College of Professional Studies Professional Projects   Dissertations, Theses, and Professional Projects

4-1-2010

Performance Measurement: Theory and Practice

Jerica Broeckling
Marquette University

Follow this and additional works at: http://epublications.marquette.edu/cps_professional

Performance Measurement Systems:

Theory and Practice

By

Jerica Broeckling, B.A.

A Professional Project Submitted to the Faculty of the Graduate School,

Marquette University,

in Partial Fulfillment of the Requirements for

the degree of Master of Arts in Public Service

Milwaukee, Wisconsin

April 2010

ABSTRACT

This project concerns the use of performance measurement in the nonprofit sector, including the growing use of this tool to increase program function, enhance decision-making, and compare organizations. It explores the implementation and modification of a small nonprofit agency's performance measurement system and includes the design and implementation of additional tools. Key staff members of the nonprofit were interviewed about their experiences regarding measurement of a federal grant, and 22 program participants completed a survey of medium-term outcomes. The data suggest several changes to enhance program delivery and increase word of mouth referrals. The present researcher designed a logic model, program dashboard, survey, and phone survey to add to the nonprofit's existing data collection tools.

Keywords: performance measurement, evaluation, comparative measurement, adaptive learning systems

TABLE OF CONTENTS

Performance Measurement Systems: Theory and Practice

Over the last decade the nonprofit sector has renewed its interest in methods for measuring organizations, programs, and their impact. This interest has many internal and external factors, including a desire for accountability, unprecedented growth and competition, pressure from funders, and an increase in the use of private sector tools. Today there are two main forms of measurement prevalent in the sector: performance measurement and evaluation.

While some practitioners use the terms interchangeably, there are important differences. Considered the "gold standard" of the sector today, evaluation uses quantitative research to determine whether a program achieves its intended results, or outcomes (Plantz, Greenway & Hendricks, 1997). Performance measurement, on the other hand, measures both social impact and organizational performance, though in a less rigorous manner (McKinsey & Company, 2008). Some argue that the preference for evaluation has skewed the field, creating a kind of tunnel vision that focuses on "*proving* whether a program or initiative works, rather than on *improving* programs" (W.K. Kellogg Foundation, 2004a, p. 4).

In recent years performance measurement has gained popularity, especially as funders and regulators look for ways to compare organizations. Performance indicators, or individual units of measurement, have the potential to break down a complex program into understandable and comparable units, offering the potential to make these cross-comparisons in a reasonable way. The purpose of this project is to investigate performance measurement's potential as a useful management tool and to document its use and evolution in a small nonprofit. It includes exploring how that nonprofit actually implemented a system for measuring performance, and the development of tools to increase that system's function without overburdening it.

The Push for Performance Measurement

The call for reform in the nonprofit sector has steadily grown over the last generation as various stakeholders grew frustrated with the lack of visible social change, public lapses of accountability, and significant growth without increased regulation. Practitioners experimented with new management, measurement, and program strategies, while stakeholders began to realize that the sector's funding structure hindered growth by starving organizational development (Letts, Ryan & Grossman, 1997; Gregory & Howard, 2010). Over that time, various nonprofit management strategies competed to achieve good governance and good results while keeping overhead costs low.

Evaluation emerged as the preferred method for determining a program's impact. However, it is not easy to apply to most nonprofit services, especially human services, leaving the potential for a skewed vision of what works for a particular social problem or population (Frederickson, 2001). Because of this and other factors, analysts say a more balanced approach, including both evaluation and performance measurement, is necessary to manage and improve programs, add to the nonprofit sector's knowledge base, and to engage donors (Radin, 2006; W.K. Kellogg Foundation, 2004a). Definitions for performance measurement vary, but practitioners are converging on a common understanding. When measures are integrated into a framework, the purpose is to "track selected performance measures at regular time intervals so as to assess performance and enhance programmatic or organization decision making, performance, and accountability" (Poister, 2003, p. 15). Poister's definition will guide this project.

Many factors led to the shift in the nonprofit sector, including both internal and external pressures. Today nonprofits must increasingly prove not only the efficacy of their programs, but also show efficient management. The donor community, including foundations, government and

individuals, increasingly looks for real and permanent change in the communities in which it

invests, and it expects nonprofit organizations to be able to show evidence of that change.

Internal Pressures

Many of the reasons for today's push towards performance measurement have internal

roots. Nonprofit organizations have a long history of measuring performance, an increased

number of professionals skilled in today's private sector business techniques, and a need to

overcome past scandals that left the sector on shaky ground with the public. On top of this, the

rapid growth in the sector has led to a need for better management in the face of fewer resources.

Growth. The number of nonprofit organizations has grown substantially over the last

two decades, with estimates today ranging from 1.4 million to 1.9 million, depending on the type

of organization (Keehley & Abercrombie, 2008; Strom, 2003). According to the National Center

for Charitable Statistics (www.nccsdataweb.urban.org/PubApps), the number of public charities

registered with the IRS grew 60 percent from 1998 to 2008. Total assets during that time grew

by 106 percent.

The huge influx of organizations did not accompany an increase in funding for the sector.

Individual giving has remained largely stable, at about two percent of the gross domestic

product, since the mid-1990's (Giving USA, 2009). The increased competition for funding

drives the need for organizations to differentiate themselves, often through the numbers

(Cunningham & Ricks, 2004).

Historical precedent. Nonprofit organizations have used measurement tools throughout

their entire history in the United States; societies founded as early as the late 1700's kept track of

the numbers of people they assisted in a given year as well as the types of assistance provided,

including food, coal and clothing (Olasky, 1992). In reviewing measurement history since the 1960's, Plantz et al. (1997) point out that before the 1990's nonprofits routinely focused on performance measures, including financial measures, program outputs and service quality measures. Outside funders and accountability mandates requested most of these metrics (Hatry, 1999).

Nonprofit organizations also routinely use and adapt those management trends in vogue in the private and public sectors, including evaluation in the 1970's, strategic planning and informed budgeting in the 1980's and 1990's (Poister, 2003), and outcomes measurement (W.K. Kellogg Foundation, 2004a) and benchmarking today (Keehley & Abercrombie, 2008). Light (2000) suggests that many nonprofits have repeatedly reformed their management strategies according to the latest trends, leading to increased stress and higher staff turnover rates without discernable improvements.

In addition, ever greater numbers of nonprofit staff are skilled in traditional private sector business practices (Blum, 2006; Cunningham & Ricks, 2004; Keehley & Abercrombie, 2008). Kevin Kearns argued in his book *Private Sector Strategies for Social Sector Success* that nonprofit executives are "among the most entrepreneurial managers to be found anywhere, including the private for-profit sector" (as quoted in Salamon, 2003). Cunningham and Ricks (2004) interviewed 10 top executives of nonprofits in Boston and New York, and they all agreed that performance measurement was a key management tool. Keehley and Abercrombie (2008) found that training in benchmarking over the last 15 years has led to its routine usage throughout the sector, albeit in an informal way.

The impact of scandals. Scandals at the United Way, American Red Cross and other organizations in the 1990s and early 2000s left a visible scar on the nonprofit sector. Excessive

executive perks and mishandling charitable donations broke some donors' trust, and the media became quick to investigate other potential abuses, real or imagined (Salamon, 2003). This sector, perhaps more than the public and private sectors, depends heavily on its reputation for trustworthiness and goodwill; the scandals revealed both the lack of transparency in decision making as well as a public misperception as to how the sector operates today. Accordingly, organizations fought to separate themselves from the scandals and focus attention on their mission and purpose, while at the same time experimenting with methods to better quantify their work and prove themselves sound investments (Light, 2000; Salamon, 2003).

In the private sector, one of the most common ways to determine a sound investment is by comparing financial information, including profit margin and stock price. Nonprofits are increasingly looking for similar metrics by which to show stakeholders their relative value (Blum, 2006; Lampkin & Hatry, 2009; McLaughlin, 2009; Meehan, Kilmer & O'Flanagan, 2004). As opposed to the profit margin, Lampkin and Hatry (2009) call this nonprofit equivalent the *outcome margin*, the measure by which nonprofits show a social return on stakeholder investment.

External Pressures

The nonprofit sector also faces increasing pressure from outside sources. The funding community has a significant impact on how the sector operates, and more and more this community demands comparative data for its own decision-making. Given the sector's explosive growth and competition, funders are leveraging their power to an ever greater extent. In addition, Congress and third party watchdogs have gotten involved in the push for transparency and accountability.

Funders - individuals. Analysts point to several catalysts in the shift towards greater accountability through measurement. The tech market bubble burst in the late 1990's forced donors to become more discerning with their money (Cunningham & Ricks, 2004). Likewise, a series of scandals involving nonprofit administrators at the United Way, American Red Cross and others led many to question the trustworthiness of the nonprofit sector (Carman, 2009; Cunningham & Ricks, 2004; Strom, 2003). The W.K. Kellogg Foundation's white paper on evaluation (2004) offers yet another reason: programs implemented during the Great Society of the 1960's failed to solve the problems they were designed to tackle, leading many in the public and private sectors to rethink the entire model for social services delivery. Quantitative research methods and corporate business metrics became the method for determining whether programs achieve their desired outcomes in an efficient way.

One of the most visible results of this re-thinking is a new kind of donor, sometimes called a *venture philanthropist* (Meehan et al., 2004): this donor not only wants to give money during his or her lifetime, but also takes an active interest in how well the money is spent. These donors, including Bill and Melinda Gates and George Soros, often tackle large-scale social problems with a desire to make a discernable impact, and they use private sector measurement tools to aid in this process (Conlin, Hempel, Polek & Dayton, 2003; Cunningham & Ricks, 2004). They also often represent the wealthiest and most prolific donors (DiMento & Preston, 2010).

Despite this new kind of donor's rise, research suggests that the majority of donors still give based on emotional connection and inherent trust of the nonprofit sector (Cunningham & Ricks, 2004; Meehan et al., 2004). This is both good news and bad news for nonprofits, which must balance their approach to meet all demands, from both the individuals giving $50 every

year and the wealthy investors looking to make an impact. Given the general movement towards

measurement, analysts increasingly argue that nonprofits must educate their donors on the new

trends and encourage their support.

Funders – foundations. Paralleling this new type of donor is a movement in the

foundation community towards a more efficient and impactful method of grantmaking.

Foundations today often focus on one or more issues and approaches to social change and choose

grantees whose programs most closely mirror their own philosophies. Like much of the funding

community, foundations are moving away from supporting organizations and towards supporting

issues or programs. They use various levels of evaluation and performance measurement to

track their own success. Kramer, Graves, Hirschhorn, and Fiske (2007) studied the changes in

foundations' approach to measurement and concluded that many have moved away from relying

on the gold-standard third-party evaluation approach, judging it to be too expensive, unwieldy

and impractical. Instead, many opt for a multi-party, iterative process that provides relevant

information in a timely manner. With data gathered from all grantees on a particular issue,

foundations adjust programs in progress, improve implementation and better assess the needs of

the community. This helps them to make better decisions for the next round of funding, plan and

implement new initiatives, and provide better grants to organizations, including operational

grants (Kramer et al., 2007).

Large, international foundations like the W.K. Kellogg Foundation, the William and

Flora Hewlett Foundation and others urge the sector to learn from its mistakes and make

information available on all sides so that the strongest organizations and models will come to the

forefront for funding (W.K. Kellogg Foundation, 2004a; Lenkowsky, 2007; McKinsey &

Company, 2008; Redstone Strategy Group, 2008; The Urban Institute & The Center for What

Works, 2006). To this end, they support research towards the development of outcomes and indicators that are comparable across organizations. They also promote and support the widespread dissemination of data collected by foundations and nonprofit organizations as a means of contributing to shared knowledge and further standardization in the field.

Despite these efforts, however, Carman (2009) found that foundation funding was not a significant predictor of evaluation or performance measurement for nonprofit organizations in the state of New York. Her study concluded that federal government contracts and the efforts of the United Way do the most to spread the implementation of these kinds of metrics. Local and state government agencies, along with foundations, relied more on descriptive reporting and monitoring with the recognition that many organizations lack the capacity to carry out large-scale evaluations and measurement systems. Lenkowsky (2007) suggests another motive in this seeming contradiction; foundations are afraid that public failure of a program will reflect badly on them and the power of philanthropy to create positive change.

Government. Following private sector trends, in the early 1990's Congress mandated performance measurement for government agencies through the Government Performance and Results Act (GPRA). The act requires federal agencies to have five-year strategic plans with performance measures tied to them. Agencies were given a period of time to implement the new systems, and starting with the 2003 budget cycle, were required to report their plans, measures and improvements (Carman, 2009; Lampkin & Hatry, 2009). The Office of Management and Budget (OMB) further implemented a process called Program Assessment Rating Tool (PART), "which attempts to link executive branch budget recommendations to the performance of specific federal programs" (Radin, 2006, p. 7), a form of informed budgeting. Because so many

nonprofit organizations contract with the government at all levels, they now find themselves facing the same regulations.

The government's interest in accountability extends to nonprofit agencies in a more direct fashion as well. Senator Chuck Grassley has expressed clear interest in a formal system of accountability (Blum & Williams, 2008), and the IRS's new Form 990, the informational tax return required for all nonprofits earning more than $25,000 in gross receipts, attempts to establish guidelines for governance, transparency and other areas of nonprofit performance (McLaughlin, 2009). As part of this larger trend, in 2005 the Panel on the Nonprofit Sector made recommendations to the Senate Finance Committee on increasing accountability for nonprofits. They suggested that nonprofit organizations create procedures for performance measurement and program evaluation based on specific goals and objectives. They also recommended more training in the sector in how to conduct evaluations (The Urban Institute & Center for What Works, 2006)

Competition. For-profit companies have entered into fields historically dominated by nonprofit organizations, such as childcare, health care, and job training. These companies win large-scale, multi-year government contracts to carry out services largely due to their experience in information technology and contract management rather than an expertise in human services (Salamon, 2003), but Light (2000) warns that the competition is here to stay, given that corporations like Lockheed Martin are achieving results. Ryan (1999) noted this trend over a decade ago, warning that "the public sector now sees business not as a pariah but as a role model" (p. 130). He predicted the monumental challenges that nonprofits have faced over the last decade as they attempt to respond to this new level of competition, including battles over mission, adequate capital, and remaining an integral part of communities while still achieving the

scale necessary to compete. Largely, nonprofits have responded by improving efficiency and output with modified private sector business tools.

Public misperceptions. One of the overarching problems is that the American public does not understand the nonprofit sector as it exists today. It is highly complex, competitive, and in many ways similar to the private sector in terms of business practices, compensation, and professionalism. Salamon (2003) describes the threat that nonprofits face today this way:

> Thanks to the pressures they are under, and the agility they have shown in
> response to them, American nonprofit organizations have moved well beyond the
> quaint, Norman Rockwell stereotype of selfless volunteers ministering to the
> needy and supported largely by charitable gifts. Yet popular and media images
> remain wedded to this older image, and far too little attention has been given to
> bringing popular perceptions into better alignment with the realities that now
> exist, and to justifying these realities to a skeptical citizenry and press (p. 79).

This older image has little room for the organizational complexity and funding required for performance measurement.

Brody (2002) lists several popular misconceptions: nonprofit managers should not receive compensation, overhead spending is a wasted investment, and fundraising costs should not exceed 15% of earnings. She states, "Clearly, a public that does not understand the demands on nonprofit organizations or that focuses on inappropriate or unrealistic considerations cannot perform effective oversight and can induce inefficient and ineffective behaviors" (p. 489). She and Salamon (2003) also join the chorus calling

for more widely available information the sector can use to educate the public about its

changing role in the American economic infrastructure.

       Third parties. Third parties are also getting involved in the accountability business, in

the form of consultants and so-called watchdogs, self-appointed agencies who evaluate the

programs and fundraising practices of other nonprofits through rating systems (Hopkins, 2005;

Preston, 2008). The Better Business Bureau's Wise Giving Alliance, Charity Navigator,

GuideStar and other third parties have their own thresholds nonprofits must reach in order to be

worthy of a donor's investment, standards that are often outside and more stringent than legal

requirements. They post their ratings, along with public information on individual organizations,

on the Internet to help donors make choices based on these standards (Keehley & Abercrombie,

2008; Meehan et al., 2004). The National Council of Nonprofit Associations and the National

Human Services Assembly responded to the trend in 2005 by rating major watchdogs and other

list-providers according to their own standards and appealing to donors and the nonprofit

community to focus on results rather than finances. Without a standardized format for

comparing results, however, it remains difficult for the donor community to take this advice.

       Over the last several years, this informal regulation has gained respect in the donor

community (McKinsey & Company, 2008); GuideStar.org alone receives eight million visitors a

year (Keehley & Abercrombie, 2008). Hopkins proposed in 2005 that many of the standards

developed by these agencies outside of the nonprofit tax code would eventually "take on the

force of law" (p. 269). The IRS reformed tax Form 990 partly in response to the increased

demand for public information and accountability. The new Form 990, while stopping short of

legal requirements, strongly suggests that many of these same standards are integral to good

governance (McLaughlin, 2009). In a very real sense, then, the nonprofit sector already faces new standardization requirements regarding governance and finance.

At the same time, many third parties today issue their ratings with open recognition that limited information makes this assessment very difficult and often impossible to do well. They suggest that the numbers are just the beginning of an organization's story; many ask those organizations they rate to add to the discussion by posting information about programs, and recommend that potential donors engage the organization in a dialogue about their financial numbers and actual program work (Preston, 2008; McLean & Coffman, 2004; Meehan et al., 2004; Waide, 2002).

Reflecting just how dynamic this trend is, Charity Navigator recently announced that it will incorporate outcomes measurement into its rating system. The organization's new CEO states that the criticisms against purely financial metrics have pushed it to incorporate new ones that measure organizations' impact, including asking whether organizations measure outcomes at all, and whether those outcomes are reasonable (Berger & Penna, 2010, March 31). This new "triad" rating system will be doubtless be difficult to implement and watched very closely by the sector. However, it is an important step towards the goal of rating organizations based on value to their communities rather than simply on the data that is publicly available.

One brand new third party rater, called Root Cause, offers a different glimpse of comparative measurement based on impact. It attempts to bridge the information gap by performing independent research on best practices and using it to rate agencies. It provides this information to venture philanthropists, or as it calls them, social impact investors, as a means to maximizing their philanthropic impact (Root Cause, 2010a). Its first report details the research on school readiness programs, the state of philanthropy on the issue, and 21 top institutions in

Massachusetts following best practices in this field. It hopes to expand the data to include larger portions of the United States in the future (Ailworth, March 31, 2010).

Independently produced research on best practices offers a significant step forward in the availability of data that investors can use to understand and support issues. The report couches its recommendations in firm data and indicators, and offers baseline components for people to look for when considering a donation to a particular organization (Root Cause, 2010b). Root Cause produced the report with funding from the William and Flora Hewlett Foundation, one of the major foundations interested in information flow and measurement in the nonprofit sector (Ailworth, March 31, 2010).

Measurement Theories and Application

Today's Performance Measurement Systems

If performance measurement is not new to nonprofit management, why did it fall out of fashion, only to be picked up again in the last ten years by funders and researchers? In part, nonprofits have the same habits as other types of organizations that get caught up in the latest management trend, letting old practices slide. On the other hand, some authors suggest that in the past performance measurement was largely funder-driven and therefore less applicable to the unique circumstances of individual programs (Plantz et al., 1997; Hatry, 1999). Others suggest that the focus on accounting-based measures led many to look for newer systems that focused on people, mission, and results (Ittner & Larcker, 1998). In other words, performance measurement was less useful to nonprofits and managers had little incentive to use measures that accurately portrayed how a program worked, problems and all. Poister (2003) says today's systems are different:

"In contrast to earlier attempts at developing performance measurement systems,

which often appeared to be less purposeful, less focused, and less well aligned

with other evaluative and decision-making processes - the current generation of

measurement systems are more mission driven and results oriented. More often,

the 'new' performance measurement is tied to a strategic framework, emphasizes

the customer perspective, measures performance against goals and targets, and

incorporates measurement systems in other management processes in meaningful

ways" (p. 9).

Another side of the equation is the change in foundations' approach to grantmaking.

With a new focus on issues over organizations, foundations find themselves using information in

specific ways, including setting baseline data for their target areas, learning from past grants,

conducting needs assessments, setting more precise goals, and improving implementation.

Finding the traditional evaluation tools inadequate, they are experimenting with their own

changes and working with grantees to improve the data collection process for all stakeholders

(Kramer et al., 2007).

Performance measurement for nonprofits starts with program managers, rather than

external evaluators. They often have the best insight as to which measures are the most relevant

in determining how the program functions and useful for decision-making (McDavid &

Hawthorn, 2006). At the foundation level, program officers may set standards and collect data

from all grantees in a particular program area, or bring grantees together to discuss these metrics

(W.K. Kellogg Foundation, 2004a). Often, individual metrics come from the organization or

program logic model and serve as a guide for managers to track the program's function and

effectiveness.

Logic models. Good performance measurement uses metrics with a real and valid link to the program's effects, and that measure what they intend to measure (McDavid & Hawthorn, 2006). To be sure the system meets these two goals, indicators are pulled from the program's logic model, the visual representation of how the program creates impact. Many models loosely describe how activities lead to the desired effect (Poister, 2004); by continuing the model into the program's processes, measurable indicators emerge. Managers can then use this data to make decisions about budgeting, cutting or expanding programs, staffing requirements, etc (Poister, 2003).

Most authors describe the process for logic model development in nearly identical terms (see, for example, Hatry, 1999; Poister, 2003; W.K. Kellogg Foundation, 2004a). The template looks something like Figure 1. A completed logic model might more closely resemble Figure 2.

Resources/inputs	Activities/processes	Outputs	Outcomes (short, medium and long term)
• Indicators	• Indicators	• Indicators	• Indicators

Figure 1. Logic model template.
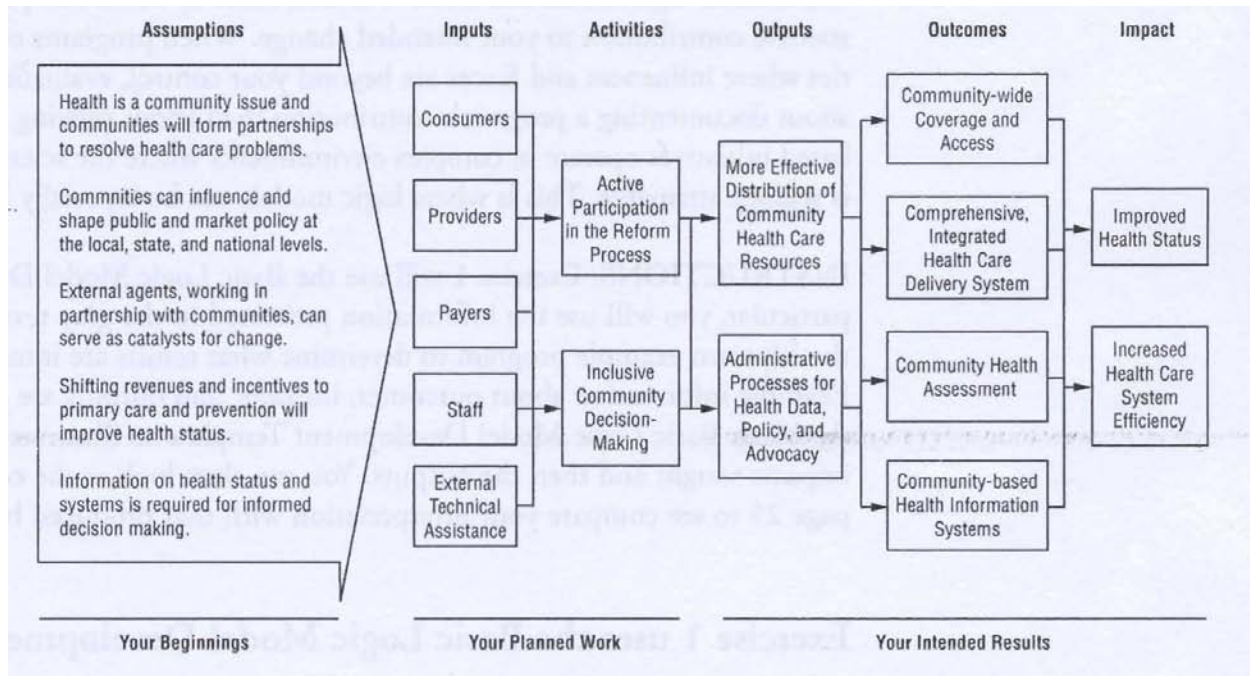
Figure 2. Completed logic model (W.K. Kellogg Foundation, 2004b, p. 11).

The key elements of a logic model are:

- *Resources*, or inputs, which includes everything used to produce the end result, whether it is a physical product, a service, or other type of work. Expenditures, employee time, and physical materials all fall under the resources category.

- *Activities*, the processes or "work steps" (Keehley & Abercrombie, 2008, p. 30) carried out to produce the result.

- *Outputs*, the individual units of work produced as a result of the activities.

- *Outcomes*, the effect of the outputs measured over the short, medium, and long term.

- *Indicators*, individual units of activity that will be measured to track performance (Hatry, 2009)

According to The Urban Institute and the Center for What Works (2006), outcomes are "the results of a program or service that is of direct interest and concern to customers of the program" (p. 6). That addition is crucial to the understanding of what can and should be included in the outcomes category: many of the results directly affecting participants lay outside of an organization's control and tracking ability. However, practitioners agree that these results must be included anyway in order to get the best perspective on a program (Poister, 2003).

Today many analysts recommend adding one or more categories to the left of the traditional model, including mission, assumptions, and external factors. By starting with the mission statement of the organization or program, managers help to ensure that the logic model stays true to the organization's philosophy and commitment to the community rather than simply funding requirements. In addition, indicators should balance so as to provide a good overall picture of how the program operates (Poister, 2003). Assumptions include the major ideas that the program is based on, including the relevant research and program models. External factors refer to the particular circumstances of the target population for a program (W.K. Kellogg Foundation, 2004a).

The extent to which an individual organization can hold itself accountable for long-term outcomes is a matter still under debate, as is the usefulness of a strict logic model (W.K. Kellogg Foundation, 2004a). Most analysts today recommend that organizations create the logic model with an understanding of its inherent limitations. One glaring flaw is that these models often lack practical transitions from theory to practice, leaving individual organizations to set outcome indicators or leave them out altogether. This is where standardization efforts come in, such as the Urban Institute and the Center for What Works' (2006) two-year effort to develop common outcomes and indicators for 14 human service program areas.

Lampkin and Hatry (2009) describe another common problem: logic models track real change over time, which sometimes occurs outside of the organization's interaction with the client. Many nonprofits are unable to follow up with clients to measure this key indicator. Sometimes this leads to the use of proxy measures, which can be problematic if they do not really align with the intended indicator. For example, training surveys often ask whether participants have increased their knowledge about a certain behavior. Change in knowledge does not equal a change in behavior, and many nonprofits do not have the ability to follow up with clients to find out whether the training affected behavior over the long term.

The Hewlett Foundations' Population Program also issued warnings in its white paper on developing logic models and frameworks for measurement and decision making throughout its programs (Redstone Strategy Group, 2008). Too-rigid logic models can oversimplify the path by which a program effects change, while loosely defined models cannot adequately guide decision-making. Identifying the wrong indicators or leaving out important ones can also skew the focus of a program's resources (Hatry, 1999). Strategists recommend treading a fine line and spending the time and effort it takes to get the logic model right.

Performance measures. Once the logic model is developed, program managers can use it to decide which performance indicators to use in a measurement system. While the logic model offers a picture of how the program creates change, the performance measurement system shows quickly and concretely how the program functions over a specific period. The number and type depend on the organization's ability to collect data, the information's usefulness to decision-makers, and the indicators' relevance to the mission of the program (Hatry, 1999; Poister, 2003). It is important to make sure that there are indicators to track every activity, that multiple parties have a hand in developing indicators (Thiel & Leeuw, 2002), and that together

the indicators support the management process of the individual organization using them

(Poister, 2003).

A list of common indicators includes *output measures*, which refer to the amount of work

performed, usually broken down into units such as number of trainings held (Poister, 2003).

Process measures look at the amount of work that comes into an organization or program. They

can include workload, such as number of clients per staff person, as well as number of activities

required by a grant (Hatry, 1999). Poister (2003) suggests that these measures only become

useful in performance measurement systems when they relate to other issues, such as a need to

investigate staffing levels or to increase productivity. *Productivity measures* are "the actual time

taken to complete a process or process step" (Keehley & Abercrombie, 2008, p 36). An example

would be number of hours spent per staff person per training, including preparation time, travel

and the actual training.

Service quality measures can include various types of data depending on the program.

Managers might track whether the activity meets grant requirements or OSHA safety standards,

for example. These measures also include data on aspects such as thoroughness, accuracy,

convenience, and whether clients received courteous and professional treatment (Hatry, 1999;

Poister, 2003). Hatry (1999) considers service quality measures intermediate outcomes.

Hatry (1999) defines *efficiency measures* as "the relationship between the amount of

input and the amount of output or outcome of an activity or program" (p. 13). These measures

receive a lot of attention from funders and government agencies today: how efficiently can

nonprofits deliver services? Hatry (1999) warns that when using output indicators in this ratio,

managers should take care not to increase efficiency at the expense of outcomes. The most cost-

effective solution is not necessarily the right one for a particular program or target population.

Effectiveness measures are the other side of the same coin. They come back to impact: is the

program producing the intended results over the short, medium, and long term (Poister, 2003)?

Customer satisfaction measures can easily fall under service quality, but Poister (2003)

recommends keeping them separate. This set of measures adds another dimension to the overall

picture of the program by asking how satisfied clients are with the services they receive.

Indicators may come from complaints, survey results, and error rates.

Application

Performance data becomes useful when it is part of a system that regularly monitors

indicators and produces reports. Managers can then analyze it in ways that bring problems to

light and suggest solutions (Hatry, 1999; Poister, 2003). As the sector moves towards greater

accountability and standardization, programs must also offer their data for comparison with

similar organizations (Lampkin & Hatry, 2009). There are several methods for analysis.

Breakouts. Breakouts are "disaggregations of the outcome data for each indicator"

(Hatry, 1999, p.103). They get into the details of the program's operations by comparing points

of interest. They can answer questions such as why high or low performance occurs in some

areas and not in others. They can also look at disparities between client groups. The basic

breakout categories include trends over time, performance per project, performance per customer

group, geographical location, difficulty of workload, type and amount of services provided, and

reason for outcome or rating (Hatry, 1999; Poister, 2003).

Dashboards. Dashboards take the most relevant indicators and put them into a visual

representation that managers monitor on an as-needed basis. For example, managers at the

Washington organization KaBoom have widgits on their computers showing information on

volunteer retention for the month (Blum, 2006). The information may take the form of graphs or pictures for a less threatening format and faster scanning (Poister, 2003). McKinsey & Company (2008) found dashboards to be information-rich but not overwhelming to nonprofits with limited infrastructure. The consulting firm considered them particularly useful in flagging potential problems, as well as highlights.

Benchmarking. Benchmarking is the process of comparing data to established measures (Keehley & Abercrombie, 2008). These measures can include previous performance, outcomes for different programs, recognized general standards, programs with different types of service delivery, pre-established targets (Hatry, 1999; Poister, 2003), or the outcomes of similar organizations or private sector companies (Keehley & Abercrombie, 2008). By comparing process and results, organizations can identify areas for improvement and find *best practices*, particularly successful techniques for achieving results. Benchmarking became a popular management practice in the 1980's in the private sector, and Keehley & Abercrombie (2008) shortened the formal technique to meet the nonprofit sector's needs through solution-driven benchmarking. This technique, instead of comparing all processes between two organizations, starts with managers identifying a *problem at hand,* then consulting with successful organizations on how to solve it through best practices.

Balanced scorecards. Balanced scorecards emerged in the private sector in the early 1990's as a framework in which to measure performance. The finished product shows a list of goals for a specific time frame next to the activities and measures used to achieve them. The goals traditionally focus on four areas: financial performance, customers, internal business processes, and innovation and learning (Poister, 2004). To be useful, it must link strategy to a set of indicators, both financial and non-financial (Chan, 2004). Chan found in his 2004 survey

that although many municipal governments adopted the balanced scorecard approach, they did so in an uneven way, focusing on some of the four target areas and leaving others, including innovation and financial effectiveness. On the other hand, management highly valued the measures that it tracked regularly.

      Integrated systems. The ideal way to implement a performance measurement system starts with training staff on how to collect data so that everyone collects the same information in the same way and enters it into the system. For individually tailored measurement systems, organizations can use databases to track information and generate reports on breakouts and key indicators. Ideally, they should be able to integrate data from different sources, such as spreadsheets generated by other databases or software (Poister, 2003). An organization can develop a new database or tailor one from stock products according to its needs and complexity. For shared systems, websites house the tracking system and each agency enters its individual data.

      When putting indicators together in a system, managers must take into account the potential abuses of the system. They should group indicators in a balanced way and tie the groups to a rewards system, in order to avoid the tendency to focus on certain indicators. Likewise, indicators must be relevant to the program (Hatry, 1999; Poister, 2003) and follow established patterns where it makes sense to do so. Measures should closely relate to the program's mission and focus on real outcomes rather than proxies wherever possible (Poister, 2004). In addition, implementation of the system should include pre-scheduled, periodic meetings to encourage their use in management decisions (McKinsey & Company, 2008).

Shared systems

One of the most desired functions of measurement is also one of the most difficult to develop. Funders, third parties, and even nonprofits themselves struggle to compare similar organizations in any meaningful way. However, considerable progress has been made in developing a common framework over the past decade (Kramer, Parkhust, & Vaidyanathan, 2009; Meehan et al., 2004; The Urban Institute and The Center for What Works, 2006). Foundations lead the effort to develop common indicators that are measurable across organizations. As the process continues, the sector slowly figures out what can and cannot be compared, where indicators and models break down, and what constitutes a common-sense approach (Kellogg Foundation, 2004; Lampkin and Hatry, 2009; McKinsey & Company, 2009; Redstone Strategy Group, 2008).

The William and Flora Hewlett Foundation funds research on how the nonprofit sector can share data in meaningful ways. One of the reports investigated ways to increase the amount of information circulating through the nonprofit sector as a way to improve performance and increase donor impact (McKinsey & Company, 2008). Their recommendations included using the information at hand more efficiently as well as generating new data in cost effective ways. They recommended that nonprofits and foundations publish or post online individual research efforts in order to reduce duplication and increase the sector's knowledge base. Many large foundations and nonprofit organizations now have research pages on their websites for this purpose (see, for example www.hewlett.org/library; www.wkkf.org/knowledgecenter; www.liveunited.org/outcomes/library).

Another report, produced by FSG Social Impact Advisors (Kramer et al., 2009), offers a glimpse at the potential for organization-driven shared performance measurement systems.

Researchers identified twenty groups of nonprofit organizations experimenting successfully with three basic frameworks: shared measurement, comparative performance, and adaptive learning systems. Hundreds of organizations are implementing these breakthrough systems by working together on various social problems across the country.

   Shared measurement. Shared measurement platforms allow organizations to choose from a common set of measures and track them using web-based tools. One of the earliest examples of progress in this type of system resulted from The Urban Institute and The Center for What Works' (2006) two-year effort to build a set of common outcomes for each of 14 program areas, along with a framework that nonprofits can use to develop outcomes for other types of programs. Researchers gathered data on nonprofit organizations, umbrella groups, and accreditation agencies in an effort to glean the best indicators used for each program area, such as family literacy or emergency shelter. The white paper on each of these program areas includes a program description, a chart describing the sequence of likely outcomes over time, and indicators that nonprofits can use to track success (The Urban Institute & Center for What Works, 2006).

   The benefits of a shared measurement system include increased credibility for individual programs, reduced measurement costs, and the benefit of extensive research by experts in the field, which leads to better indicators and benchmarks. For example the Success Measures system, created by a professional development organization over the span of two years, costs individual organizations around $2500 per year to use, with a one-time training fee of between $7000 and $10,000, significantly less than most independent evaluators charge (Kramer et al., 2009). These systems also allow organizations a great degree of independence to personalize the

metrics according to their unique programs. They can also keep their results private or share

them as they see fit (Kramer et al., 2009).

Funders can use shared systems with their grantees, reducing their own costs and need to

develop individual measurement systems. The similarities between the data also help them to

make better decisions about the allocation of resources. The result for this process is better

programs with greater impact (The Urban Institute & Center for What Works, 2006).

Comparative performance. The next level of integration identified by Kramer et al.

(2009) is comparative performance systems, where organizations and funders agree to use the

same measures and methodologies to track results and make comparisons. These systems

basically integrate benchmarking into a formalized process across many organizations. The

federal government's GPRA measures could be considered a kind of comparative performance

system. Nonprofit organizations must agree to pre-set measures when they contract for services,

and the government agency uses the data to track how well it is achieving its own goals (Carman,

2009). However, there is more leeway in data collection for GPRA than in the systems described

by Kramer et al. (2009). Some funders take the process a step or two further, requiring

participating nonprofits to take part in cluster evaluations to determine best practices across a

program area (Carman, 2009), thereby increasing the level of shared knowledge within a

comparative framework.

The danger of a system like this is that nonprofits could find themselves being compared

without the proper context, a situation they constantly try to avoid given the highly competitive

nature of funding. However, according to Kramer et al. (2009) the key to these systems is

recognition by all parties that the purpose of comparison is mutual learning, not competition.

The benefits of such a system include sharing true best practices for improved performances,

tracking field-level impact, and better assessment of grantee performance by funders. In addition, individual organizations benefit when multiple funders agree to use the same measurement requirements, thereby reducing workloads and redundancy.

Adaptive learning systems. The highest level of integration is adaptive learning systems, which "involve highly structured, long-term processes that build the capacity of participating organizations to collectively define, measure, learn from, coordinate, and continuously improve their efforts" (Kramer et al., 2009, p. 16). The systems usually involve one independent organization that coordinates the efforts of many others around an issue, such as public education. Agencies working on issues all along the spectrum come together and develop metrics through a facilitated process. The independent organization tracks the overall progress of the learning system and facilitates continued learning and improvement (Henderson, 2009).

For example, the Strive initiative in Cincinnati uses 15 "action networks" for agencies working on education, including early childhood, tutoring, and college counseling. The agencies use a modified version of the Six Sigma process, which includes defining the working group and its goals over time, developing indicators and tracking progress towards them, analyzing data and "establishing local evidence of effectiveness" (Kramer et al., 2009, p. 43), developing a plan of action, and working on a continuous improvement plan. Individual agencies coordinate within their networks, but also agree to overarching goals for the larger initiative. Strive reports progress and learning each year to the community (Kramer et al., 2009). The reports take the form of a dashboard of sorts, easily understood graphic representations of results. See Figure 2 for an example. The Strive model has been so successful in its first three years that it is already in the planning stages for a four-city expansion through Living Cities, a collaboration of foundations (Henderson, 2009).

Goal 1: Every child will be PREPARED for school

	Current percentage	Current benchmark	Change since recent year	Change since baseline year
Cincinnati	48%	59%	↑	↑ (4 pt)
Covington	71%	75%	↑	n/a
Newport	62%	85%	↑	↑ (2 pt)

The arrows have been shaded to help discover meaningful changes and visually set them apart in this format. Green shading is used for changes of three or more percentage points up, red shading is used for changes of three or more percentage points down, and changes that are less than plus or minus three percentage points are left white.

Arrows not within a circle represent numerical changes, not percentage point changes.

Goal 2, 3 & 4: Every student will be SUPPORTED, SUCCEED academically and ENROLL in college

Cincinnati Public Schools

	Current average	Current benchmark	Change since recent year	Change since baseline year
Develop. assets	46%		No trend data available	
4th grade reading	60%	75%	↑ (3 pt)	↑ (4 pt)
8th grade reading	65%	79%	↓ (3 pt)	↑ (1 pt)
4th grade math	55%	74%	↑ (2 pt)	↑ (2 pt)
8th grade math	54%	58%	↓ (3 pt)	↑ (2 pt)
Graduation	80%	95%	↑ (3 pt)	↑ (2 pt)
ACT composite	19.0		↓ 0.5 pt	↓ 0.7 pt
College enrollment	64%	70%	↑ (2 pt)	↓ (1 pt)

Covington Independent Schools

	Current average	Current benchmark	Change since recent year	Change since baseline year
Develop. assets	33%		No trend data available	
4th grade reading	54%	61%	↑ (2 pt)	↑ (4 pt)
8th grade reading	43%	59%	↑ (4 pt)	↑ (3 pt)
4th grade math	51%	42%	↑ (6 pt)	↑ (4 pt)
8th grade math	27%	37%	↑ (4 pt)	↑ (2 pt)
Graduation	90%		↑ (1 pt)	↓ (3 pt)
ACT composite	17.8		↓ 0.2 pt	↓ 0.04 pt
College enrollment	45%		↓ (2 pt)	↓ (1 pt)

Figure 2. Strive Report Card to the Community. (Strive, 2009, p. 5)

Complete College America, formed in 2009, applies comparative measurement and adaptive learning for increased impact. The organization focuses on dramatically increasing the number of people graduating from college or certificate programs in the United States. It does this through its Alliance of States, a coalition of state governments that have agreed to set goals for increasing college completion and to collect and report on common metrics (The Alliance of States, nd). To accomplish their goals, states receive technical assistance from experts. Five major foundations recently agreed to help fund this major collaboration (Foundations, states working to significantly increase college completion, 2010).

The goal of adaptive learning is to tackle large, multi-pronged social problems, called adaptive challenges, through systemic change. Because of this, it requires input and commitment from multiple sources and a flexible but rigorous evaluation process. Kramer et al. (2007), in studying this process across the country, note that "evaluation must serve to enable ongoing experimental 'corrections' in both foundation interventions and stakeholder actions" (p. 23). Measurement, then, becomes an integral part of the intervention process, aiding in the group's learning and ultimate success.

The benefits of adaptive learning systems include a structured process of measurement, professional facilitation, collaborative problem solving, and an integrated approach to social problems, ensuring that all aspects of the problem receive the proper attention. Kramer et al. (2007) call this approach collaborative inquiry. Organizations operating in these systems also gain the benefit of scale, which gives them the power to demand information and participation from larger institutions, including government (Kramer et al., 2009).

Kramer et al. (2009) identified several elements for a successful adaptive learning system, including multi-year, substantial funding, engagement at a broad level with clear expectations in place, voluntary participation, measures set independent of funders, and use of web technology. Organizations also need to make use of feedback for continual improvement and adequately train and facilitate the implementation of the process. These systems offer the potential for increased efficiency, increased knowledge, and increased impact on social problems.

The Limitations of Performance Measurement

Despite these advances, the nonprofit sector still has a long way to go before a clear, comparable framework of data becomes the norm rather than the exception. Given these benefits, one must ask what is holding the sector back. The real crux of the measurement problem is two-pronged and deeply entrenched. In many ways, it boils down to funding and improper tools.

Funding

The same pressures that push nonprofits to quantify and prove effectiveness also leave them with no way to pay for measurement.   Despite the changing mindset of some foundations and donors, the funding community still largely wants to finance programs, not management or evaluation (Cunningham & Ricks, 2004; Keehley & Abercrombie, 2008; Meehan, Kilmer, and O'Flanagan, 2004). Performance measurement systems require increased infrastructure, consultant expenses, and maintenance to produce the results desired by managers (Poister, 2004). Program evaluations are even more expensive, costing tens of thousands of dollars a year for external validation (Kramer et al., 2009). Many donors and foundations see them as a drain on resources better spent elsewhere, especially given that the techniques have not been perfected (Carman, 2009).

Some individual donors take this attitude in another direction, believing that all nonprofit organizations are equally worthy of support, so that metrics cease to matter (Cunningham & Ricks, 2004). Still others believe that spending on capacity means stealing from programs; this attitude shows up when third party watchdogs fault organizations for 'excessive' spending on overhead and building up financial reserves (Waide, 2002). Though they pertain to programs,

some functions of performance measurement and evaluation, such as the information technology

required to run them, become overhead expenses, which can throw off an organization's

financial balance (Brody, 2002; Ryan, 1999).

The result is that organizations with larger budgets, mandates, and sources willing to fund

evaluation are more likely to carry out large-scale evaluations than smaller organizations with

limited funding streams (Carman, 2009; Lampkin & Hatry, 2009). Because of the value placed

on expensive evaluations, however, programs and organizations that do not use them face

increased scrutiny and may potentially lose respect. Complex interventions are also less likely to

be evaluated because of the difficulty and cost involved, leaving them vulnerable to funding cuts

(W.K. Kellogg Foundation, 2004a).

Today's brand of performance measurement is becoming more popular and credible

(Preston, 2008) but is still less likely to be funded. However, organizations willing and able to

incorporate performance measurement techniques find themselves in a better position to

maneuver the organization and its programs for maximum efficiency and impact (Blum, 2006;

Lampkin & Hatry, 2009; Meehan, Kilmer, and O'Flanagan, 2004).

Inherent limitations

It is important to recognize that the slow adoption of performance measurement and

common frameworks stems from the nature of nonprofit programs, which often focus on social

problems. Evaluators, researchers and funders want to see common points of interest across

programs, but many program managers and funders are understandably wary of adopting such

measures wholesale. Most measurement techniques assume a level of control and comparability

that are simply unavailable in this setting. Interventions do not take place in laboratories and

external factors are often the largest variable in a participant's success. Because of this,

traditional metrics are not the best tool for ascertaining success (Frederickson, 2001; W.K.

Kellogg Foundation, 2004a).

The nature of social problems and their complex intervention strategies make them

inherently hard to measure in and of themselves, much less to compare to other intervention

strategies. Organizations do not apply the same strategies in the same ways to the same sets of

clients: underlying philosophies, management structures, resources, and clientele are different for

each one. Because of this, no comparison is truly an "apples to apples" situation, but only an

approximation. One performance measurement system applied to two organizations will likely

produce uneven results without the proper input. Because of this, analysts recommend that

organizations research and tailor performance measurement tools to fit their unique

circumstances (Poister, 2003; Radin, 2006; Redstone Strategy Group, 2008; The Urban Institute

& The Center for What Works, 2006; W.K. Kellogg Foundation, 2004a).

In addition, detractors rightly argue that focusing on numbers can reduce the quality of

programming rather than increase it. Brody (2002) argues that focusing on outputs can lead to

quotas, where the goal of a program is to reach target numbers rather than to provide solid

services. She also notes that focusing on outcomes can hold nonprofits hostage to social factors

outside of their control, given that external factors affect such a large proportion of social service

programs. Frederickson (2001) adds that the public and nonprofit sectors deal with society's

hardest problems, so failure may not signify a poor program, but simply reflect the intractability

of the problem. Even though some organizations try to measure exactly this, he argues "it is

often impossible to know how much would have been achieved had the program not existed" (p.

39).

Radin (2006) observes that where government agencies focused on measures to the exclusion of other factors, they have seen an increase in "gaming the numbers," which includes symbolic behavior and teaching to the test, methods of improving the data without improving the situation for the target population. With symbolic behavior, the numbers describe the correct activities and processes, but they are not really taking place. Teaching to the test is a problem exacerbated by No Child Left Behind, where school curriculum shifted to focus on problems covered in standardized testing rather than providing a more complete study of subjects. This happened because the legislation added financial rewards and penalties to test results but did not measure other aspects of education deemed equally important by many in the field.

The question of innovation is another that several authors have explored, namely the phenomenon that performance measurement restricts innovation rather than spurring it. Frederickson (2001) refers to Max Weber's iron cage theory, where bureaucracy grows so powerful that it becomes self-perpetuating. When this happens, programs cease to work for positive change and instead look for ways to stay in business. Frederickson also sees measurement as a top-down method for controlling a program, rather than a bottom-up approach to improvement. Bottom-up approaches are more innovative because stakeholders have shared responsibility, history and recognition for success. This environment is inherently more innovative than a top-down approach that focuses on numbers. An organization like Strive starts to overcome this tendency by involving all parties in collaborative innovation and metrics for shared learning.

Frederickson (2001), Light (2000) and Radin (2006) also note the tendency for organizations to become more alike through practices like benchmarking, rather than becoming more successful. Programs are viewed through lenses of science, reputation, and bureaucracy,

forcing them to fit into a uniform framework in order to receive consideration. When this happens, competition revolves around slight differentiations rather than real impact. Completely new techniques for achieving results have no place in these systems.

Nonprofits themselves understandably shy away from the many attempts to quantify and judge the sector based on data of questionable quality. They also raise concerns that the public is not yet sufficiently aware of the sector's complexity and breadth and so may misunderstand any standardized measurement system put in place (Brody, 2002). Tools like standardized disclosure forms and formal benchmarking are viewed with suspicion, though Keehley and Abercrombie (2008) argue that the sector has used informal benchmarking for a decade or more as a method to learn and implement best practices. They further suggest that nonprofit organizations largely do not face the kind of reluctance over information sharing experienced in the private sector. Because these organizations see themselves as collaborators against social problems they are more inclined to share results, although they do compete for resources and funding.

Even with these detractions, most agree that the future of nonprofit management includes performance measurement. To answer the very real questions of one-size-fits-all indicators, it is important to avoid a top-down approach, led by funders. As the stakeholder with the most experience in programs and affected populations, the nonprofit sector must lead on this issue. Kramer et al. (2009) described several ways to do this, all of them promising. Nonprofits must also work hard to educate the public on both the uses and limitations of these systems. The literature suggests that loose standards are preferable to strict comparisons for organizations not participating in adaptive learning systems. It also suggests that some types of data will remain out of reach for the foreseeable future. Given this, for most organizations measures should be

developed by practitioners and be strict as to the individual measurement but loose enough to avoid restricting the program or its pathways of change.

At the same time, many organizations already know and collaborate with others in the field on common social problems. Strive offers a promising structure to solve a common phenomenon: several organizations working across the spectrum of an issue, competing instead of working together to push the needle forward. Given the initiative's success, more influential organizations should consider adapting the model to fit their own environment and collection of organizations.

With foundations taking on the task of social change, there is the potential for increased learning and funds to aid in this process. First steps by the Hewlett Foundation and others provide good learning tools for the entire sector, and research on best practices is uncovering many promising models for organizations and foundations to consider. The literature suggests that some funders and organizations are moving closer to agreement on the best uses of measurement; the rest of the sector should study the issue closely and take a role in this process, or risk having the task completed without their input.

The research on nonprofit performance measurement and evaluation offers many resources for individual organizations looking to improve their own metrics. Despite the disagreements within the literature about the importance and uses of measurement, organizations who wish to thrive in the nonprofit sector of the 21st century must have a system for tracking their own performance and impact, and for determining success. By implementing proven practices, organizations can set themselves up for improved performance, increased impact, and better relationships with funders.

From Theory to Practice: Study and Modification of a Mid-size Nonprofit Agency's

Performance Measurement Framework

Methods

The present researcher worked with a nonprofit agency over the course of seven months to evaluate and add to its performance measurement system. The researcher studied the organization's system through a combination of interviews, data collection and sampling. After determining strengths and challenges, the present researcher helped the organization to identify new measurement indicators, developed new breakout categories, and administered a survey to help the organization set baseline data for its new indicators. Finally, the present researcher identified areas for future planning and measurement and made suggestions as to how to use the data in the organization's decision-making.

Organization overview

The organization works on education issues in the state of Wisconsin by promoting parental engagement from birth throughout children's academic careers. As part of this mission the organization promotes and runs three programs for parental engagement across the state, all using nationally recognized program models that have been externally evaluated. It also works on policy at the state level, advocating for a unified infrastructure of government, nonprofit and private services for children and families so that everyone has access to the same high quality education and protections.

Because the organization focuses on high quality programs, it values independent research and evaluation of the program models it supports. After receiving its designation as an

independent 501(c) 3 nonprofit agency four years ago, the organization began developing its

infrastructure to handle the requirements accompanying its new status. One of the first

developments was a database to store and process the data collected internally and externally on

its programs. The organization currently works to respond to new and varied data requests from

its funders by collecting all pertinent program information within its database system.

Current metrics. The organization currently contracts with three consultants to carry

out external evaluations of its three major program models. Two of these evaluations focus on

determining causality; they use quasi-experimental, longitudinal designs to determine whether

the programs work as intended. The third evaluation focuses largely on outputs, whether the

program is providing the services specified in its contract.

Because of these evaluations, the organization already uses a considerable portion of its

resources on program analysis and evaluation, including employing a full time information

technology (IT) staff person to design, maintain and enter data into databases created for

evaluation. External evaluators are likewise contracted to carry out rigorous quantitative studies

that look into the effectiveness of various program models in changing client behavior over the

short, medium, and long term. They not only make use of the data the organization collects, but

also provide their own data at regular intervals.

The organization operates a large federal grant that it has maintained for 15 years, called

here Program A. A major component of Program A is trainings for parents and education and

early childhood professionals. In fiscal year 2008, the organization held nearly 100 trainings and

presentations at conferences for these groups on topics such as discipline and the development of

parent-centered action teams in schools. The long-term goal of these services is better-educated,

more parent-centered service providers, as wells as parents who are more fully engaged in their

children's learning.

When the original grant was written in 1993 logic models were not the preferred method

for program planning, and so the program has always functioned according to the goals table

designed for the original grant. The table has changed with each subsequent grant cycle, but the

grant writers did not produce a logic model. Performance indicators developed for the goals

table during the proposal stage still guide data collection practices today.

In addition, the federal agency that administers the grant requires all contractors to track

eight GPRA measures. In recent years, the GPRA measures have become increasingly

important, to the point where they are now the federal agency's major focus. According to the

program director and executive director, they are also the only common outcomes that Program

A grantees track across the country; each grantee also carries other activities with various

indicators, creating a wide variety in nature of work associated with Program A (personal

interview, March 24, 2010).

The organization developed Program A's current database in 2007 after adding two new

staff members, including a new director for the program and an administrative staff member with

extensive IT knowledge. Before this time, staff collected data largely by capturing numbers in

their daily calendars and entering it into a spreadsheet once a year for the grant's reporting cycle.

With the combination of new staff members and several new volunteers carrying out activities

that must tracked, the organization decided to computerize the system. An internal committee

developed and implemented the new system over a period of two months with the help of a

database consultant. Today staff members use an Access database that can integrate or separate

the two major components of Program A (see Figure 3), and a data collection form they fill out

every time an activity takes place. Staff members enter data on an ongoing basis so that the

program's progress on any indicator can be determined at any point in the year merely by

clicking a button. The database also ties indicators back to the original grant objectives so that

the program director can monitor whether Program A's activities are fulfilling the larger goals of

the organization.
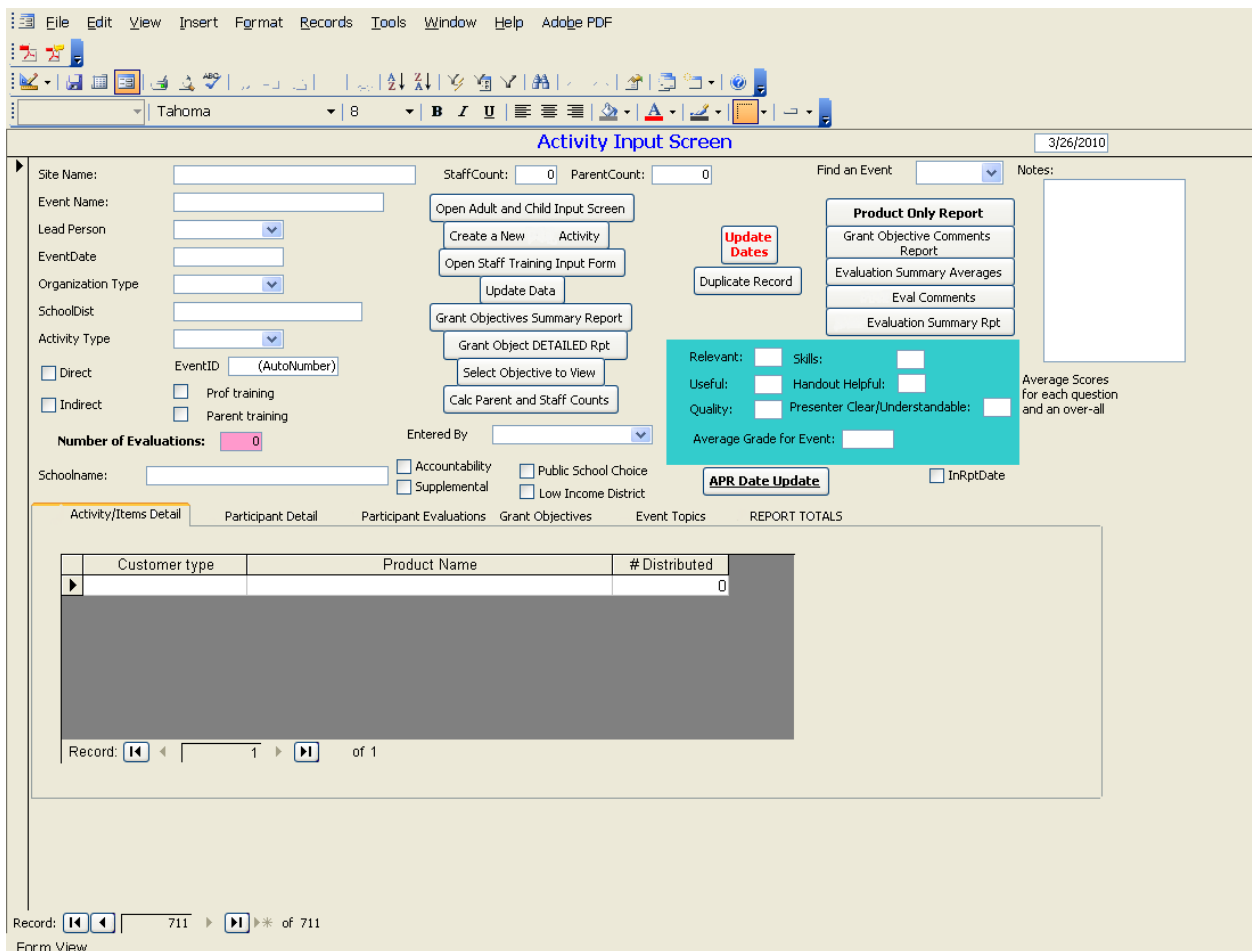
Figure 3. Screen shot of program database.

Comparative measurement. The organization also supports standardizing data

collection in the field as a way to provide stronger programs and achieve greater impact. It has

focused these efforts in the field of early childhood home visitation. Home visitation programs

involve trained staff members visiting families with young children in their homes at regular

intervals. They generally focus on supporting parents by teaching them activities and techniques

that build children's cognitive abilities, conducting developmental screenings for children, and

providing community resource networks to address any other needs the families may have

(Ramey & Ramey, 1998).

Research on home visitation programs has been uneven, and the findings are mixed

(Ramey & Ramey, 1998). Partly because of this, the organization is a founding member of a

comparative performance initiative for home visiting programs in the state. Several

organizations came together in 2001 to develop a common set of outcomes, a unified data

collection method, and one reporting format. Practitioners from all partner organizations enter

data into the state's public health data collection system, called SPHERE, so that other service

providers have access to and can act upon the information (Early Years Home Visitation

Outcomes Project of Wisconsin, 2009). The organization's executive director believes that

common outcomes such as these can help the nonprofit sector by providing the necessary data

for improved programs and proof of impact. She says, "If you're not having a positive impact,

what's the point?" (personal interview, March 24, 2010).

Needs. The organization is already forward-looking when it comes to evaluation and

measurement, as evidenced by its mature measurement system and participation in comparative

performance. Even with acceptance and use of these tools however, there are still questions that

go unanswered. Some of these questions look more towards outcomes, including whether

training participants make use of training material in their own work and home environments

after their sessions have finished. Other needs relate more to the management side and include

such issues as client satisfaction and trainer evaluation. For Program A in particular, the

organization deals with changing emphasis from its funders and the need for indicators that track more than just grant requirements.

In addition, the organization wants to begin tracking the source of its referrals as it continues a period of sustained growth. A good proportion of the organization's clients contact them because past training participants recommended it or because they attended a conference featuring the organization's presentations. The organization spends almost no money on marketing, communications or fund development, so this type of word of mouth is one of the only ways it maintains its visibility in the community. In order to maintain high quality word-of-mouth referrals, the organization needs to know whether current trainees are recommending its work to colleagues and friends, or whether its training program fails to meet client expectations.

Interview

The present researcher interviewed the organization's executive director, program director, and information technology manager about performance measurement and data collection processes. Because the organization operates a federal grant with GPRA measures, recently switched to a computerized data collection process, and has extensive experience with evaluation and performance measurement, its experience over the last few years in this area follows many of the trends noted in the literature. The purpose of the interview was to document and explore each participant's decisions and participation in the process, use of data today, and her attitudes towards performance measurement in the nonprofit sector. The present researcher followed the interview protocol, and added additional questions according to the answers given by each participant. Answers were then analyzed through the lens of the trends noted in the literature as well as each participant's role and perspective.

The participants agree that the computerized tracking system is far superior to the old method of tracking data, which was not very reliable. They cite the ease with which they can produce high quality data at any point in the grant cycle. As opposed to spending days working together to account for all activities tracked by hand, the IT director can now gather all the data for the yearly report "in about half an hour."

In addition, reports group the data according to individual indicators or in aggregate, allowing the program director to reallocate resources to fulfill unmet grant requirements. The IT director says the new system allows them to "see if we're hitting our target." She adds that staff members often express surprise at the amount of work completed and information disseminated in the reports, suggesting that the old method of tracking may have dramatically underreported the amount of work completed. When the reports look suspicious, the IT director can produce more in-depth sets of data that that show every piece of work completed by each staff member throughout the reporting period, going back to 2007 when the system was implemented. This either reveals errors in data entry or collection, or reaffirms the amount of work completed.

The reasons for moving to the new system vary according to participant. The program director and IT director were both new staff members at the time, and both had worked with computerized systems in the past. They both pushed strongly for a database after going through the hand counting process early in their tenure. The executive director cites several reasons for the switch, including the organization's growth, both in staff members and in reporting requirements, the importance of improving data quality, and the federal agency's increased focus on GPRA measures.

The organization considered several ways to revamp its performance measurement system in 2006, including contracting with a private entity to create an entirely new system of

indicators and a database. This option was prohibitively expensive, despite the availability of

several contractors with expertise in this area. At the time, the executive director had no

knowledge of shared measurement systems to purchase and tailor to Program A, and indeed few,

if any, existed at that time. The organization also considered a multi-site system, with each of

Program A's partner sites entering its own data. The executive director felt this option would

mean the loss of control over data quality, and did not pursue it further. She appreciates the

organization's ability to test the data and provide continuous training to data entry staff to ensure

the process maintains the highest quality.

In the end, the new staff members were able to use the organization's existing technology

to design a system they can monitor easily at a relatively low cost. While a computerized system

is considerably more expensive than tracking a few indicators by hand, the increased usability of

the data justifies the costs. Program A's grant does not specify costs for this method, so the

organization was able to make the change without getting permission from its funder, thus

bypassing one of the larger hindrances for many organizations considering such investments.

Respondents were glad they implemented the system when they did because the federal

agency's reporting requirements have become more specific in each of the last three years, partly

due to a greater reliance on GPRA measures and partly due to changes in administration. Instead

of a completely smooth reporting process, staff members have had to comb through the data to

track indicators that received little or no attention in the past. With the database, they have

access to the information being requested, which might not be the case otherwise. Likewise, the

program manager plans to continue adding new indicators in anticipation of further changes in

reporting requirements.

Respondents cited various uses for the data according to their roles in the organization. The IT director is responsible for maintaining, testing and adding data to the system, as well as generating reports. She spends time working with field staff to tweak the system so that data collection tools ask exactly the right questions and the system reports accurate results. The executive director uses the system to monitor contract compliance with Program A's partners, to secure future funding based on results, to report to the Board of Directors, and for marketing purposes. The program director uses the data mainly for reporting purposes and finds it limited for decision-making.

Both the executive director and program director agree that the restrictive nature of the grant precludes them from using the data to make changes in staffing or program activities. They see these types of decisions as largely controlled by funding and budgets set for the five-year grant cycle, rather than by needs suggested through ongoing data collection. The performance measurement process more or less breaks down at this point, when the results cannot be tied back to program changes that affect the budget. This may be a factor of the long grant cycle in which detailed budgets are set five years in advance, or more generally of the federal agency's handling of the grant.

Design and Implementation of Measurement Tools

Logic model. The organization collects some information based on a set of goals designated in the grant proposal, but never developed a full model that includes expected outcomes over the short, medium, and long term. The present researcher developed such a model (see Appendix A) from the working documents used by staff members and the external evaluator. Based on these working documents, the headings are somewhat different from

standard form, but designed to provide easy cross references between grant requirements and current database functions. The headings include Assumptions, Inputs, Goals, Process/Activities, Indicators, and Outcomes (Short, Medium and Long Term).

According to Redstone Strategy Group (2008), which documented the implementation of outcome-based programming for one of the Hewlett Foundation's issue areas, the logic model should be explicit enough to adequately describe the program's theory of change and to suggest indicators by which to measure its success. Those indicators may be difficult or impossible to measure by the particular organization or for the particular population, but they should still be included in the discussion. Program designers use substitute indicators that get as close to the ideal measurement as possible; these are called proxy measures (Redstone Strategy Group, 2008). In addition, developers should recognize that very explicit logic models could lead to an over-simplification of the problem and its solutions. This reasoning, based in practical experience, informed the development of the organization's model.

Indicators. The Urban Institute and The Center for What Works' (2006) research into outcomes development offers several useful suggestions. First, researchers suggest starting with a small number of indicators, and adding more as staff members become more comfortable and adept in the measurement process. Key in the development of indicators is choosing those that are intrinsic to the mission or program outcome, rather than those that are easiest to track.

The organization currently tracks such information as the lead presenter, date and type of training; the number, type and organizational affiliation of participants, amount and type of information distributed, and the individual grant objectives achieved by each activity. In addition, participants fill out surveys immediately after the training rating its quality and suggesting improvements.

Based on discussions with staff members, the present researcher developed an additional set of indicators to measure, including:

- *Productivity measure:* the number of trainings held per month, graphed over time.

- *Output measures:* the number of parents and staff members trained.

- *Effectiveness measures:* the percentage of respondents who state that, because of the training, they improved their knowledge of the subject material, increased their ability to perform their jobs, improved their parenting skills, and/or improved their ability to advocate for a child.

- *Customer satisfaction measures:* satisfaction with the presenter three months after the training, and the percent of respondents satisfied with the handouts and with the training. These measures contain both new and old data that were previously aggregated and not available as a separate report.

- *Word of mouth:* the percentage of respondents who used or passed on the information they received, the number of respondents who scheduled a follow up training, and the percentage of respondents who stated that they would recommend the organization to others.

- *Service quality measures.* The present researcher added several new measures and changed others in this category. New measures include three-month surveys tracking participants' satisfaction with the trainers' knowledge and presentation of the material, the thoroughness of the training, and whether a translator was or should have been provided. The organization previously aggregated service quality data into an average performance score for each training. The present researcher broke the data out into percentages of participants who expressed levels of satisfaction in six unique categories in the survey

given immediately after trainings. In addition, comments and suggestions were

categorized according to type and provided as a separate breakout. See Table 1 for results.

- *Efficiency measure*: the amount of money spent on trainings per month. The present

  researcher discussed this commonly used measure with the program director, who

  ultimately decided not to include it. The organization provides trainings free of cost and

  beyond the numbers required by its grant. Likewise, knowing the costs per month would

  not affect any decisions regarding staffing or program activity structure.

Survey. The present researcher designed a brief survey with staff input to capture data

that was not already being collected. The survey was sent to everyone who participated in a

training from September 1, 2009 through November 30, 2009. This period is typically a busy

time for staff, as it covers the beginning of the school year. The total population of 158 training

attendees, 131 women and 17 men, was invited to participate, including parents, school staff, and

early childhood staff. Twenty-two people participated in the survey.

Methodology. The survey design followed the guidelines set by Newcomer and Triplett

(2004), who recommend testing all clients to avoid sampling concerns, collecting data on a

continual basis, and producing a professional survey instrument to increase the likelihood of

response, especially by former clients. With the rising prevalence of internet communications,

Newcomer and Triplett recommend that nonprofits incorporate email or web-based surveys

where possible.

With these factors in mind and working with program staff, the present researcher pilot-

tested the survey in two delivery formats, online and through the mail. To increase response

rates, the letter assured participants that the organization's staff would not see their responses

and offered a small incentive, a five-dollar gift certificate for completing the survey. The

executive director signed the cover letter to increase its importance level (see Appendix B).

Trainers informed participants of the organization's grant requirements as they signed in to the

training, and told them they would be contacted for a follow up survey within three months.

Approximately three months after their training, participants received up to two invitations

to take the survey based on the contact information they gave staff at the training, either email or

mail addresses. With mail surveys, the present researcher sent follow-up requests 10 days after

the initial mailings. Email follow-ups arrived between five and ten days after the initial request.

Results. Twenty-two participants took the survey of medium-term outcomes for trainings,

with twenty responding by email and two responding by mail. Of the participants, 13 were staff

members and nine were parents. Twenty-one were female and one was male. While the parent

to staff ratio was roughly equivalent, females were overrepresented in the sample.

Results were categorized according to the indicators previously set, and can be seen in

full in Appendix F. Responses overall were positive. Ninety-five percent reported increased

knowledge in the subject matter after the training; 91% increased their ability to advocate for a

child, defined as defending or supporting a cause or a child's interests. All respondents were

satisfied with the trainers' knowledge and presentation of the material, and all stated that they

would recommend the organization to others. Respondents also gave useful suggestions for

improving future trainings.

Limitations. There are several limitations to the pilot test the organization must consider

as it moves forward. It is notable that for three trainings, held for parents at schools, no

participants from the particular training responded to the email. This accounts for 28 people, or

17% of the total population. The most likely reason for this lack of response is that these three

trainings were the first in which trainers informed participants about the survey, and there may have been some confusion in the messaging.

With a 14% response rate, the data has a high probability of nonresponse bias, a skewing of the data caused by a low response rate. Because of this, it is hard to know whether the results represent a true baseline for the organization to measure against in the future. Staff members and the current researcher believe that the close nature of the early childhood industry in Milwaukee may have prevented some training participants from responding due to concerns about evaluating colleagues. There is also the possibility that survey respondents elected to participate because they had something positive to say. Without more respondents, it is impossible to know for certain that the views are representative of all trainees.

Despite the limitations, however, the survey represents the first attempt to gather data beyond short-term outcomes, a move that is encouraged in the literature and by leading organizations such as the United Way. Many nonprofits have a difficult time maintaining contact with former clients and must rely on the data they have, keeping in mind its limited value. Rather than a strong baseline, the data form a snapshot that gives managers a hint about what questions to ask in the future. Forthcoming attempts to survey the population can include adjustments designed to increase the response rate, such as sending more reminders, adding reminders to the training packets that participants receive, and asking for two contact methods in case one fails (Newcomer, 2004).

Discussion

While the data from the survey alone cannot be used to establish a baseline, it can be used in combination with other data to help the organization better understand how training

participants perceive it. By adding new data, reorganizing old data, and putting collected

information together into categories, the organization's management staff found several new

tools for decision-making. Useful breakouts include suggested improvements for future trainings

(Table 1), and word of mouth tracking (Table 2).

Table 1

Breakout of Suggested Improvements by Type

Suggestion type	Number requesting	
	Parents	Staff
More in-depth information	3	4
More information	3	2
More professionally presented	0	5
Longer presentation ("more time")	7	8
Shorter presentation	0	1
Improvements to handouts/materials	2	2
Interpreter requested	4	
Administrative improvements (including presenters' phone numbers, scheduling breaks, bringing food, handling participation and interest levels, more or less time given to certain activities)	4	7

Managers had been considering potential improvements to the training curricula in their

effort to remain relevant and useful to the target community. The table combines data from

surveys taken immediately after trainings with data from the three-month interval survey. The

first survey asks generally for suggestions regarding the training. The three-month interval

survey asks whether and in what ways respondents used the training's material in the last three

months. By collating the data, three major themes become apparent: a desire for more time, and

better handling of the schedule so that participants feel no time is wasted, and a desire by

educators to have a more professional presentation. Another breakout lists specific suggestions

for additional subjects to cover (See Appendix F).

The many requests for more time led to a discussion among training staff and managers

as to how to accommodate the request. The realities of scheduling for participants, hosts, and

trainers often limit sessions to one hour. However, trainers may be able to offer a series of

sessions to particular groups when initial surveys come back with several requests for "more

time."

The organization is also beginning to look at marketing and communication efforts to

increase growth and information dissemination. Managers suspect that the major reason the

organization gets new clients is word of mouth from conferences and trainings, but had no data

to confirm this. The three-month interval survey asked participants whether they would

recommend the organization to others and why. Table 2 shows that 100% of participants would

recommend the organization, and that three respondents, or 13.6%, requested a follow up as part

of their comments. This suggests that managers may be correct in their assumptions. The

present researcher also designed a phone survey to use with people who call the organization

requesting materials or trainings to increase the amount of data on this particular question (See

Appendix G). With this information, managers can look at ways to increase word of mouth,

such as passing out business cards to every participant, improving the quality of certain

materials, or including information on other trainings in handouts.

Table 2

Word of Mouth Breakout

Participants who:	Percentage	Number
Passed materials to others	50	11
Requested a follow up training	13.6	3
Would recommend organization to others	100	22
Reasons	Percentage	Number
Presentation gave good information	32	7
Presentation is good for parents	41	9
Presentation is good for staff	9	2
I learned something I didn't know	9	2

Another useful breakout was the number of people trained by type. This number previously was part of an aggregate, or reduced to number of trainings by type. By giving this information its own indicator, managers were able to use this in combination with information from Table 1 to highlight an issue they suspected but had no data on. The program model lends itself more easily to training staff than to training parents, as evidenced by the number of staff members trained. At the same time, most trainings are tailored to parents, as evidenced in the word of mouth breakout and the staff requests for a more professional presentation. After viewing the information in this light, the program manager considered ways to modify the program implementation, such as designing a second set of trainings that offer the same subjects with more depth and practical tips for service providers. Clients could then choose which method would better appeal to their particular group when scheduling a training with organization staff.

Conclusion

Performance measurement and evaluation remain fluid concepts as the nonprofit sector continues to embrace new tools and needs for data. With many options available to individual organizations, it is useful to explore their relative value and potential, especially new and innovative ideas such as adaptive learning systems and other forms of comparative measurement. With a better knowledge of the systems that exist today, organizations can make informed decisions about their own measurement needs.

It is also useful to explore how organizations carry out the work of measuring performance, especially given the changing nature of the sector. This process is almost always messier than the best practice scenario and involves real challenges created by funders, inexperienced staff members, inadequate resources, reluctant clients, and changing standards over time. This project studied a nonprofit's efforts to revamp its performance measurement system and aided in that process. The result is a useful framework for the program itself, a set of additional tools by which to track valuable indicators, and a larger pool of information to guide the organization in implementation.

The areas of comparative and adaptive measurement offer great potential for future research. More case studies are needed on successful attempts at these breakthrough systems, along with more data on their potential use for increasing the flow of information throughout the nonprofit sector, not only about individual organizational performance, but also on increasing impact on entrenched social problems. With this kind of information, organizations will be empowered to replicate or adapt successful systems to their own communities' needs. When this happens, the anticipated shift towards common outcomes measurement will be more likely to occur.

References

Ailworth, E. (2010, March 31). Firm's ratings give donors new tool in picking worthy charities. *The Boston Globe.* Retrieved from http://www.boston.com/

Berger, K. & Penna, R.M. (2010, March 31). Charity Navigator responds. Message posted to The Hauser Center for Nonprofit Organizations at Harvard University Viewpoint blog: http://hausercenter.org/iha/archives/338

Blum, D. (2006, October 12). Checking the dashboard. *The Chronicle of Philanthropy, 19*(1), 15. Retrieved from http://philanthropy.com/section/Home/172

Blum, D. E., & Williams, G. (2008, May 15). Putting charity to the test. *Chronicle of Philanthropy, 20*(15), 13. Retrieved from http://philanthropy.com/section/Home/172

Brody, E. (2002). Accountability and public trust. In L. Salamon (Ed.), *The state of nonprofit America* (pp. 471-498). Washington, D.C.: Brookings Institution Press.

Carman, J. G. (2009). Nonprofits, funders, and evaluation: Accountability in action. *The American Review of Public Administration, 39*, 374-390. doi: 10.1177/0275074008320190

Chan, Y. L. (2004). Performance measurement and adoption of balanced scorecards: A survey of municipal governments in the USA and Canada. *International Journal of Public Sector Management, 17*(2/3), 204-221. doi: 10.1108/09513550410530144

Conlin, M., Hempel, J., Polek, D., & Dayton, T. (2003, December 1). The top givers: Today's philanthropists aren't leaving the good works to future generations - they're making their mark now. *Business Week* (3860), 78-84. Retrieved from http://www.businessweek.com/

Cunningham, K., & Ricks, M. (2004). Why measure? *Stanford Social Innovation Review, 2*(1), 44-51. Retrieved from http://www.ssireview.org/

DiMento, M., & Preston, C. (2010, February 7). The Philanthropy 50: A slow year for big gifts

　　spurs creativity by wealthy donors. *The Chronicle of Philanthropy.* Retrieved from

　　http://philanthropy.com/section/Home/172

Early Years Home Visitation Outcomes Project of Wisconsin (2009). *2008 Outcome Attainment*

　　Report Executive Summary. Retrieved from Children's Hospital of Wisconsin website:

　　http://www.chw.org/display/PPF/DocID/45172/router.asp

Foundations, states working to significantly increase college completion (2010, March 3).

　　Philanthropy News Digest. Retrieved from

　　http://foundationcenter.org/pnd/news/story.jhtml?id=286400014

Frederickson, H.G. (2001). First there's theory, then there's practice. *Foundation News and*

　　Commentary, 42(2), 37-41.

Giving USA Foundation (2009). *Giving USA 2009: The annual report on philanthropy for the*

　　year 2008. Indianapolis: Author.

Gregory, A.G., & Howard, D. (2010). The nonprofit starvation cycle. *Stanford Social Innovation*

　　Review, 7(4), 49-53.

Hatry, H. (1999). *Performance measurement: Getting results.* Washington, D.C.: Urban Institute

　　Press.

Henderson, F. (2009, May 12). Strive Education makes a giant leap into national spotlight.

　　Cincinnati Soapbox. Retrieved from http://soapboxmedia.com/

Hopkins, B.R. (2005). *Nonprofit law made easy.* Hoboken, New Jersey: John Wiley & Sons.

Ittner, C.D. & Larcker, D.F. (1998). Innovations in performance measurement: Trends and

　　research implications. *Journal of Management Accounting Research, 10*, pp. 205-238.

Keehley, P. & Abercrombie, N. (2008). *Benchmarking in the public and nonprofit sectors: Best practices for achieving performance breakthroughs.* San Francisco: Jossey-Bass.

Kramer, M., Graves, R., Hirschhorn, J., & Fiske, L. (2007). *From insight to action: New directions in foundation evaluation.* Retrieved from FSG Social Impact Advisors website: http://fsg-impact.org/app/content/actions/item/177

Kramer, M., Parkhurst, M., Vaidyanathan, L. (2009). *Breakthroughs in shared measurement and social impact.* FSG Social Impact Advisors. Retrieved from Hewlett Foundation website: http://www.hewlett.org/library

Lampkin, L.M., & Hatry, H. (2009). Measuring the nonprofit bottom line. In J. Cordes & C.E. Steuerle (Eds.), *Nonprofits & Business* (pp. 239 -269). Washington, D.C.: The Urban Institute Press.

Lenkowsky, L. (2007, August 9). Evaluations can be dangerous. *Chronicle of Philanthropy, 19*(20), 35. Retrieved from http://philanthropy.com/section/Home/172

Letts, C., Ryan, W., & Grossman, A. (1997). Virtuous capital: What foundations can learn from venture capitalists. *Harvard Business Review, 75*(2), 36-44.

Light, P. (2000). *Making nonprofits work: A report on the tides of nonprofit management reform.* Washington, D.C.: Brookings Institution Press.

McDavid, J.C., & Hawthorn, L.R.L. (2006). Performance measurement as an approach to evaluation. In *Program evaluation & performance measurement: An introduction to practice* (pp. 281-304)*.* Thousand Oaks, CA: Sage.

McKinsey & Company. (2008). *The nonprofit marketplace: Bridging the information gap in philanthropy*. Retrieved from the Hewlett Foundation website: http://www.hewlett.org/library?page=2

McLaughlin, T.A. (2009). *Streetsmart financial basics for nonprofit managers.* Hoboken, New Jersey: Jon Wiley & Sons.

McLean, C. & Coffman, S.E. (2004, June). *Why ratios aren't the last word.* Retrieved from GuideStar website: http://www2.guidestar.org/rxa/news/articles/2004/why-ratios-arent-the-last-word.aspx

Meehan, W., Kilmer, D., & O'Flanagan, M. (2004). Investing in society. *Stanford Social Innovation Review, 1*(4), 34-44. Retrieved from http://www.ssireview.org/

National Council of Nonprofit Associations & National Human Services Assembly. (2005). *Rating the raters: An assessment of organizations and publications that rate/rank charitable nonprofit organizations.* Retrieved from National Human Services Assembly website:

http://www.nassembly.org/nassembly/documents/Rating_the_Raters_Final_3%20.pdf

Newcomer, K. and Triplett, T. (2004). Using surveys. In J.S. Wholey, H. Hatry & K. Newcomer (Eds.), *Handbook of practical program evaluation* (2nd ed., pp. 257-291). San Francisco: Jossey-Bass.

Olasky, M. (1992). *The Tragedy of American Compassion.* Washington, D.C.: Regnery.

Plantz, M.C., Greenway, M.T., & Hendricks, M. (1997). *Outcomes measurement: Showing results in the nonprofit sector.* Retrieved from United Way website:

http://www.liveunited.org/Outcomes/Library/ndpaper.cfm

Poister, T. (2003). *Measuring Performance in Public and Nonprofit Organizations.* San Francisco: Jossey-Bass.

Poister, T. (2004). Performance Monitoring. In J.S. Wholey, H. Hatry & K. Newcomer (Eds.), *Handbook of practical program evaluation* (2nd ed., pp. 98-125). San Francisco: John Wiley & Sons.

Preston, C. (2008, November 13). Making a measurable difference. *Chronicle of Philanthropy, 21*(3), 4. Retrieved from http://philanthropy.com/section/Home/172

Radin, B. (2006). *Challenging the performance movement: Accountability, complexity and democratic values.* Washington, D.C.: Georgetown University Press.

Ramey, C., & Ramey, S. (1998). Early intervention and early experience. *American Psychologist*, *53*, pp. 109–120. doi: 10.1037/0003-066X.53.2.109.

Redstone Strategy Group, LLC. (2008). *Making an impact: Formalizing outcome-driven grantmaking: Lessons from the Hewlett population program.* Retrieved from Redstone Strategy Group website:

http://redstonestrategy.com/documents/Making%20an%20Impact.pdf

Root Cause. (2010a). *Investing for social impact: A guide to the Social Impact Research methodology.* Retrieved from Social Impact Research website:

http://www.socialimpactresearch.org/sites/socialimpactresearch.org/files/SIR%20method ology%20FINAL.pdf

Root Cause. (2010b). *Social issue report: School Readiness.* Retrieved from Social Impact Research website: http://www.socialimpactresearch.org/sites/socialimpactresearch.org/ files/School-Readiness-Social-Issue-Report-Final.pdf

Ryan, W. (1999). The new landscape for nonprofits. *Harvard Business Review*, *77*(1), 128-136.

Salamon, L. (2003). *The resilient sector: The state of nonprofit America*. Washington, D.C.: Brookings Institution Press.

Strive. (2009). *Striving Together: Report Card.* Retrieved from Strive website:

http://www.strivetogether.org/documents/ReportCard/2009StriveReportCard.pdf

Strom, S. (2003, November 17). New equation for charities: More money, less oversight. *The New York Times*. Retrieved from http://www.nytimes.com

Thiel, S. & Leeuw, F.L. (2002). The performance paradox in the public sector. *Public Performance & Management Review, 25,* 267-281.

W.K. Kellogg Foundation. (2004a). *Evaluation Handbook.* Retrieved from W.K. Kellogg Foundation website: http://www.wkkf.org/knowledge-center/Resources-Page.aspx

W.K. Kellogg Foundation. (2004b). *Logic Model Development Guide*. Retrieved from W.K. Kellogg Foundation website: http://www.wkkf.org/knowledge-center/Resources-Page.aspx

Waide, K. (2002, April 15). *Evaluating charities: Why the numbers count.* Retrieved from Charity Navigator website:

http://www.charitynavigator.org/index.cfm/bay/content.view/catid/68/cpid/65.htm

Urban Institute and The Center for What Works. (2006). *Outcome indicators project; Building a common outcome framework to measure nonprofit performance.* Retrieved from Urban Institute website: http://www.urban.org/url.cfm?ID=411404

Assumptions	**Inputs**	**Objectives**	**Activities/Processes**

Assumptions

-Integrating intervention efforts leads to greater access.
-Coordination at the state level can increase impact, reduce duplication, standardize programs, and ensure that all who need assistance receive it.
-Home visitation and parental engagement have proven track records in increasing childhood academic success.
-Positive parenting practices increase children's cognitive abilities and school readiness.
-Improved literacy and language skills lead to improved academic skills.
-Education and early childhood staff often serve as the main point of contact with parents.
-Parents who feel welcome in their children's school will become more involved.
-Parental engagement is strongly correlated with children's success in school.
-Training provided to parent and community leaders will be disseminated to others.

External Factors

-Low income children often attend schools considered to be of lower quality.
-Low income parents often face many barriers to becoming involved in their children's education, including work conflicts, poor interactions with school staff, and distress over their own school failures.
-Early childhood and school staff and school districts often are not equipped with the tools to engage low income parents.

Inputs

Collaboration with 8 WI agencies with extensive knowledge of the target population, early childhood intervention, and parent engagement.

Nationally recognized, research-based models for early childhood development and parental engagement: Parents as Teachers (PAT) and Family-School-Community Partnerships.

Fifteen years of experience in successfully implementing programs for the target population.

Experienced trainers who are certified in the relevant training curricula.

Objectives

1. Advocacy and Public Policy

Coordinate and expand program collaboration with the Wisconsin Department of Public Instruction (Including Title I and, No Child Left Behind) and other key public and non-profit organizations that promote early childhood education, improved student learning and parental/family involvement in urban and rural areas throughout.

2. Parental Engagement

Develop and strengthen the relationship and partnerships between schools and parents/families as a means to improve children's academic achievement.

Activities/Processes

Maintain a membership in 9 key statewide committees. A1

Give 10 presentations at national and statewide conferences to increase knowledge and enlist support of key educators and parent advocates. A2

Participate in conference planning committees for 4 key statewide conferences. A3

Conduct 35 workshops for low income and limited English parents. B1

Disseminate 30,000 pieces of materials to target parents to educate and inform them of their rights on school choice and supplemental services. B2

Provide 4 training workshops for MPS parent coordinators, other leaders. B3a

Provide 2 UW Milwaukee certification programs for MPS, parent leaders. B3b

Provide 11 workshops for tribes on parent involvement, special needs. B4

Increase visits to PPI website by 20% annually through site restructure. B5

Give out 5,000 copies annually of Tips on Talking to School Staff booklet. B6

Outputs/Indicators	Outcomes— 1. Public Policy and 2. Parent Engagement		
	Short Term	**Medium Term**	**Long Term**
Number of state committees org. serves on.	80% of conference attendees increase their knowledge of the subject matter.	30% of new statewide early childhood education initiatives include a home visiting component.	State government coordinates and funds high quality early childhood programs, especially for at-risk families.
Percent of conference attendees who find a presentation relevant.	70% of conference attendees believe that presentation material is relevant.	Home visiting is promoted by leading institutions as a respected program delivery model.	Urban and rural families have access to high quality early childhood programs designed to increase future academic success.
Percent of conference attendees who increase knowledge of subject matter.	10% of conference attendees either schedule follow up trainings or pass material along to someone who does.	30% of new statewide education initiatives include a parental engagement component.	
Number of requests for information on parent engagement.	80% of workshop attendees believe that presentation material is relevant.	Early childhood and education leaders express support for good early childhood education as a key to academic success.	An increased number of "family friendly" schools.
Number of requests for information on home visitation.	80% of workshop attendees increase their knowledge of the subject matter.	Early childhood and education providers express support for parental engagement as a key to academic success.	An increased number of children achieve academic success.
Number of home visitation programs operating in the state.			
Number of workshops provided for target population.	70% of workshop attendees increase their ability to perform their jobs.		
Percent of parents who report an increase in their parenting skills.	60% of professionals attending workshops use the organization as a resource at their schools.	70% of parents attending workshops report an increase in their parenting skills.	
Percent of service providers who report an increase in their ability to perform their jobs.	10% of workshop attendees either schedule follow up trainings or pass material along to someone who does.	70% of workshop attendees report an increase in their ability to advocate for a child.	
Number of informational materials disseminated throughout the state.	70% of parents report the Tips booklet informs them of their rights as parents.	Yearly, an increased number of parents enroll their children in SES services for low-performing schools.	

Assumptions	Inputs (cont.)	Objectives (cont.)	Activities/Processes (cont.)

Assumptions

-Intervention efforts are more effective when they start earlier in life and address multiple risk factors.
-Improved school readiness at kindergarten entry leads to increased academic achievement over the long term.
-Children with developmental delays who receive intervention early can reduce the delays and/or their impact.
-Positive parenting practices, including reading with children, increase children's cognitive abilities and school readiness.
-Improved literacy and language skills lead to improved academic skills.
-Education and early childhood staff often serve as the main point of contact with parents.
-Parents who feel welcome in their children's school will become more involved.
-Parental engagement is strongly correlated with children's success in school.
-Low income parents may need alternative forms of access to schools due to work schedules and other inhibiting factors.

External Factors

-Low income children often experience less stability in their home lives, which affects academic achievement.
-Children from low income families have less exposure to the tools that prepare them for education, including language development, reading, organizational and social skills.

Inputs (cont.)

Partnership with state education agency, DPI, to provide full-time volunteers in individual Milwaukee schools, along with other resources.

Five training staff members with extensive experience and multiple certifications in partnership models, early childhood, ethics and other training subjects.

Partnership with state college to provide trainings through an extended network.

Nationally recognized program model for engaging parents in schools.

Evidence-based and standardized training courses.

Objectives (cont.)

3. School-Family-Community Partnerships

Assist schools to design and implement effective parental involvement policies, programs, and activities that improve children's academic achievement.

Activities/Processes (cont.)

Present 12 Parent Involvement: Policy to Practice workshops to school staff.

Conduct at least 3 parent involvement self-assessment workshops for elementary school educators. C1

Plan and present annual Family-School-Community Partnerships conference with DPI. C2

Train 8-10 VISTA volunteers annually to assist schools and organizations in implementing family-school-community partnerships. C3

Publicize and attend 20 NCLB Supplemental Services enrollment and provider fairs for Milwaukee parents. C4

Provide 10 workshops on understanding poverty for school and community organization staff. C5

Provide 24 Parents with Voices trainings to target parents in schools and community organizations. C6

Distribute information packets annually to every WI school district on org. services, NCLB information and parent involvement strategies. C7

Outputs/Indicators	Outcomes— 3. School-Family-Community Partnerships		
	Short Term	**Medium Term**	**Long Term**
Number of workshops held for school staff.	80% of conference attendees increase their knowledge of the subject matter.		
Number of people attending signature partnerships conference.	70% of conference attendees believe that presentation material is relevant.	All Milwaukee schools join the National Network for Partnership Schools.	
Percent of conference attendees who increase knowledge of subject matter.	10% of conference attendees either schedule follow up trainings or pass material along to someone who does.		All school districts identified for improvement under NCLB integrate effective parent engagement into their infrastructure.
Number of VISTA volunteers trained.			
Number of SES events attended.	Annually, 3-5 additional schools join the National Network of Partnership Schools.	30% of eligible children attend SES in Milwaukee.	Urban and rural families seek supplemental education services designed to increase academic success.
Number of materials given to parents on SES.			
Percentage of children enrolled in SES vs. the percentage eligible.	Annually, 3% increase in the number of eligible children enrolled in SES.	40% of workshop attendees (staff) report implementing new strategies for increasing parental involvement.	
Number of poverty workshops provided.	70% of workshop attendees (staff) learn new strategies to use with parents.		
Number of parent empowerment workshops provided.		40% of workshop attendees (parents) report using new strategies in their parenting.	An increased number of parents report positive interactions with their children's school(s).
Percent of parents who report an increase in their advocacy skills.	70% of workshop attendees (parents) learn new strategies in working with school staff.		
Percent of staff who report an increase in advocacy skills.	Annual increase in number of schools or districts outside Milwaukee who contact org for services.	Increase services annually to schools and districts outside of Milwaukee through the use of Internet technology.	An increased number of children achieve academic success.
Number of school districts who follow up on informational packets.			

Assumptions

-Intervention efforts are more effective when they start earlier in life and address multiple risk factors.
-Improved school readiness at kindergarten entry leads to increased academic achievement over the long term.
-Successful early intervention programs must address several intertwining risk factors.
-Children with developmental delays who receive intervention early can reduce the delays and/or their impact.
-Positive parenting practices, including reading with children, increase children's cognitive abilities and school readiness.
-Improved literacy and language skills lead to improved academic skills.
-Parents who receive training in positive parenting practices on a long term basis are more likely to implement them.

External Factors

-Low income children often experience less stability in their home lives, which affects academic achievement.
-Children from low income families have less exposure to tools that prepare them for education, including language development, reading, organizational and social skills.
-Children in the Milwaukee school system experience higher levels of mobility than other school systems.
-There are often birth disparities between low-income and middle class children, including low birth weight, lack of maternal nutrition and prenatal care, and premature birth.

Inputs (cont.)

Evidence-based home visiting model that has been shown to reduce child abuse and increase school achievement.

More than a decade of experience as the State Office for Parents as Teachers in Wisconsin.

Training staff members who are also experienced and current parent educators.

Partnership with four Milwaukee area agencies to provide intensive combined services to high-need clients.

Ongoing longitudinal evaluation of program impact for Milwaukee area clients.

Objectives (cont.)

4. Early Childhood Parent Education

Further the developmental progress of Wisconsin's children by establishing, expanding or operating Parents as Teachers (PAT) Home Visiting services.

Activities/Processes (cont.)

Train and assist 33 PAT staff in collaboration with the Milwaukee Comprehensive Home Visiting Project. D1

Provide 4 annual professional development opportunities for home visitors, child care providers and their supervisors. D2

Disseminate the new PAT quality standard report to 35 PAT programs. D3

Provide relevant information through monthly e-newsletters, PPI website. D4

Convene 2 annual meetings of the State Advisory Committee. D5

As State Office for PAT, provide 4 annual PAT trainings. D6

Fund PAT services for at least 150 WI families through Family Resource Center Network. D7

Provide Parents Are Key program to at least 60 Milwaukee families annually. D8

Outputs/Indicators	Outcomes— 4. Early Childhood Parent Education		
	Short Term	**Medium Term**	**Long Term**

Outputs/Indicators:

- Number of PAT programs operating in the state.

- Yearly, number of PAT trainings provided in Wisconsin.

- Number of professional development trainings provided.

- Percent of PAT program managers who report having received the new PAT quality standard report.

- Monthly, number of e-newsletters sent to PAT email list.

- Monthly, number of other communications sent to PAT list.

- Yearly, number of Advisory Committee meetings held.

- Yearly, number of Wisconsin families receiving PAT services.

- Yearly, number of families receiving PAT services funded by org.

- Percentage of PAK participants who report increase in positive parenting skills.

Short Term:

- Yearly, add 1-2 PAT programs operating in WI.

- 70% of training attendees report an increase in their ability to assist families.

- 70% of training attendees report that handouts and materials are useful.

- 50% of PAT staff members feel up-to-date on research and policy regarding their field.

- Annually, ensure that 150 families receive unbroken PAT services through org. funding.

- Increase PAK services to reach 100 Milwaukee families.

- 70% of PAK parents report increased knowledge of positive parenting practices.

- 70% of PAK parents report spending time every day reading to their children.

Medium Term:

- PAT programs operate in all 72 Wisconsin counties.

- 100% of PAT programs follow model guidelines for the highest quality programs, including monthly home visits.

- 70% of participating children identified with possible developmental delays will receive services.

- Participating families maintain nurturing home environments that are conducive to learning.

- PAK children will have improved relationships with parents, cognitive abilities, pre-academic and early literacy skills, and social-emotional development.

- 70% of PAK children will demonstrate increased school readiness.

- *Children* will:
  • Develop key social skills.
  Parents will increase:
  • Ability to communicate with children's school.
  • Participation in school.
  • Increase ability to advocate for their children.

Long Term:

- Urban and rural families will have access to proven early childhood programs that increase cognitive abilities and reduce incidences of child abuse.

- Improved academic performance for participating children.

- *Children* will have:
  • Fewer unidentified developmental delays,
  • Increased access to services,
  • Fewer placements in special education.

- *Children* will:
  • Demonstrate positive social/emotional skills.
  Parents will demonstrate:
  • Increased involvement in children's education.
  • Improved relationships with teachers.

- Greater opportunities for success for participating children.

Brooks-Gunn, J., Berlin, L., & Fuligni, A. (2000). Early childhood intervention programs: What

about the family? In J. P. Shonkoff & S. J. Meisels (Eds.), *Handbook of early childhood*

intervention (2nd ed., pp. 510-548). Cambridge, MA: Cambridge University Press.

Brooks-Gunn, J., & Markman, L.B. (2005). The contribution of parenting to ethnic and racial

gaps in school readiness. *The Future of Children, 15*(1), 139-168. Retrieved from

http://www.futureofchildren.org/

Campbell, F.A., & Ramey, C.T. (1995). Cognitive and school outcomes for high-risk African-

American students at middle adolescence: Positive effects of early intervention.

American Educational Research Journal, 32, pp. 743-772. DOI:

10.3102/00028312032004743

Currie, J., & Thomas, D. (2000). School quality and the longer-term effects of Head Start.

Journal of Human Resources, 35, pp. 755–774. Retrieved from

http://www.ssc.wisc.edu/jhr/

Duncan, G.J., & Magnuson, K.A. (2005). Can family socioeconomic resources account for racial

and ethnic test score gaps? *The Future of Children, 15*(1), 35-54. Retrieved from

http://www.futureofchildren.org/

Gorey, K. (2001). Early childhood education: A meta-analytic affirmation of the short- and long-

term benefits of educational opportunity. *School Psychology Quarterly, (16)*1, 9-30. DOI:

10.1521/scpq.16.1.9.19163

Magnuson, K.A., & Waldfogel, J. (2005). Early childhood care and education: Effects on ethnic and racial gaps in school readiness. *The Future of Children, 15*(1), 169-196. Retrieved from http://www.futureofchildren.org/

Pigott, T.D., & Israel, M.S. (2005). Head Start children's transition to kindergarten: Evidence from the Early Childhood Longitudinal Study. *Journal of Early Childhood Research, 3*, 77-104. DOI: 10.1177/1476718X05051347

Reynolds, A.J., & Robertson, D.L. (2003). School-based early intervention and later child maltreatment in the Chicago Longitudinal Study. *Child Development, 74*, pp. 3-26. DOI: 10.1111/1467-8624.00518

Reynolds, A.J., Ou, S., & Topitzes, J.W. (2004). Paths of effects of early childhood intervention on educational attainment and delinquency: A confirmatory analysis of the Chicago Child-Parent Centers. *Child Development, 75*, pp. 1299-1328. DOI: 10.1111/j.1467-8624.2004.00742.x

Reynolds, A.J. (2005). Confirmatory program evaluation: Applications to early childhood interventions. *Teachers College Record, 107*, pp. 2401-2425.

Rouse, C., Brooks-Gunn, J., & McLanahan, S. (2005). Introducing the issue. *The Future of Children*, *15*(1), 5-14. Retrieved from http://www.futureofchildren.org/

Schweinhart, L., & Weikart, D. (1998). High/Scope Perry Preschool Program effects at age twenty-seven. In J. Crane (Ed.), *Social programs that work* (pp. 148-1620). New York: Russell Sage Foundation.

Yoshikawa, H. (1994). Prevention as cumulative protection: Effects of early family support and

education on chronic delinquency and its risks. *Psychological Bulletin, 115*(1), 28-54.

DOI: 10.1037/0033-2909.115.1.28

Zigler, E., Pfannenstiel, J.C., & Seitz, V. (2008). The Parents as Teachers program and school

success: A replication and extension. *Journal of Primary Prevention 29*(2), 103-120.

DOI: 10.1007/s10935-008-0132-1

Subject line: Org. training follow up

Dear Name,

Thank you for participating in a recent training given by_____, (training title) on (date). We appreciate the opportunity to serve you, and we are working to improve our programs so that we can provide the best service possible. We would like your feedback to help us in that process. Please complete the survey <u>here</u> to let us know how we are doing. ***We know your time is valuable, so if you choose to complete the survey we will send you a $5 gift certificate to Pick N Save as a token of our thanks.***

We will use your information in a few ways. First, we will use your answers to gauge our own performance and to see where we can improve on our services to parents and professionals in the field of early childhood and education. We are also required to report our outcomes, the results of our programs, to our funders. We may also use specific quotes in printed materials, but we will remove your identifying information before we do so.

In addition, this survey was designed by a Marquette University graduate student who is working with our agency this year. She will use it as part of her Professional Project for graduation, which will focus on the design and implementation of an internal evaluation. Because of this, a consent form will precede the survey. This form gives permission to use your data ***without your identifying information*** in her published report to Marquette University. We understand the language in the form may be difficult, so if you have any questions, we will be happy to go over it with you.

Follow the link below to access the survey.
<u>Link to survey</u>

If you have any questions or concerns about the form or the survey, please give Jerica Broeckling a call at xxx-xxx-xxxx. You can also call our Director at xxx-xxx-xxxx. Thank you again for your time and participation.

Jerica Broeckling

Letterhead

Date

Participant Name
Mailing Address

Dear Name,

Thank you for participating in a recent training held by _____, (training title) on (date). We appreciate the opportunity to serve you, and we are working to improve our programs so that we can provide the best service possible. We would like your feedback to help us in that process. Please complete the enclosed survey to let us know how we are doing. ***We know your time is valuable, so if you choose to complete the survey we will send you a $5 gift certificate to Pick N Save as a token of our thanks.***

We will use your information in a few ways. First, we will use your answers to gauge our own performance and to see where we can improve on our services to parents and professionals in the field of early childhood and education. We are also required to report our outcomes, the results of our programs, to our funders. We may also use specific quotes in printed materials, but we will remove your identifying information before we do so.

In addition, this survey was designed by a Marquette University graduate student who is working with our agency this year. She will use it as part of her Professional Project for graduation, which will focus on the design and implementation of an internal evaluation. Because of this, we ask that you read the enclosed consent form. This form gives us permission to use your data ***without your identifying information*** in her published report to Marquette University. We understand the language in the form may be difficult, so if you have any questions, we will be happy to go over it with you.

If you have any questions or concerns about the form or the survey, please give Jerica Broeckling a call at xxx-xxx-xxxx. You can also contact _____ at xxx-xxx-xxxx. You can also complete the survey online at http://www.zoomerang.com/Survey/WEB229ZJE8R3K8/?jobfunction=[jobfunction_value].

Sincerely,

Name
Executive Director

___

Survey of Medium-Term Training Outcomes

Page 1 - Heading

MARQUETTE UNIVERSITY
AGREEMENT OF CONSENT FOR RESEARCH PARTICIPANTS
Internal Evaluation of the _____.
Jerica Broeckling, Marquette University College of Professional Studies
You have been invited to participate in this research study. Before you agree to participate, it is important that you read and understand the following information. Participation is completely voluntary. Please ask questions about anything you do not understand before deciding whether or not to participate.
PURPOSE: The purpose of this research study is to determine how effective _____ trainings are in helping parents and professionals to implement best practices. This survey is part of an evaluation of the _____. You will be one of approximately 10 participants in this research study.
PROCEDURES: You will complete a brief survey about your past training with _____.
DURATION: Your participation will take approximately five to ten minutes.
RISKS: The risks associated with participation in this study are minimal and no more than you would encounter in daily life, when discussing trainings. You may skip questions if you feel uncomfortable. Your answers will remain confidential.
BENEFITS: _____, Inc. will use this study to evaluate and improve upon its training programs for parents and professionals in the areas of parent involvement and early childhood development. You will directly benefit from your participation in this study if you choose to participate in a future _____ training.

Page 1 - Heading

CONFIDENTIALITY: All information you reveal in this study will be kept confidential. All your data will be assigned an arbitrary code number rather than using your name or other information that could identify you as an individual. When the results of the study are published, you will not be identified by name. The data will be held indefinitely and may be used in future research. Your research records may be inspected by the Marquette University Institutional Review Board or its designees, _____ and (as allowable by law) state and federal agencies.

COMPENSATION: A $5 gift certificate will be mailed to your designated address after you complete this survey.

VOLUNTARY NATURE OF PARTICIPATION: Participating in this study is completely voluntary and you may withdraw from the study and stop participating at any time without penalty or loss of benefits to which you are otherwise entitled. Once the survey is submitted, you cannot withdraw from the study.

CONTACT INFORMATION: If you have any questions about this research project, you can contact Jerica Broeckling at 414-755-8714 or at jerica.broeckling@marquette.edu. If you have questions or concerns about your rights as a research participant, you can contact Marquette University's Office of Research Compliance at (414) 288-7570.

By completing this survey, you are agreeing to the above procedures and conditions.

Page 1 - Question 1 - Name and Address (U.S)   [Mandatory]

Please enter your full name and the address where you'd like to receive your $5 gift certificate.

- ✎ Name
- ✎ Company
- ✎ Address
- ✎ City
- ✎ State
- ✎ Zip

Page 1 - Question 2 - Open Ended - Comments Box

Please enter the training you attended (see the original email or letter to get this information).

Page 2 - Question 3 - Yes or No   [Mandatory]

Have you used any of the information provided in your training during the last 3 months?

- ○ Yes
- ○ No
- ○ If you answer yes, please provide an example.

Page 3 - Heading

On a scale of 1 to 5, where 1=no increase and 5=major increase, how much did the training increase:

Page 3 - Question 4 - Rating Scale - One Answer (Horizontal)   [Mandatory]

Your knowledge of the subject

No increase	2	3	4	Major increase	Does not apply
☐	☐	☐	☐	☐	☐

Page 3 - Question 5 - Rating Scale - One Answer (Horizontal)   [Mandatory]

Your ability to perform your job

No increase	2	3	4	Major increase	Does not apply
☐	☐	☐	☐	☐	☐

Page 3 - Question 6 - Rating Scale - One Answer (Horizontal)   [Mandatory]

Your skills as a parent

No increase	2	3	4	Major increase	Does not apply
☐	☐	☐	☐	☐	☐

Page 3 - Question 7 - Rating Scale - One Answer (Horizontal)   [Mandatory]

Your ability to advocate for a child
(We define an advocate as someone who defends or supports a cause or someone's interests)

No increase	2	3	4	Major increase	Does not apply
☐	☐	☐	☐	☐	☐

Page 4 - Heading

On a scale of 1 to 5, where 1=not at all satisfied and 5=very satisfied, how satisfied were you with:

Page 4 - Question 8 - Rating Scale - One Answer (Horizontal)

The trainer's knowledge of the material

Not at all satisfied	2	3	4	Very satisfied	Does not apply
☐	☐	☐	☐	☐	☐

Page 4 - Question 9 - Rating Scale - One Answer (Horizontal)   [Mandatory]

The trainer's presentation of the material					
Not at all satisfied	**2**	**3**	**4**	**Very satisfied**	**Does not apply**
☐	☐	☐	☐	☐	☐

Page 4 - Question 10 - Rating Scale - One Answer (Horizontal)   [Mandatory]

The thoroughness of the training					
Not at all satisfied	**2**	**3**	**4**	**Very satisfied**	**Does not apply**
☐	☐	☐	☐	☐	☐

Page 5 - Question 11 - Yes or No

Was a translator provided for your training? (If you answer yes, please skip the next question and go straight to question 13.)

❍ Yes
❍ No

Page 5 - Question 12 - Rating Scale - One Answer (Horizontal)

Would you have liked a translator to be present at the training?		
Yes	**No**	**Does not apply**
☐	☐	☐

Page 5 - Question 13 - Rating Scale - One Answer (Horizontal)

Did the translator help you to understand the material being presented?		
Yes	**No**	**Does not apply**
☐	☐	☐

Page 6 - Question 14 - Open Ended - Comments Box   [Mandatory]

What, if anything, do you feel could have improved the presentation?

....
....
....
....

Page 6 - Question 15 - Open Ended - Comments Box   [Mandatory]

Is there information would you like to receive on this topic that was not covered during your training? If so, what?

_____

_____

_____

_____

_____

Page 6 - Question 16 - Open Ended - Comments Box   [Mandatory]

Would you recommend a _____ training to colleagues or to other parents? Why or why not?

_____

_____

_____

_____

_____

Page 7 - Heading

Thank you for your participation. Your answers will be kept confidential, and your identifying information will be removed before this survey is analyzed by _____ staff.

Page 7 - Question 17 - Open Ended - One Line

What is your job function?

_____

Thank You Page

Redirect: organization website

Screen Out Page

Redirect: organization website

Over Quota Page

(Standard - Zoomerang branding)

Survey Closed Page

(Standard - Zoomerang branding)

MARQUETTE UNIVERSITY
AGREEMENT OF CONSENT FOR RESEARCH PARTICIPANTS
Performance Measurement System Implementation: Theory and Practice (Interview)
Jerica Broeckling
College of Professional Studies

You have been invited to participate in this research study. Before you agree to participate, it is important that you read and understand the following information. Participation is completely voluntary. Please ask questions about anything you do not understand before deciding whether or not to participate.

PURPOSE: The purpose of this research study is to document a nonprofit organization's experience with implementing a performance measurement system as a way of understanding how real-life practice relates to theory. You will be one of approximately three participants in this research study.

PROCEDURES: You will be asked to participate in one interview about your efforts to design and implement a performance measurement system. You will be audio taped during the interview to ensure accuracy. The tapes will be destroyed after three years after the completion of the study. For confidentiality purposes, your name will not be recorded. The researcher may contact you to follow up on your responses.

DURATION: Your participation will consist of approximately 45 minutes of interview time and approximately 15 minute of follow-up.

RISKS: The risks associated with participation in this study are minimal, and no more than you would encounter in everyday life when discussing your organization and its practices. If any questions make you uncomfortable, you are free to skip them.

BENEFITS: The benefits associated with participation in this study include gaining a deeper understanding of the decision-making practices within your organization, especially as they related to the best practices in the field.

CONFIDENTIALITY: All information you reveal in this study will be kept confidential. All your data will be assigned an arbitrary code number rather than using your name or other information that could identify you as an individual. When the results of the study are published, you will not be identified by name. Data will be kept in the principal investigator's home office in a locked file cabinet. The data will be destroyed by shredding paper documents, deleting electronic files, and erasing audio tapes three years after the completion of the study.

VOLUNTARY NATURE OF PARTICIPATION: Participating in this study is completely voluntary and you may withdraw from the study and stop participating at any time without penalty or loss of benefits to which you are otherwise entitled. If you wish to stop participating,

please let the principal investigator know as soon as possible. Once the research project is complete, you may no longer withdraw from the study.

CONTACT INFORMATION: If you have any questions about this research project, you can contact Jerica Broeckling at 414-334-8646 or at jerica.broeckling@marquette.edu. If you have questions or concerns about your rights as a research participant, you can contact Marquette University's Office of Research Compliance at (414) 288-7570.

I HAVE HAD THE OPPORTUNITY TO READ THIS CONSENT FORM, ASK QUESTIONS ABOUT THE RESEARCH PROJECT AND AM PREPARED TO PARTICIPATE IN THIS PROJECT.

_____   _____
Participant's Signature   Date

_____
Participant's Name

_____   _____
Researcher's Signature   Date

1. Please describe the organization's process for collecting data before the development of the database.
2. Please describe the organization's process for tracking and reporting performance before the development of the database.
3. How and when did the organization decide to change its practices?
4. Please describe the process you went through in developing the current structure.
5. Why did you choose the database program you currently use as opposed to one that is pre-made?
6. What are two of the major differences between the old system and the new system?
7. Does the new system meet the organization's needs better? Why or why not?
8. What are cost differences between the old and the new system?
9. Are there instances where the new system does not provide the information you need?
    a. If so, how do you handle those instances?
10. How is the data used today? (for example, decision making, grant reporting, financial determinations, evaluation, etc)

Appendix F
Performance Measurement Dashboard

Productivity

Trainings Sept-March
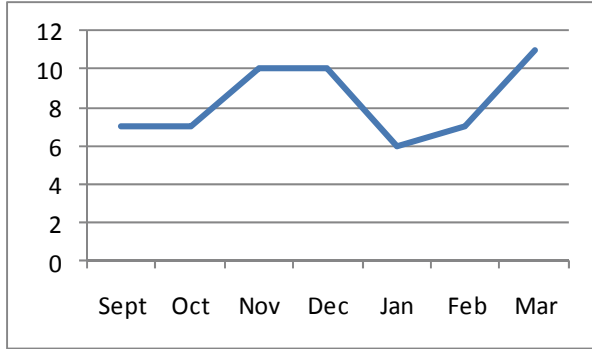
Trainings by Group Sept-March
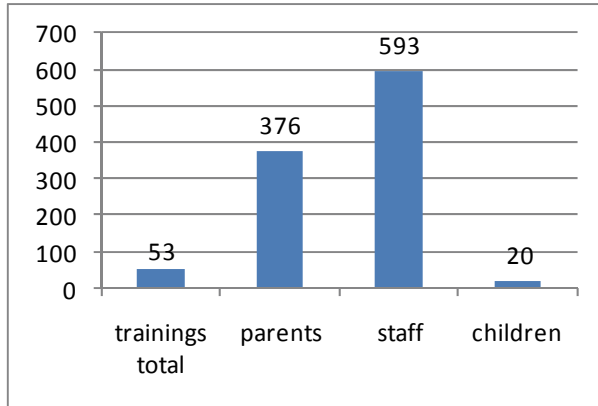
Effectiveness

3 Month Outcomes

Participants who:	%	#
Increased knowledge of subject	95	21
Increased ability to perform job	82	18
Increased parenting skills (excluding "does not apply")	94	17 (4DNA)
Increased advocacy skills	91	20

Customer Satisfaction

	% (#)	
	Parents (74)	Staff (85)
Initial survey: Percent who agree/strongly agree:		
that presentation was clear/ understandable	97 (72)	98 (83)
that presentation was of high quality	92 (68)	93 (79)
that the handout and materials were useful	89 (66)	94 (80)

|	% (#)	
3 month interval: Percent:	Parents	Staff
Satisfied with trainer's knowledge	100 (22)	100 (22)
Satisfied with trainer's presentation	100 (22)	100 (22)
Agree/Strongly agree that training was thorough	100 (22)	100 (22)

Word of Mouth

Participants who:	%	#
Passed materials to others	50	11
Requested a follow up training	13.6	3
Would recommend organization to others	100	22

Reasons	%	#
Presentation gave good information	32	7
Presentation is good for parents	41	9
Presentation is good for staff	9	2
I learned something I didn't know	9	2

Service Quality

Presentation Quality

Initial survey: Percent who agree/strongly agree that:	% (#)	
	Parents (74)	Staff (85)
The material was relevant	92 (68)	96 (82)
The information was useful	91 (67)	93 (79)
The presentation helped increase their skills	86 (64)	92 (78)

Spanish Interpretation
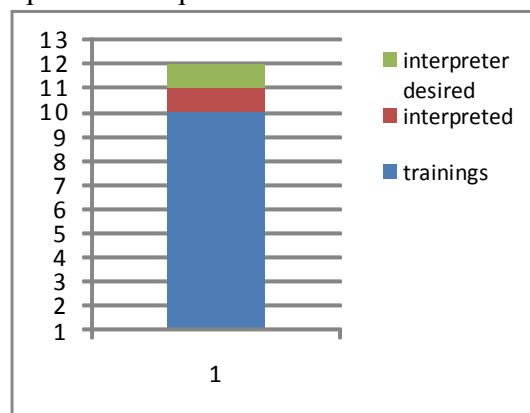
Suggested Improvements by Type

Suggestion type	# requesting	
	Parents	Staff
More in-depth information	3	4
More information	3	2
More professionally presented	0	5
Longer presentation ("more time")	7	8
Shorter presentation	0	1
Improvements to handouts/materials	2	2
Interpreter requested	4	0
Administrative improvements	4	7

Specific information requests

Parents	Staff
Science Fairs	Helping children with reading
Correlation between childhood development and behaviors (parent)	More in-depth information (from Guiding Children Toward Success)
Information on children's medical issues (Parents w/ Voices)	Advocacy/self-esteem for parents/students (From Parent Coordinator training)
More resources for getting ahead (parent)	Improving communication between staff and parents
More information on listening (Joyce Epstein)	Helping more students graduate
Parents Night ideas and home school meeting	
More ideas for games at home	

Phone Script

Date:_____

Time:_____

Staff taking call:_____

Caller's name:_____

Caller's geographic area (city, county, or school district):_____

Request:_____

Thanks for contacting us. We're happy to provide you with the materials you requested. I'd also like to ask you a question. We're trying to track how our information moves through the community. Can you tell me how you heard about our organization?

- o   From a colleague: (Name, if possible) _____
- o   From another parent: (Name, if possible) _____
- o   From my child's teacher/school staff (Name of school, district) _____
- o   From a workshop I attended
    - o   Provide name/date if possible: _____

Have you visited our website?

- o   Yes

    Were you able to find the information you needed there?

- o   No

Thank you. If you need more resources or want to check out our available trainings, please don't hesitate to call me back or check out our website (give web address).