly defined, some staff felt that additional clarification of staff roles was needed. The area of communicating staff expectations did not appear to be an issue at the other sites. Staff at the large Midwest Y and the medium Y in the Mid-Atlantic area felt that expectations were clearly communicated to them, either at hiring or through targeted trainings. Staff at these sites felt their roles were clearly explained and had the necessary support as well. Both of these sites described conducting staff trainings as evaluation duties were expanded or new programs are implemented.

Despite some challenges, it was evident based on the site visits that all of the Ys strive to be mission-driven organizations and find that evaluation can support their mission. This was discussed by all the sites in context of the Y pillars of youth development, healthy living, and social responsibility. The leadership at the site in the Midwest expressed that evaluation is critical to achieving their mission and demonstrating that their programs impact these areas.

Using Data to Inform Ongoing Work

All sites reported using data to guide their work and make improvements. Cited most often was using data to improve the membership experience. Audits, SEER survey, and market research are the typical methods used to improve the membership experience, which were largely related to enhancing facilities, staff behaviors, and membership services. The large Mid-Atlantic Y also looks at process data to establish trends in membership behavior and to ensure programs operate at peak efficiency as does the large Y in the Midwest. Tracking membership trends (enrollment and retention) has been a priority for the Midwest site as they've seen their membership fall drastically due to the economy. This site works closely with the branches to track different aspects of member usage on a monthly basis.

Sites are also using data to make programming decisions. The small Mid-Atlantic Y has collected information from members to gauge interest in potential programs. They recently partnered with an area school to determine what afterschool activities are preferred for this group of students. They surveyed parents to help them make those planning choices and found that parents are interested in swimming programs and Friday after school programs.

Several sites have used evaluation to improve staff performance and management. As previously mentioned, the large Mid-Atlantic Y performs regular onsite audits to assess program quality, which are used to spur course correction in terms of program improvements as well as identify appropriate areas for staff training. This was also reported by staff at the Y in New England where evaluation findings are useful for staff improvement plans.

Capacity—Staff and Systems

When making hiring decisions, evaluation knowledge and experience are not typically considered. There were mixed opinions as to whether staff had the knowledge and skills to carry out evaluation; there were individuals at each site that believed their current staff knowledge of evaluation is not sufficient. Some felt that funding was the primary barrier and that inadequate resources resulted in insufficient staff time to support evaluation activities.

None of the sites provide general training on evaluation. Typically, staff receive project-specific or task-specific training. In Midwest site, evaluation training is provided to new staff or as staff assume new positions within the Y and are trained on specific data collection systems and instruments. Minimal training was provided by the medium and large Mid-Atlantic sites and the New England site. At the medium Mid-Atlantic Y, staff learn survey development “on the job” and have the support of their team and supervisor when needed. This is consistent with the training approach described during the New England site visit where training is provided on an ad hoc basis and was characterized as “trial by fire.” The large Mid-Atlantic site also provides evaluation training as needed, such as for a new grant-funded program. If a new project requires considerable evaluation needs, then sometimes new staff with the appropriate background is hired.

When asked about systems in place to support data management, SEER and Zoomerang were the most common responses. All sites work with SEER Analytics to collect, analyze, and report data but none of the sites described an effective process or system for storing, managing, and analyzing their data.

Evaluation Design and Implementation

With the exception of the large Y in the Midwest, all the sites discussed connecting evaluation to their strategic plan. The large Mid-Atlantic Y described how strategies are linked to specific indicators and outcomes in their programmatic areas. The strategic plan also includes sample data collection instruments to collect this information. At the medium Mid-Atlantic site, they review their plan to ensure that activities connect to the mission and three areas of focus (healthy living, youth development, social responsibility). Staff at the small Mid-Atlantic Y discussed their strategic plan at great length because they completed their plan over the summer and it is currently an organizational priority. Most of the site visit was focused on discussing their strategic planning experience, which was facilitated by one of the Y-USA Resource Directors over a 6-month period. This plan includes six strategic objectives and each person on the management team is assigned specific annual objectives they are overseeing.

When asked about whether sites are using conceptual frameworks to guide evaluation work, sites for the most part, did not use conceptual frameworks in their evaluation design. There were only a few instances where logic models were developed as part of a grant requirement. The large Mid-Atlantic Y staff have found logic models to be useful but do not routinely use them to design their evaluations. Both the New England and Midwest Ys felt that increasing their use of conceptual frameworks could strengthen their evaluation and is something that should be utilized more. A program director at the Midwest site also thinks that a logic model could be useful as a communication tool to staff as well. None of the staff at the small or medium Mid-Atlantic Ys have used these types of frameworks in evaluation and in fact, some staff at the latter did not really understand how a conceptual framework might be useful.

All the sites administer program-specific surveys to help them assess overall program satisfaction and solicit suggestions for improvements but not for the purposes of assessing program outcomes.

Across all sites, assessing program outcomes was the weakest evaluation area. Several of the sites described their desire to improve in this area and collect better data. As one respondent described, they want to go beyond ‘what’ was provided and understand ‘how’ and ‘why’ it was effective or not. If a program is successful or is failing, evaluation can help them understand why, so that they can make course corrections in cases of failure, or capitalize on efforts that are successful.

All the sites mentioned that the programs with the strongest evaluation components were those that are grant funded, such as by the United Way. Of all the sites, the large Y in the Midwest is implementing the most rigorous evaluation activities as part of the 21st Century afterschool program funded by the Department of Education.

Surveys are the most utilized data collection method across all the sites. All sites find them useful but some admit that they have a tendency to over-survey their members and can be challenged with low response rates at times. Sites administer surveys for their program areas but also use direct observation and focus groups. Other than the formal methods, all the sites use informal feedback in some way as well.

Data analysis was a big weakness across all the sites. Both the medium and small Mid-Atlantic sites described conducting minimal analysis of the data collected. The staff at the medium Mid-Atlantic Y do not conduct analysis beyond the results generated through Zoomerang. The small Mid-Atlantic Y does not have a system in place for evaluating their data and staff describe reading through surveys to identify strengths and areas for improvement.

Lessons Learned

Overall, the site visit process went well and resulted in insightful data. The visits helped us better understand the local Y environment and provided context to data we collected through the YMCA Evaluation Capacity Assessment Survey. We also identified a number of evaluation practices in the various topic areas that can be useful to other Ys in developing their capacity.

Despite the positive outcomes from the site visits, we did identify opportunities for improvement. One of the first challenges we faced was related to the selection process. While we felt the site visits yielded valuable information, there were questions after the fact as to whether all the sites were properly vetted and considered to be high capacity. As the most expensive method used in this assessment, we wanted to ensure that all sites included in the site visits were good examples of high capacity Ys. As discussed earlier, sites were vetted by national and regional Y-USA staff familiar with the local Ys that were under consideration. This input informed the selection process. In hindsight, the Altarum team felt we should have conducted brief interviews with each site as part of the vetting process to make certain these sites were in line with our expectations as to the attributes of high capacity Ys.

Once Ys were selected, Altarum made initial contact to schedule the site visits but encountered some resistance from a few of the sites because they were unfamiliar with Altarum and were suspicious of our interest in visiting their sites. In hindsight, we may have had better cooperation at the outset of our process if Y-USA staff had sent the initial initiations to sites. This would have demonstrated that Y-USA is aware of and supports the site visits.

Overall, the process used to coordinate and conduct the visits worked well. It was very helpful to have a person at each Y coordinate with other staff although we should have provided some briefing materials that could have been shared with other interviewees. Although we were able to research the sites prior to the visits, it would have been useful had we requested additional information, such as an organizational chart or

other helpful background documents. Most questions on the interview protocol worked well for our purposes but some questions were less successful. Those related to evaluation design and implementation assumed a base knowledge of and experience conducting outcome-based evaluation. Because most sites were not conducting such rigorous evaluation examining program outcomes, these questions were not pertinent and less useful.

Summary of Lessons Learned

Through our evaluation capacity assessment process with the Y, we learned many lessons that may serve useful to others embarking on a similar process with a large nonprofit organization. In Table 9, we summarize the facilitators and barriers for each step of our process, and offer potential solutions for overcoming barriers.

Table 9. Facilitators and barriers encountered during the Y evaluation capacity assessment process, by step.

Facilitators	Barriers and Potential Solutions
Step 1: Becoming Acquainted with Y Culture, Structure, and Evaluation Needs	
<ul style="list-style-type: none"> ▲ Hosting a face-to-face kickoff meeting with project staff to build rapport and learn about the organization’s structure and culture. ▲ Identifying systems, resources, staff, and structures that can support the assessment process. ▲ Identifying potential pitfalls that could negatively influence the evaluation capacity assessment process, so that processes can be adjusted as necessary to minimize these pitfalls. ▲ Defining evaluation capacity to ensure the organization and evaluators share a common understanding of what it means to possess this capacity. ▲ Identifying a conceptual framework to guide the process, and ensure important evaluation capacity attributes are assessed. ▲ Developing a logic model with the organization that shows how the assessment process will operate and what the end products will be. 	<ul style="list-style-type: none"> ▲ Time constraints may make it difficult to spend adequate learning about organization structure and culture before beginning the assessment process. In these situations, we recommend utilizing a structured agenda that incorporates planning activities to focus on building common understanding among the organization and evaluators and abstracting as much documentation about evaluation capacity attributes as possible to avoid spending too much time on any one topic. ▲ Organization and evaluators located in different geographic regions mean in-person meetings throughout the process may not be possible. Utilize regularly scheduled meetings to keep one another abreast of progress while continuing to learn about the organization. Use webinars and video conferencing if available for more interaction and face time.
Step 2: Evaluation Capacity Assessment Survey	
<ul style="list-style-type: none"> ▲ Surveys are a low-cost, easy-to-implement method for assessing evaluation capacity across a large organization. ▲ Creating or using scoring rubrics that assign levels of evaluation capacity helps to segment a large population into manageable groups from which to sample for further analyses. 	<ul style="list-style-type: none"> ▲ Currently, there are no evaluation capacity assessment survey instruments in the field that have been validated or undergone reliability testing. If time allows, we recommend conducting some pre-testing and reliability testing with the organization prior to administering the tool. ▲ Ability of survey respondents to assess their organization’s evaluation capacity. Those with less evaluation knowledge and experience may overestimate their capacity, while organizations with high capacity may more readily identify weaknesses. This is difficult to avoid; however, offering clear definitions of key terms may be useful. If time allows, a webinar or training on evaluation could be provided ahead of time to clarify meaning around key terms.

Facilitators	Barriers and Potential Solutions
	<ul style="list-style-type: none"> ▲ Asking open-ended questions in an evaluation capacity survey that will be completed by a large number of individuals may not yield the type of data desired. Surveys are often unable to validate open-ended response options, so respondents are able to write any type of response they wish. Sorting through numerous correct and incorrect responses to such questions and analyzing this data can be cumbersome and time consuming. We recommend not including or minimizing the number of open-ended survey questions used, and if used, ensuring they are very well defined so participants know how to answer the question. An even better solution is to use qualitative data collection methods (e.g., interviews, focus groups) which offer opportunity to more thoroughly discuss questions.
Step 3: Virtual Focus Group	
<ul style="list-style-type: none"> ▲ Virtual focus groups can be a highly effective strategy for engaging geographically diverse organizations in dialogue around evaluation. ▲ Using a webinar platform that supports presentation slides, real-time polls, and other interactive elements can help maintain participant engagement. 	<ul style="list-style-type: none"> ▲ Avoid presenting overly complicated information during the focus group, such as logic models. These are too detailed to digest quickly and can detract from the flow of conversation.
Step 3: Key Informant Interviews	
<ul style="list-style-type: none"> ▲ One-on-one interviews are an effective tool for obtaining high-quality, detailed information about evaluation activities and environments. ▲ When conducted in conjunction with a survey, interviews can help to provide contextual and clarifying information about survey responses. 	<ul style="list-style-type: none"> ▲ Interview questions were difficult for respondents to understand. If detailed questions are required, provide the respondents with a copy of the questions ahead of time to review and refer to during the interview. ▲ Respondent unfamiliarity with evaluation terminology. Provide respondents with a sheet of definitions prior to the interview, and go over key definitions and answer any questions before the interview questioning begins. Include diagrams or pictures of information if relevant to articulate meaning. ▲ All questions are not relevant to respondents because of differing roles. If staff in multiple roles will be interviewed, consider tailoring questions for each audience to ensure they are relevant, or flag questions that they are only to be asked of certain respondents. Alternatively, select one audience to focus on for interviews to ensure comparisons across interviews is consistent.
Step 4: Site Visits	
<ul style="list-style-type: none"> ▲ Site visits are an effective tool for obtaining detailed information about local evaluation activities and environments. ▲ When conducted in conjunction with a survey, site 	<ul style="list-style-type: none"> ▲ When using site visits with high capacity organizations, consider a thorough vetting process and establish criteria to ensure appropriate organizations are selected.

Facilitators	Barriers and Potential Solutions
<p>visits can help to provide contextual and clarifying information about survey responses.</p> <p>▲ Identify someone at each site that can serve as the lead and work directly with site visit participants to schedule interviews.</p>	<p>▲ Have the client send the initial communication to sites. This worked well for the virtual focus group but we failed to do this in scheduling the site visits, which resulted in delays in scheduling some visits.</p>

Conclusions from the Evaluation Capacity Assessment Process

This chapter presents key conclusions with respect to the Y's evaluation capacity as well as the process that was used to complete the assessment. Although some conclusions are based on specific Y data, they have broad applicability to other nonprofits similar to the Y.

Y Specific Conclusions

Ys can capitalize on current organizational practices and Y culture

A consistent finding across methods is that Ys are strong in terms of having organizational practices and commitment necessary to support evaluation. A central attribute of the Y culture is the focus on mission and the importance on meeting community needs. A number of assessment participants discussed the Y mission in context of the pillars of youth development, healthy living, and social responsibility. This foundation lends itself well to promote evaluation as critical to a mission-driven organization, especially when there is a high level of leadership support.

There are concerns about whether current capacity and systems can support evaluation

There were concerns among Ys as to whether they have the appropriate capacity, systems, and expertise in place to ensure that high-quality evaluation. This was the case even among the high capacity sites that described antiquated or inadequate data systems, which were identified as a major barrier. For most Ys, upgrading data systems are cost prohibitive and will not be a feasible option. Instead, they will have to rely on low cost approaches such as using Microsoft Excel for data management and analysis but could benefit from additional training or guidance.

Large amounts of data are collected, but use is not being maximized

Across all the methods, it was clear that Ys are data oriented and at minimum administer member surveys on a regular basis. During the focus group with low capacity Ys, participants raised the issue of investing time to collect data but then failing to use it. This practice was also observed during the site visits within the high capacity Ys, which begs the question as to whether current data is being maximized. There is an opportunity to improve evaluation, especially in the childcare programs, which seem to have higher volumes of data collection due to program requirements (e.g. Head Start) or related to accreditation (e.g. NAEYC) but data are not being analyzed to document outcomes. At one site where childcare program is tracking developmental outcomes over the course of the year, staff did not recognize these as program outcomes.

Ys recognize the value in undertaking rigorous evaluation but need additional support

Based on the assessment results, most Ys seem comfortable assessing membership and satisfaction, especially with the additional support provided by SEER Analytics. Assessing changes in members' knowledge, attitudes, or behaviors was much less common. Even among the high capacity sites, it was clear that they are also challenged in measuring program outcomes. Several sites described their desire to improve in this area and collect better data but may lack sufficient knowledge or skills to be able to determine relevant outcomes and indicators, and then design appropriate data collection instruments. Beyond data collection, data analysis was an identified weakness across sites. Even though the majority of survey respondents (77% agreed or strongly agreed) were confident that their staff had a basic understanding of evaluation, the site visits revealed that little to no training on general evaluation concepts are provided to staff. There appears to

be a baseline of knowledge that needs to be established before Ys are expected to undertake more rigorous evaluations.

Ys are challenged in identifying additional resources to support evaluation

While most Ys would agree their organization could benefit by expanding evaluation activities, most point to a lack of resources as a primary barrier. Preparing staff can take significant resources and the various aspects of evaluation (design, planning, training, implementing, analyzing, and reporting) require substantial staff time. The programs with the strongest evaluation were those that were grant funded (e.g. United Way) because they were able to designate resources to support these activities. All methods revealed that Ys do not have the ability to allocate financial support or include a budget line item to ensure ongoing evaluation, especially in the current economic climate.

Conclusions About the Process Overall

Establishing open and frequent communication with Y-USA supported completion of the assessment in a collaborative manner

The time allotted for the assessment was approximately 7 months so there were times when immediate feedback from Y-USA was needed to keep evaluation activities moving forward. We set up weekly calls to check-in, plan for each step of the assessment, and address any potential barriers. Y-USA was also really helpful in facilitating contact with the sites when needed, such as when organizing the virtual focus group.

It is critical to define terminology to be used in the assessment

This was established early in the assessment, beginning at the kick-off meeting between Altarum and Y-USA staff. This was important for ensuring the Y and Altarum shared a common understanding about what evaluation capacity entailed for the Y, and what aspects of evaluation capacity were considered most relevant. While we established common terminology with Y-USA staff, we were less successful in doing this with the assessment participants, such as those who participated in the key informant interviews and the site visits. We found that some individuals were unfamiliar with the evaluation terms being used, particularly when discussing outcome evaluation. This made it challenging when probing for further information. We could have improved this aspect of the assessment by sharing more definitions with participants as well as relevant Y examples, to help them understand the type of information we were trying to gather.

The use of multiple methods worked well and is recommended for this type of assessment

Altarum proposed the use of multiple methods because we felt it was the most appropriate approach to gather information from both large and small samples of Ys, to triangulate information, and capture both quantitative and qualitative data. We felt the multiple methods complemented each other and increased the depth of analysis overall. The YMCA Evaluation Capacity Assessment Survey was a good tool for getting a broad view of the organization's capacity in a relatively short period of time. The survey results helped us to segment the audience into groups of Ys with similar evaluation capacity. This broke down a large population into more manageable groupings upon which we could conduct further assessment. The use of subsequent methods including the virtual focus group, key informant interviews, and site visits, then were used to explore the findings that emerged through the survey. These methods improved our understanding of the contextual factors contributing to evaluation capacity.

References

BTW Consultants, Inc. (n.d.) Evaluation capacity diagnostic tool. Retrieved from www.btw.informingchange.com.

Bruner Foundation. (2011). Modified evaluative thinking assessment tool.

Capacity 4 Health. (2011). Monitoring and evaluation capacity assessment tool. Retrieved from <http://library.capacity4health.org/category/topics/monitoring-and-evaluation-me/evaluation-basics/evaluation-capacity-assessment-tool>.

Connolly, P., & York, P. (2002). Evaluating capacity-building efforts for nonprofit organizations. *Organizational Design Practitioner*, 34(4), 33–39.

Mackay, K. (2002). The World Bank's ECB experience. *New Directions for Evaluation*, 93:88-99.

Major, D., Brennan, K. (2011). Expanding the impact of grantees: how do we build the capacity of nonprofits to evaluate, learn, and improve? Briefing Paper Topic 7. Washington, DC: Grantmakers for Effective Organizations.

McKinsey & Company. (2011). McKinsey organizational capacity assessment tool. Retrieved from <http://ocat.mckinseysociety.com/about-ocat/>

Nielsen, S. B., Lemire, S., & Skov, M. (2011). Measuring evaluation capacity—Results and implications of Danish study. *American Journal of Evaluation*, 32(3), 324–344.

Preskill, H., Boyle, S. (2008). Building evaluation capacity research: executive summary. Claremont Graduate University. Funded by the James Irvine Foundation.

Stockdill, S., Baizerman, M., Compton, D.W. (2002). Toward a definition of evaluation capacity building process: a conversation with the evaluation building literature. *New Directions for Evaluation*, 93:7-26.

Urban Institute and The Center for What Works. (2006). Building a common outcome framework to measure nonprofit performance. Urban Institute: Washington, DC.

Welsh, M., Morariu, J. (2011). Evaluation capacity building: funder initiatives to strengthen grantee evaluation capacity and practice. Washington, DC: Innovation Network.

Appendix A. Evaluation Capacity Assessment Survey



FOR YOUTH DEVELOPMENT
FOR HEALTHY LIVING
FOR SOCIAL RESPONSIBILITY

ONLINE SURVEY TOOL

YMCA EVALUATION CAPACITY ASSESSMENT SURVEY

GREETINGS

Thank you for participating in the YMCA Evaluation Capacity Assessment Survey! This survey is designed to assess the evaluation capacity of your local YMCA. The survey will take approximately 15-20 minutes to complete.

We are looking for YOUR OPINIONS; there are no right or wrong answers. You should answer based on your perceptions of what your local YMCA actually does, not what you think we might want to hear.

We will be sending this survey to YMCA staff across the country that are in positions that may involve evaluation activities. As such, multiple staff within your YMCA may receive the survey, and we hope all staff complete the survey individually so we can learn as much as possible about your evaluation capacity.

All data will be in aggregate; your individual responses or those responses from your YMCA will not be identified by name in any public reports.

We will be using the responses from this survey to learn more about YMCAs' capacity to conduct evaluations so that we can learn what is working well, and what areas could be improved through technical assistance, training, and other strategies. Your honest responses are important for helping us shape evaluation strategies across the Movement, and we greatly appreciate your taking the time to complete this survey.

If you have any questions about the survey, please contact: [*name redacted*]

IMPORTANT NOTE

Evaluation concepts sometimes mean different things to different people. So that we are all on the same page and you can answer consistently with your fellow YMCA peers, please use the following definition:

Evaluation is the process of examining and rating something based on important features. You can evaluate a program, process, policy or initiative.

Program Evaluation typically includes outcome and process evaluation.

- *OUTCOME evaluation* assesses changes in knowledge, attitudes, and behaviors as a result of participating in the program.
- *PROCESS evaluation* monitors program environment, as well as if activities were completed and how well they were implemented.

ORGANIZATIONAL CONTEXT

Organizational Culture & Practices Around Evaluation

	Strongly Agree	Agree	Disagree	Strongly Disagree
Our local YMCA sees evaluation as a tool that is integral to our work .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA models a willingness to be evaluated by ensuring that evaluations, both their process and findings, are routinely conducted and visible to others within and outside of our organization.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Results of program evaluations including findings about participant outcomes (as appropriate) are shared with leaders, staff, board members and funders.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The organization involves program staff and organizational leaders (as appropriate) in meaningful ways in evaluation planning, implementation, and discussion findings.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA values learning , as demonstrated by staff actively asking questions, gathering information, and thinking critically about how to improve their work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Organizational Commitment & Support for Evaluation

	Strongly Agree	Agree	Disagree	Strongly Disagree
Organizational leaders support and value program evaluation and evaluative thinking.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA has established clear expectations for the evaluation roles of different staff.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA ensures that staff have the information and skills that they need for successful participation in evaluation efforts (e.g., access to evaluation resources through Web sites and professional organizations, relevant training).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA allows adequate time and opportunities to collaborate on evaluation activities , including, when possible, being physically together in an	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

environment free from interruptions.

Our local YMCA provides **financial support** (beyond what is allocated for evaluation through specific grants) **to integrate evaluation** into program activities.

Our local YMCA has a **budget line item** to ensure ongoing evaluation activities.

Our local YMCA has **existing evaluation data collection tools and practices** that we can apply/adapt to subsequent evaluations.

Our local YMCA has **integrated evaluation processes** purposefully into ongoing organizational practices.

In our local YMCA, **sound data management** for evaluation is practiced, including collecting actionable data, ensuring consistency among different sets of data, and safeguarding data quality.

In our local YMCA, we have the **appropriate capacity, systems, and expertise** in place, internally or through partnerships, to ensure that high-quality, purpose-driven evaluation and measurement are practiced.

In our local YMCA, the evaluation work we are doing is part of a larger **organizational vision** to demonstrate the difference that we make and continuously improve our work.

Using Data to Inform Ongoing Work

Strongly Agree

Agree

Disagree

Strongly Disagree

Our local YMCA **modifies its course of action based on evaluation findings** (e.g., changes to specific programs or organizational-wide changes).

Evaluation findings are integrated into decision making when deciding what program options and strategies to pursue.

Managers **use evaluation data to set staff goals** and evaluate staff performance.

YMCA EVALUATION CAPACITY ASSESSMENT SURVEY

EVALUATION EXPERIENCE OF STAFF

Existing Evaluation Knowledge & Experience

	Strongly Agree	Agree	Disagree	Strongly Disagree
Our local YMCA has staff that have a basic understanding of evaluation (e.g., key evaluation terms, concepts, theories, assumptions).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The organization insures that there are staff members whose jobs or components of their jobs are dedicated to evaluation .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA can identify which data collection methods are most appropriate for different outcome areas (e.g., changes in norms require determining what people think about particular issues, so surveys, focus groups and interviews are appropriate).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA has staff with experience developing data collection tools and collecting data utilizing a variety of strategies, such as focus group sessions, interviews, surveys, observations and document reviews.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA has staff that know how to analyze data and interpret what the evaluation data mean .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA has staff that are knowledgeable about and/or experienced at developing recommendations based on evaluation findings.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Developing a Conceptual Model for Designing Outcome Evaluations

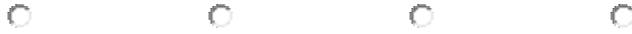
	Strongly Agree	Agree	Disagree	Strongly Disagree
Our local YMCA has articulated how we expect change to occur among individuals in our programs , and how we expect program specific activities to contribute to this change (such as through a logic model).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA has clarity about what outcomes want to accomplish in the short term (e.g., one to three years) and what success will look like.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA has articulated how our evaluation goals connect to a broader strategic plan .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA's evaluation design has the flexibility to adapt to changes in the environment and our related work as needed (e.g., benchmarks and indicators can be modified as the project evolves).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA has tools and methods for evaluating outcomes of our work .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Defining Benchmarks & Indicators

	Strongly Agree	Agree	Disagree	Strongly Disagree
Member satisfaction is regularly assessed .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Member outcomes (such as changes in knowledge, skills, attitudes, behaviors) are regularly assessed .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA can identify outcome indicators that are important/relevant for our work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA has identified what indicators are appropriate for measuring the impact of our work (e.g., changes in attitudes, changes in behaviors, or increase in knowledge).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our local YMCA can identify what indicators are appropriate for	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

measuring how we do our work.

Since some program goals, such as change in behaviors, can take years to achieve, our local YMCA **identifies and tracks interim outcomes that can be precursors of change, such as changes** in knowledge, attitudes, and beliefs that tell us if we are making progress and are on the right track.



Open Ended Questions

We are interested in learning more about how you are evaluating your work. Please complete the table below. In the first column, list any and all programs that you are currently evaluating. In the second column, describe the methods you are using. In the third column, list any and all data collection tool(s) that you are using to collect evaluation data. Please be as specific as possible, using actual names of programs and tools.

	Program Currently Being Evaluated	Evaluation Methods (e.g. surveys, interviews, focus group)	Data Collection Tool(s) Being Used
a.			
b.			
c.			
d.			
e.			

What types of program participant data management systems do you have (e.g., nFocus, Cayen, CitySpan)? Please list:

Please tell us about the systems you use to gather, store, and analyze data.

a. Explain the systems/software available to help you **collect evaluation data** (for example, a web-

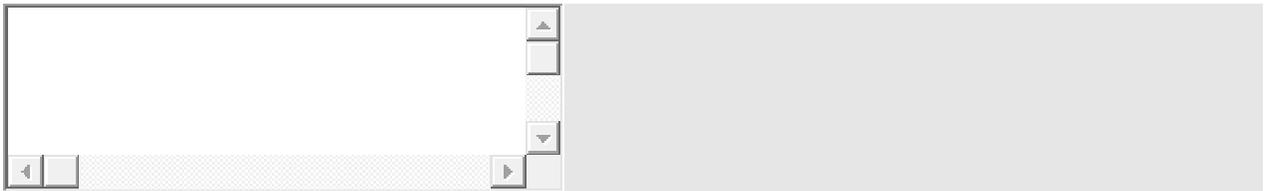
based survey account).

b. Describe the systems/software available to you to **store evaluation data** (for example, are data electronic or in paper format? Is this data stored online, shared drives, or in file cabinets? etc).

c. Describe the systems/software available to you to **analyze evaluation data** (for example, Microsoft Excel, SAS).



When you are conducting programs, do you typically measure how closely the program is implemented according to plan (i.e., measure fidelity)? Please explain.



Do you use logic models for any of your programs?

Yes No

If Yes, are these projects for or related to United Way?

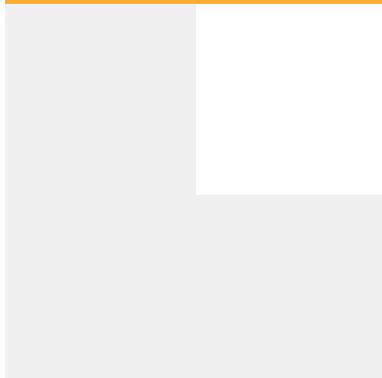
Yes No



Do you evaluate adult membership beyond satisfaction?

No

Yes, please explain:



Appendix B. Virtual Focus Group Protocol

Focus Group Moderators Guide

YMCA Focus Group on Evaluation Capacity

BACKGROUND

Welcome to our group discussion. Thank you for taking the time to participate in today's discussion about the evaluation capacity.

My name is [name redacted] and I work for Altarum Institute. We are a research and consulting organization and are working with the Y-USA to help them better understand the different factors that can impact evaluation capacity of local YMCAs. Today's discussion is building on information that we collected in a recent survey, which you completed in the past month. We wanted to bring a group together to get additional information about some of the responses that have emerged from the survey. So that's why you all have been invited to participate.

Today, we are going to use a focus group format. How many of you have participated in a focus group before? [take poll: yes/no] Because of the virtual nature, we will have to make some adjustments.

- Like in a typical focus group, I have a list of questions that I will be asking. We'll spend a few minutes on a question and then move onto the next.
- Only one person should talk at a time. Because I can't see you, we are going to use one of our webinar features—and ask that you raise your hand when you would like to respond. You'll notice on the top of your screen describe how to access feature. This will help me manage the conversation. Let's practice now—can everyone raise their hand?
- We'll also use the poll feature from time to time. Here's an example of how this will work-- [take poll: yes/no].
- We ask that you try to focus on the questions. We know it would be easier if we were all in person but please do your best to not multi-task and to follow the discussion.
- We would like everyone to participate. But, you each don't have to answer every question. We specifically selected you all because as a group, you have a range of experience with evaluation. So, you may feel that some questions may not apply to you, and that's okay. I may call on you though, if I haven't heard from you in a while.
- It is also ok to disagree with one another. We want to hear everyone's point of view. If you disagree, please do so respectfully.
- And lastly, we really want your honest opinions. There are no right and wrong answers. Please feel free to share your thoughts, whether they are positive or negative. Your comments will be used to shape evaluation strategies across the Movement.

Today's discussion should take about an hour and a half so let me know if you need to leave before then.

Are there any questions before we start?

INTRODUCTIONS/ICEBREAKER

Let's get started. I'd like to start by having each of you share a little about yourself.

1. My name is _____.
2. I work for the _____ YMCA located in _____.
3. My experience with evaluation has been _____.

INTERVIEW QUESTIONS

Since all of the questions today are related to your evaluation capacity, let's review what we mean by this.

Evaluation capacity is the ability of your Y to conduct evaluation. **Evaluation** is the process of examining and rating something based on important features. You can evaluate a program, process, policy or initiative.

Program Evaluation typically includes outcome and process evaluation.

- *OUTCOME evaluation* assesses changes in knowledge, attitudes, and behaviors as a result of participating in the program.

- *PROCESS evaluation* monitors program environment, as well as if activities were completed and how well they were implemented.

Today, we're going to discuss some of the key issues that were raised in the assessment survey. We know that evaluation looks very different from Y-to-Y. For example, evaluation may occur at the branch level, but some of the evaluation resource and strategy decisions may happen at the association level. You can determine what level makes most sense as you're responding to the questions.

A. Organizational Commitment & Support for Evaluation.

Let's start by talking about Organizational Commitment & Support for Evaluation from the assessment survey you took a few weeks ago. There were a number of questions in this section with high rates of disagreement among respondents. So, we'll go through these and see what you all think may be some of the reason behind the scores. [pull up score]

1. As you can see, around 40% of respondents did not feel that there are clear expectations for evaluation roles. Is there anyone that had this experience?
 - In thinking about the future, how could your local Y improve on making evaluation responsibilities clear to staff?
 - Within your local Y, what staff do you think should be responsible for evaluation? Generally, what position is typically responsible for evaluation?
 - Do you think there should be a job position dedicated to evaluation? [take poll: yes/no] Why or why not?
 -
2. We heard that many respondents felt there was not enough time to collaborate on evaluation activities. [pull up score]
 - How are evaluation activities carried out? Is there one person who does this isolation? Or, as a group?
 - How well does this process work?
 - What are the barriers to collaborating on evaluation?

B. Capacity

If you're going to conduct evaluation, you obviously need staff that can do this work. Let's think about what kind of knowledge and experience is necessary to do this work.

3. There were a few questions about whether local staff have the knowledge and experience to make appropriate evaluation decisions. [pull up scores] For example, you can see that almost half of respondents disagreed with the first statement about local Y staff being able to identify appropriate data collection methods based on the desired outcome. Almost half also disagreed with the second statement about not having experience in developing data collection tools and using different methods, such as surveys and interviews.
 - In thinking about your local Y, what might be some of the reasons behind this?
 - Do you think this is an issue of training? In other words, if staff received additional training on data collection methods, would this be helpful? [take poll: yes/no]
 - Are there other ways to improve knowledge about data collection methods?

Beyond having the right staff, we know there are other items you need to conduct high quality evaluations.

4. [pull up scores] The next area where respondents disagreed, was related to whether local Y's have the capacity and systems in place to conduct evaluation. What is preventing local Y's from having these in place?
 - Let's start with the capacity to conduct high-quality evaluations. What factors are impacting your Y's capacity to do this?
 - Next, let's talk about systems to conduct high-quality evaluations. For example, some systems might include What systems do you need to accomplish this? What systems do you have that are working or not working?

5. The largest issue in the section was related to the financial support & having a budget line item to support evaluation activities. [pull up score]
 - Do you typically have funds allocated to evaluate your programs?
 - If no, why do you think sufficient funds are not allocated?
 - Why do you think evaluation is not a funding priority?
 - What is the effect on your work of not having sufficient evaluation funding?

C. Developing a Conceptual Model for Designing Outcome Evaluations

So, we've talked about what kind of resources you need to carry out an evaluation—staff, time, funding. Now, let's talk about how evaluation can be helpful.

6.
 - [pull up score] About half of respondents did not think their local Y really made the connection between program activities and the long term changes you expect to see among your members. When we talk about making a connection between activities and changes, what does this mean to you?
 - In thinking about the program we just saw, does your Y use logic models or something else, describing the changes you're expecting to see? What has worked well?
 - Are these shared with all the staff?
 - How are they useful?

D. Defining Benchmarks & Indicators

We have a few questions left. Let's talk specifically about the kinds of changes that you might be tracking.

7. [pull up scores] Almost 40% of respondents think that member outcomes are not regularly assessed—and this would mean looking at whether your programs are changing knowledge, skills, attitudes & behaviors. Are your Y's looking at these kinds of changes? [take poll: yes/no]
 - For those that answered yes, can you give us some examples of the changes you are tracking? How do you use this information?

- For those that answered no, why do you think your Y is not tracking these changes? What makes it difficult?
8. In thinking about the bigger picture, in what ways can your local Y track how you do your work? In what ways can this be improved?
 9. Before we close today's discussion, are there any other issues or questions you'd like to raise?

IV. CLOSING

Thanks again for participating in today's focus group. We have learned a lot about the issues that impact your evaluation capacity.

Thank you again for your time and the information you provided.

Appendix C. Key Informant Interview Guide

Key Informant Interview Guide

YMCA of the USA Evaluation Capacity Assessment

[*Note: Instructions to interviewers are in brackets and italics. This information is for the interviewer only and should not be read to the respondent.*]

Interviewee Name:
YMCA Name:
Survey Respondent Number:
Interviewer:
Date:

Introduction and Consent

Thank you for agreeing to participate in this interview. My name is [name redacted]. I work for Altarum Institute, a non-profit health systems research organization. We are working with YMCA of the USA (Y-USA) to conduct an evaluation capacity assessment of YMCAs.

This interview is designed to help us learn more about your evaluation practices and capacity to inform potential evaluation strategies, systems, and models that other YMCAs like you might be able to adopt to improve their evaluation capacity. You were selected to participate in this interview based on your responses to the YMCA Evaluation Capacity Assessment Survey that you completed over the summer. We will be talking about your responses to this survey so that I can get more contextual information around your responses, and we will also talk about your experiences and feelings about evaluation more generally. Your input will help us increase our understanding of how Ys are implementing evaluations and their capacity to conduct outcome evaluations, as well as identify areas where Ys are at increased need for resources, training, or technical assistance on evaluation.

This interview should last approximately 45-60 minutes. Your organization's name and location and your general job title may be identified in reports prepared for this study and in files given to Y-USA. However, none of your responses will be released in a form that identifies you or any other interviewees by name. Participation in this interview is voluntary, and you may decline to answer any question or stop the interview at any time.

I would like to tape record this interview to ensure that we are accurately recording your responses and to check against our notes. Do we have your permission to do that? [*IF PERMISSION IS GRANTED, BEGIN RECORDING. IF PERMISSION IS NOT GRANTED, SAY, "I understand," AND PROCEED WITH THE INTERVIEW, RELYING ON YOUR NOTES TO CAPTURE RESPONSES.*]

Do you have any questions before we begin?

Introductory Questions

I would like to begin by asking you a few background questions to learn more about you and your work at the Y.

1. What is your role at the Y?

2. How long have you served in this role?
3. What types of evaluation activities do you do in your role?

Organizational Commitment and Support for Evaluation

Now I want to gather more information from you based on your responses to the YMCA Evaluation Capacity Assessment Survey. First, I would like to talk about your Y's commitment to and support for evaluation.

4. Do you feel that your Y ensures that staff have the information and skills that they need to successfully participate in evaluation efforts? For instance, do you hire staff with evaluation skills, or do you provide existing staff evaluation resources and relevant training?
 - a. *[If yes]* What types of resources and training do staff have access to? How are these useful to your evaluation efforts?
 - b. *[If no]* What types of information and skills are lacking among staff? Do you have any suggestions for how this could be improved?
5. Does your YMCA give staff members time and opportunities to collaborate on evaluation activities?
 - a. *[If yes]* How frequently does this occur? Can you provide any examples of how Y staff have collaborated on evaluation activities?
 - b. *[If no]* Do you think having specific time set aside to collaborate on evaluation activities with other staff would be useful for building your Y's evaluation capacity? Why or why not?
6. Are there financial resources allocated to evaluate program activities?
 - a. *[If yes]* Where does this funding come from? Is there a line item in the budget for evaluation? How is the amount of this funding determined? Is this amount of funding sufficient to evaluate program activities?
 - b. *[If no]* Why do you think there is no funding for evaluation? How does this affect your Y's efforts to evaluate programs?
7. On the survey, we asked whether your Y has integrated evaluation processes purposefully into ongoing organizational practices and you *[insert survey response]* with this statement.
[Question 13 on survey]
 - a. *[If agree/strongly agree from survey]* Can you provide examples of how your Y has done this? (Probe: do you have examples for whether or not you have documented strategies and policies in place for applying evaluation?)

- b. *[If disagree/strongly disagree from survey]* Do you have thoughts on why your Y has not done this? (Probes: lack of time, staff training, funding)

- 8. In the survey we asked you about whether your YMCA has established clear expectations for the evaluation roles of different staff. You answered *[insert survey response]* that your YMCA did/did not do this. [Question 7]
 - a. *[If agree/strongly agree from survey]* What are the expectations for evaluation roles of staff at your Y? How is this communicated to staff?
 - b. *[If disagree/strongly disagree from survey]* How does not establishing clear expectations for evaluation roles hinder your Y's evaluation capabilities?

- 9. In the survey, we asked about whether your Y had sound data management for evaluation. This includes collecting actionable data, ensuring consistency among different sets of data, and safeguarding data quality. You answered that you *[insert survey response]* with this statement. [Question not in analysis, AP]
 - a. *[If agree/strongly agree from survey]* Tell me more about your storage, analysis and use of evaluation data. What types of system(s) do you have in place? Where do you store data? Does your evaluation data “talk to”/work in concert with your other data systems, such as membership data? How is sensitive data safeguarded?
 - b. *[If disagree/strongly disagree from survey]* Which components of sound data management does your Y struggle with? (Probes: collecting actionable data, ensuring consistency among different sets of data, using data from different systems together, safeguarding data quality, and/or data security)

- 10. In the survey, we asked whether your Y has the appropriate capacity, systems, and expertise in place, internally or through partnerships, to insure high-quality, purpose-driven evaluation and measurement practices. You answered that you *[insert survey response]* with this statement. [Question not in analysis, AQ]
 - a. *[If agree/strongly agree from survey]* Does your Y have capacity, systems, and expertise in place internally or through partnerships? How does your Y ensure that evaluations are high quality? Purpose-driven?
 - b. *[If disagree/strongly disagree]* What capacity, systems, or expertise are most lacking at your Y? Have you ever tried to fill these gaps externally through partnerships?

Existing Evaluation Knowledge and Experience

Now I would like to transition to some questions about the existing evaluation knowledge and experience at your Y.

- 11. Are there specific people in your Y whose job or components of their job is to work on evaluations?

- a. *[If yes]* About how many positions? What roles do these staff serve?
 - b. *[If no]* How does not having dedicated staff members impact your Y's ability to conduct program evaluations?
12. In the survey, we asked whether your Y can identify which data collection methods are most appropriate for different outcome areas (e.g., changes in norms require determining what people think about particular issues, so surveys, focus groups and interviews are appropriate). You *[insert survey response]* with this statement. [Question 19]
- a. *[If agree/strongly agree from survey]* Can you tell me more about the process you use at your Y to determine which data collection methods are most appropriate to measure outcomes?
 - b. *[If disagree/strongly disagree]* How does not knowing which data collection methods are appropriate impact your Y's ability to measure program outcomes?
13. Do you or does anyone at your Y have experience developing data collection tools for a variety of methods, such as focus groups, interviews, surveys, and document reviews?
- a. *[If yes]* Which types of tools (Probe for all methods)? Have you ever used any of these methods to collect data? Please provide examples.
 - b. *[If no]* Where do you typically obtain data collection tools (Probes: partners, Y-USA, internet/websites). Which data methods have you used to collect data in the past (Probe for all types)?

Developing a Conceptual Model for Designing Outcome Evaluations

For the next set of questions, I would like to ask you about your Y's experience in designing outcome evaluations.

14. Does your Y articulate how they expect change to occur among the individuals that participate in your program, and how program activities contribute to this change?
- a. *[If yes]* How is this articulated to you? How does this help to improve program evaluations?
 - b. *[If no]* How does this hinder your ability to conduct program evaluations?
15. Have you ever created a logic model for your program before?
- a. *[If yes]* For which types of programs? What are your thoughts on logic models? Are these useful to you?
 - b. *[If no]* Have you ever seen a logic model that someone else has developed before? What are your thoughts on logic models? Would they be useful to you in planning and conducting program evaluations?

16. Does your Y define what success looks like for its programs? That is, what outcomes will look like in 1-3 years?
 - a. [If yes] Please explain how your Y does this.
 - b. [If no] Does not knowing what success looks like for your programs impact your ability to measure program outcomes? If yes, in what ways?
17. Does your Y connect program evaluation goals to a broader strategic plan?
 - a. [If yes] Can you give me some examples?
 - b. [If no] Would connecting evaluation goals to strategic plans improve your ability to evaluate programs? Why or why not?
18. When you are evaluating a program, are you able to be flexible and modify benchmarks as the project evolves?
 - a. [If yes] Can you give specific examples of how you have changed program benchmarks or indicators to adapt to changing work or environments?
 - b. [If no] What makes it difficult to be flexible? Does this impact your ability to evaluate programs?

Defining Benchmarks and Indicators

My next questions are about designing benchmarks and indicators for evaluation.

19. At your Y, do you regularly assess member outcomes such as changes in knowledge, skills, attitudes, and behaviors?
 - a. [If yes] How often is this done? Can you give specific examples?
 - b. [If no] Why do you think your Y does not regularly assessment member outcomes? How does this impact your Y's ability to assess member changes?
20. Are you familiar with the term 'indicators' used in evaluation? What does this mean to you? (If unknown or for clarification: An indicator is a metric that provides information to measure performance, that is, it is a sign that something exists or is true.) Is your Y able to identify outcome indicators that are important/relevant to your work?
 - a. [If yes] What is the process your Y uses to identify the indicators that are most relevant to your work? Has your Y already identified the indicators that are appropriate for measuring the impact of your work (e.g., changes in knowledge, attitudes, or behaviors) for key programs? Which programs has this been done for?

- b. [If no] If you knew the indicators that were most relevant to your programs, would this improve your ability to evaluate your programs? Why or why not?
21. Some program outcomes, such as changes in disease state, can take years to achieve. Does your Y identify and track interim outcomes that are precursors to change— such as change in knowledge, attitudes, and beliefs—that tell you programs are making progress and are on the right track?
- a. [If yes] Can you provide any specific examples?
 - b. [If no] Would tracking interim outcomes participants achieved as a result of participating in your programs help to tell you whether programs are on the right track? Why or why not.

Summary and Closing

We are almost finished. I just have a couple additional general questions for you.

22. When you think about your Y's evaluation capacity as a whole, where do you feel your Y is strongest? Weakest?
23. What resources, training, or technical assistance are most needed by your Y to improve your evaluation capacity? What would be the best way to provide these (Probes: indicators/logic models for key programs, Y Exchange, in-person trainings, webinars, one-on-one TA)?

I have now reached the end of my questions. Is there anything else you would like to add that you feel would help us to better understand the evaluation capacity at your Y? Do you have any final questions for us? Thank you so much for your time today!

Appendix D. Site Visit Protocol

Interviewer Guide

YMCA Site Visit on Evaluation Capacity

A. Introductory Questions

1. What is your role at the Y?
2. How long have you served in this role?
3. What is your role in evaluation at your Y?

B. Organizational Commitment & Culture around Evaluation

Leadership

4. What does evaluation mean to you? Evaluation capacity? In terms of having the appropriate capacity and systems to carry out evaluation activities, how would you describe your Y?
5. In general, how do you view evaluation in the context of your Y? How is evaluation viewed by others?
6. Do you consider evaluation integral to your organizational mission? How does evaluation support your mission or implementation of your organizational strategic plan? Please explain.
7. What prompted you to undertake evaluation? How do you institutionalize evaluation as an organizational practice?
8. Are financial resources allocated to evaluation? How do you ensure there are sufficient resources for evaluation? (grant funding, line item, % of funding for each program)
9. Is there a dedicated position(s) for evaluation? What job responsibilities does this person(s) have?
10. Are staff involved in evaluation-focused discussions, in terms of planning or implementing? If so, please explain. How collaborative is this process? How well does this process work? What are the opportunities and barriers to collaborating on evaluation?
11. What are the expectations for staff in terms of evaluating programs? How do you ensure that staff expectations around evaluation are clearly communicated? What staff are typically responsible for evaluation?

Other staff

12. In your opinion, do you think evaluation is seen as integral to your organizational mission? Is this communicated to you? If yes, how?

13. What prompted you to undertake your current evaluation activities?
14. How do you feel leadership supports evaluation? Please explain.
15. Do you feel that you have adequate funds to ensure sufficient resources for evaluation? Please describe some of the factors that make this easy or challenging.
16. Are staff involved in evaluation-focused discussions, in terms of planning or implementing? If so, how well does this process work? Do you feel your feedback is valued?
17. Do you feel that leadership has established clear expectations for staff roles? Please explain.
18. In terms of having the appropriate capacity and systems to carry out evaluation activities, how would you describe your Y?

C. Using Data to Inform Ongoing Work

Leadership

19. Do your evaluation findings influence how you do your work? If yes, please share some examples of how evaluation can inform your decisions as a leader at your Y (e.g. pursuing funding; evaluate staff performance; determine strategic direction).
20. Are evaluation findings shared, internally and externally? How often is this done? Please share some examples. In what ways is this important?

Other Staff

21. In what ways do your evaluation findings influence how you do your work? Please share some examples of how evaluation can inform your decisions as _____ (e.g. set program goals; evaluate staff performance).

Are evaluation findings shared, internally and externally? How often is this done? Please share some examples. In what ways is this important?

D. Existing Evaluation Knowledge & Experience

Leadership

22. When hiring new staff, how important is evaluation experience? Do you look for evaluation experience and skills in jobs candidates? If so, what types of experiences and skills?
23. Is training on evaluation provided to staff? What types of training are provided and how often? Are staff encouraged to seek evaluation training? If so, how are they encouraged?

24. For those staff responsible for evaluation, do you feel they have an adequate understanding of evaluation? Are they able to determine appropriate methods and develop data collection instruments, if necessary? Please describe. If not, then what adjustments are made?
25. Overall, do you feel your staff have the information and skills necessary to carry out evaluation activities? In what ways do you support them? What makes this easy or challenging?

Other Staff

26. For those staff responsible for evaluation, do you feel they have a basic understanding of evaluation?
27. Is training on evaluation provided to staff? What types of training are provided and how often? Is this training sufficient? How are you notified of training opportunities? Are you encouraged to attend evaluation training?
28. Are you ever tasked with identifying appropriate methods and developing data collection instruments? If so, how often? How do you approach these tasks? What has worked well? What has been challenging?
29. Overall, do you feel staff have the information and skills necessary to carry out evaluation activities? In what ways does leadership support this development?

E. Evaluation Design & Implementation

Leadership

30. To what extent do you evaluate adult membership beyond satisfaction? What do you assess? How do you use this information?
31. For Y programs, have you described how activities will lead to short-term and long-term outcomes, and eventually achieve your program goals? Have you done this through logic models? If so, how important are logic models in conducting evaluation work? Can you share some examples with me? Are these theories of change regularly shared with your staff? How are they discussed? Who is included in those discussions? To what extent do staff understand them?
32. Are you assessing outcomes in any of your programs? If so, what types of outcomes are you assessing—changes in knowledge, skills, attitudes, behaviors? Please describe by program. How did you determine these?
33. Are you assessing fidelity or how programs are being implemented (i.e. process evaluation)? If so, please describe examples (and share relevant documents) of programs and how processes are being tracked.
34. What data collection methods does your Y use? What factors influence the selection of these methods? What staff preparation is necessary to use these methods?

35. What data collection instruments did you use? Who developed these instruments? Can you please share these? Do you feel these instruments captured the necessary data?
36. Is there a typical data collection process at your Y? If so, what is it? What worked well? What was challenging?
37. Once data collection was completed, who was responsible for the analysis? What has worked well? What has been challenging?
38. What system did you use to manage and analyze the data (e.g. CitySpan, Nfocus)? How did you select this system? Can you walk me through your system?
39. In hindsight, what changes would you make (e.g. methods, instruments, system)?

Other Staff

40. To what extent do you evaluate adult membership beyond satisfaction? What do you assess? How do you use this information?
41. For Y programs, have you described how activities will lead to short-term and long-term outcomes, and eventually achieve your program goals? Have you done this through logic models? If so, how important are logic models in conducting evaluation work? Can you share some examples with me? Are these theories of change regularly shared with your staff? How are they discussed? Who is included in those discussions? To what extent do staff understand them?
42. Are you assessing outcomes in any of your programs? If so, what types of outcomes are you assessing—changes in knowledge, skills, attitudes, behaviors? Please describe by program. How did you determine these?
43. To what extent are you assessing fidelity or how programs are being implemented (i.e. process evaluation)?
44. What methods did you choose (for outcome or process monitoring)? What factors influenced the selection of these methods? What staff preparation was necessary to use these methods?
45. What data collection instruments did you use? Who developed these instruments? Can you please share these? Do you feel these instruments captured the necessary data?
46. Can you describe the data collection process? What worked well? What was challenging?
47. Once data collection was completed, who was responsible for the analysis? What has worked well? What has been challenging?

48. What system did you use to manage and analyze the data (e.g. CitySpan, Nfocus)? How did you select this system? Can you walk me through your system?
49. In hindsight, what changes would you make (e.g. methods, instruments, system)?

F. Closing

50. When thinking about evaluation, what does your Y do well? What could be improved?
51. What types of evaluation advice, materials, or training would be useful for you to improve upon your Y's evaluation activities?
52. What evaluation advice would you share with other Y's? Is there anything else you'd like to share?