idealware

# The Reality of Measuring Human Service Programs:

## RESULTS OF A SURVEY

January 2014

# TABLE OF CONTENTS

# ABOUT IDEALWARE

Idealware, a 501(c)(3) nonprofit, provides thoroughly researched, impartial, and accessible resources about software to help nonprofits make smart software decisions. Nonprofits maintain a complicated relationship with technology. Most know that software can streamline their processes and help fulfill their missions more efficiently and effectively, yet lean staffing and tight budgets mean they're unable to devote the time necessary keep up with new technologies and find the right tools.

From the most basic questions (like how to use software to help manage emailing hundreds of people at once), to the more complex (like understanding the role of social networking and mobile phone text-messaging in fundraising strategy), organizations need a trusted source for answers. Idealware provides an authoritative online guide to the software that allows U.S. nonprofits—especially small ones—to be more effective. By synthesizing vast amounts of original research into credible and approachable information, Idealware helps nonprofits make the most of their time and financial resources.

# AUTHORS AND CONTRIBUTORS

## Rachel Christian, Research and Training Intern

As the 2013 Summer Research and Training Intern, this report was Rachel's primary project. Rachel has been involved in human service nonprofits as an active volunteer since 2007. She attends Loyola University Maryland where she studies writing and sociology. Rachel has worked for the Catholic Campaign for Human Development whose mission is to end poverty and injustice in the United States where she created and executed a social justice speaker series for college campuses. Rachel has received recognition awards for her dedication to community service and her research on homelessness in Baltimore. She brought her passion and experience for nonprofit work into her work with Idealware.

## Laura S. Quinn, Executive Director

Laura has been working in the software sector for more than 15 years. As Idealware's Executive Director, she directs Idealware's research, writing, and training to provide candid reports and articles about nonprofit software. Prior to Idealware, Laura founded Alder Consulting, where she helped nonprofits create internet strategies, select appropriate software, and then build sophisticated websites on limited budgets. She has also selected software, designed interfaces, and conducted user research for multimillion dollar software and website implementations with such companies as Accenture and iXL. Laura is a frequent speaker and writer on nonprofit technology topics.

## Kyle Andrei, Research Analyst

As Research Analyst, Kyle is responsible for researching software through demos, interviews, and surveys, and using that information to create Idealware's reports and articles. In addition, Kyle also draws on his broadcast experience to produce Idealware's Ask Idealware videos. Outside of Idealware, Kyle volunteers with the Maine League of Young Voters as chair of the Civic Guide Committee, providing nonpartisan guides to the civic process in Portland. Kyle is a graduate of Indiana State University, where he studied broadcasting, managed the student radio station, and volunteered on local election campaigns.

## Elizabeth Pope, Director of Research and Operations

Elizabeth Pope directs Idealware's software research projects, conducting interviews and product demos, writing articles and reports, and helping to design project methodology. She earned an M.S. in Library and Information Science from the University of Illinois at Champaign-Urbana, where she honed her research, instruction, and tech skills. Before moving to Maine, Elizabeth worked in archives and libraries in New York City for several years. Her background also includes fundraising and content development for nonprofit organizations.

# EXECUTIVE SUMMARY

In the summer of 2013, Idealware created and distributed a survey to learn how human service organizations from our own mailing list are actually using technology to measure and evaluate the outcomes of their programs. The survey looked at a general overview of outcomes measurement and program evaluation topics, from how frequently they look at data and how much time they spend doing so to what types of metrics the organizations were tracking.

To further understand the realities of measuring program effectiveness, Idealware conducted a site visit and interview of three human service organizations in Portland, Maine. The results clearly show that the respondents are struggling to measure their programs. For example:

**Many nonprofits are only tracking the most basic data to measure programs.** While the majority of organizations were tracking the activities they conduct and participation in their programs, well less than half of survey respondents were using short-term satisfaction metrics—only 42 percent of the 120 surveyed were tracking short-term client satisfaction with the program, and only 35 percent were tracking how many program participants went on to reach a specific milestone, like graduating high school or finding employment.

**Few are tracking any kind of long-term assessment metrics.** Only 23 percent said they were tracking and analyzing either client satisfaction in the long term, or concrete measures of long-term program success.

**Organizations are confused about what it means to measure their impact on the community.** Of the organizations that said they were tracking their attributable impact on the community, over 60 percent weren't actually looking at the measures that would help them determine that—like comparing their results to those of a control group or pulling in supporting data from other sources. Without these pieces of information, it's almost impossible to separate your own results from those of other services.

**Most nonprofits lack the software they need.** Over half of the organizations surveyed were tracking their program data in Excel spreadsheets instead of a database, which may serve them in the short term but will limit them as their evaluations grow more complex.

**Only a minority of respondents feel their measurement methods are working.** Only 45 percent of respondents said they knew how well their programs were working, and only about 42 percent felt that their evaluations accurately reflected their impact.

**Nonprofits are limited by time and resources.** Over 25 percent of respondents indicated that the lack of time, money, or software was their biggest hurdle in evaluating their programs.

Based on the findings from this survey, we suggest that while there is an abundance of theories and terminology around program evaluation in the sector, nonprofits lack the practical framework to make sense of them and to put the process into practice.

# THE QUEST TO MEASURE PROGRAM EFFECTIVENESS

Nonprofits hoping to improve their programs' services or reach must first understand their own effectiveness. Measurable numbers—how many meals served at a soup kitchen, how many students in a mentoring program graduate high school, what percentage of the target population does not have access to affordable housing—are critical data that can help organizations identify where they can improve their programs.

The public conversation about measuring programs is often theoretical, focusing on what organizations *could* or *should* do rather than on tactics for success. At Idealware, we were curious about the opposite—what specific tactics are organizations using to measure their programs? What are the barriers? To find out, in the summer of 2013 we created and distributed a survey to human service organizations on our own mailing list—generally small to mid-sized organizations; see Appendix A for the demographics of the survey respondents—to learn how they were actually using technology to measure and evaluate the outcomes of their programs.

Of the more than 24,000 subscribers on the Idealware mailing list, 120 organizations completed the survey. The survey looked at a general overview of program evaluation topics, from how frequently they measured their programs and how much time they spent doing so to what types of metrics they were tracking. To augment the survey, we selected three organizations in Portland, Maine, where Idealware is based, and conducted site visits and interviews with them to understand their strategies for measuring programs.

The results from this survey clearly do not statistically represent the whole of the nonprofit sector. But there's no reason to suspect that the picture it presents—one of a sector struggling with measurement—is more negative than the reality. In fact, as the Idealware mailing list consists of organizations who found Idealware and have specifically signed up to hear more about technology topics, including technology to help with programs and program evaluation, it's reasonable to suspect that nonprofits on the Idealware mailing list are *more* savvy when it comes to strategic topics like program evaluation than the sector at large.

> "Every program has its share of challenges and lessons learned. These are important to know and share."

As is true with most emerging fields, program evaluation is riddled with conflicting and inconsistent terminologies and theories. For example, as a sector we often cannot agree on what specific terms mean—program evaluation vs. outcomes measurement, or output vs. outcomes vs. impact. Frequently, these terms have subtly different definitions for different organizations, consultants, or even experts.

But whatever we actually call it, nonprofits need to measure their programs. And many need a more-detailed tactical framework than is widely available to put this process into practice.

Idealware created the following model to try and reconcile much of the terminology into a format that resonates with what organizations actually do on a day-to-day basis. The types of measurements that form the basis of most-effective strategies are on the bottom of the pyramid; more sophisticated metrics are toward the top.

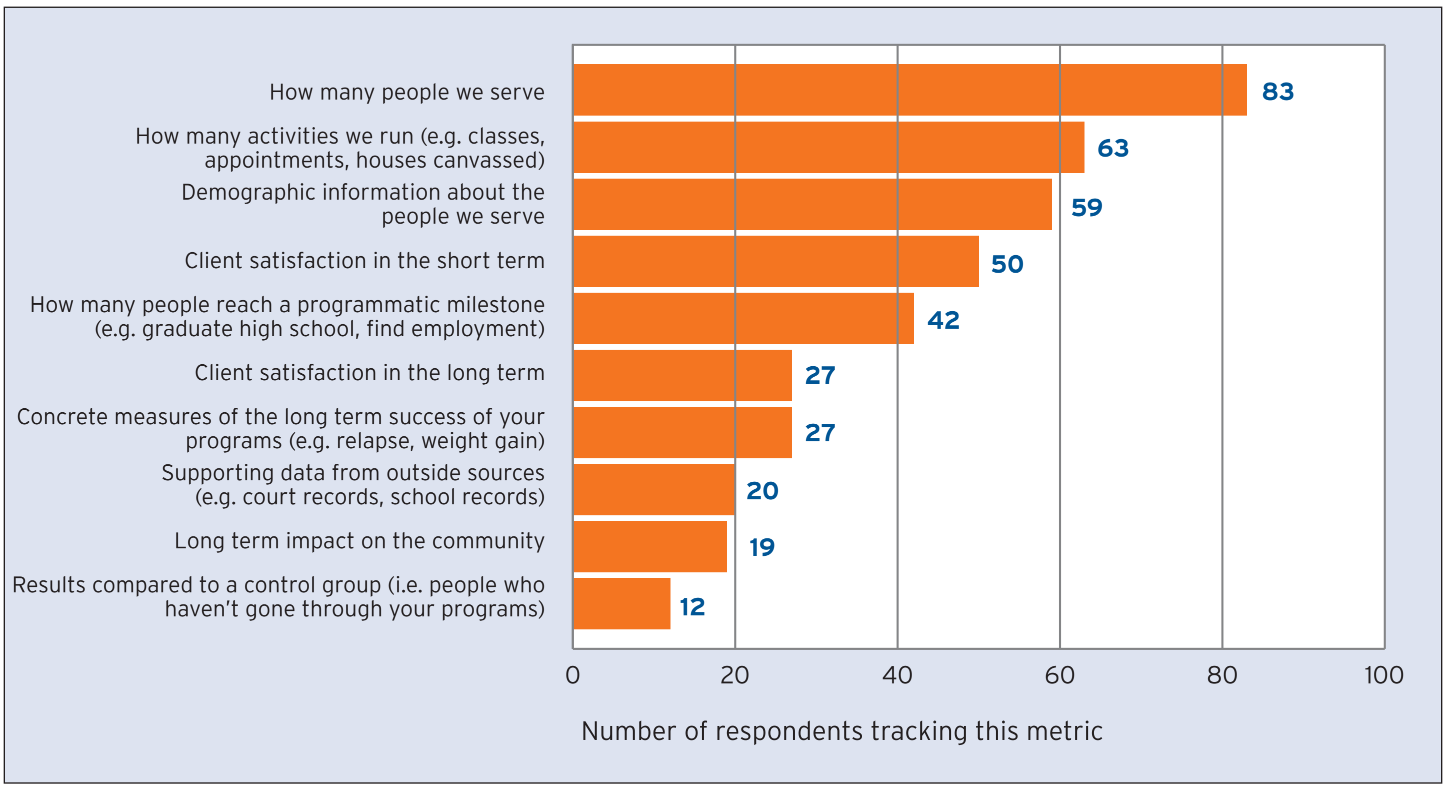
Most organizations will benefit from working from the bottom up—skipping up to one of the upper tiers of long-term success before you have a firm understanding of your basic program metrics will result in confusion and inaccuracy. How can you know the long-term impact of your programs on the community without first knowing how many clients you serve?

# THE NONPROFIT REALITY

How are the nonprofits that responded to the survey measuring the effectiveness of their programs? A full list of possible metrics would be difficult to define and include in a survey, but we presented participants with a list of 10 generalized metrics that would apply to most human service organizations and asked them to select the level to which they were tracking and analyzing each metric, if at all. This list was not intended to be definitive, but to present options that would be relevant to a broad range of organizations.

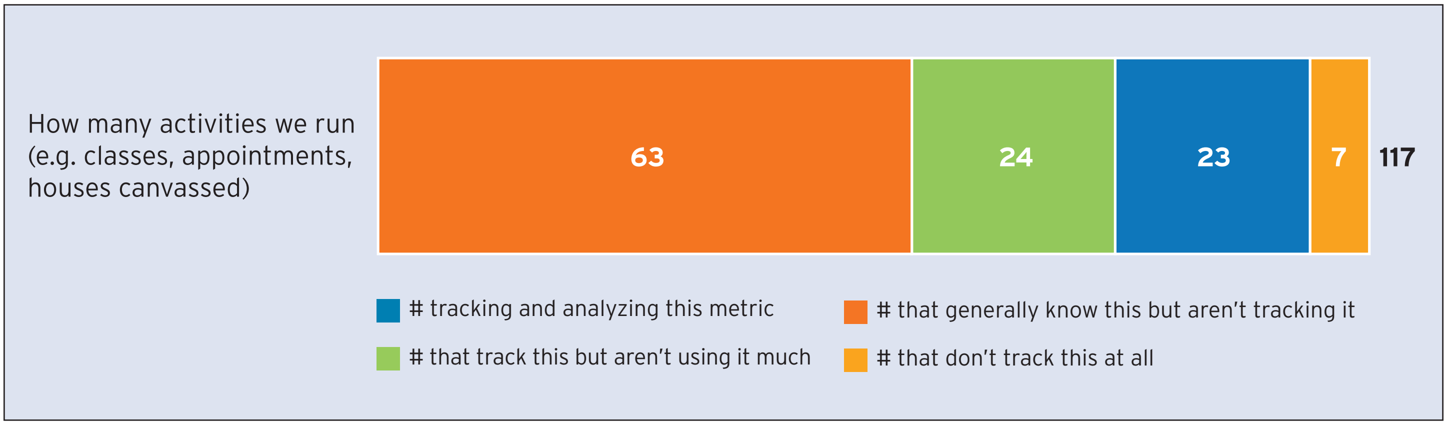
Overall, survey respondents were far more likely to be tracking the most-direct, tangible metrics rather than short-term results, let alone long-term results. These metrics also corresponded to the pyramid we described earlier, with more organizations tracking metrics at the bottom of the pyramid but relatively few tracking those toward the top.

Let's look at metrics from each tier of the pyramid individually.



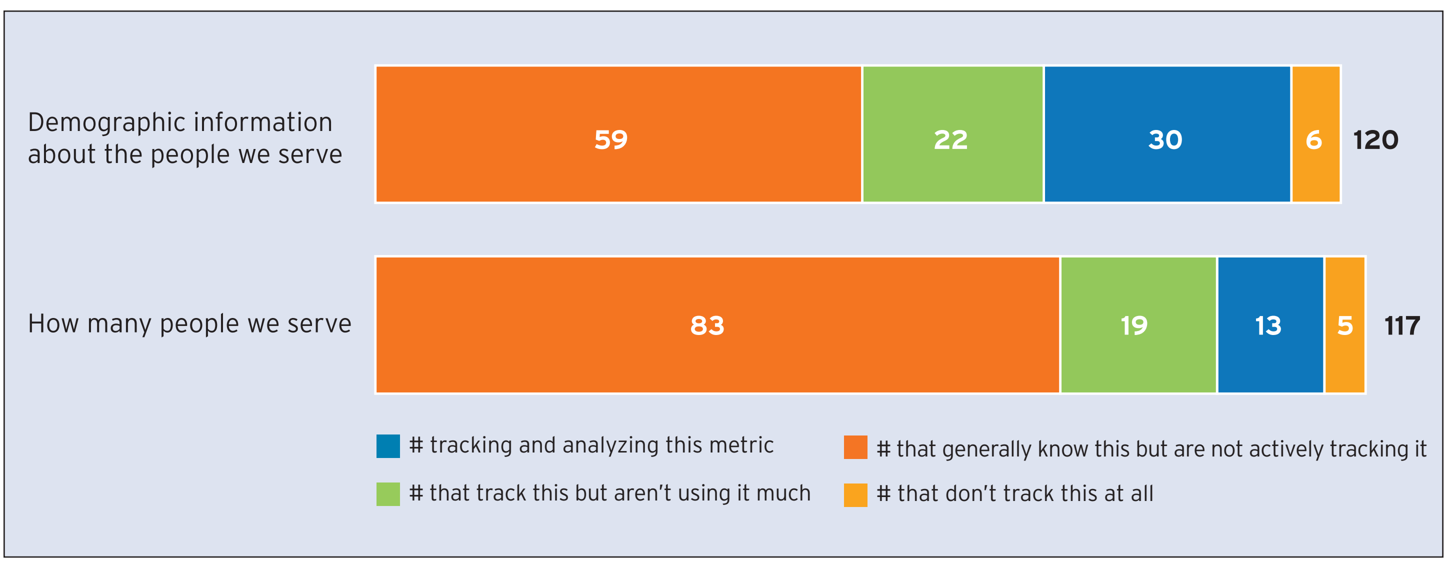
## Your Own Activity

At the bottom of the pyramid are metrics to track your organization's own activities. Metrics at this level are typically more straightforward to obtain, since they come entirely from your own organization and do not depend on outside data as much as other measures. Just about half of all participants reported that they were tracking the number of activities they run. It's interesting that notably more respondents are tracking participation in activities (see the next section) than are tracking the number of activities they actually run—it may be that some organizations do not think of the quantity of their services as a critical measure, or simply that a number of respondents did not feel the word "activities" described their program delivery.




## Participation

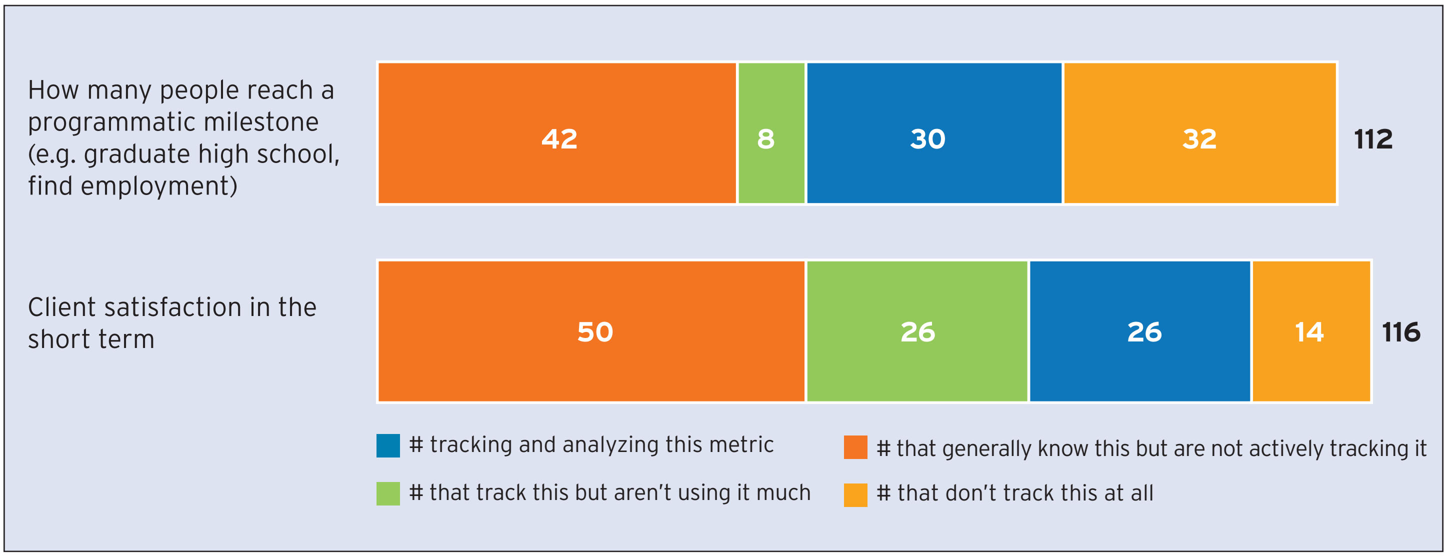
Moving up the pyramid, we come to metrics that pertain to program participation. Of the 120 organizations surveyed, 83 (around 75 percent) were tracking the number of people they served, the highest of any of the 10 metrics included in the survey. Only 59 (just under half) said they were tracking demographic information about those people.



## Initial Perceived Satisfaction or Success

The next tier looks at the initial satisfaction with or success of the program. Metrics in this tier are concerned with immediate outcomes of an organization's programs and can be directly measured through exit interviews or surveys of program participants. For example, if a program is aimed at helping unemployed clients find a job, a metric of initial program success could be the number or percentage of participants who find employment within six months, while a survey of program participants could determine their satisfaction with the program.

Less than half of survey respondents were tracking metrics at this tier of the pyramid—only 50 of the 120 surveyed were tracking short-term client satisfaction with the program, and only 42 were tracking how many program participants went on to reach a specific milestone like graduating high school or finding employment.



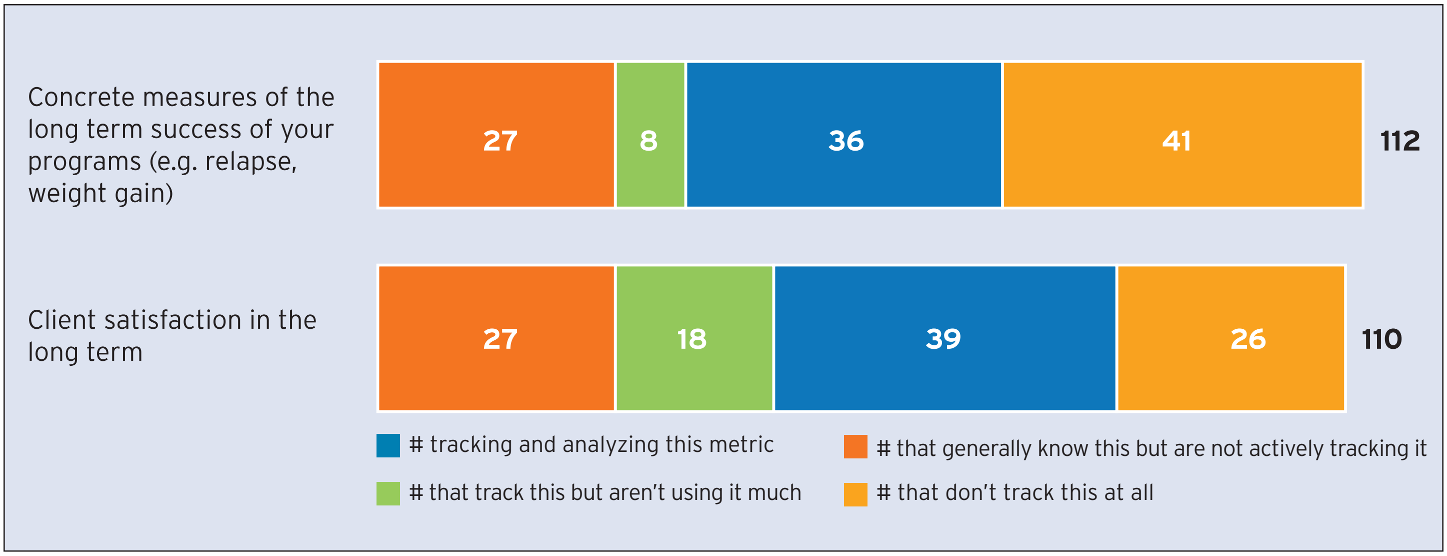
## Case Study: The Immigrant Legal Advocacy Project

Prohibitive resources can make accurate program measurement a challenge, as the Immigrant Legal Advocacy Project (ILAP) learned. ILAP helps immigrants in Maine keep their families together, gain protection from persecution and domestic violence, attain residency and work authorization, and become proud U.S. citizens. The number of people ILAP serves is limited by its eligibility requirement—clients must be at or below 200 percent of the poverty level of income—but even so, the number of people who meet that requirement is considerably more than ILAP has the capacity to help.

Intake for new clients is handled with paper applications, which are then entered in the system later. "Ideally, there would be tablet for every client to fill out forms on, or for volunteers to be using while meeting with clients," said Susan Roche, the interim Executive Director. But tablets may not be realistic, or even desirable—in addition to the organization's limited technology budget, there is a desire within the organization to emphasize the personal interaction with new clients and a concern that technology may get in the way of that relationship. Furthermore, many of the trained volunteers who handle intake are not as tech-savvy, preferring to stick to paper forms.

Data integrity is also an issue for the organization. Currently, it is using two separate databases to manage client information for different types of cases—one of them quite outdated—with no integration between the two. There is concern at ILAP over the quality of the data; with intake forms being processed by many different volunteers into both systems, reporting on that data is unpredictable, and staff members are often unsure of what data is new and what is out-of-date.

Another hurdle in program evaluation is the gap between what outcomes funders want to see and what outcomes ILAP can realistically track. Funders are interested not just in numbers served and client demographics, but also in eventual employment and other long-term outcomes. However, the organization doesn't have an easy means to track what happens to clients after their cases close. Staff often has trouble keeping in touch with clients after ILAP's service is complete, and some clients view attempts to stay in touch as an invasion of their privacy, because ILAP's mission is not to help immigrants get jobs or housing—it's to help them become legal citizens.




## Longer Term Satisfaction or Activity

While initial or short-term results are often relatively straightforward to measure, it can be much more difficult to measure satisfaction or activity in the long term. Past clients may be less likely to fill out a survey six months or longer after the program ended. The amount and quality of data you can collect at this point is often dependent on the data you collected during the program as well as the initial or short-term satisfaction.

Relatively few respondents were tracking these long-term outcomes. Only 27 of 120 (around 23 percent) said they were tracking and analyzing either client satisfaction in the long term or concrete measures of long-term program success.

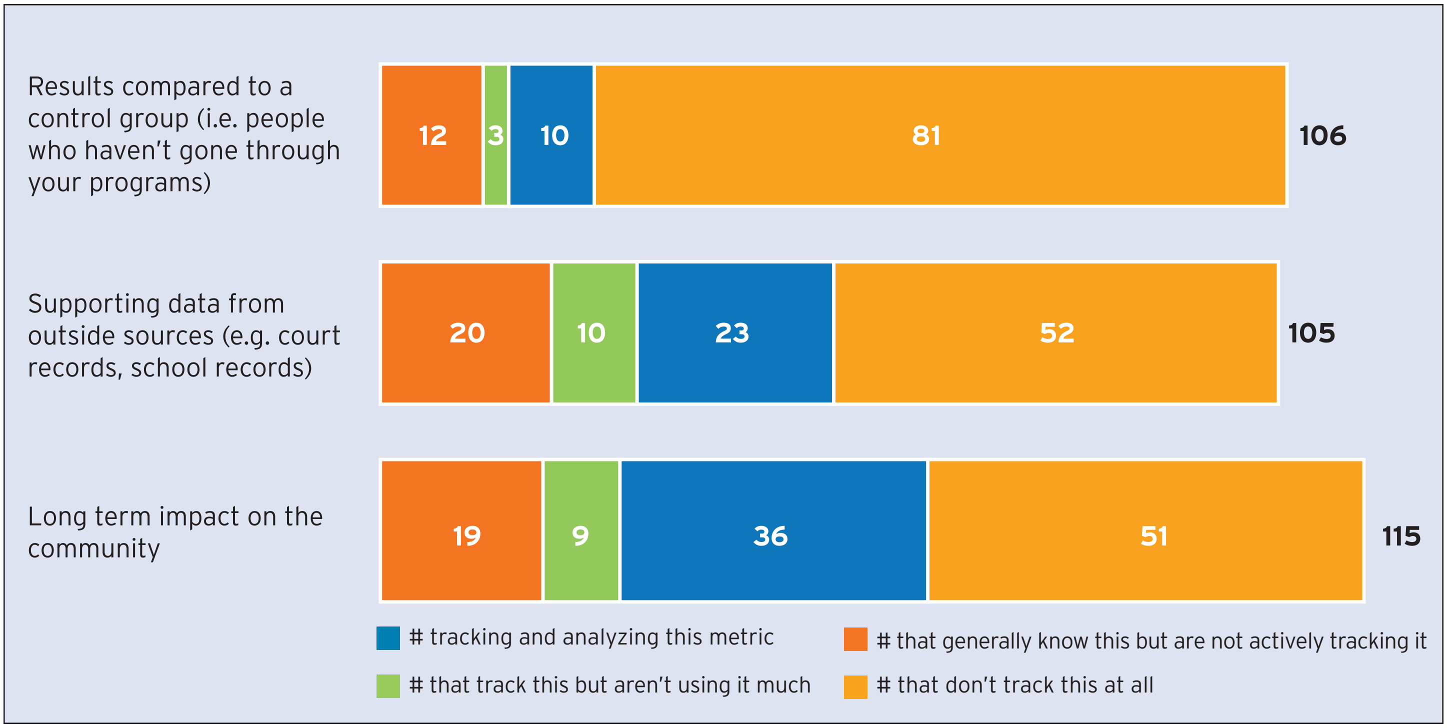
## Attributable Impact

The top tier of the program evaluation pyramid contains metrics that help measure the impact on the community that can be attributed to your organization's programs—these are the most difficult to collect and measure of the metrics included in this survey. Supporting data from outside sources like court or school records are often highly desirable, but difficult to acquire in a way that can be matched up to your organization's own records, based on privacy concerns and proprietary issues around data. Likewise, comparing your results to a control group can often be the only way to truly know what results your programs specifically caused, but can be very complicated and expensive to set up.

The last metric, long-term impact on the community, will require at least one of the other two measures. It's very difficult to actually determine the long-term impact of your programs compared to another organization's programs that affected much of the same population—or even just change that would have happened anyway. Understanding your attributable impact generally requires a rigorous level of research beyond even what many mid-sized to larger organizations are capable of doing.

Unsurprisingly, these metrics were the least-tracked by survey respondents. Only 20 (about 17 percent) reported tracking or analyzing supporting data from sources outside the organization, and a mere 10 percent were measuring results against those of a control group of people who had not gone through their programs.

Long-term impact on the community is quite possibly the most difficult metric from this survey for an organization to track accurately, as it requires not just an understanding of the direct outcomes of a program but also of the community at large. For example, in order to determine the long-term impact on the community of a program that aims to reduce dropout rates in area high schools, you would need to know not only the dropout rate of students that had gone through the program but also for each school




served—for years before the program, as well as in the years following. Of course, while this would allow the organization to separate the impact attributed to their program from any overall trends, it's still not enough information to find how much of that impact came from that one program as opposed to other similar programs in the area (which would require sharing data with other organizations serving the same population).

We're confused by the 19 survey respondents who said they were tracking the long-term impact of their programs on the community. Of those, over 60 percent indicated they were not also tracking supporting data from outside sources or results compared to a control group. Without these supporting sources of data, it is highly unlikely that these respondents were in fact measuring the impact directly attributable to their own programs.

And what to make of the 37 percent of respondents that said they generally knew their impact but were not actively using the data? To Idealware, this simply indicates that the word "impact" has lost almost all true meaning when it comes to measuring programs—organizations understand that they "should" be measuring it, but have no tactical understanding of what it would mean to actually do so.

## How Much Time is Spent on Evaluation?

When people talk about the theory of program evaluation, they tend to discuss it either as a continuous, ongoing process or as something that's done retrospectively—measuring outputs and outcomes is often defined as an ongoing process, while performance evaluation is typically thought of as happening regularly, but after the fact.

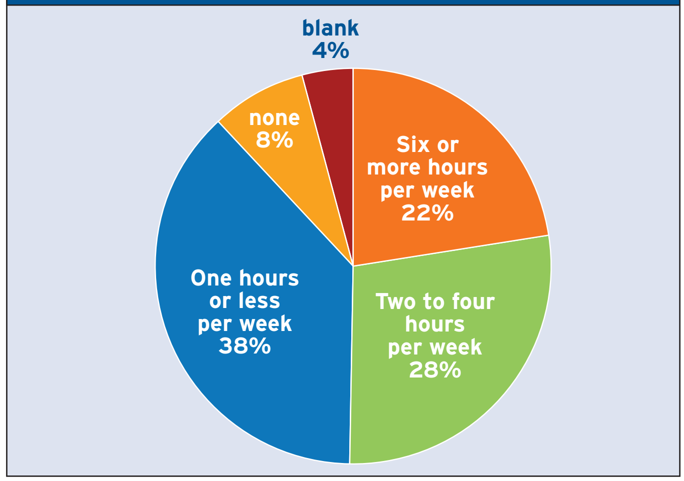
In reality, most of the organizations surveyed were conducting their evaluations retrospectively. Overall, program evaluation was not a regular occurrence for the organizations who took the survey, but likely occurred only when reports were due to foundations or other funders—71 of the 120 respondents (just under 60 percent) said that they were doing so quarterly at most, and over half of those organizations were only evaluating annually. Only about a third of respondents were evaluating on a monthly or weekly basis, but less than 10 percent reported never conducting program evaluation.

| How frequently do you evaluate your programs? | | |
|---|---|---|
| Quarterly or less | 71 | 59% |
| Monthly or more | 39 | 33% |
| Never | 8 | 7% |

When surveyed organizations were actively evaluating their programs, almost 30 percent (33 of 120) reported spending about two to four hours per week doing so, and just under a quarter reported spending six or more hours per week. Almost 40 percent of respondents were spending one hour per week or less on program evaluation. Overall, time was the obstacle to conducting program evaluation most frequently mentioned by survey participants.

While it could be expected that an organization that evaluates its programs more frequently would have to spend less time per week doing so than one that evaluated only quarterly or annually, the data showed no trends to suggest such a correlation.

**About how many hours per week are spent on program evaluation?**



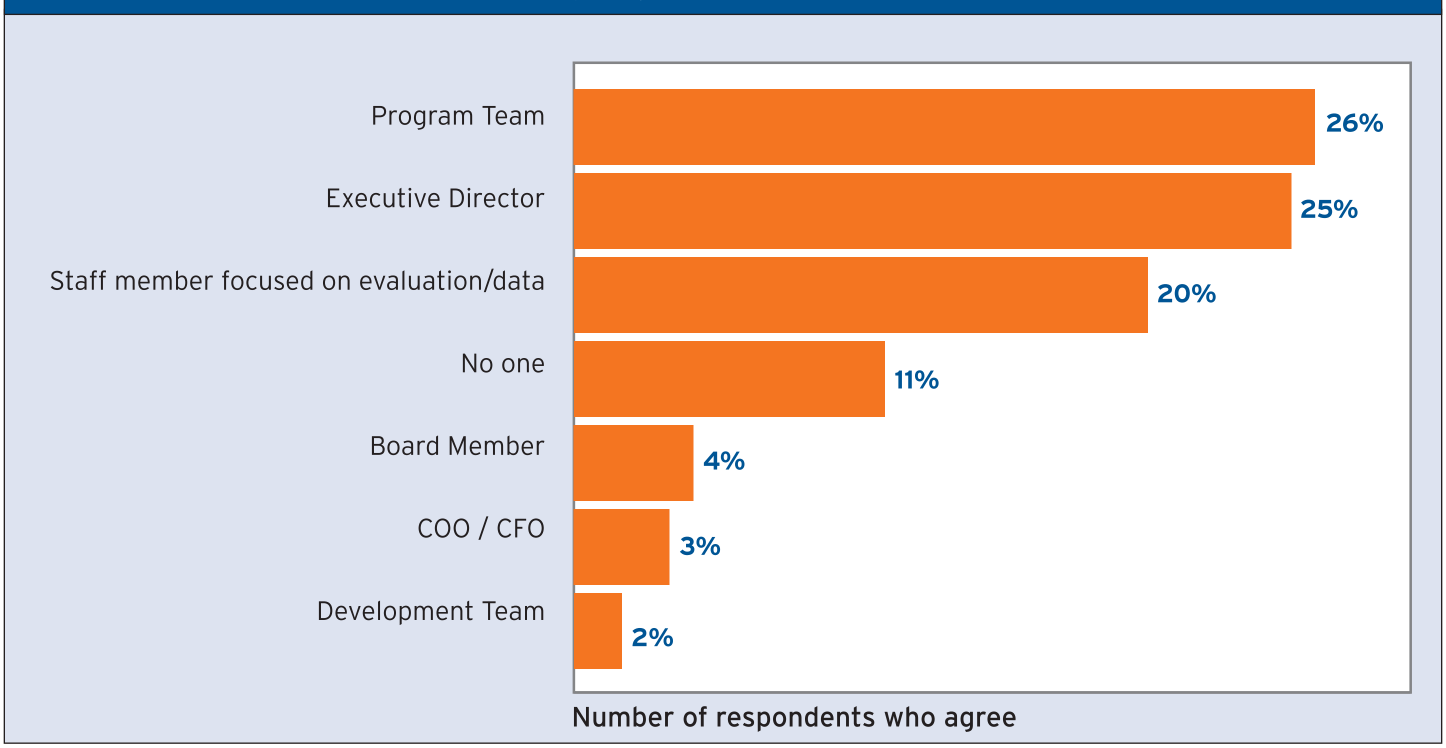

## What Staff Members are Responsible for Program Evaluation?

Nonprofit organizations may need more than an executive director to evaluate their programs. Ideally, the whole staff has a stake in evaluating success. As one of our survey respondents said, the process "needs to include line staff—not just supervisors and managers—to get buy-in and to help staff recognize and feel good about the impact they are having."

In our survey, about 25 percent of respondents said their executive director was in charge of program evaluation.

Conversely, some organizations have not designated someone to be in charge of program evaluation. A concerning 11 percent of the people in our survey said there is no one responsible for program evaluation. Without someone to give direction, there will likely be some amount of confusion as well as a lack of accountability.

**Who on your organization's staff is most directly responsible for program evaluation?**



“[The] process needs to include line staff—not just supervisors and managers—to get buy-in and to help staff recognize and feel good about the impact they are having.”

Some nonprofits reported hiring a staffer specifically focused on directing evaluation. Program team leaders also take on responsibility. About 26 percent of the people surveyed said the program team is responsible for evaluation.

## Nonprofits Lack Effective Software for Evaluation

While program directors might have the ability to devote more of their time and resources to evaluating their program outcomes, when the responsibility falls on executive directors or board members, they may only be able to do so infrequently. Of the 30 organizations that reported that their executive directors were responsible for conducting program evaluation, almost 75 percent were conducting their evaluations quarterly or annually. While we are unable to determine a direct correlation, this data suggests that evaluation might be taking place less frequently because executive directors' schedules don't allow for more regular evaluation. Organizations where the program team was responsible for evaluation were about equally likely to do so quarterly or annually as monthly or more frequently—approximately 58 percent and 42 percent, respectively.

While program evaluation is more a strategy than a process dependent upon any one tool or type of software, a good database and reporting tool can greatly ease the process of collecting, recording, and analyzing the output and outcome data an organization is tracking. Similarly, the lack of flexible tools can hinder an organization's effectiveness and add unnecessary time to the evaluation process.

Over half of survey respondents reported using spreadsheets as their primary tool for managing and reporting on program data. While spreadsheets can be useful to move data in or out of your database, or to help make sense of relatively straightforward data, they are not an ideal way to manage all your data—especially not data about your clients or constituents. Unlike databases, spreadsheets track "flat files," or non-relational data. This means that without multiple entries, a spreadsheet doesn't really understand how a constituent fits into your organization.

Another 22 participants—18 percent—were using a custom-built system. Idealware generally recommends that organizations refrain from building custom systems to track constituent data, as the many available software packages are typically cheaper and easier to maintain in the long run. The fact that only 20 percent of respondents reported using *any* packaged constituent management software is concerning however, and could imply either a sizable gap in the marketplace for the type of software that could help, or a sizable lack of knowledge of the software that's available.

| **What kind of software does your organization use most often to manage program related data?** | |
|---|---|
| Spreadsheets | 65 |
| Custom Database (e.g. MS Access, File-maker) | 22 |
| Constituent Relationship Management Software | 9 |
| Case Management System | 5 |
| Homelessness Management Information Systems (HMIS) | 3 |
| Learning Management System | 3 |
| Membership Management System | 2 |
| Electronic Medical Record Software | 1 |
| Legal Case Management System | 1 |
| none | 3 |

Overall, survey respondents see the value in program evaluation—over 70 percent said program evaluation

## Case Study: A Preschool in Portland

One organization we visited, a Portland-based preschool for children up to five years old with disabilities and their peers, exemplifies some of the technology challenges reported in the survey. As part of a national organization, the school enjoys the benefit of having an IT department—but as IT serves the entire region and is located in another state, it is difficult to get in-person help when it's needed.

As part of the clinical strategy employed by the school, data is collected rigorously throughout each school day. Teachers are expected to record student behavior, activities, and skill development every 15 minutes. In a busy classroom, this frequency can be difficult, resulting in quickly recorded handwritten notes in order to reduce the time spent away from students. As a result, most data entry occurs after the school day, when the over 60 staff members must share only five computers to enter their handwritten notes into a spreadsheet. Tablets might allow staff to automatically input data into a spreadsheet, but there is no budget for such a change. Funders are mostly focused on increasing the time spent on direct interaction with students, providing very little time or budget to help teachers with data entry.

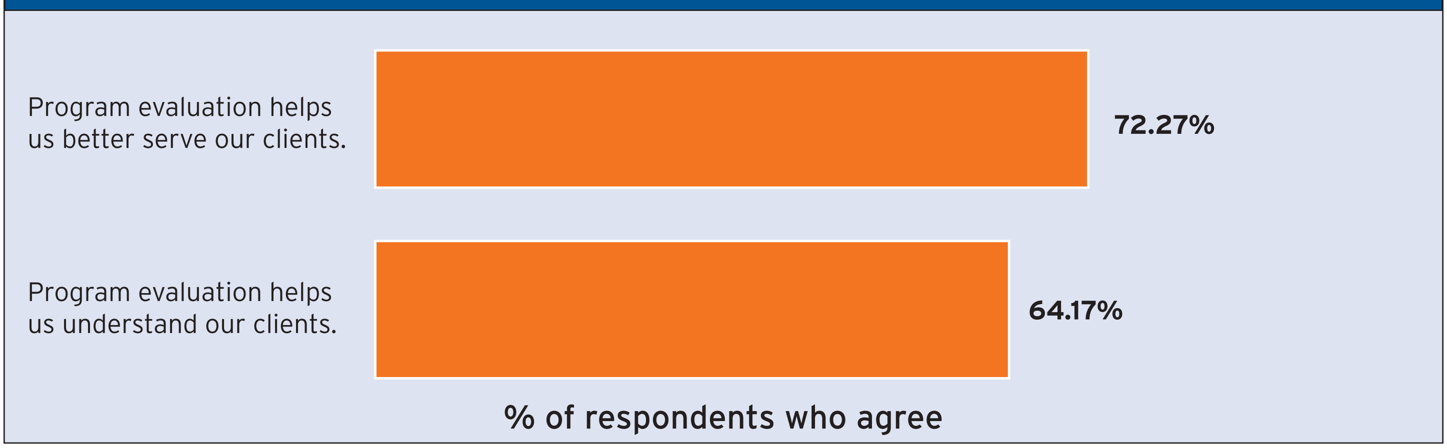
The data the teachers and other staff collect is subjective information regarding behavior and skill acquisition, designed to track the individual progress of particular students. This focus makes it difficult for them to use many existing case management systems. Instead, the organization depends on self-designed spreadsheets for data collection and evaluation. While staff looks at some aggregate measures of program success across students, they focus on initial and short-term outcomes, and don't attempt to follow the success of students after they move on from preschool.

# ARE NONPROFITS HAPPY WITH THIS REALITY?

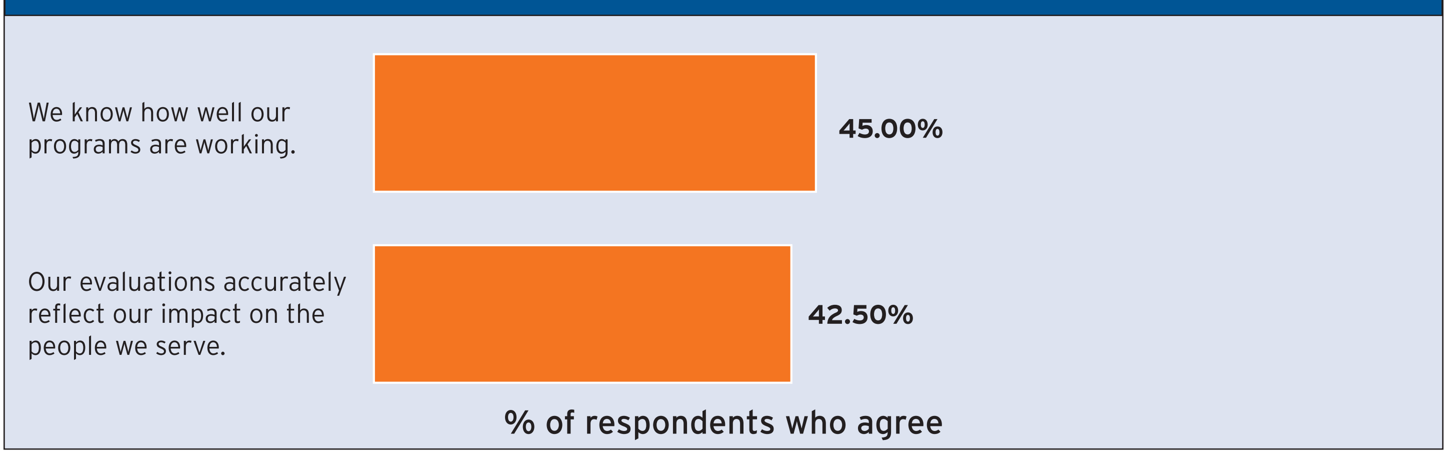
helps them better serve their clients, and over 60 percent felt that it helped them understand the clients they work with. However, it's unclear as to how many of them were actually measuring their programs and finding it valuable and how many merely felt that, theoretically, program evaluation would be valuable if they were doing it.

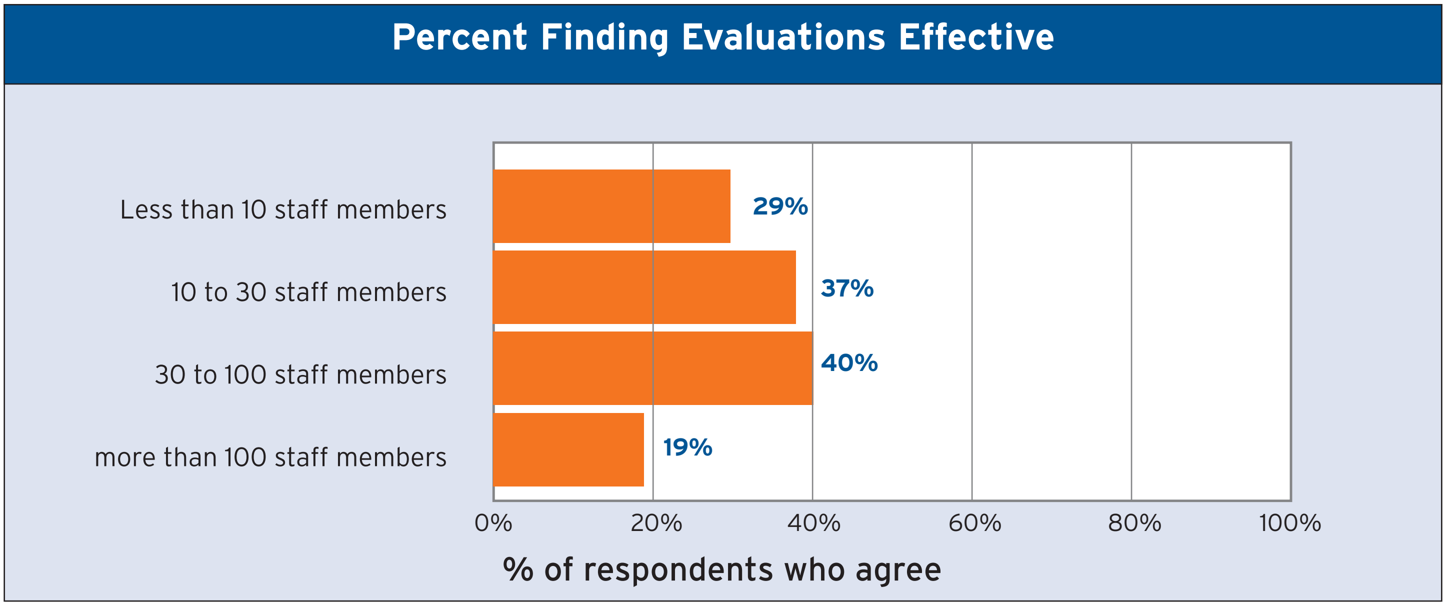
While the organizations surveyed certainly see the inherent benefits of program evaluation, they seem to still struggle with the process of evaluation. Only 45 percent knew how well their programs were working, and just over 42 percent felt their evaluations accurately reflected their impact on their constituents. While program evaluation seems to be helpful for gaining a high-level understanding of programs and clients, it doesn't seem to go all the way to understanding the root causes of program success or impact.

**Perceptions of Evaluation**




**Effectiveness of Evaluation**






While there appears to be a trend suggesting that larger organizations were finding their evaluations more effective, we lack sufficient data to say so with certainty. The survey sample was generally skewed toward smaller nonprofits with fewer than 30 staff members, preventing us from making accurate statements about the habits or effectiveness of larger organizations.

| How would you describe the size of your organization? | |
|---|---|
| Less than 10 staff members | 48 |
| 10 to 30 staff members | 35 |
| 30 to 100 staff members | 20 |
| More than 100 staff members | 16 |

Most of the organizations surveyed identified themselves as providing Educational, Youth Development, or Health Care or Promotion services. While it can be expected that certain types of human service organizations would have different obstacles to effective evaluation—for example, a health clinic might have difficulty reporting on patient data due to HIPAA and other privacy regulations—or have longer histories evaluating their programs, there was no clear trend among the organizations surveyed to suggest that certain focus areas were more or less effective than others. This is likely due, at least in part, to both the diversity of responses to this question and the low number of responses to each option. Survey respondents also had the option of selecting multiple focus areas, further complicating our ability to draw conclusions based on organization type.

| Focus Area of Organization | Total |
|---|---|
| Educational Services | 46 |
| Youth Development | 30 |
| Health Care/ Promotion | 27 |

## Obstacles to Effective Evaluation

This struggle to establish an effective regimen of program evaluation is likely hampered by a number of obstacles, both internal and external. Less than half (approximately 48 percent) of respondents reported that their staff members understood the importance of data collection for program success, and less than 15 percent found the process of program evaluation easy.

By-and-large, foundations have led the push toward better program and outcomes evaluation among grantee organizations. Fortunately, the overwhelming majority of organizations we surveyed did not seem to view evaluations as merely a way to ensure funding—just under 12 percent reported evaluating exclusively to obtain grants.

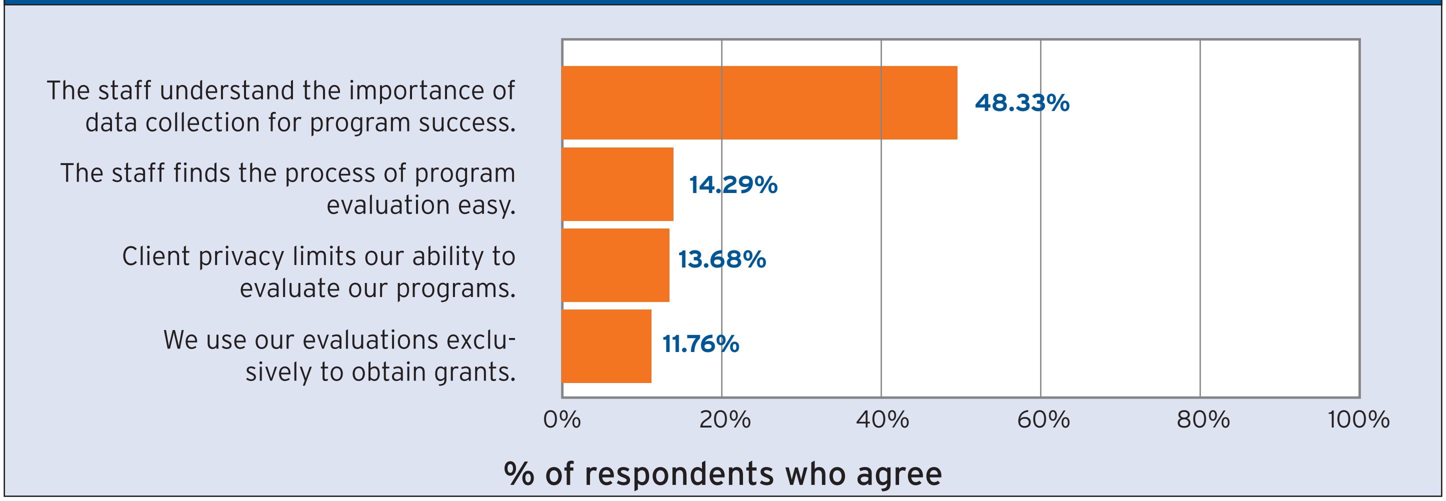
While we initially expected many organizations to be limited in their ability to evaluate their programs by client privacy—like HIPAA regulations for health clinics—less than 14 percent of those surveyed indicated that this was the case.

When asked through a free text field to identify the biggest hurdle or hurdles to effective program evaluation, most participants cited either a lack of time to conduct the evaluations or the lack of the physical resources—money or tools—to do so. About 25 percent said their biggest hurdle was with the assessment itself, from defining the right metrics to measure to evaluating the data and even communicating their results effectively.

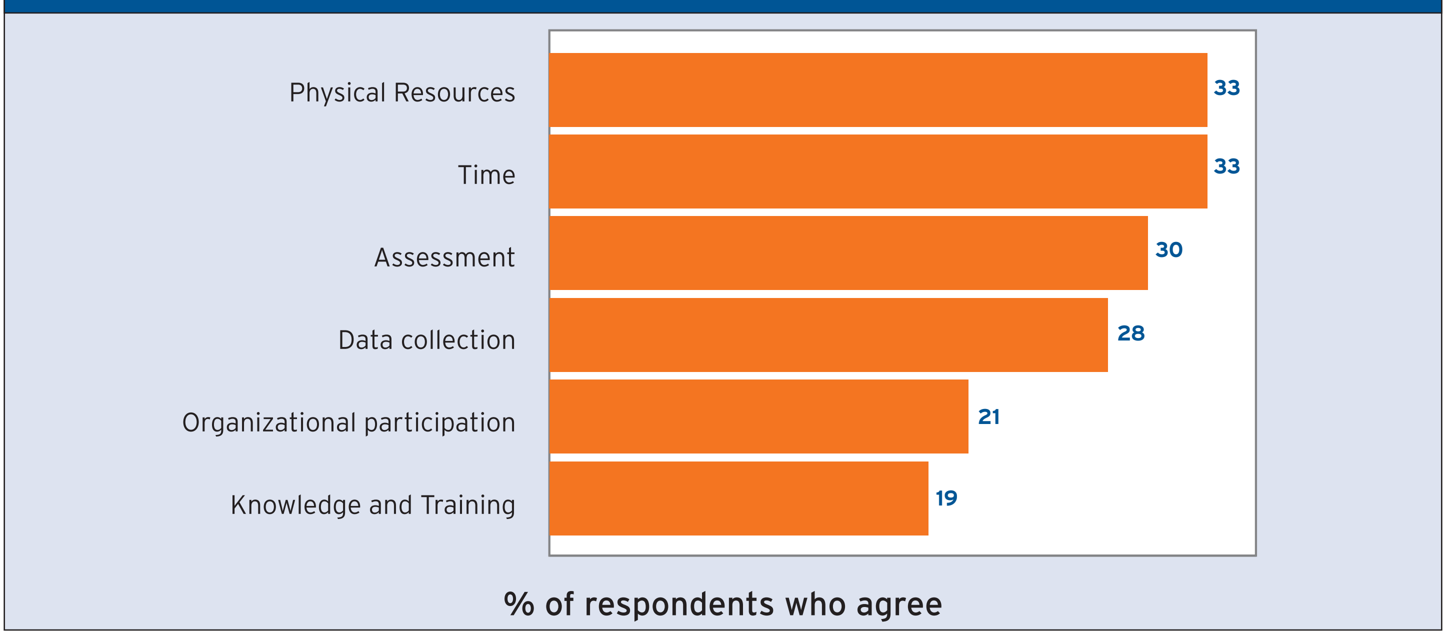
Making sense of program data is difficult if data is not collected consistently and regularly. Data collection was the fourth most cited hurdle to evaluation, including difficulty in accessing external data, getting clients to complete follow-up surveys, and the overall quality of the data that is collected.

Of our 120 survey respondents, 21 reported that their biggest hurdle was participation by staff—over half of those cited a lack of consistency in the evaluation process. Finally, almost a fifth of the organizations surveyed said their biggest hurdle was a lack of knowledge or training on how to effectively evaluate their programs.

**Obstacles to Evaluation**




**What is your biggest hurdle when it comes to program evaluation?**



## Case Study: The Center for Grieving Children

For young people who have suffered the loss of a parent, sibling, or grandparent, the Center for Grieving Children is a refuge. Located just off a busy street in greater Portland, the cozily decorated Center takes up an entire building, and is devoted entirely to helping kids, teens, and young adults work through grief. A satellite campus in York County, about an hour south, serves that community. The organization—one of the first of its kind in the country—recently marked its 25th anniversary, and reports serving close to 70,000 people with a staff of about 10 FTEs and about as many interns and volunteers. The Center's programs include Bereavement Peer Support groups to help those recovering from loss work through their emotions together and a Tender Living Care program for family members of those diagnosed with terminal illness. All services are completely free.

The organization's approach to program measurement has evolved over the past several years from an anecdotal model to a more data-focused one. Outcomes measurement is one of the main responsibilities of Susan Giambalvo, the Center's Director of Programs. With the help of a capacity-building grant from the Maine-based Unity Foundation, Susan recently presided over the successful transition of the Center for Grieving Children's constituent- and program-tracking database from a custom-built solution in Microsoft Access to Social Solution's Efforts to Outcomes (ETO). The frustrations with the outdated custom solution were many—the most critical issue was that the database was difficult to modify in order to track new or different data points, so the system had little relation to the Center's revamped logic models.

Susan worked with the software's vendors and a consultant to tailor the Center's implementation of ETO so that it helped track data points critical to the organization's program evaluation strategy in a central location, such as demographics of the participants, membership and attendance, electronic case notes, outreach and training, and volunteer hours. She also worked closely with the development and administrative teams at the Center to make sure that the database tracked information important to them, and served as the cheerleader and primary project manager in the migration process. Her leadership was critical to the success of the changeover. "Somebody has to be excited about it. This kind of transition is not something to do half-heartedly," she said.

The Center is just beginning to roll out the ETO solution, but staff hopes that this revamped data collection and evaluation tool will ease the process of annual reporting for funders and the board and more clearly demonstrate the quantitative effects that the Center for Grieving Children has on constituents. One recent success demonstrates the importance of external data in a program evaluation strategy. The Center's Multicultural program works with Portland public schools to help young refugees deal with the trauma they've experienced. The school system provided grade, attendance, and behavior data about participants in the program, and the Center's analysis demonstrated a positive correlation of statistical significance between improved academic performance and participation in the program. In other words, the Multicultural program helps kids, and the Center now has the data to show it.

Still, challenges about—even for an organization with a sophisticated program evaluation strategy. For its clinical survey instrument, The Center for Grieving Children uses the Developmental Assets Profile, which can be administered online. However, the Center's peer programs are mostly moderated by volunteers, who administer the survey information through paper forms that staff enter into the database later—there simply aren't enough computers for all the volunteers. Although the Center tracks attendance data and short-term client satisfaction, it's difficult for them to track long-term progress from people who have gone through the program. Some clients may not want to be reminded of such a painful period, and it's also a challenge to track down the young people served by the Center into their adulthood. A more philosophical challenge arises as the Center tries to measure the success of programs focused on something so personal and individualized as healing from grief, as well as to protect the minors they serve. The staff at the Center generally steers clear of individual reporting and monitoring and look more at trends from the surveys collected to understand the experiences of certain age groups or other clusters of clients.

When asked why the Center for Grieving Children conducts evaluation of its programs, Susan said it's because the organization is committed to delivering the best possible service to the families that visit. In this sense, the nonprofit's program evaluation strategy is an unequivocal success: The data collected so far demonstrates that the children and families served by the Center report greater resiliency, positive identity, and self-worth while participating in its programs.

# WHAT TO MAKE OF ALL OF THIS?

This survey was not intended to provide a scientifically accurate portrait of the nonprofit sector as a whole, but overall, our findings seem to reinforce much of what we have suspected all along—that nonprofits are generally struggling to measure their programs. No theory or strategy around program evaluation can help if organizations don't have the knowledge and training to successfully measure their programs.

Fortunately, these results also show some ways organizations just starting to look at their programs or struggling with evaluations can make effective use of their time. Clearly, good program evaluation starts with a strong base of good data. Paying closer attention to collecting data on your programs rather than the end goal will go a long way toward forming a strong foundation of data for future evaluations. As the survey results suggest, organizations must first have solid data on their programs and outputs before they'll be able to effective analyze their impact.

There is a wide range of data for organizations to collect, but for all nonprofits—especially those just starting to evaluate their programs—it's best to start with the data closest to home by measuring what you are actually doing in your programs: such things as program attendance, number of programs, and demographic information about your clients. With this solid foundation, your organization can successfully build a program evaluation strategy.

# APPENDIX A: METHODOLOGY

In the summer of 2013, Idealware conducted a survey of staff members at human service nonprofit organizations and asked them about their experience with program evaluation. The survey asked specifically for staff members involved in the process of program evaluation for their organizations. (See Appendix B for the full survey questionnaire).

We distributed the survey invitation to Idealware's email list and through the Idealware Twitter and Facebook accounts. It's difficult to judge how many people saw an invitation to take the survey, since some people passed it on to others, but we estimate around 30,000 people. Not everyone who saw the invitation fell within the target audience of human service organization staff involved in program evaluation, but it is clear that only a small percentage of those who saw the invitation responded.

Our sample was targeted to small nonprofits, and our sample matched our objective. The majority of respondents came from nonprofits with fewer than 30 staff members. See the chart below for more detailed information.

| STAFF SIZE | # |
|---|---|
| Extremely Small (all volunteer) | 3 |
| Very Small (less than 10 staff) | 45 |
| Small (10 to 30 staff) | 35 |
| Medium (30-100 staff) | 20 |
| Large (100 to 250 staff) | 8 |
| Very Large (250 to 500 staff) | 4 |
| Extremely Large (more than 500 staff) | 4 |

This survey was not intended to be representative of the nonprofit sector as a whole, but only an informal sample of small human service organizations. The respondents were reasonably representative of the different kinds of human services, but there was a slight bias toward educational services and youth development. The reason for the large number of educational services may be due to the wide interpretation of education, from public awareness to child development. Many respondents provided multiple kinds of services, so the numbers do not correspond to the survey's sample size. See the chart below for more detailed information.

| TYPES OF SERVICES | # |
|---|---|
| Educational Services | 46 |
| Youth Development | 30 |
| Health Care/ Promotion | 27 |
| Homeless Services | 19 |
| Job Training/ Placement | 16 |
| Crisis Services | 15 |
| Food and Nutrition | 15 |
| Mental Health Services | 15 |
| Community Center | 14 |
| Family Counseling | 13 |
| Senior Assistance | 12 |
| Sports/ Recreation | 11 |
| Child Welfare | 9 |
| Disability Assistance | 8 |
| Economic Development | 7 |
| Religious Services | 5 |
| Arts and Culture | 5 |
| Environmental Services | 4 |
| Immigration Services | 4 |
| Substance Abuse Treatment | 4 |
| International Services | 3 |
| Legal Aid | 3 |
| Advocacy | 3 |

Following the close of the survey, Idealware identified three of the participating organizations in the vicinity of Portland, Maine, where we are based, and interviewed them in person for the creation of the case studies included in this report. These organizations were selected because we felt their survey responses indicated that they had implemented and established a regimen of program evaluation. Following the interview, each organization was compensated for their time with pro bono resources from Idealware, including registration for an Idealware webinar, printed copies of Idealware reports or articles, and/or impartial advice on the organization's specific technology needs.

# APPENDIX B: FULL TEXT OF THE SURVEY

## Program Evaluation for Human Services

1. What types of services does your organization provide? (choose all that apply)

- Mental Health Services
- Food and Nutrition
- Child Welfare
- Job Training/ Placement
- Senior Assistance
- Homeless Services
- Youth Development
- Crisis Services
- Sports/ Recreation
- Foster Care
- Family Counseling
- Community Center
- Health care/ promotion
- Substance abuse treatment
- Immigration Services
- Educational Services
- Developmental Disability
- Other (please specify)

2. How would you describe the size of your organization?

- Extremely Small (all volunteer)
- Very Small (less than 10 staff)
- Small (10 to 30 staff)
- Medium (30-100 staff)
- Large (100 to 250 staff)
- Very Large (250 to 500 staff)
- Extremely Large (more than 500 staff)

3. Who on your organization's staff is most directly responsible for program evaluation? (e.g. collects, measures, and monitors the outcomes and effectiveness of your programs)

- No one
- Staff member focused on evaluation/data
- Development Team
- Program Team
- IT/ Technology Team
- Board Member
- Executive Director
- COO / CFO
- Other (please specify)

4. About how often does your organization conduct program evaluation?

- never
- once a year or less
- quarterly
- monthly
- weekly
- continuously

5. When your organization conducts program evaluation, about how many hours per week are spent?

- none
- less than one hour per week
- about one hour per week
- about two hours per week
- about four hours per week
- about six hours per week
- about eight hours per week
- about 10 hours per week
- more than 10 hours per week

6. What kind of software does your organization use most often to manage program related data?

- Spreadsheets
- Constituent Relationship Management Software
- Case Management System
- Volunteer Management Software
- Electronic Medical Record Software
- Learning Management System
- Legal Case Management System
- Homelessness Management Information Systems (HMIS)
- Membership Management System
- Custom Database (e.g. MS Access, Filemaker)
- Other (please specify)

7. On a scale from 1 to 5, do the following statements describe your organization?

| | 1 (No, not at all) | 2 | 3 | 4 | 5 (Yes, absolutely) |
|---|---|---|---|---|---|
| Our organization collects, measures, and monitors the outcomes and effectiveness of its programs (i.e. program evaluation) | | | | | |
| We know how well our programs are working. | | | | | |
| The staff understand the importance of data collection for program success. | | | | | |
| The staff finds the process of program evaluation easy. | | | | | |
| We use our evaluations exclusively to obtain grants. | | | | | |
| Program evaluation helps us understand our clients. | | | | | |
| Program evaluation helps us better serve our clients. | | | | | |
| Our evaluations accurately reflect our impact on the people we serve. | | | | | |
| Client privacy limits our ability to evaluate our programs. | | | | | |

8. What types of data does your organization track about its programs?

| | We don't track this at all | We generally know this but are not actively tracking it | We track this but are not using it much | We track and analyze this but could be using it better | We strategically track and analyze this | Not applicable |
|---|---|---|---|---|---|---|
| How many activities we run (e.g. classes, appointments, houses canvassed) | | | | | | |
| How many people we serve | | | | | | |
| Demographic information about the people we serve | | | | | | |
| How many people reach a programmatic milestone (e.g. graduate high school, find employment) | | | | | | |
| Client satisfaction in the short term | | | | | | |
| Client satisfaction in the long term | | | | | | |
| Concrete measures of the long term success of your programs (e.g. relapse, weight gain) | | | | | | |
| Supporting data from outside sources (e.g. court records, school records) | | | | | | |
| Results compared to a control group (i.e. people who haven't gone through your programs) | | | | | | |
| Long term impact on the community | | | | | | |

9. What is your biggest hurdle when it comes to program evaluation?

10. Is there anything you have learned from evaluating your organization's programs that you would like to share?

11. After this survey, we are doing additional research into the factors that help organizations make effective use of technology for program evaluation. We are looking for case studies to include in our final report. Would you be interested in further dialogue so that your organization could become a case study?

- Yes
- No Thanks

12. If you answered "Yes" to the question above, please provide contact information so we can reach you. Please include your name, organization, and email address.