UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

SYSTEMATIC EVALUATION FOR NONPROFIT ORGANIZATIONS: PROBLEMS,
PROSPECTS, AND RECOMMENDATIONS

A Dissertation

SUBMITTED TO THE GRADUATE FACULTY

In partial fulfillment of the requirements for the

Degree of

Doctor of Philosophy

By

Valerie R. Thompson

Norman, Oklahoma
2005

UMI Number: 3174278

UMI®

SYSTEMATIC EVALUATION FOR NONPROFIT ORGANIZATIONS: PROBLEMS,
PROSPECTS, AND RECOMMENDATIONS

A Dissertation

APPROVED FOR THE DEPARTMENT OF POLITICAL SCIENCE

_____
Professor Thomas James, Dissertation Director
and Major Professor

_____
Professor Kelly Damphousse

_____
Professor Aimee Franklin

_____
Professor Cindy Rosenthal

_____
Professor Jill Tao

TABLE OF CONTENTS   Page

vi

vii

viii

List of Tables

Illustrations

ABSTRACT

The multitude of nonprofit agencies continues to grow as new challenges and more unmet needs are being identified. Government, at all levels, has continued to increase its reliance on nonprofit organizations to fill in the gaps that the public and private sector leave behind. With limited funding and volunteers, social nonprofit organizations traditionally have been innovative and committed to meeting community needs that go unfilled by the public or the private sector.

Evaluation is essential to every organization. Nonprofit organizations typically devote little time to program evaluation but concentrate their efforts on service delivery. Even less time is devoted to a systematic approach to program evaluation. Yet, the pressure for nonprofits to become more accountable, effective and efficient has increased in recent years (Hatry and Lampkin, 2001). Most nonprofits have far too little time, money, or in-house skills for developing evaluation systems and maintaining them over time (Murray, 2001).

The purpose of this study is to assess the current status of evaluation activities and the utilization of their results for the nonprofit sector in the United States. The dissertation will seek to discover how much evaluation is done by nonprofits, what types of evaluations are being conducted, and how evaluations are being used. If evaluations are not being used, this dissertation will try to answer why nonprofit organizations do not use evaluations. Additionally, the dissertation will

define conditions or factors that facilitate or inhibit evaluations and recommend how to promote the systematic evaluation of nonprofit organizations in light of the challenges and conditions in which they operate.

Replicating the voluntary sector portion of the Canadian Centre for Philanthropy's study, this dissertation seeks to answer questions concerning evaluation practices for nonprofit organizations in the United States. Specifically, the study will examine the evaluation practices of over 100 nonprofit organizations in the United States that are affiliates of the National Urban League.

CHAPTER 1

EVALUATION PRACTICES FOR NONPROFIT ORGANIZATIONS

Introduction

Since the late 1900s, nonprofit organizations have increasingly become an important part of America's society. Federal, state and local government has become more reliant upon nonprofit organizations to provide social services while resources from the government, foundations and private funders to provide client services continue to diminish. The increasing demands on nonprofit agencies have forced rigorous examination on how to provide more services with limited funds. Like public and some private sectors, nonprofit organizations are faced with increasing demands for services without corresponding increases in funding.

The multitude of nonprofit agencies continues to grow as new challenges and more unmet needs are being identified. Government, at all levels, has continued to increase its reliance on nonprofit organizations to fill in the gaps that the public and private sector leave behind. With limited funding and volunteers, social nonprofit organizations traditionally have been innovative and committed to meeting community needs that go unfilled by the public or the private sector.

The increasing demand on nonprofit organizations to provide more services with limited and sometimes decreasing funding has created challenges for

nonprofit organizations and their funding sources. Questions have arisen concerning the need to impact more people with fewer dollars, which persuade nonprofit organizations and their funding sources to find ways to become more effective and provide greater social and economic impact.

According to Margaret Plantz, Martha Taylor Greenway and Michael Hendricks (1997), the nonprofit sector has been measuring certain aspects of performance for twenty-five years or more. During that period, the scope of measurement has expanded to address such issues as financial accountability, inputs, cost, program products or outputs, adherence to standards of quality in service delivery, participant-related measures (i.e. demographic characteristics), and client satisfaction. The authors suggest that these measures yield critical information about the services nonprofits are providing, but the measures reveal nothing about whether the nonprofit's effort made a difference. Funders increasingly have demanded more accountability of the programs they fund. Not just fiscal accountability, but accountability for results. Funders want to know that their funds are being used in a way that makes a difference. In other words, was anyone better off as a result of the service from the nonprofit organization?

Consequently, nonprofit organizations must engage in effective planning and management. This planning and management requires systematic assessments of past activities and their results as well as current and projected community needs.

The evaluation of nonprofit funded programs can provide high quality, timely information on program process and activities, short and intermediate term outcomes, and program costs as they relate to activities and outcomes. Unfortunately, most nonprofits do not engage in or require comprehensive evaluations of programs they administer. The limited evaluation activities performed by nonprofit organizations typically focus on measures of output (e.g. number of clients served, number of workshops held, number of pamphlets/brochures distributed) with little or no attention to the potential outcomes associated with those activities or the cost required to produce them. A study conducted by Joanne Carman (2003) of 175 nonprofit organizations in New York City found that most nonprofits gathered data on program expenditures and the number people they served. Less than one quarter of the organizations (23%) reported using standardized tests or data collection instruments. Less than half (47%) gathered data on best practices or benchmarks, and only 11% gathered control or comparison data.

Improving performance and building capacity for nonprofit organizations continue to be growing concerns. Paul Light (2000) notes the transition by nonprofits to look at performance measurements recently transpired. He states:

> "Whether the sector's glass is half full or half empty, unanimous consensus has emerged that nonprofit organizations have to improve their performance. Much of the "lean and mean" rhetoric that so preoccupied private firms and

3

government agencies during the 1980s and early 1990s has now filtered over to the nonprofit sector."

However, Light (2000) also notes that the sector has begun to take interest in the value of outcome measurement. He states,

> "Much of the stress in the sector is caused by what some might characterize as over empowerment: too much work, not enough money, and not enough operating support. Nevertheless, the nonprofit sector and its funders have embraced at least two of the key elements of what Vice President Al Gore called liberation management: outcomes measurement and competition."

Nonprofit organizations typically devote little time to program evaluation but concentrate their efforts on service delivery. Even less time is devoted to a systematic approach to program evaluation. Carman's (2003) research also found that only 41%, of the nonprofits she surveyed, used a performance measurement system. Her research reported that very few organizations designed program logic models (17%) or participated in accreditation processes (16%) and just 3% used a balanced scorecard management system. Nonprofit organizations usually are focused on fiscal oversight, management or human capital. Sawhill and Williamson (2001) interviewed thirty well-known nonprofit organizations and found that every organization could rattle off its basic statistics for financial activity, fundraising, membership and so forth. But a systematic method of evaluating a program by examining its process, outcomes and costs did not exist.

4

The pressure for nonprofits to become more accountable, effective and efficient has increased in recent years (Hatry and Lampkin 2001). The first efforts aimed at conducting evaluations by nonprofits focused on counting inputs and formal evaluations after project completion. Now, nonprofit organizations are being encouraged to develop relevant measures of outcomes and to collect data on a regular basis as part of ongoing projects. Some funders are requiring on-going measurements to help evaluate the impact of services and programs. However, program evaluation is viewed by nonprofit organizations as a luxury or a nonessential. Carman (2003) notes that few organizations have the luxury of using an external evaluator or maintaining separate evaluation funding. Most nonprofits have far too little time, money, or in-house skills for developing evaluation systems and maintaining them over time (Murray 2001).

Purpose of the Study

The purpose of this study is to assess the current status of evaluation activities and the utilization of their results for the nonprofit sector in the United States. The dissertation will seek to discover how much evaluation is done by nonprofits, what types of evaluations are being conducted, and how evaluations are being used. If evaluations are not being used, this dissertation will try to answer why nonprofit organizations do not use evaluations. Additionally, the dissertation will

5

define conditions or factors that facilitate or inhibit evaluations and propose recommendations to promote the systematic evaluation of nonprofit organizations in light of the challenges and conditions in which they operate. The study draws information from a network of affiliated nonprofit agencies located throughout the United States of America.

Research Questions: Canadian Centre for Philanthropy

A study conducted by the Canadian Centre for Philanthropy (2003) sought to address the following questions:

1. To what extent are nonprofit organizations attempting to evaluate their performance?

2. What methods are organizations using to evaluate their performance?

3. What are funders' expectations regarding evaluation and what support do they provide for evaluation activities?

4. What do organizations perceive to be strengths and their weaknesses in evaluations?

5. What are the perceived needs for resources to assist with evaluation?

6. What resources are needed to assist with better evaluations?

The study found that nonprofit evaluations and performance assessments are common features of organizational life in the nonprofit sector in Canada, and expectations for evaluations have been increasing. They also found that nonprofits perceived barriers to evaluation to be lack of internal capacity, such as staff or time, and lack of money. The nonprofits also saw unclear direction from funders about what is expected in an evaluation and lack of skills and knowledge in conducting evaluations as barriers.

Nonprofit organizations need to conduct systematic evaluations that assess the process, outcomes and the cost of the programs they administer. These organizations need information to replicate good programs, determine outcomes, measure the cost of a success, and address areas that can be improved. Evaluations for nonprofit organizations need to be as comprehensive and rigorous as possible.

Despite the benefits of evaluations, why are there not more nonprofit organizations that systematically conduct evaluations and use their results? Replicating the voluntary sector portion of the Canadian Centre for Philanthropy's study, this dissertation seeks to answer questions concerning evaluation practices for nonprofit organizations in the United States. Specifically, the study will examine the evaluation practices of 58 nonprofit organizations in the United States that are affiliates of the National Urban League.

Research Questions: National Urban League Affiliates

Evaluation Practices Revealed

The answers to the following research questions will reveal how nonprofit organizations are conducting evaluations within the National Urban League affiliate network. The study will reveal the expectations and support of the funders. The study seeks to unveil some barriers that prevent nonprofit organizations from conducting evaluations. To understand the extent of nonprofit evaluation several questions were asked of more than one hundred nonprofit organizations affiliated with the National Urban League. These questions include:

1. To what extent are nonprofit organizations attempting to evaluate their performance?

2. How do nonprofit organizations evaluate their performance?

3. For those nonprofits who are doing evaluations, how do they use the results?

4. For nonprofits who are not conducting evaluations, why are they not performing them?

8

Funders' Expectations and Support

The research seeks to answer questions concerning funders' expectations on nonprofit organizations and the type of support the nonprofit organizations received from their funding sources to conduct evaluations. One of the research questions in this study asks, what are funders' expectations regarding evaluation and what support does the funder provide for evaluation activities?

Evaluation Barriers for Nonprofit Organizations

The research also seeks to find if there are any barriers that prevent nonprofit organizations from conducting evaluations. The following is a list of several key questions that seek to gather information about nonprofit organizations' strengths and weaknesses relating to evaluation:

1. What do nonprofit organizations perceive to be their strengths and weaknesses in evaluations?

2. What are the perceived needs for resources to assist with evaluation?

3. What resources are needed to assist with better evaluation?

4. What will help nonprofits who do not currently conduct systematic evaluations as a core component of their operations, begin to systematically evaluate their programs?

The data gathered to answer these questions will come from an analysis of evaluating activities of 58 National Urban League affiliates who responded to a survey[1].

This research will make theoretic and practical contributions. This study will contribute to the theoretic understanding of evaluation utilization of nonprofit organizations by examining the use or non-use of systematic evaluations by examining the National Urban League affiliates. It also will make a contribution to the practical application of evaluation by providing recommendations on how to promote the systematic evaluation of nonprofit organizations in light of the challenges and conditions in which they operate.

Dissertation Organization

Chapter One provides an overview of the study and outlines the research questions to be answered in the dissertation. Chapter Two discusses the importance of evaluation. The chapter defines evaluation and why nonprofit organizations should care about it. It also provides a summary of the recent developments in nonprofit evaluations. Chapter Two also provides information on evaluation utilization, which is a subject that gets a great deal of attention in the evaluation literature.

A historical overview of the National Urban League and its affiliates is provided later in this document.

10

Chapter Three discusses the challenges and opportunities for nonprofit organizations that perform or do not perform nonprofit evaluations. It provides a summary of the nonprofit literature with an emphasis on the evaluation literature. The chapter reviews the literature that addresses nonprofit evaluation and highlights several themes that emerge from the discussion. It examines the problems nonprofit organizations encounter in conducting systematic evaluations.

Chapter Four provides a historical overview of the National Urban League and its affiliates in the United States. A discussion of this ninety-five year old organization and its relationship with the affiliates provides the background information on the population surveyed in this study. This chapter discusses current evaluation practices and the National Urban League's Performance Assessment tool. The chapter concludes with a summary on the extent of systematic evaluations performed by the National Urban League affiliates.

Chapter Five provides an explanation of the methodology used to answer the research questions in this study. The chapter discusses the survey method, data collection, and provides information on the population surveyed. This chapter discusses the characteristics used to analyze the agencies concerning evaluation practices.

Chapter Six presents the findings of a survey of affiliate agencies of the National Urban League. The chapter begins with a discussion on the

characteristics of the survey respondents. The National Urban League affiliates or agencies were chosen for this study because they are diverse in size but also because the agencies are dispersed in more than one hundred cities around the United States. The affiliates of the National Urban League provide services based on the geographic need of its target population.

The chapter follows with the findings from the survey questions. This chapter explains the extent that nonprofit organizations are attempting to evaluate their performance. It discusses which nonprofit organizations are doing evaluations and why they are performing evaluations at their agencies. The chapter informs the reader what nonprofit agencies are evaluating the most.

Chapter Seven reviews the findings in context. It provides a comparison of the findings from the Canadian Centre for Philanthropy with the findings from the survey from National Urban League agencies. It highlights their commonalities and shows areas for improvement to increase systematic evaluations. The findings in both studies support the literature provided in the dissertation.

The final chapter provides an interpretation of the research findings along with recommendations for nonprofit organizations that were drawn from the answers to the research questions. It provides recommendations for nonprofit organizations to increase systematic evaluations in the nonprofit sector. The chapter recommends specific action for the National Urban League and its

affiliates to increase quality systematic evaluations. Chapter Eight also provides

recommendation for future research on systematic evaluations for nonprofit

organizations.

CHAPTER 2

EVALUATION: WHAT IS IT AND WHY SHOULD WE CARE?

This chapter explains what evaluation is and why nonprofit organizations should care about it. The chapter highlights three recent developments in the nonprofit sector in evaluation. There is information on evaluation utilization, which is a subject that gets a great deal of attention in the evaluation literature. We begin this chapter by defining of program evaluation.

Evaluation Defined

What is program evaluation? There is no single definition for program evaluation on which scholars completely agree. However, there are some general themes that emerge when there is a discussion on the definition of program evaluation:

- Program evaluation is a systematic.

- Evaluation is ongoing.

- Evaluation involves the collection of data and performing analysis on the data.

- Evaluation involves the application of social research methods.

- Evaluation consists of an investigation or analysis that is measured against a standard to determine or arrive at a conclusion about the activity or activities.

- Evaluation gives stakeholders tools to make judgments or decisions.

- Evaluation is the use of information to improve activities or programs.

Patton (1997) characterizes program evaluation as "the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming." Similar to Patton, Mattessich (2003) suggests that program evaluation is "the systematic process for an organization to obtain information on its activities, its impacts, and the effectiveness of its work, so that it can improve its activities and describe its accomplishments." In both definitions, a systematic process and an analysis or judgment take place.

Carol Weiss (1998) offers what she calls a provisional definition of program evaluation and states that there are key elements of her definition which includes a systematic assessment and a standard for comparison. Weiss (1998) defines program evaluation as "the systematic assessment of the operation and/or the outcomes of a program or policy, compared to a set of explicit or implicit standards, as a means of contributing to the improvement of the program or policy." Yet others refer to program evaluation as the systematic description and

judgment of programs and, to the extent feasible, systematic assessment of the extent to which they have the intended results (Wholey, Hatry, and Newcomer 1994).

Rossi, Lipsey and Freeman (2004) suggest that program evaluation is a social science activity directed at collecting, analyzing, interpreting and communicating information about the workings and effectiveness of social programs. They also suggest that program evaluation is the use of social research methods to systematically investigate the effectiveness of social intervention programs in ways that are adapted to their political and organizational environments, and are designed to inform social action to improve social conditions.

Fitzpatrick, Sanders and Worthen (2004) offer a similar, but more complex definition. They define evaluation as "the identification, clarification, and application of defensible criteria to determine an evaluation object's value (worth or merit) in relation to those criteria." My definition of program evaluation is simple. I define program evaluation as a systematic process of analyzing and interpreting the results of a program's implementation, activities and costs measured against the program's intended purpose for the improvement of a program. It also should be noted that there are some scholars that differentiate between evaluation research, program evaluation and performance monitoring. However, for the purpose of this research, I will use them interchangeably.

Brief History of Nonprofit Evaluation

Man has counted and measured things from the beginning of time. The tenure of nonprofit organizations cannot be marked from the beginning of time, but it has been a stable fixture for decades throughout American history. Since the early decades of the American republic, nonprofit organizations have played a critical role in helping people in need by providing education, training, residences, counseling, and in-kind and cash support (Lipsky and Smith 1993). Following World War II, numerous major federal and privately funded programs were launched to provide urban development and housing, technological and cultural education, occupational training, and preventive health activities. It was also during this time that federal agencies and private foundations made major commitments to international programs for family planning, health and nutrition, and rural development (Rossi, and Freeman 1999).

Although human beings have been attempting to solve social problems using some kind of rationale or evidence for centuries, program evaluation began with the ambitious, federally funded social programs of the Great Society initiative during the mid-1960's. As resources poured into these programs to try to solve the complex problems, the problems did not disappear. Instead societal problems grew and became more complex. The public grew more cautious and there was increasing pressure to provide evidence of the effectiveness of specific initiatives

in order to allocate limited resources. During this period "systematic evaluation was increasingly sought to guide operations, to assure legislators and planners that they were proceeding on sound lines and to make services responsive to their public (Cronbach et al. 1980)."

It soon became apparent to those in the trenches and those who were allocating the dollars that there were not enough resources to solve all of our social problems. Michael Patton (1997) describes the birth of program evaluation and its impact on the nonprofit sector. He states the following:

> "Program evaluation as a distinct field of professional practice was born of two lessons from this period of large-scale social experimentation and government intervention: First, there is not enough money to do all the things that need doing; and second, even if there were enough money, it takes more than money to solve complex human and social problems. As not everything can be done, there must be a basis for deciding which things are worth doing. Enter evaluation."

As government funding of nonprofit organizations grew, the pressures on government officials to maintain accountability over public funds increased as well. Over time, government officials responded to the increased reliance on private agencies by instituting new regulations, changing contract requirements, and increasing administrative oversight (Lipsky and Smith 1993). Since the Great Society era, pressure to demonstrate the worth of social programs has increased. Limited resources, increasingly complex and layered social problems, the

changing political climate, and a seeming shift in public opinion about the extent to which government and other institutions should support have shifted the balance to an almost exclusive focus on accountability or prove that it works, versus quality of the program or work to improve (W.K. Kellogg Foundation 1998).

Common Approaches and Developments in Nonprofit Evaluation

There are several methods and approaches to evaluations that nonprofit organizations are attempting to implement. Each of the methods attempts to answer questions about the nonprofit's organization or program. Rossi and Freeman (1999) state that program evaluation is essentially an information-gathering and interpreting endeavor that attempts to answer a specified set of questions about a program's performance and effectiveness. They note that evaluators have developed relatively distinct conceptual frameworks and associated methods to address each type of question. Rossi and Freeman (1999) categorize the evaluation methods as the following:

- Needs assessment. Answers questions about the social conditions a program is intended to address and the need for the program;

- Assessment of program theory. Answers questions about program conceptualization and design

- Assessment of program process (or process evaluation). Answers questions about program operations, implementation and service delivery.

- Impact assessment (impact evaluation or outcome evaluation). Answers questions about the programs outcomes or impact;

- Efficiency assessment. Answers questions about program cost and cost-effectiveness.

Posavac and Carey (2002) suggest that there are four general types of evaluation. They are evaluation of need, evaluation of process, evaluation of outcome and evaluation of efficiency.

Recent Developments in Nonprofit Evaluation: Participatory, Theory-based, &
Outcome Evaluations

Recent developments in the nonprofit sector in evaluation have directed the sector's attention towards three common approaches. The most common developments and approaches to evaluation among nonprofit organizations are participatory evaluation, theory-based evaluation and impact or outcome evaluation. The following information provides a brief summary of each of the approaches and how the approaches are being used in the nonprofit sector.

20

Participatory Evaluation

Participatory evaluation is a process where the evaluator's perspective is given no more priority than the other stakeholders, including program participants. This type of evaluation makes the evaluation process and its results relevant and useful to stakeholders for future actions. Participatory evaluation attempts to be practical, useful, and empowering to multiple stakeholders and helps to improve program implementation and outcomes by actively engaging all stakeholders in the evaluation process.

The Sabo and Fusco (2002) define participatory evaluation as collaboration between evaluator and client. They suggest that participatory evaluation:

- Increases the likelihood that evaluation results are accurate and relevant.

- Ensures that program staff will be motivated and prepared to use evaluation findings.

- Builds the program's capacity to continue to design and conduct quality evaluations with diminished reliance on outside assistance.

- Stimulates deep thinking about programmatic issues, often leading to refinement in the program itself.

- Gives staff new tools for communicating their program to others.

This approach in nonprofit evaluation is very effective for nonprofit organizations because of the many stakeholders that are usually involved in

nonprofit programs. Another recent development in the nonprofit sector is the focus on theory-based evaluations

Theory-based Evaluation

Every evaluation should be based on a program theory or program logic model. Training on the relevance of program theory in nonprofit evaluations and training for nonprofit organizations on the development of program logic models is becoming more prevalent. Theory-based or program theory evaluation starts with the premise that every social program is based on a theory, and therefore the key to understanding what really matters about the program is through identifying this theory.

Theory refers to the contextualized understandings of how, why, under what conditions, and with what meaningfulness human behavior occurs. Greene (2001) suggests that an evaluation's relationship with theory is based upon the idea of evaluation as knowledge production or knowledge construction.

In theory-based evaluations, evaluators seek to combine outcome data with the understanding of the process that led to those outcomes to learn about the program's impact and its most influential factors. In theory-based evaluations, evaluators develop logic models to draw a picture or describe how a program works. Evaluators use program theory to develop key interim outcomes that will lead to ultimate long-term outcomes in the logic model. Developing interim

outcomes provides a tool to track short-term achievements and modify the program to increase the potential for achieving long-term impacts.

Program theory can help nonprofit organizations track the linkages from a program's activities to the expected outcomes. Weiss (2000) states that theory-based evaluation offers many advantages to the evaluator who conducts the study and to the program individuals who receive the results. Theory-based evaluations test the links between what programs assume their activities are accomplishing and what actually happens at each small step along the way.

The major perspective in theory-based evaluation is to build an evaluation model or conceptual framework that takes account of pertinent assumptions and mechanisms underlying a program, such as what must be done in terms of treatment and implementation environment, what kind of causal processes are involved, and/or what intended and unintended consequences are likely to be generated. The exact features in a program theory to be included in the evaluation model are dependent upon key stakeholders' needs, resources available for research, and evaluators' judgments. After these factors have been taken into consideration, the evaluation model will provide ideas on what kinds of data should be collected and what procedures are needed to analyze the data. Pertinent methods then are selected and applied to serve these needs (Chen 1994).

Rogers, Petrosino, Huebner and Hacsi (2000) provide an overview of program theory evaluation. They define program theory evaluation as an explicit theory or model of how the program causes the intended or observed outcomes and the evaluation is at least partly guided by this model. They state that federal research funders now require discussions of program theory in applications submitted for evaluation support. They also confirm that many not-for-profit agencies have followed the United Way's lead in developing performance measures based on a generic causal model of inputs, processes, outputs, and outcomes. The authors' review showed diversity in theory and practice on how program theories are developed and how they are used to guide evaluations.

Rogers, Petrosino, Huebner and Hacsi (2000), argue that for some evaluations, the program theory has been developed largely by the evaluator. In other evaluations, the program theory has been developed primarily by those associated with the program, often through a group process. Most program theories are summarized in a diagram showing a casual chain. At its simplest, a program theory shows a single intermediate outcome by which the program achieves its ultimate outcome. More complex program theories show a series of intermediate outcomes sometimes in multiple strands that combine to cause the ultimate outcomes. The third type of program theory is represented by a series of boxes labeled inputs, processes, outputs and outcomes.

24

There are two main uses for program theory evaluations. The main purpose is to test the program theory to identify what it is about the program that causes the outcome. This type of evaluation is used most commonly in large, well-resourced evaluations focused on answering summative questions. The other type of program evaluation is often done for small project level evaluations by or on behalf of project managers and staff. This type of evaluation tends to be more formative and guides the organization in making decisions concerning day-to-day activities. Many of these evaluations have been developed in response to increasing demands for programs and agencies to report performance information, and to demonstrate their use of evaluation to improve their services (Rogers et al. 2000).

Theory-based evaluation is regarded by many practitioners as a solution to many different problems in evaluation. Theory-based evaluations can provide some clues to help answer the questions of why some programs succeed when others fail. It can also provide an agenda for the next program and evaluation. This provides a road map to keep others in the field from falling into similar programmatic pitfalls. Theory-based evaluations are attractive for large-scale projects or organizations, but for small nonprofit organizations, it may not prove as successful. This is because small nonprofit organizations may not have the time or skills to collect and analyze data in ways that either test the program theory or

provide useful information to guide decisions or actions. Additionally, if program theory is used to develop accountability systems, there is a risk that staff will seek to achieve targets and stated objectives at the cost of sustaining the program.

Although there are benefits to theory-based evaluations, there are also costs. One problem Weiss (2004) identifies with theory-based evaluation is that there is not a consensus in the evaluation field on the definition of theory. For some evaluators, theory refers to the social science version of theory. For other evaluators, it means a road map or sequence of events that take place along the journey of a program or project. Despite this problem with theory-based evaluation, Weiss suggests that this method of evaluation has become popular for three reasons. First, it provides a logical framework for planning data collection. The second reason is that it helps evaluators try to attribute causality in evaluating complex program where randomized assignments are impossible. Lastly, it helps the evaluator communicate why and how the program works.

Chen and Rossi (1992) also point to potential problems in using theory-based evaluations. They suggest that theory-driven evaluations require evaluators to construct a program theory underlying a program. A program theory then provides a guide for evaluation activities, including the selection of research method(s) for data collection and data analysis. In this sense, in evaluating the same program, theory-driven evaluations require extra effort and time for

specifying program theory and for gathering additional data. All other things being constant, theory-driven evaluations are more expensive.

One of the approaches that have been derived from the theory-based evaluation that has recently gained momentum among funders and agencies in the nonprofit sector is outcome measurement or outcome evaluation. The following section provides a summary of this approach to evaluation and its recent affect on the nonprofit sector.

Outcome Evaluation

Many state and local government agencies, foundations, managed care systems, and accrediting bodies have added outcome measurement to the list of performance and accountability measures they require of nonprofit organizations within their sphere (United Way of America 2000). Outcome measurement, outcome evaluation or impact evaluation assesses the short and long-term results of a project and seeks to measure the changes brought about by the project. This type of evaluation measures the extent to which the program's stated goals and objective were achieved and determines any unintended consequences of the program and whether these were positive or negative (Formative Evaluation Research Associates 2005). Outcome evaluations focus not only on the ultimate outcomes expected, but also attempts to discover unanticipated or important

27

interim outcomes. Outcome evaluation helps determine how well a project is progressing in an effort to improve the well-being of the target population.

Nonprofit organizations have focused their resources on outcome evaluations, despite the barriers for some nonprofit organizations to conduct this method of evaluation. Rossi, Lispey and Freeman (2004) suggest that to conduct an outcome evaluation, the evaluator must design a study capable of establishing the status of the program recipients on relevant outcome measures, and also capable of estimating what their status would be had they not received the intervention. They note that outcome evaluation is very demanding of expertise, time and resources and it is difficult to set up properly within the constraints of routine program operations. They insist that outcome evaluations are most appropriate for mature, stable programs with a well-defined program model and a clear use for the results that justifies the effort required. Despite this warning, many foundations are requiring outcome evaluations from nonprofit organizations, and most nonprofit organizations have no choice but to commit their limited funds to the organization or the end product produces a scaled-down version of an outcomes evaluation.

The United Way of America (2000) released the findings of a survey of approximately 400 agencies and found that while there were challenges, outcome measurement produced rewards for agencies that implemented it carefully and used it as a management tool. Because, they stated, their focus was on results, the

feedback that outcome measurement provides and the data it produced offered two distinct benefits. Outcome measurement helped increase effectiveness in the agency's services and it helped communicate the value of what the agencies did in their operations.

Importance of Evaluation

Why should we care about program evaluation? Program evaluation is important for several reasons. First, it can provide valuable information about a program's process, costs and results. It can provide feedback on a program's successes, failures, strengths and weaknesses. It can help determine if the allocation of resources is being expended in an appropriate manner, or it can provide direction for a better use of funds. Second, it can provide information on the impact of the intervention on the program's clients or the intended user. Program evaluation can give program staff information about what works, and it can document unanticipated results. Third, evaluation can be used for oversight and compliance to funding sources. It also can be used for organizational and program improvement.

Evaluation is a means to strengthen development, whether it is human, economic, or other forms of development (Sanders 2002). It is agreed by directors of nonprofit agencies that funders and government agencies often mandate evaluations of supported programs. This is done because legislatures, or in the

case of foundations, boards of directors, require it for accountability. But the question remains: Is evaluation gaining value as a practice? Sanders (2002) suggests that the frequent mandating of evaluations by government institutions and other funders does not indicate a widespread valuing of evaluation that will continue to appear even in the absence of legislative or board mandates. Nonprofit organizations need to work with stakeholders (board, staff and community) and clients to demonstrate the importance of evaluation, even when it is not a mandate.

Hatry et al. (1996) give several reasons why nonprofit organizations should conduct program evaluations. Their primary reason for conducting evaluations is that outcome measures provide a process "to see if programs really make a difference in the lives of people." The authors note that outcome measures help programs improve services by providing a learning loop that feeds information back into programs to determine how well the program is doing. They also suggest that outcome measurement can help nonprofit organizations do the following:

- Recruit and retain talented staff

- Enlist and motivate able volunteers

- Attract new clients or participants

- Engage collaboration

30

- Garner support for innovative efforts

- Win designation as a model or demonstration site

- Retain or increase funding; and to help gain favorable public recognition

The authors also list benefits to program managers of nonprofit organizations. As a management tool, the benefits to outcome measures include the ability to strengthen existing services; to help target effective services for expansion; to help identify staff and volunteer training needs; to help develop and justify budgets; to prepare long-range plans; and to help focus board members' attention on programmatic issues. Despite Hatry's et al. (1996), compelling list of benefits that underscore why nonprofit organizations should use evaluations to improve their services, most nonprofit organizations have failed to systematically use program evaluations. Even Murray (2001) concludes from his research that there still appears to be a long way to go before there will be available a tried and tested evaluation system that can be applied by most nonprofit organizations to reveal a valid picture of how well they are performing.

The W.K. Kellogg Foundation (1998) suggests that evaluation is important in determining how well a program is operating, but also it has a responsibility to tell how the program is impacting society. The foundation suggests that demonstrating effectiveness and measuring impact are important and valuable. They suggest that it is equally important to focus on gathering and analyzing data,

31

which will help improve our social initiatives. The foundation reports that when the balance is shifted too far to focus on measuring statistically significant changes in quantifiable outcomes, scholars miss important parts of the picture. They contend that this ultimately hinders our ability to understand the richness and complexity of contemporary human service programs.

Evaluations for nonprofit organizations are usually subjected to two basic constraints: time and money. As a result of these constraints, many of the basic principles of impact evaluation design in which most nonprofits use to assess their programs are often sacrificed. Bamberger, Rugh, Church, and Fort (2004) suggest that most policy makers, managers, and funding agencies need answers to relatively straightforward questions that do not require great methodological sophistication. The authors state that the pressures of working under budget and time constraints has often resulted in a lack of attention to sound research design, with limited attention being given to identifying and addressing factors affecting the validity of the findings. The authors offer a Shoestring Evaluation approach to assist evaluators in conducting methodologically sound evaluations when operating with budget and time constraints, and with limitations on the types of data to which they have access. The Shoestring Evaluation includes,

1. **Planning and scoping the evaluation.** This involves having a clear understanding of the nonprofit organization's priorities and information

32

needs to eliminate unnecessary data collection and analysis, hence reducing the cost and time of the evaluation

2. **Addressing Budget Constraints**. This includes reviewing alternative evaluation designs and considering possible ways to reduce sample size while ensuring the required level of statistical precision. The authors also suggest a review of all possible options for data collection to identify the most economical methods consistent with quality concerns. They also suggest looking for reliable secondary data to reduce data collection costs.

3. **Addressing Time Constraints**. A consideration for nonprofits is to conduct an internal study by the agency prior to the arrival of a time-constrained external consultant. The authors also suggest formatting project records to include and facilitate the analysis of critical data for impact analysis.

4. **Addressing Data Constraints.** A consideration for nonprofits is to identify and assess possible secondary data which could potentially provide baseline data on project or control groups. Also, nonprofits should identify and assess the extent to which project records can be used to provide baseline data. However, nonprofit organizations need to consider the need for special data collection methods to obtain sensitive data or to include difficult-to-reach groups.

5. **Identifying Threats to the Validity and Adequacy of the Evaluation Design and Conclusions.** The authors recommend assessing the importance of each threat for the purposes of the evaluation and for the validity of the conclusions and recommendations.

6. **Addressing Identified Weaknesses and Strengthening the Evaluation Design and Analysis.** The authors recommend considering possible measures that can be taken to eliminate or reduce each threat to the validity and apply the most appropriate measure within the given scenario. They also recommend disclosing the threat if it cannot be corrected in the evaluation report with a discussion of the implications for the conclusions and recommendations of the evaluation.

Evaluation is an important tool for nonprofit organizations and should be an integral part of the organization's operations. Aaron Wildavsky (1972) said it best when he said, "The ideal organization would be self-evaluating. It would continuously monitor its own activities so as to determine whether it was meeting its goals or even whether these goals should continue to prevail." This is especially true for nonprofit organizations. Evaluation is an important part of the operation, but few nonprofit organizations conduct evaluations unless they are coerced by the funder.

There are many reasons why nonprofit organizations should evaluate their programs. Michael Scriven (1991) argues the importance of evaluation in pragmatic, ethical, social and business, intellectual and personal terms. He states the importance of evaluation in pragmatic terms by saying that bad products and services cost lives and health, destroy the quality of life, and waste the resources of those who cannot afford waste. He states the importance of evaluation in ethical terms by saying evaluation is a tool in the service of justice. Scriven states the importance of evaluation in social and business terms by saying that evaluation directs efforts where it is most needed and endorses the new and better way when it is better than the traditional way, and the traditional way where it is better than the new high-tech way. He explains the importance of evaluation in intellectual terms when he says it refines the tools of thought. Lastly, he demonstrates the importance of evaluation in personal terms by saying that it provides the only basis for justifiable self-esteem.

The discussion on the importance of evaluation has provided several reasons why nonprofit organizations should evaluate their programs. I advocate that the importance of evaluations for nonprofit organizations should be to improve the programs and services to better serve the clients and stakeholders the programs and services are intended to impact.

There are several types of evaluations that nonprofit organizations can conduct. Choosing a method of evaluation is important, but the most important thing for consideration is that they evaluate. Nonprofit organizations do conduct evaluations. But, do they use the evaluation information and how do they use it? The next section informs the reader of evaluation utilization.

Evaluation Utilization

The lack of systematic evaluations conducted and used by nonprofit organizations may be tied to the limited value placed on evaluation reports by nonprofit organizations. Along with assessing the value and worth of a program, evaluations should aid decision-making, support change management, and improve processes (Hubelbank 2001). There must be a high level of value placed on evaluation in order for nonprofit organizations to begin to institutionalize evaluation practices.

In fact, the discussion of utilization of evaluation has been most noted by Michael Quinn Patton (1997) who coined the term utilization-evaluation. Patton states that intended users are more likely to use evaluations if they understand and feel ownership of the evaluation process and findings. Patton also notes that utilization-focused program evaluation is evaluation done for and with specific, intended primary users for specific, intended uses.

36

Traditionally, evaluation utilization is thought to have four components: instrumental, conceptual, process and symbolic (Johnson 1998). Instrumental use occurs when a decision or action follows from the evaluation. Instrumental use is results-based use. Conceptual use is the use of evaluations to influence thinking about issues in a general way. Conceptual use is the development of understandings, opinions, and attitudes toward a program, project or evaluation based on the experiences of past evaluations.

Process use occurs when there are changes in individual thinking or behavior, and program or organizational changes in procedure and culture among those involved in the evaluation. Process use is apparent when learning occurs and behavior changes within the individual and/organization. Symbolic use occurs when results are used for political self-interest, such as to legitimize a preset opinion or earlier decisions.

Nonprofit organizations that conduct evaluations use their results in various ways. Hatry and Lampkin (2001) suggest that for nonprofit organizations that track the outcomes of their work collect the data to respond to funder requirements. Some nonprofits have begun to use the information in their marketing efforts, but few actually use the data to improve their own services. Hatry and Lampkin advocate the use of outcome measurement data to make informed management decisions about ways to allocate scarce resources and

methods and approaches of delivering services that will help nonprofit practitioners, researchers, and funders improve their outcomes.

Nonprofit organizations must find significant value in evaluations to consider the trade-off in staff, time and funding that is directed to program implementation for an administrative report. Nonprofit organizations must view evaluation and evaluation use as a necessary component to providing services to the client as the service itself. This may require changing the organization's culture and involving numerous stakeholders in the process. Dibella (1990) states that evaluation reports are inanimate objects, and it takes human interest and personal action to use and implement evaluation findings and recommendations. The implications of evaluation must be transferred from written page to the agenda of program managers. The utilization of evaluation, through which program lessons are identified, usually demands changed behaviors or policies. This requires the shifting of priorities and the development of new action plans for the operational manager.

Four problems can potentially emerge in planning an evaluation that could possibly prevent the evaluation from being useful for the intended user. Joseph Wholey (1994) identifies these problems as the lack of agreement on the goals, objectives, side effects and performance criteria to be used in evaluating the program. The second problem is that program goals and objectives may be

38

unrealistic, given the resources that have been committed to the program and the program activities that are underway. The third problem is the unavailability of relevant information on program performance. The fourth potential problem in creating a useful evaluation is the inability of policy makers or managers to act on the basis of evaluation information. Wholey suggests the evaluator develop an evaluability assessment to help evaluators and program managers understand the expectations of those who have the most important influence over the program.

Getting nonprofit organizations to view evaluation as a tool for learning instead of a mandate from a funder or an additional administrative function is a challenge. Rossi, Freeman and Lipsey (2004) state that evaluation results are rarely intended by evaluators or evaluation sponsors to be "knowledge for knowledge's sake." Rather, they are intended to be useful, and to be used, by those with responsibility for making decisions about the program, whether at the day-to-day management level or at broader funding or policy levels.

The Canadian Centre for Philanthropy (2003) found that the majority of nonprofit organizations they surveyed used their evaluation information. Twenty-three percent of the nonprofits surveyed reported that they made very effective use of the evaluation information and 50% of the nonprofit organizations found effective use of the evaluation information. Twenty-one percent of the nonprofit organizations thought that they somewhat effectively used their evaluation results

39

and one percent found no use for the evaluation information. The study found that the greater the board's involvement in the evaluation process, the more likely it is that the evaluation information will be used effectively.

The study also found that nonprofit organizations used evaluation information to a great extent to improve programs and services (68%), and for strategic planning (55%). The survey found that evaluation information was least likely to be used for fundraising purposes or information sharing. The study reported that larger organizations with annual revenues of $1.5 million or more were more likely than other organizations to use evaluation information to improve programs and services, and for strategic planning purposes.

Nonprofit organizations, regardless of size, should use systematic evaluations in their strategic planning. Aimee Franklin (1999) writes that strategic plans have been described as the keys to creating high performing governmental organizations. She also suggests that evaluation use is important because it furthers the alignment of strategic management systems. Like governmental agencies, nonprofit organizations have focused on creating better organizations by seeking to develop or improve their strategic plans. Most nonprofit organizations focus on the process of strategic planning more than systematic evaluation and evaluation use. Evaluation use and strategic planning should co-exist in the nonprofit sector.

40

Brett, Hill-Mead and Wu's (2000) examination of evaluation use of a nonprofit organization demonstrated the complexities and challenges in evaluation and evaluation use. The organization they examined was fortunate to have an evaluation department with a professional director of evaluation on the staff. Most nonprofit organizations rarely have this opportunity. This organization had staff time devoted to evaluation, expertise and funding for evaluation, but found it difficult to implement evaluation use at the national and site levels. The authors suggest that establishing a culture of evaluation and evaluation use may be a gradual process so that individuals may find uses for data in their work and embrace the process that led them to the data. They also note that evaluation is a key capacity for an organization that wants to be both a learning organization and one that is "built to last."

Fitzpatrick, Sanders and Worthen (1994) give several examples of evaluation uses in public and nonprofit sectors. They suggest evaluation can be used to:

- Decide whether to implement an urban development program

- Establish the value of a job-training program

- Decide whether to modify a low-cost housing project's rental policies

- Improve a recruitment program for blood donors

- Determine the impact of a prison's early-release program on recidivism

- Gauge community reaction to proposed fire-burning restrictions to improve air quality

- Determine the cost-benefit contribution of a new sports stadium for a metropolitan area

Process use is a very effective type of use for nonprofit organizations. According to Patton (1997), a major development in evaluation in the last decade has been the emergence of process use as an important evaluative contribution. Process use is distinguished from findings use and is indicated by changes in thinking and behavior, and program or organizational changes in procedures and culture stemming from the learning that occurs during the evaluation process. This is very important in the nonprofit sector. Not only is evaluation used to determine if program and projects are working as they are intended, but also learning takes place within the organization and possibility the community.

Patton (1997) notes that the evidence of process use is represented by the following kind of statement after an evaluation: "The impact on our program came not just from the findings, but also from going through the thinking process that the evaluation required." For many nonprofits, this statement is replayed many times with staff and their board when evaluation is implemented and the process is complete. Therefore, evaluations can have dual tracks of impact. The first track of impact for the nonprofit organization is that it uses the findings of

the evaluation to improve its program or organization. The second track is that the evaluation process becomes a learning tool for the organization. It helps people in the programs learn to think and engage each other evaluatively.

Patton (2004) suggests that it is not enough to have trustworthy and accurate information. People must also know how to use information to weigh evidence, consider contradictions and inconsistencies, articulate values and examine assumptions. This is particularly true for nonprofit and community-based organizations.

Andrew Mott (2003) also suggests that the dual evaluation utilization track is the best approach for community based organizations, particularly grassroots organizations. Mott states that evaluations of grassroots efforts must be designed to help organizations learn and build capacity. He notes that for foundations that fund community organizing and other grassroots efforts, an emphasis on internal learning and capacity building is crucial. He suggests that without strong, increasingly knowledgeable, and competent organizations to take the lead, foundation grants cannot lead to the desired impact. Like Patton, he suggests that while funders must continue to be concerned about tracking and assessing performance, the funders must become equally concerned about designing evaluation systems that build grantees' capacity and help them learn. Thus, the

dual track of findings use and process use must be integrated within the evaluation.

Along the same lines of thought, Jeanne Hubelbank (2001) states that along with assessing the value of a program, evaluations should aid in decision-making, support change management and improve processes. She states that for evaluation to have effective use, it has to be a shared educational collaborative between the evaluator and those involved in the program. Hubelbank (2001) proposes three conditions that should occur in an evaluation to affect change and move the evaluation report from the file drawer to active use. The three conditions for evaluation are informing, collaborating and mixing. She suggests that the evaluator should inform participants about evaluation thinking, approaches and procedures. Second, she suggests that collaboration with users should occur during all stages of the evaluation. Lastly, she suggests that there should be a mixture of evaluation approaches and methods.

Changes in the environment in which evaluators now operate have forced evaluators to think differently on who their end users are and how they will use the information. Ray Rist (2004) states that for nearly three decades, evaluators have debated the variety of uses for evaluation. He notes that an evaluation has been generally understood to be a self-contained intellectual or practical product intended to answer the information needs of an intended user. Therefore, the

44

debate about the use of evaluations has depended upon what evidence can be developed to support evaluations' direct, instrumental impact or enlightenment. He also acknowledges that recent developments on evaluation use have focused on process use or influence.

Rist (2004) sounds the alarm and calls attention to the fact that users of evaluative knowledge are now confronted with growing rivers of information and analysis systematically collected through carefully built monitoring systems. Users are fed with streams of information from the public, private, and nonprofit sector in country after country across the globe. He proclaims that we see governments, nongovernmental organizations, and the private sector all using new means of generating real-time, continuous flow of evaluative knowledge for management and corporate decisions. These new realities, Rist (2004) explains, blur and make obsolete the distinctions between direct and indirect use, between instrumental and enlightenment use and between short and long term use.

CHAPTER 3

CHALLENGES AND OPPORTUNITIES FOR NONPROFIT

ORGANIZATIONS

This chapter discusses the challenges and opportunities with respect to evaluation in the nonprofit sector. It provides a summary of the nonprofit literature with an emphasis on the evaluation literature. This chapter also reviews recent studies that have been conducted on nonprofit evaluation. It examines the problems nonprofit organizations encounter in conducting systematic evaluations. From this discussion, several research questions are proposed that guided the study.

Challenges Faced by Nonprofit Organizations

The literature on evaluations in the nonprofit sector has been limited. In general, when nonprofit organizations have engaged in their evaluation efforts, they have focused on output measures. Knowing how many people were served has been the focal point in evaluation measures used by nonprofits. De Vita, Fleming and Twombly (2001) suggest that there is a growing need among nonprofit organizations to evaluate their programs. They note that:

> "There is a persistent call for nonprofit organizations to demonstrate that their products and services are making a difference to society and that they are effectively using their resources heightens the need to measure and evaluate these

products and services. Funders and community leaders want to know how well a program is working and what it has accomplished".

In their assessment of nonprofit evaluation methods they state that nonprofit organizations are more adept at measuring outputs than outcomes and are only beginning to explore how to develop outcome measures. Corporate foundations that give to the nonprofit sector often look for something similar to business performance measures to make comparisons among programs, projects, and other funding opportunities. Many key business performance measures do not work for most not-for-profit organizations. For example, the "bottom line" measurement of profit or loss indicates how effective a business is at achieving its goal of generating profits for the owners. However, generating profits is not a goal for nonprofit organizations (Henderson, Chase, Woodson 2002).

Since most key business performance measures do not work for nonprofit organizations, Henderson, Chase and Woodson (2002) suggest that these groups need to develop specialized measures to help them achieve their goal-driven strategic plans. They suggest that nonprofit organizations are very accustomed to reporting input measures, but are still in the early stages of developing output and outcome measures. They note that one important consideration in developing performance measures is to make sure the organization has the time, resources and availability to track these output/outcome resources. They conclude that some

small nonprofit organizations may not have the time, resources and availability to track output/outcomes resources.

Another problem identified in the literature is the fact that nonprofit organizations have a multidimensional problem in performance measurement (Brooks 2002). On the one hand, nonprofit organizations must be cognizant of their bottom line or they risk bankruptcy and disappearance. On the other hand, the not-for-profit designation under the Internal Revenue Code is tied to performance of a public mission. Nonprofit organizations must develop multidimensional assessments of performance to address their multidimensional objectives.

Malcolm Macpherson (2001) also makes a distinction between performance indicators in business compared to nonprofit performance indicators. He states that in the industrial and commercial environments, process management dominates. Employee and customer opinion, captured in satisfaction surveys and organizational climate surveys, may be tracked and published, but the link to purpose, process and product is often weak. He asserts that in nonprofit organizations, the reverse tends to be the case. Customer/client data exists and is linked to purpose and outputs. This data is rarely linked to outcomes and process information is not usually gathered.

Bruce Glasrud (2001) identified problems nonprofit organizations have in conducting systematic evaluations. He states that to relieve people's anxiety about social change, nonprofits are under pressure from funders to apply outcome-based assessment measures to their mission and methods. He states that part of the push for outcome-based accountability comes from the performance-based corporate mindset that has been sweeping over the funder community. What is helpful for nonprofit organizations is a better connection between their inputs and results that is useful for planning.

The evaluation process that nonprofit organizations and their funders have placed into practice may set up forces counterproductive to their goals of benefiting the community. With funds for evaluation, staff, data-crunchers, and survey tools, outcome-based evaluation is an uncertain affair, even in the largest nonprofit organization. Glasrud (2001) states that even with sufficient resources and evaluation systems in place, what is measured is not always what needs measuring. Without adequate grounding in research methods, many nonprofit organizations have trouble coming to terms with what outcome is actually produced by their program or service. The problem with most current outcome measurement is that it is geared for immediate gratification. That need has tainted those who fund our human and social systems and the type of programs and services they fund.

Another problem for nonprofit organizations is their motivation for conducting

the evaluation may be different from that of the funder. Carson (2000) notes that

if a major motivation for conducting outcome evaluation is to direct funding to a

project with proven success, there is little evidence that this happens with any

frequency. Carson (2000) points out that a continuing source of tension between

the grant makers and nonprofit organizations is that there is seldom an agreement

beforehand about what benchmarks are important to measure and how the results

will be used.

Foundations have struggled with the concept of systematic evaluations. With

an endowment budget of more than $8 billion, the Robert Wood Johnson

Foundation (RWJF) continues to struggle and seek improvement in their

evaluation system. In 1993, the foundation embarked upon a journey to design

and implement an evaluation system. Like so many other organizations, it was

met with resistance from staff and the RWJF had the problem of managing

institutional change. In a case study that outlined the organization's challenges

and successes with implementing their evaluation system, Kevin Bolduc and Phil

Giudice (2004) describe the foundation's evaluation process. They state,

> "Over the years, the performance system has evolved from a
> simple summary of selected successes and challenges for RWJF to
> a sophisticated, comprehensive, board-level articulation of
> progress against the foundation's theories of change and indicators
> of performance. Even as it continues to evolve, the performance
> assessment system provides the staff and board a crucial common

ground for discussing the fundamental question of how RWJF is performing and how it could continue to learn and improve."

Weiss (1998) suggests that from the point of view of the funders who are taking a chance on an untried project, it is reasonable to require that there be some evidence on the extent to which the project is working. To the operators of a project, the demands of starting up and running the new program take priority. Plagued as they often are by immediate problems of staffing, budgeting, logistics, community relations and all the other trials of pioneers, they tend to neglect the evaluation. They see it mainly as a ritual designed to placate the funding bodies, without any real usefulness to them.

Vic Murray (2001) says the sad truth is that regular and systematic evaluation of programs, functions and organizations in this sector still is relatively uncommon. This is partly because when evaluation efforts do not produce the value for the money, they are quickly abandoned.

Nonprofit organizations that lack evaluations are limited in their ability to demonstrate positive change or results and do not have information necessary to make informed decisions about day-to-day management on future programming and resource allocations. The lack of evaluations performed by nonprofit organizations not only affects the nonprofit organization, but also the clients or customers they serve. Clients of nonprofit organizations are provided some unit of

51

service that usually is measured in hours served or product provided to the client. However, the measurement of the impact of change on the client is rarely known. Do the hours or service or products provided to the client result in an improvement to their lives?

The majority of nonprofit organizations resist conducting evaluations unless they are required by the funder. The reasons nonprofits give for not conducting evaluations range from lack of money, limited staff, it is time intensive, just to name a few. Nancy Rosenbaum (2003) asserts that saying evaluation is too expensive is a myth. She explains that as a director of a nonprofit organization, she was also reluctant to engage in evaluation until it was demanded by a funder. The result was a four-year research initiative that proved to be a very positive experience for the organization. Rosenbaum (2003) explains that the evaluation process provided an opportunity for staff to reflect critically on their work instead of functioning exclusively in program operations mode. Rosenbaum (2003) also suggests that beyond helping nonprofits improve programs, evaluation can leverage an organization's ability to attract support from the foundation community. She contends that positive evaluation results, when it comes from an independent evaluation, provide objective evidence that what you do works. An additional incentive for nonprofit organizations to conduct evaluations is that it

demonstrates an ongoing dedication to organizational improvement that can distinguish the organization as a leader in the field.

Rosenbaum (2003) acknowledges that there are costs to evaluation, but suggests that nonprofits should view this cost as an opportunity cost or the "value of what must be given up to obtain something else" (p.2). She admits that funders want nonprofit organizations to be able to demonstrate the effectiveness of the programs they fund, but few are willing to subsidize the work. She explains that this cost can be managed and provides some practical solutions. First, she suggests that nonprofits allocate a percentage of their general operating budget for research and evaluation that fits within the agency's strategic plan. Second, she advises building the evaluation into the program's budget and general operating budgets as a fixed cost line item. Her third suggestion involves obtaining interns from local universities to conduct a lost cost evaluation. Perhaps most importantly, she suggests getting started, even on a small scale, by evaluating a single program or a discrete set of activities with the plan to expand the evaluation efforts at a later date.

Some nonprofit organizations understand the value of systematic evaluations but have difficulties making the organizational transition. Paul Light (2000) offers a metaphor to describe the changes that are underway for nonprofit organizations. He describes the evaluation changes for nonprofit organizations as the tides of the

ocean, and the nonprofit sector is caught in the middle of an unrelenting contest between competing philosophies and advocates of reform, all of which produce significant motion back and forth across different reform ideologies. The problem, Light (2000) suggests, is that the nonprofit sector has yet to develop the knowledge base to help individual organizations choose the reform approach that benefits them the most. The sector, he adds, has yet to reach an agreement on any outcome measure at all for reform.

The discussion so far has focused on the pressures from foundation funders on the nonprofit sector to implement evaluation. Not only foundations funders, but the government sector also is demanding systematic evaluations from the nonprofit sector. Steven Rathgeb Smith (2000) notes that government and foundations are demanding that nonprofits pay greater attention to the outcomes of their services, even conditioning funding on specific performance targets. Smith (2004) notes that this emphasis on efficiency, outcomes measures and more competition has pushed nonprofits to behave more like for-profit, market organizations. Smith (2004) suggests that this shift may have positive benefits, especially given the long-standing inattention to outcomes and efficiency considerations on the part of many nonprofits. He cautions, however, that the push for efficiency may have a long-term cost for the community and the nonprofit organization.

A similar warning was given by Giudice and Bolduc (2004) about the need for nonprofit organizations to evaluate their programs. Good intentions and anecdotes do not systematically provide data that can allow for learning and improvement nor demonstrate clear accountability to key stakeholders (Giudice and Bolduc 2004).

Nonprofit organizations need more training in program evaluation. In order for nonprofit organizations to have the capacity to conduct systematic evaluations, the agencies must have staff trained in evaluations. Most nonprofit agencies are limited in conducting evaluations because of the limited staff that are trained in evaluation. The also are limited because of the lack of funds to maintain trained evaluators on staff or consultants. The lack of capacity for nonprofit organizations to implement and maintain systematic evaluations limits the nonprofit organization in its ability to implement effective services.

McKinsey and Company (2002) shed light on the critical importance of fiscal management and program evaluation for nonprofit organizations. They suggest that we should envision an organization that has a demonstrated record of success in delivering a particular program, but has very limited skills in such areas as financial management or program evaluation-- a common combination in the nonprofit sector. They note that this skill gap inherently compromises the ability to improve and expand services to more clients. Donors and government agencies,

for example, will be reluctant to dedicate significant resources to an enterprise with weak financial controls. Similarly, organizations that do not rigorously evaluate and measure the effectiveness of their programs have a hard time demonstrating the kind of tangible results that inspire donors. The authors challenge nonprofit organizations to be visionary and place emphasis on developing ways of quantifying the social impact of their organizations through evaluation and research programs.

De Vita, Fleming and Twombly (2001), in an effort to help nonprofit organizations become stronger and better organizations, developed a capacity building framework. Within the author's framework, they discuss the importance of evaluation in building and making strong nonprofit organizations. The framework consists of five components that are commonly found in all organizations and intermediary structures: vision and mission, leadership, resources, outreach, and products and services. The authors suggest that the five factors are interrelated and mutually dependent on one another. Each factor reinforces and bolsters the other factors in the model.

Their focus on evaluation is discussed within the products and services component. They note that conceptually organizational outputs and outcomes are the product of the multiple and cumulative interactions of vision and mission, leadership, resources and outreach. These components work together to create

effective output and outcomes, driving the framework and helping to shape the quality of the end product. The output and outcomes provide the feedback loop to the other elements in the framework and can enhance or diminish the nonprofit's availability and capacity. They state that poorly delivered products or services may result in fewer resources coming to the organization or signal the need to change leadership. In contrast, high-quality products and services can increase access to resources, create greater networks, give more visibility to the organization and strengthen leadership.

Funders play an important role in the implementation and maintenance of systematic evaluations for nonprofit organizations. Funders can also create other barriers for nonprofit organizations when conducting evaluations. Funders are usually the initiators of evaluations for most nonprofits. For foundations, requesting evaluations from nonprofit organizations is a touchy but inevitable subject. According to a report from the Grantmakers Evaluation Network (GEN) and Grantmakers for Effective Organizations (GEO) conference (2000), foundations enter the gray area of ethical uncertainty when they begin making grants for evaluation and organizational effectiveness. Whether the funder underwrites the cost of a consultant, the usual arrangement, or pursues the less common practice of directly advising its grantees, the situation is rife with

57

potential conflict in terms of power, boundaries, inadvertent messages and unintended consequences.

Lack of funds to train staff on evaluations or to hire trained evaluators is a problem for nonprofit organizations. Another problem for nonprofit organizations is knowing what to report and how to report the information to multiple funders. Paul Light (2000) suggests that those of us who advocate for nonprofit evaluations should not sell short our hopes for evaluation despite several countervailing trends and conditions. One of the most frequent conditions facing nonprofits is the trend to provide more information to multiple funders. He states,

> "Nonprofits now face a flood of well-intended reforms similar to
> those that have swamped government. In the name of efficiency,
> or waste-saving, or outcomes, or openness – each an undeniable
> virtue, and all subject to fanatical excess – the nonprofit sector
> must shrug through wave after wave of mandated changes
> originating with foundations and government agencies."

Despite the problems nonprofit organizations may face in conducting systematic evaluations, the literature suggests that nonprofit organizations must find ways to conduct evaluations. Reid Zimmerman (2003) gives some practical advice to nonprofits about evaluations. When asked, "why do nonprofits need to evaluate their programs," he responded by saying that public opinion is the strongest evaluator a nonprofit organization faces and could be the only one if the organization does not have data to back up its claims of community value. When

58

asked about evaluation use for nonprofits, he states that evaluation is meaningless when it is done simply to fulfill a requirement or to satisfy the board of directors.

Unfortunately for most nonprofit organizations, the majority perform evaluations for those two reasons. Zimmerman (2003) states that evaluation is most useful when it begins and ends with an organization's clients, customers and constituents in mind. He further states that the results of a well-done evaluation will challenge an organization to change and continue to evaluate its offerings. He also suggests that evaluation can help inform the organization about where to invest its resources. These improvements can have significant impact on the number of people served or the total resources expended to provide a particular service.

Zimmerman (2003) advises nonprofit organizations on the importance of conducting evaluations by stating that choosing not to evaluate a program or service will ultimately serve as its demise. He states that change is the constant in the nonprofit sector. Knowing what changes to make, when, and to what degree will help the organization be better informed after the evaluation. Change resulting from an evaluation can only improve a nonprofit organization's future.

As stated previously, nonprofit organizations have looked toward the business world for models to improve their organizations. Sawhill and Williamson (2001) shared that the lessons learned from the private sector have proven useful to

nonprofits, particularly in such areas as strategic planning, marketing, finance, information systems and organizational development. One area in the business sector that nonprofits have not been able to master is the duplication of the crisp, straightforward way that businesses measure their performance.

The authors suggest that nonprofits should measure their success by measuring their impact, activity and capacity. They found that nonprofits, rather than pouring resources into measuring mission, had more luck in setting a measurable, mission-oriented goal and then assessing progress against the goal. They also suggested that nonprofit organizations should develop measures that can be easily understood and communicated. This was as important as keeping the measures simple. Measures that require little interpretation or that can be presented graphically allow staff members, donors and key stakeholders to see how the organization is doing (Sawhill and Williamson 2001).

In an effort to uncover key predictors of quality in outcome evaluation plans, Poole, Davis, Reisman and Nelson (2001) offer advice to nonprofit organizations to improve their evaluation plans. First, they suggest that nonprofits recognize that outcome evaluations are here to stay. They state that funders have recognized the value of program accountability that goes beyond outputs and units-of-service reporting. They suggest that nonprofit management may best respond by planning to build capacity in program delivery by assigning

responsibility for collecting and reporting outcome data as part of staff duties and responsibilities. Second, nonprofit management must ensure that staff members have training and technical resources to design and implement outcome evaluation systems.

The authors make two additional suggestions relating to resources that are less than practical for most nonprofit organizations. First, they suggest that if management does not know how to design and implement these systems, then they should hire staff or consultants with the requisite knowledge and skills. The other suggestion is that management purchase the appropriate technology to store, retrieve and analyze outcome data. Both suggestions would be embraced by most nonprofit organizations if funding was available. In most instances, funders will not underwrite the cost for either expenditures.

Third, they suggest that management provide leadership and tangible support to help their agency make the transition by communicating a clear message of the overall importance of outcome evaluation to the agency, allocating resources to support planning and implementation processes, and building evaluation into performance appraisal systems for program coordinators and staff. Finally, they suggest that management involve staff and board members in the decision making process to improve practicality and feasibility of outcome evaluation plans.

Understanding that many nonprofit organizations may not have the resources to conduct some performance measures, and understanding that nonprofit organizations are under increasing demands to demonstrate they are making a difference, Vic Murray (2001) asks if there is a set of performance indicators that many different kinds of charitable organizations can easily apply and will yield the clear picture of impact that stakeholders want. Additionally, he asks if there is a universally agreed-on best practices applicable to all nonprofits that have been proven to increase performance. He answers both questions by saying not yet. However, he recommends the United Way's program outcomes, The Balanced Scorecard and the Canadian Comprehensive Auditing Foundation Framework for Performing Reporting as some processes that nonprofit organizations and their stakeholders can adopt that could help improve the analysis of performance.

Susan Sanow (2003) offers some generic aspects of what it takes to be a well-managed nonprofit organization. One of the aspects involves evaluation. She states that systematic ways of evaluating how well the organization is serving its various constituencies, responds to changes in client need, and addresses environmental/external changes is a key aspect of what it takes to be a well-managed nonprofit organization.

Are there tools and resources to help nonprofits conduct evaluations? Sandra Bozzo (2000) found that although there were numerous tools designed for use by

nonprofit organizations, many of these tools exist in unpublished forms and are not widely distributed. Additionally, many of the tools fail to take into account the diversity of nonprofit organizations in terms of their skill sets of their staff and volunteers, their financial resources, and the types of programs they deliver. Many nonprofit organizations with little or no internal capacity to undertake sophisticated data collection methods may find these resources intimidating.

Bozzo (2000) suggests that the manuals on evaluations for nonprofit organizations lead one to believe, through the manner in which the technical material is presented, that nonprofit organizations may not be able to conduct an evaluation without the help or support of an external evaluator. She states that there is room for substantial improvement in the type of information on evaluation that is made available to nonprofit organizations. Bozzo (2000) recommends that nonprofit organizations focus their evaluation efforts and have a clear idea of what they would like to achieve and who they would like to include before embarking on an evaluation process. She states that there is a large need to develop evaluation capacity in the nonprofit sector through education, training and skill building. By providing clear and uniform information, she suggests that subsets of the nonprofit sector can help to coordinate efforts in this area and accelerate the dissemination of information regarding approaches.

Abramson's (2003) report from a discussion among 24 nonprofit leaders echoed some of the resource concerns outlined by Bozzo. The report suggested that many nonprofit organizations are so busy doing their work that they do not take time to identify the outcomes they want. As a result, the nonprofit organizations may not know when they have been successful. The participants in the discussion reported that they need resources to undertake evaluations which they felt would affect efficiency in a positive manner. The participants in the discussion also stated that there was a need to understand the impact of various stakeholders in determining choice of outcomes.

Abramson reports that often at the core of these competing outcome expectations are differences in values. Different stakeholders that may include funders, the public, program staff, program directors, boards and clients, may value different outcomes. Since there are multiple preferred outcomes, organizations may always appear to be inefficient and ineffective to some stakeholders, regardless of how others prize the organization. The differences in values become even more acute as nonprofit organizations compete for funds. A profound but simple statement arises out of the discussion: "societal needs cannot be compared".

Doug Easterling (2000) shares the practical limitations to outcome evaluations for nonprofit organizations. He states that for nonprofit organizations, outcome

evaluations can be expensive; there can be difficulties in assessing outcomes precisely; and there can be fear in nonprofit organizations to conduct outcome evaluations. Easterling (2000) also validates previous authors' statements concerning nonprofits lack of resources and skills to conduct outcome evaluations. He states that measuring outcomes and impacts requires questionnaires, interviews, tracking clients after they leave the program, data management and analysis-tasks that impose a real burden on staff and clients. He further suggests that few nonprofit organizations have the expertise or staff to design rigorous studies, identify valid instruments and carry out sophisticated statistical analyses.

The United Way of America became one of the catalysts that propelled many nonprofit organizations into the systematic evaluation realm. In 1999, the United Way of America (2000) reported that more than 280 United Way agencies were working with agencies they fund on program outcome measurement. In January 2000, they reported that more than 400 United Way agencies across the country were asking agencies they fund to identify and measure program outcomes. They also reported that a small but growing number of United Ways were seeking to expand on program evaluation to measure their impact on the community. They admitted that there was a struggling relationship between program outcomes and

community outcomes and that achieving community change would require more than funding programs (United Way of American, 1999).

Warren Logan, President and CEO of the Chattanooga, Tennessee Urban League (Logan 2005) stated that his agency had a paradigm shift in thinking about how they measured what they did at their agency from a conversation he had with their local United Way representative. He states as follows:

> "A United Way staff person stated that they did not care if they served 50 or 5000 people, the question was did the service they provided change the lives of the people they served? This statement made me think differently about how we should do business."

Peter Hall (1994) has a pessimistic view about how foundations and nonprofits conduct and use evaluation information. Hall suggests that foundations paid more attention to the call to evaluate when foundations and nonprofits were being heavily scrutinized three decades ago. Hall proposes that today, because public sentiment is more favorable to nonprofits and foundations, the pressure is not as great to conduct rigorous evaluations. He states as follows:

> "The difference is that evaluation, while framed with the same rhetoric of rationality and purposiveness, in practice has taken on a very different function. Result-oriented boards demand proof of foundation efficacy, but are indifferent to evaluation findings. A foundation management pressures staff to do evaluation, but does not use the information it generates in planning. Foundation staffs do evaluation, but generally lack the resources or the competence to do it with any rigor. Grantees are compelled to participate in

66

evaluation, but – in instances where they have access to its products – seldom find it useful."

Is there hope for the nonprofit sector with limited resources and skills? How will management in the nonprofit sector cope with its deficiencies? Daniel Forbes (1998) suggests that nonprofit managers should educate themselves on ways to be effective and look for tools of effectiveness. He states as follows:

> "Dealing with questions of effectiveness can provoke managers to rethink their understandings of critical issues at every level of their organizations, leading, in some cases, to the reaffirmation of existing organizational practices and, in other cases, to search for new ways of doing things. In navigating through these individual judgments, contemporary managers will be called on to be reflective practitioners who possess an understanding of their organizations that transcends the limited knowledge contained in effectiveness measures and yet are capable of using such measures as reference points in understanding and communicating about their organizations. More practically, contemporary managers will also do well to be familiar with the criteria and measures of effectiveness that are relevant to the domains in which they operate so that they can critique, modify, and perhaps invent effectiveness measures as it is appropriate in the course of their work."

For nonprofit organizations to be successful in evaluation, they must realize that evaluation is both art and skill. There are many types of evaluation. The key is choosing the right type of evaluation for the program or organization that fits the need at a particular time. Many nonprofit organizations realize evaluations could provide information that they could use in future decision making. Several

67

scholars have noted that evaluation takes time, money and requires a certain amount of expertise to implement it successfully.

Vic Murray (2001) states that there are many tough choices to make in designing and implementing an evaluation system in nonprofit organizations. He suggests that evaluations for nonprofits can be categorized along four dimensions: level, focus, method, and pattern of involvement.

The first dimension is by level. This dimension looks at who or what to evaluate. Should the evaluation be about an individual, program or the organization? The second dimension he says nonprofit organizations should make the decision on how to evaluate is by focus. He suggests that the organization should decide if it should measure inputs, outputs, process or outcomes. Murray states that most of the evaluation that has been done in the nonprofit sector has been on measuring inputs and outputs. He also states that the bulk of the evaluation that occurs within the nonprofit sector occurs at the level of programs or projects.

Murray (2001) advises that there is nothing wrong with focusing evaluation at the program level, provided one does not need to make decisions that involve comparing one program's performance with another. When programs are being compared to other programs in the organizations, the agency should conduct an

organizational level evaluation. Murray (2001) cautions that it is quite possible that a program is working well but the organization is in trouble.

However, evaluation at the organizational level is suggested by Murray to be the most difficult to undertake. He says that evaluation at this level requires that there be a well worked out strategic plan in place which lays out the main strategic priorities that are the keys to success in the face of opportunities and threats in the organization's physical environment. The evaluation for the organization becomes a critical piece of the planning process that attempts to measure how efficiently the priority objectives are being met. Evaluation of the total organizational performance is complex, costly and requires a clear sense of organizational objectives as its starting point.

Today, there is a paradigm shift occurring in the nonprofit sector that is directing their attention to the actual results produced by a program or organization. The shift in evaluation focus from measuring inputs and outputs to measuring outcomes has been slow. This is because measuring outcomes for nonprofit organizations can sometimes be very difficult to do in practice, and it is more costly and time consuming than input or output evaluations.

Nonprofit organizations have a tendency to shy away from measuring outcomes because some objectives can be vague or difficult to measure. Also, some objectives may be very long term so that it would take years to yield

outcome data. For nonprofit organizations to be effective in outcome evaluations, they should consider measuring short-term outcomes that might logically lead to long-term outcomes. Even when short-term and long-term outcomes are clear, they may be too costly or time consuming to measure. In this case, Murray (2001) suggests choosing a measure in some combination of inputs and processes/outputs.

More importantly, he suggests building a logic model that is clear about its assumptions and specifies how the inputs and outputs are supposed to produce the desired outcomes. The logic model is the underlying theory for a program or project that links actions to results. Murray (2001) states that, unfortunately, many nonprofit organizations fail to think through the logic model behind their evaluation system, so they run the risk of measuring things that are not useful.

The third dimension Murray (2001) identifies as a type of evaluation is method. This dimension involves the "how to" of the evaluation. Will the evaluation involve professional evaluators who apply a variety of methods bringing special expertise and tools such as questionnaires, focus groups, interviews and the gathering of various statistics and report? Or, will the evaluation consist of more commonly used and fewer less sophisticated measurement methods which are typically less costly? Murray (2001) suggests that the irony of choosing the appropriate evaluation method for nonprofit

70

organizations is that many of those who are calling for more and better evaluation in the nonprofit sector have in mind what we call "full scale" evaluation. Yet most funders and many boards of directors are not willing to put up the considerable funds that this would cost. The choice of methods for evaluation will always be a matter of trade-offs between high quality, costly approaches and the lower quality, cheaper ones. Murray (2001) offers looking at two criterion when choosing an evaluation method: level of information for decision making and funders. That is, choosing an evaluation method to provide evaluation data to demonstrate the importance of understanding the past in order to make changes in the future. Or, choosing an evaluation method because of the level of competition for funds vis-à-vis a given funder. In other words, choosing the method based upon the amount of high quality information for making key internal policy decisions and for influencing powerful external stakeholders (funders).

The last dimension Murray (2001) outlines is conducting evaluations by pattern of involvement. This method takes into account the level of stakeholder participation. Who will be involved in the development of the tools and the interpretation of the results? Murray (2001) suggests that every effort should be made to involve those who will be affected by evaluation in the process of design and implementation. He offers that a full communication plan covering the reasons for the evaluation and how it will be implemented may be necessary. He

71

further suggests that if this is not possible, the evaluator should be aware of the risks caused by the strong tendency of evaluatees to want to "look good and avoid blame" and "manage to the measure," thereby negating the value of evaluation.

The role of evaluation for nonprofit organizations will depend on if the evaluator is part of the agency staff versus an external evaluator who works with the agency for special projects. Fitzpatrick, Sanders and Worthen (2004) suggest that nonprofit organizations will increasingly require the expertise of evaluators to help them meet the demands of government contracts, even as they work to improve their capacity to do their own evaluation. However, they note that many small nonprofit organizations will continue to rely on external evaluators for special projects.

The organizational structure and the relationships within the nonprofit organization are different for evaluators who work in the public sector. Understanding the relationships that exist between the director of the nonprofit organization, the board of directors and the community can assist the evaluator in the evaluation process.

Recent Studies on Conducting Evaluations in the Nonprofit Sector

Several studies on the nonprofit sector have focused on the sector's trend of directing its attention to evaluations. The Robert Wood Johnson Foundation

funded a study conducted by the Innovative Network on nonprofit evaluation (Fine, Thayer, and Coghian 1998). The research study was initially designed to look at two broad approaches to program evaluation: (1) the traditional model, where a neutral external evaluator collects and interprets quantitative and qualitative findings, and (2) the participatory model, which engages program staff, clients and other stakeholders in the evaluation process. The researchers' preliminary research revealed that nonprofit organizations used a number of evaluation approaches that did not fall neatly into the two categories. Unfortunately, the researchers shifted their focus from finding out how nonprofits conducted their evaluations and why they used a particular evaluation process, to the role of stakeholder participation in program evaluations. The researchers viewed this shift as a key difference between traditional and participatory approaches. The new focus sought to determine whether the degree of stakeholder participation was associated with particular organizational characteristics, evaluation characteristics and/or evaluation outcomes. The study also focused on how organizations used evaluation results and whether the uses were associated with different levels of stakeholder participation.

The Fine study (1998) found that the participation of stakeholders increased satisfaction with an evaluation and improved its design and usefulness. According to the study, the most useful evaluation began with a focused design, documented

73

successes, provided recommendations for program improvement, and contributed to an agency's planning effort. The study found that a recent trend in evaluations by nonprofit organizations focused on outcome measurements. The researchers also found that nonprofit organizations conducted evaluations primarily for current funders, followed by program staff and board members. Their research noted that the evaluations conducted included a mixture of quantitative and qualitative data collection methods.

The Fine study (1998) also concluded that organizations conducting an evaluation for volunteers were more likely to have high-participation evaluations than low participation. Organizations designing an evaluation primarily for strategic planning were more likely to have low participation than high participation. The researchers also concluded that respondents with high participation evaluation were more likely to think that their staff, board and funders found the evaluation highly credible than respondents with low participation evaluations.

The Fine study (1998) credited increased stakeholder involvement with improving evaluation design. Respondents believed that more involvement by stakeholders ensured that relevant questions were asked, which would lead to appropriate measures being selected and that resources were available to implement the evaluation and its recommendations. Participation was also

credited in increasing both stakeholders' understanding and appreciation of an agency and their understanding of the evaluation. The study also cited that most respondents stated the evaluations conducted by an external contractor were perceived to be more credible. The main reasons given for the credibility were objectivity and expertise of consultants.

Another recent study looks at how foundations are measuring performance (Putnam 2004). This study is important because how foundations measure themselves directly impacts how they require nonprofit organizations to measure their performance. The Putnam research sought to conduct a scan of the field to determine how foundations are conducting foundation-wide evaluation; what measures are being used; and what issues a foundation should consider before conducting a foundation-wide evaluation. The study defined foundation-wide evaluation as the "process through which foundations examine the overall value of their philanthropic activities." Foundation-wide evaluation examines the performance of the organization as a whole where the unit of analysis is the organization. Similar to nonprofit organizations, this research suggests that because of the real and perceived challenges involved in foundation-wide evaluations, few foundations engage in foundation-wide evaluations.

The Putnam study (2004) found six approaches to foundation-wide evaluation. The approaches include the balanced scorecard, the performance dashboard,

performance assessment, mission-level evaluation, theory of change and results-based accountability, and foundation-wide evaluation based on grant or program outcomes.

The balanced scorecard is a management system based on goal congruence as a means of improving performance (Dinesh 1998). The balanced scorecard is most used by corporations and Fortune 500 companies. This measurement system challenges organizations to translate their strategies into objectives, measures, targets, and initiatives in four "balanced" areas: customer, internal process, learning and growth, and financial. Putnam suggests that public agencies, health organizations, nonprofit organizations and foundations have modified the balanced scorecard with differing results (Putnam 2004).

The performance dashboard visually presents evaluation data in a brief document with an appendix of more detailed information for each measure. A performance dashboard includes the performance and evaluation information of greatest interest to a foundation, such as a summary of grant characteristics, investment performance, strategic alignment, grant highlights, administrative expenses, grantee perception, and other measures of impact. Putnam (2004) suggests that the performance dashboard is often used with a balanced scorecard.

The third approach to foundation-wide evaluation discovered in the study is the performance assessment, which includes a framework of four types of measures:

- Optimizing governance—Accountability, stewardship and active engagement

- Managing operations—Consistency with objectives, grantee selection process, grantee interactions, staff recruiting, review and retention, administrative expense, and endowment investment performance

- Setting the agenda and strategy—focus areas, goals and approach

- Achieving impact—program objectives, field effects, strengthening grantees, funding influence and leverage, and field effects

The fourth approach to foundation-wide evaluation identified in the Putman study (2004) is mission-level evaluation. This approach starts with clarifying the foundation's mission and creating an evaluation design tailored to the foundation to determine how well foundation strategies, grant making and operational processes contribute toward fulfilling the mission.

The fifth approach to foundation-wide evaluation is the theory of change and results-based accountability. This approach is also referred to as a theory of philanthropy and theory of grantmaking. These approaches encourage foundations to clarify the theory of change at the foundation level. This theory of change is

77

based on determining the foundation's intended results and its planned work for reaching the results.

The sixth approach found in the Putman study (2004) is the grant or program outcomes approach. This approach assesses the impact of grants, initiatives, and/or program areas as their approach to assessing foundation-wide impact. This means combining the performance of grantees, initiatives, and programs by examining the overall performance of the foundation's portfolio. This approach to evaluation is being strongly encouraged by foundations as the method for nonprofit organizations to use to conduct evaluations.

The scan in this study uncovers 20 different measures being used by foundations conducting foundation-wide evaluation. The study notes that no foundation uses all of these measures, and the utility of each depends on the foundation and what it wants to learn. Also, the study suggests that the foundation's use of a measure does not necessarily mean the measure is part of an overall strategy for foundation-wide evaluation.

An interesting finding from the Putman study (2004) revealed that foundations have difficulties finding resources for conducting evaluations like the nonprofit organizations they fund. Putnam noted that it was difficult for almost all foundation staff members and experts interviewed to quantify staff, time and financial resources needed to engage in foundation-wide evaluation. Several of

the foundations commented that the resource considerations were significant, especially at the front end of evaluation planning, but well worth the costs for clarifying mission, goals, strategies, operations, and programs. Putman reported that one of the evaluation experts estimated that a foundation should have three to five years to get everything in place.

According to a recent Center for Effective Philanthropy survey (2002) of 77 chief executive officers from 225 of the largest foundations in the United States, measures most commonly used to assess foundation performance in achieving social impact and operational goals included: grant, grantee and program evaluations (72 percent), administrative costs (37 percent), investment performance (22 percent), human resource-related measures (16 percent), strategic review (13 percent), changes in the field (13 percent) and changes in public policy (12 percent).

Herman and Renz (2002) studied a group of local nonprofit charitable organizations from 1993 to 2000 in Missouri to learn more about organizational effectiveness for nonprofit organizations. The authors offer nine propositions to help understand and explain nonprofit organization effectiveness. However, the nine propositions do not directly speak to nonprofit evaluation. The authors offer some advice to the reader by saying that nonprofit leaders and executives need to be very careful in their use of program outcomes assessments to judge nonprofit

organization effectiveness. They suggest that there are very few circumstances under which program outcomes could legitimately be considered to equal organizational effectiveness. They contend that such a conclusion would be valid, for example, in situations where the nonprofit organization conducts only one program, and there are no possible other explanations for outcomes (such as the effect of other programs or events), and when all core stakeholders are rational, objective, reasonably intelligent and have all the relevant data. The authors suggest that other forms of program evaluation have more utility. They suggest that there is a place for experimental and quasi-experimental program evaluations, but they usually are difficult, expensive and time consuming, and other forms of program evaluation are likely to have more relevance.

Bozzo and Hall (1999) conducted an assessment of the availability of tools that can help nonprofit organizations conduct evaluations and then appraised the adequacy of the tools. Their analysis was based upon a review of the evaluation literature that focused on materials that are tailored to the needs of nonprofit organizations and, in particular, practical evaluation tools such as guides and manuals. They discovered that there are a number of evaluation tools that have been designed specifically for nonprofit organizations. However, most of these tools are not readily accessible and many fail to take into account the diversity of nonprofit organizations in terms of the skill sets of their staff and volunteers, their

financial resources and the types of programs they deliver. The researchers found twenty-two evaluation tools and/or resources with only fourteen of the tools/resources having practical use for nonprofit organizations. They concluded that only five of the resources reviewed showed promise as a practical resource, but each of the five also had a number of shortcomings.

Bozzo and Hall (1999) learned from their discussions with practitioners and researchers that there is a larger focus on outcome measurement in the nonprofit sector, particularly in the United States. This is strongly related to the move of nonprofit organizations towards accreditation. For this reason, organizations are using program logic models to predict outcomes and to collect large volumes of information about their programs. They state that the manuals on outcome measurement are geared towards helping organizations generate detailed program information. They further explained that most of the manuals the authors reviewed raised doubts as to the ease with which nonprofit organizations were able to apply program logic models to evaluate their programs.

Bozzo and Hall (1999) report that very few of the resources they reviewed presented processes that could be entirely sustained from within an organization. They suggest that many of the resources referred to the external support that would be needed to either help design the evaluation, collect or analyze the data or assist in presenting findings. Most of the manuals they reviewed made little or

no reference to organizational capacity or resources that are required to conduct an evaluation.

O'Sullivan and Lesky (2000) explore evaluation efforts of nonprofit organizations that partnered to deliver program services. The study was designed to learn if the outcomes of partnerships were being evaluated, and if not, why not. The authors contend that without evaluation, ineffective programs may persist simply because they are subject to loose standards of accountability. The researchers conclude that grant requirements determined if and what efforts at evaluation took place. They note that if a grant required measurable outcomes, outcomes were written. If a grant required the collection and reporting of specific information, the information was collected and reported. No one reported collecting data on program components beyond what was required.

In a later study, O'Sullivan and Miller (2002) study the effects of multiple evaluation requests from multiple funders on nonprofit evaluations. The study uncovers how common multiple evaluation requests are; whether they vary by funder; what approaches to evaluation predominate; how grantees cope with multiple reporting requirements; and how the grantees use the evaluation information. The researchers studied projects that were funded for two years by the U.S. Department of Labor in collaboration with the Department of Justice Office of Juvenile Justice and Delinquency Prevention. The projects were to help

the funding agencies identify effective strategies to break the cycle of youth violence and to move youth offenders and at-risk youths into productive employment. The authors suggest that these projects represent the complex environment of contemporary nonprofits.

By examining the actual evaluation requirements, how programs comply with competing demands, and how programs use evaluation findings, the researchers shed light on the effectiveness of current evaluation approaches and practices for nonprofit organizations. The researchers examine the demands of the funders for information on program operations, effectiveness, efficiency, or adequacy. They look for implications that the information would be used to make program decisions and that the information was generated systematically. The researchers include demands based on accountability, which called for explicit information on performance.

The authors conclude that multiple funders meant collecting a considerable amount of data. Grants that are specifically related to educational and occupational attainment track similar outcomes. Other grants have different requirements, depending on the specific nature of the grant. The grantees view these data requirements as a cost of doing business. The degree of burden seems to depend on how a program manages its grants.

The researchers also find that the federal funders' demand for performance measures seem to have been developed from legislation that required an annual performance report from each federal agency. The annual reports compare program results to program goals. At the federal level, the authors suggest, the aggregation of the data eliminates the probability that the federal government will sanction a program with poor results. They also find that nonprofit organizations report making limited use of the performance data. Additionally, both performance measurements and process evaluation data quality are a major problem. The lack of consistent operational definitions and invalid data eliminated the possibility that funders or the nonprofit organizations could compare projects. The researchers note that while each project is unique in respect to its client population and community culture, comparing projects can help in setting benchmarks and linking variations to clients or services to outcomes.

Joanne Carmen (2003) studies nonprofit organizations that provide human services in the state of New York. She uses a framework of five organizational theories to understand program evaluation use and practice among nonprofit organizations. Her five theories include rational choice, strategic management, agency theory, resource dependence theory and institutional theory. Carmen suggests that each of the theories articulates different assumptions about how nonprofit organizations relate to their environment, and offers different

84

predictions for why nonprofit organizations would conduct program evaluation

and how they would use program evaluation.

Table 1 – Reasons Nonprofits Evaluate

Theoretical Explanations					
	Rational Choice	Strategic Management	Agency Theory	Resource Dependency	Institutional Theory
NP's Relationship with Environment					
Frequency	N/A	Episodic	Ongoing	Ongoing	Ongoing
Nature	N/A	Proactive	Reactive	Proactive	Reactive
Character	N/A	Reflective	Reflective	Linking	Legitimizing
Motivation/Purpose					
Drive	Internal Forces	Internal Forces	External Forces	Internal Forces	External Forces
Basis	Normative Reasons	Survival	Coercion	Survival	Normative Reasons
Primary Purpose	Gather information to make decisions	Gather information to make decisions	Gather information for funders	Gather information to secure resources	Gather information to seek legitimacy
Nonprofit Use					
	Make decisions that improve services & help reach goals	Make strategic choices about internal operations & external opportunities	Use it to report to convey information to funders --OR-- conceal information from funders	Use it as promotional or political tool; to create linkages; seek additional funding	Use is routine, --OR-- Use is symbolic

2

Carmen finds that evaluation is being conducted by nonprofit organizations on some level. Her findings suggests that evaluation is being done by nonprofit organizations internally, by the executive management staff, relying on a wide range of strategies, which range from simple reporting to more formal evaluation and performance measurement strategies. She concludes that the majority of the nonprofit agencies in the study identify their primary reason for doing evaluation corresponds to the traditional view that evaluation is a rational planning tool to be used for decision making purposes. A sizeable proportion of the organizations identify other reasons to explain why they engage in program evaluation and how they use the results. Some of the other reasons acknowledged other motives including responding to environmental opportunities and threats; the need to secure scarce resources; the need to fulfill funder requirements; and institutional cultures and rituals.

In 2000, a national study was conducted by the Canadian Centre for Philanthropy (2003) which assessed the evaluation practices, strengths, limitations and needs of Canadian nonprofit organizations and their funding sources. The survey presents findings of the first national survey of the evaluation practices, strengths, limitations, and needs of Canadian voluntary organizations and their

funders. The Canadian Centre for Philanthropy (CCP, 2003)) study surveyed 1,965 nonprofit organizations and 322 funders on a variety of aspects of evaluation including their evaluation practices, changing expectations for evaluation information, satisfaction with evaluation, and perceived problems and needs for assistance. This study is a partial replica of the CCP study. A review of the study's findings along with a comparison of this study's findings is provided in chapter seven.

Chapter Three outlined challenges and opportunities facing nonprofit organizations in conducting systematic evaluations. It provided a review of the nonprofit literature and the evaluation literature. The chapter also discussed many problems that nonprofit organizations face in conducting evaluations. The chapter proposed several research questions that guided this study.

The chapter highlighted three recent evaluation developments in the nonprofit sector. It also provided information on evaluation use. The following chapter gives a history of a national nonprofit organization, the National Urban League, and a discussion on its performance assessment tool.

CHAPTER 4

THE NATIONAL URBAN LEAGUE

This chapter provides a historical overview of the National Urban League and its affiliates in the United States. A discussion of this ninety-five year old organization and its relationship with the affiliates provides the background information on the population surveyed in this study. This chapter discusses current evaluation practices and the National Urban League's Performance Assessment tool. The chapter concludes with a discussion on the extent of systematic evaluations the National Urban League affiliates perform.

Description of the National Urban League (NUL)

The following information provides historical information on the National Urban League and its affiliate organizations. The affiliates of the National Urban League consists of over 100 nonprofit organizations with varying clientele, programs, budgets and personnel that operate as independent nonprofit organizations.

The Urban League movement was founded in 1910. It is the nation's oldest and largest community-based movement dedicated to empowering African Americans to enter the social and economic mainstream. The National Urban

League, headquartered in New York City, leads this non-profit, nonpartisan, community-based movement with over 100 affiliates in 34 states and the District of Columbia.

Playing a pivotal role in the 20th century Freedom Movement, the National Urban League grew out of a spontaneous grass roots movement for freedom and opportunity that came to be called the Black Migration. When the U.S. Supreme Court declared its approval of segregation in the 1896 *Plessey v. Ferguson* decision, the Court rapidly transformed what had been a trickle of African Americans northward into a flood. African Americans who traveled North soon discovered they had not escaped racial discrimination. Excluded from all but the most menial jobs in the larger society, victimized by poor housing and education, inexperienced in the ways of urban living, many lived in horrific social and economic conditions.

To capitalize on the opportunity to successfully adapt to urban life and to reduce the pervasive discrimination, Blacks needed help. That was the reason the Committee on Urban Conditions Among Negroes was established on September 29, 1910 in New York City. A year later, the Committee merged with the Committee for the Improvement of Industrial Conditions Among Negroes in New York (founded in New York in 1906), and the National League on Urban

Conditions Among Negroes. In 1920, the name was shortened to the National Urban League (New Pittsburg Courier, UL-4).

In its beginning, the organization counseled Black migrants from the South, helped train Black social workers and worked in various other ways to bring educational and employment opportunities to Blacks. Its research into the problems Blacks faced in employment opportunities, recreation, housing, health and sanitation and education spurred the League's quick growth. By the end of World War I, the organization had staff working in 30 cities.

In the Annals of the American Academy of Political and Social Science, Mr. Whitney M. Young Jr., (1965), Executive Director of the National Urban League, wrote about the National Urban League and its strategy. Young (1965) explained that the National Urban League was started with a small office and two staff members. But by 1965, the national organization had grown to more than 70 affiliates across the nation. Like today, the National Urban League's focus was to work to improve the education, employment, housing and health of African Americans.

In his article, Whitney Young (1965) stated that the National Urban League's programs were designed to bring about positive changes in the status of the African American population. "The League," Young states, "provides a unique

machinery for effective communication between white and Negro citizens to avoid social misunderstanding and strife."

Early in its development, the National Urban League set itself apart from other ethic organizations. Young (1965) explained that the League's method did not call for sit-ins and freedom rides, but described itself as an "action agency" with programs to eliminate barriers. It also was recognized as a vital part of the struggle for human rights and joined the 1963 March on Washington. In his address at the Lincoln Memorial, Young (1965) states:

> "The National Urban League is honored to be a participant on this historic occasion. Our presence not only reflects increased awareness of the Urban League's role but most important, it says and I hope loud and clear, that while intelligence, maturity and strategy dictate that as civil rights agencies, we use different methods, we are united as never before in the goal of securing first-class citizenship for all Americans now. This is the real significance of our March on Washington today... our march is a march for America."

Open its history books, research its archives, peruse its annual reports and journals, attend its annual conferences, review its research and it will be clear that the 90-plus year old National Urban League and its local affiliates have devoted themselves to improving the social and economic condition of African Americans through fair and equitable access to employment and economic justice (Spriggs and Bergeron 2002).

91

Today, the organization's mission is to enable African Americans to secure economic self-reliance, parity and power, and civil rights. The organization has a three-pronged strategy for pursuing the mission that includes ensuring that urban children are well-educated and equipped for economic self-reliance in the 21st century; helping adults attain economic self-sufficiency through good jobs, homeownership, entrepreneurship and wealth accumulation; and ensuring civil rights by eradicating all barriers to equal participation in the economic and social mainstream of America[3].

The Relationship between the National Urban League and its Affiliates

Affiliates are organized and incorporated individually and separately from the National Urban League as prescribed by the Terms of Affiliation (contract between the National Urban League and the affiliates) and in accordance with policies and standards of operation as set by the National Urban League. The National Urban League's operation and function is closely tied to the activities of the professionally staffed affiliates working in local communities in close relationship with the National Urban League staff. However, autonomy is maintained between the two entities.

---

[3] For additional history on the National Urban League see "The State of Black America" series published by the National Urban League.

Through this affiliation, the National Urban League carries out its mission through programs and services of the local affiliates. In return, the National Urban League provides each affiliate with research, public relations, fiscal consultation and other technical assistance, including but not limited to program consultation, training, as well as access to participation as subcontractors in National Urban League sponsored federal, foundation, and corporate funded national programs.

The National Urban League movement consists of several interrelated and mutually dependent components, chief among them is the National Office and the affiliates, which are independent corporations that subscribe to the policies, principles, and strategies of the National Urban League movement. It is through the efforts of the National Urban League and the affiliates working in concert that the goals of the movement are achieved (NUL 2002).

The National Urban League is a private, nonprofit corporation with a board of trustees responsible for establishing policy and direction of the movement. The National Urban League accomplishes its responsibilities to the movement with the advice of affiliate leadership. One of the roles of the National Urban League is to serve, support and develop the affiliates.

According to the *Policies, Standards and Procedures Manual for Affiliates of the National Urban League* (2002), the National Urban League's responsibilities to the affiliates include, but may not be limited to the following:

93

- Creating a structure for the governance of the movement

- Establishing terms and conditions for affiliation

- Establishing for the National Office and the affiliates standards and accountability, measures and review processes sufficient to ensure fulfillment of the mission

- Providing oversight and technical and programmatic support to affiliates

- Supporting the work of the Delegate Assembly (annual meeting of the membership of the National Urban League)

- Establishing and maintaining the Urban League Affiliate Manual and other resources

- Directing programmatic thrusts and public relations activities that reinforce the mission and image of the National Urban League movement

- Obtaining financial and other forms of support to enable the National Urban League and the affiliates to carry out mission related work

- Supporting the affiliates in mission related activities and in developing and strengthening affiliate infrastructure

- Implementing compliance procedures to protect the brand and viability of the movement

The affiliates and the National Urban League share an interdependent relationship in that both rely on the maintenance and strength of its name brand,

owned by the National Urban League. On the strength of this name brand, the affiliates in over 100 cities and the National Urban League obtain federal, state, and local funding.

NUL Performance Assessment

The board of directors of each affiliate must review their own accomplishments and progress annually, but every three years, an in-depth review is conducted through the National Urban League Performance Assessment. To meet their own accountabilities to the Delegate Assembly (elected representatives from each affiliate) and the National Board of Directors, the National Urban League, through the Programs and Affiliate committees, has developed administrative criteria and operational standards for an effective Urban League affiliate.

The criteria cover broad areas of an affiliate's responsibility for its internal operations. The performance measures were developed within the context of three criteria: organizational soundness, organizational vitality, and implementation of the mission. The Performance Assessment is intended to take a snap-shot of an affiliate's administrative operations. The Performance Assessment, however, does not evaluate local affiliate programs.

The National Urban League's Performance Assessment evaluates three distinct and interrelated components: strategic planning, tactical planning and review, and appraisal. The goal of the performance assessment is to provide a tool to assess the effectiveness of the affiliates. Organizational soundness requires that the affiliate conduct its operations consistent with the requirements, policies and standards of the National Urban League. This criteria reviews the internal structure of the organization and evaluates or measures organizational and compliance issues. Examples of data collection tools used to assess performance include affiliate bylaws, board minutes, board surveys, President/CEO reports, volunteer surveys, affiliate strategic plan, needs assessment, etc.

Organizational vitality requires that the affiliate has sufficient resources and assumes responsibility for managing them in order to ensure the continuation and expansion of the Urban League in the affiliate community. This criteria seeks to identify the level of fiscal and administrative capacity. Examples of the data collection tools to assess performance include, delegate lists, training records, committee lists, lists of collaborative groups, affiliate personnel policies, budget and salary structure, financial audit, financial reports, fund development strategies, etc.

Implementation of mission requires that the affiliate programs are delivered to the diverse population in its community. Although this criterion specifies the need

for the affiliate to implement programs within its community, it does not address if the affiliates programs are working effectively, nor does it evaluate if the programs are impacting the community it serves. This criterion seeks to understand if the affiliate attracts and maintains membership and serves its constituents from areas of its community. This criterion also monitors if the affiliate is conducting some level of evaluation in compliance with funding source requirements. Data collection tools include community profiles, strategic plans, client evaluations, reports to funders, etc.

Systematic Program Evaluation for Affiliates

Interviews with Senior Executive Vice President for Affiliate Development for the National Urban League, Annelle Lewis (2002, 2004), provide some insight into the current evaluation system used by National Urban League affiliates. While the National Urban League affiliates have measurement tools in place to report to funding sources and to demonstrate community needs within each program, a nationwide systematic approach to program evaluation is limited.

Ms. Lewis stated that approximately 25% of the National Urban League affiliates, who have active strategic plans in place or under development, systematically evaluate their programs with an internal or external evaluator that looks at program process, program impact, and program costs. She says, "the

number of affiliates who will begin performing program evaluations will increase with the National Urban League's performance assessment, Sarbanes-Oxley legislation and other funding requirements." Ms. Lewis states, "Most affiliates measure past program activities, the costs, and the results for their programs based upon whatever measurements the funder requires." She thinks the finance committee's use of cash flow projections and other financial analysis tools were used when management in the affiliate operations realized projects/programs were not paying for themselves and the general fund becomes over extended.

Chapter Six will provide the results to the survey that was distributed to the National Urban League affiliates concerning their practices in systematic evaluation and their evaluation utilization.

The following discussion provided a brief history of the National Urban League and its relationship with its affiliates. The chapter also provided a discussion on the national organization's assessment tool and its limitation in providing a systematic evaluation method for its affiliates.

The next chapter provides an explanation of the methodology of the study. The chapter discusses the survey method used in this study. Information is also provided about the population surveyed and how the data was collected.

CHAPTER 5

METHODOLGY

Chapter Five provides an explanation of the methodology used to answer the research questions in this study. The chapter discusses the survey method, the data collection process and provides information on the population surveyed. This chapter discusses the agency characteristics that were examined that helps in shedding light on the evaluation practices on nonprofit organizations.

This study of the National Urban League affiliates was designed to answer the following questions:

1. To what extent are nonprofit organizations attempting to evaluate their performance?

2. How do nonprofit organizations evaluate their performance?

3. What are funders' expectations regarding evaluation and what support do they provide for evaluation activities?

4. What do nonprofit organizations perceive to be their strengths and weaknesses in evaluations?

5. What are the perceived needs for resources to assist with evaluation?

6. What resources are needed to assist with better evaluation?

7. For those nonprofits who are doing evaluations, how do they use the results?

8. For nonprofits who are not conducting evaluations, why are they not performing them?

9. What will help nonprofits who do not conduct systematic evaluations as a core component of their operations begin to systematically evaluate their programs?

Survey Method

Partial Replica of the Canadian Centre for Philanthropy's Study

The survey instrument used for the National Urban League evaluation study was designed as a partial replica of the Canadian Centre for Philanthropy's study (2003) that sought to uncover nonprofit organization evaluation practices, strengths, limitations and needs of the Canadian voluntary organizations and their funders.

There were three components to the Canadian Centre for Philanthropy (CCP) research:

1. An initial set of focus groups was conducted with representatives of voluntary organizations and funders in twelve communities across Canada.

2. A national survey was conducted of 1,965 nonprofit organizations and 322 funders to validate the initial focus group findings and determine the extent to which they represented the views of nonprofit organizations and funders in Canada. The response rate from the survey for the nonprofit organizations was 40.4%.

3. A final set of consultations was conducted in the fall of 2001 with representatives of nonprofit organizations and funders to discuss the survey findings and to develop recommendations for developing tools and strategies to assist nonprofit organizations with their evaluation activities.

Acknowledging that nonprofit organizations are very diverse in nature, the Canadian Centre for Philanthropy's study (2003) looked at nonprofit organizations according to three main organizational characteristics: type of organization, as defined by an organization's major purpose or activity (e.g. Health, Housing, Arts and Culture, Social Services); size, as defined by annual revenues; and geographic location. The CCP study was intended to answer the following research questions:

1. To what extent are nonprofit organizations attempting to evaluate their performance?

2. What methods are organizations using to evaluate their performance?

3. What are funders' expectations regarding evaluation and what support do they provide for evaluation activities?

4. What do organizations perceive to be their strengths and their weaknesses in evaluation?

5. What are the perceived needs for resources to assist with evaluation?

6. What resources are needed to assist with better evaluation?

National Urban League Evaluation Study

As a partial replica of the CCP study, the National Urban League evaluation study only surveyed nonprofit organizations. This study did not seek to incorporate a funders' survey. A survey was distributed to 105 nonprofit agencies located in 105 different cities who are affiliates of the National Urban League.

The CCP study conducted their analyses around three main characteristics: type of organization, as defined by an organization's major purpose or activity (e.g. Health, Recreation, Arts and Culture, Social Services); size as defined by annual revenues and geographic location. The CCP study did not state how the regions were defined for the study

This study examines evaluation practices around three main characteristics: type of organization as defined by an organization's major purpose or activity (e.g. Housing, Education, Health, Employment Services, Economic Development,

and Civil Rights); agency size; and geographic location by region (Western, Central, Southern or Eastern). Agency size was determined by annual revenues and the number of full-time equivalents. Revenue size was grouped in the CCP study into four categories: agencies with revenue less than $125,000; agencies with revenues between $125,000 and $499,000; agencies with revenues between $500,000 and $1.5 million and agencies with revenues of more than $1.5 million. The agencies in this study are group by revenue size into three categories: agencies with revenue of less than $1 million; agencies with revenues between $1 million and $3 million; and agencies with revenues of more than $3 million.

Examining the organizations by major purpose or activities, the study sheds light on the type of programs or activities that are evaluated. Examining the organizations and size by annual revenues, and number of full-time equivalents, provides information on the fiscal capacity for evaluation practices. The study also looks at the geographic location of the agencies for trends across the country in nonprofit evaluation based on regions and examined the proximity of resources between nonprofits in order to look for opportunities to share evaluation expertise and resources. Geographic location was predetermined by the National Urban League. A map of the affiliates in each region is included in the Appendix A.

The Research Approach for NUL Affiliates

To explore the extent and use of evaluations in the nonprofit sector, the researcher selected the 105 National Urban League affiliates across the country to be the population for this study. The researcher sent a pre-survey letter to the President of the National Urban League, the Executive VP for Affiliate Development, the President of the Association of Executives (President of the affiliate CEOs), and to each of the Regional Presidents of the AOE. The purpose of the pre-survey letters was to explain the purpose of the study, the importance of the study and to ask for time on each of the Regional business meetings to explain and distribute the surveys to the affiliate CEOs. The researcher received endorsements from the National Urban League's President/CEO, Marc Morial and Annelle Lewis, Senior Executive Vice President for Affiliate Development for the National Urban League.

The survey tool was a questionnaire that the chief executive officers (CEOs) completed during the 2004 National Urban League Conference. See Appendix D for a copy of the survey instrument. The first distribution of the surveys took place at the four (4) regional Association of Executives (AOE) business meetings (Western, Southern, Central and Eastern regions) in July 2004 in Detroit at the National Urban League Conference. The researcher was placed on the agenda at

each of the business meetings to explain the purpose of the research and asked for cooperation from the Executives in returning the completed surveys.

The researcher also gained the assistance and support of the National Urban League office in the collection of the surveys by allowing a drop-off box to be conveniently placed at the CEO conference registration sites at the hotel and convention center. Approximately one-half of the surveys were completed and returned at the National Urban League assembly. Those who did not complete and return their surveys were asked a second time to participate by mail. Fifty-four completed surveys were returned from the 105 affiliate agencies. A list of all of the conference attendees was obtained from the National Urban League staff in order to conduct follow-up interviews, send survey information to CEOs who attended the conference but did not complete the survey, and to send survey information to affiliate CEOs who did not attend the conference. CEOs that did not complete the survey at the conference were sent a cover letter and questionnaire via email or fax. For those who requested the questionnaire by mail, a paper version of the questionnaire with a self-addressed postage-paid envelope was sent. A follow-up reminder letter was sent via E-mail, fax or mail approximately two weeks after the original mailing. Four more surveys were received as a result of follow-up correspondence. A final attempt to reach the

non-respondents conducted through a small email survey. The affiliates were asked:

1. Did you evaluate your programs within the last 12 months?

2. Did you conduct outcomes evaluations within the last 12 months?

3. How many full-time equivalents (FTEs) do you have?

4. Estimate your annual revenue for your affiliate?

5. Did your affiliate receive evaluation training in the last 12 months?

Two affiliates responded to the email. A total of 60 affiliates participated in the survey which yielded a response rate of 57%.

The population surveyed provides a good representative sample of nonprofit organization operations around the country because of the agencies' diverse size and services and the dispersion of geographic locations.

Six follow-up interviews were conducted with a purposive sample of respondents to clarify responses. The researcher interviewed affiliates who vary in evaluation expertise, budget size and agency characteristics. The follow-up interviews assisted in providing additional information that may not be apparent through the survey data.

Also used in this study were the observations by the researcher as a survey participant in the study and the observations gleamed from two years of participation in nonprofit evaluation training and implementation. As President

and Chief Executive Officer of the Urban League of Greater Oklahoma City and the researcher in this study, I included the Urban League of Greater Oklahoma City's training and learning experiences in evaluation in the study.

This chapter provided information on the methodology used to answer the research questions in this study. The chapter discussed the survey method, the data collection process and provided information on the population surveyed. This chapter discussed the agency characteristics examined that were used to shed light on the evaluation practices on the nonprofit organizations in this study.

CHAPTER 6

DATA ANALYSIS

Chapter five provided information on the methodology of the study including the survey method used, the population surveyed and the method of data collection. This chapter will present the characteristics of the population surveyed and present the findings on evaluation practices for nonprofit organizations. This chapter will also explore the extent of nonprofit evaluation. It will reveal nonprofit funders' expectations and how funders support nonprofit evaluations. This chapter will provide an analysis of the nonprofit organization's perceived strengths and weaknesses along with uncovering the needs that nonprofits may have to conduct evaluations. This chapter will also discuss evaluation utilization for nonprofit organizations.

Characteristics of Agencies Surveyed

Nonprofit organizations vary in many different ways. The organizations vary mostly in size, by type of services provided and geographic location. This dissertation analyzes the data based upon three agency characteristics: agency sized defined by annual revenue, and number of full-time equivalents, types of services provided and geographic location. The size of the agencies is analyzed in

several ways. For this study, size is defined as the agencies annual revenues and the number of its full-time employees (FTEs). Another determinant of agency size is the annual budget of the program, the budget of the product or the budget of the services that is evaluated.

The National Urban League affiliates or agencies were chosen for this survey because they are diverse in size, but also because the agencies are dispersed in more than one hundred cities around the United States. The affiliates of the National Urban League provide services based on the geographic need of its target population. The agencies are affiliated based on their agency brand, but their services and the size of the affiliates vary depending on local community needs. Many programs are funded through a variety of sources including, federal government, state government, local government, and foundations.

There are agencies of the National Urban League whose mission is to deliver housing in distressed neighborhoods, while another's focus is to educate children in areas where schools are not performing. The agencies who participated in this study were asked to list their top three programs or services they provide that supported their mission.

Not all Urban League affiliates deliver the same services in every city. Because the National Urban League was founded in the early 1900s to assist African Americans in civil rights and equality issues, many people believe the

agencies service function is primarily focused on civil rights. The top three service areas by National Urban League affiliates in this survey were education, employment and housing. Services pertaining to civil rights were ranked last.

Table 2 – Most Important Service Provided Based on Organization's Mission by Percentage (N = 56)

Education	54%	32%	9%	5%
Employment	36	38	16	11
Housing	5	14	21	59
Health	2	2	13	84
Economic Development	0	7	9	84
Civil Rights	0	5	31	60

Table 2 summarizes the results and shows that more than half of the affiliates (54%) respond that education is their top priority, and more than thirty percent (32%) indicate that education is the next important service they provide. Educational services provided by the National Urban League affiliates include, but are not limited to Early HeadStart programs, HeadStart programs, after school programs, adult literacy programs including GED services, and computer literacy programs.

Employment is the second service area that affiliates say they focus their mission. More than seventy percent of the agencies say that this is a major focus of their mission. This program specializes in locating employment for individuals who are the "hardest to employ." Employment services include, but are not

110

limited to job skills training; resume building and management; internet access and posting; interview skills training; educational counseling and assistance; college application and information assistance; college tours; financial aid assistance; scholarships; and internship opportunities.

Other service areas provided by the affiliates of the National Urban League are housing services, health services, economic development services and services pertaining to civil rights. Housing services include, but are not limited to providing quality and affordable homes to low-to-moderate income families which includes multi-family housing development, homebuyer education, homeownership counseling, rehabilitation of homes and new home construction for low-income families.

Health services provided by the National Urban League affiliates also vary. These services include, but are limited to referrals for health assistance, mobile health screenings and testing, health education and information to African Americans concerning cancer prevention, hypertension, sickle cell, heart disease, HIV/AIDS prevention and diabetes.

Economic development services provided by the National Urban League affiliates have primarily focused on minority small business development. The services provided include management and technical assistance to minority clients who wish to start a business or expand an existing business. Agencies develop

111

business plans, marketing plans and feasibility studies. Economic literacy workshops are frequently conducted to expand the minority community's knowledge about business loans and business and financial management. The affiliates services have recently expanded to not only include services for start-up businesses, but also the focus is being directed to wealth creation for African Americans businesses through new market tax credits.

Civil rights issues were the catalyst for the development of many of the Urban League programs and services. Ironically, sixty percent of the affiliates responded that civil rights issues were not their main focus as it related to their mission. Civil rights issues for most National Urban League affiliates focused on equality in pay, police brutality issues, equal access to jobs, racial profiling, etc.

The size of an agency based on annual revenues was an important characteristic in this study. Carol Weiss (1998) reminds us that nonprofit organizations are often plagued by immediate problems of staffing and budgeting so they tend to neglect evaluation. An analysis of agency revenue helps access if larger agencies are more effective than smaller agencies in conducting evaluations. The fifty-eight agencies in this study were grouped into three categories according to the size of their annual revenues. Fifteen agencies or twenty-six percent of the respondents had annual revenues of less than $1 million.

Twenty-seven agencies (48%) had revenues between $1 million and $3 million. The remaining agencies (26%) had annual revenues of more than $3 million.

The agencies were also asked to report on the budget size of their program, project or services they evaluated. The program budget size was grouped into three categories and was categorized based on their self reported program, project or services budget. Thirty-nine percent of the agencies reported that they had program, projects or services budgets of less than $100,000, and thirty-one percent reported they had program budgets between $100,001 and $499,999. Also, there were thirty percent of the agencies who reported they had program budgets of $500,000 or more.

This characteristic is important in determining if the agency has the capacity to conduct the evaluation based upon the size of the program budget. It is also important in determining if program budgets could be a future resource for funds to conduct systematic nonprofit evaluations. Foundations and governmental agencies are now being more flexible when granting program dollars to include funds for evaluation. Nancy Rosenbaum (2003) says that nonprofit organizations should allocate a percentage of their general operating budget for evaluation that fits within the agency's strategic plan. She also recommends that agencies should build the evaluation into the program's proposal budget as a fixed cost line item.

Therefore, the size of the agency's annual revenues and the size of the program budget are important characteristics in this study.

The size of an agency's staff is also a good indicator of agency capacity. Most of the research points to the lack of staff capacity to conduct evaluations. Putnam (2004) noted that it was difficult to quantify staff, time and financial resources needed to engage in evaluation. Staff size in this study is measured in full-time equivalents (FTEs). The FTEs are grouped into three categories: agencies with less than ten FTEs; agencies who had between eleven and twenty-five FTEs; and agencies with more than twenty-five FTEs.

The geographic location of the agency is an important characteristic of this study for several reasons. First, the National Urban League affiliates are dispersed geographically in more than one hundred cities and urban communities around the nation. This geographic dispersion provides an indication about what nonprofit organizations are doing concerning systematic evaluation around the country. It also will help identify where certain regions of the country are conducting evaluations and where there may be opportunities for share resources in other regions to increase systematic evaluations. Geographic location was grouped into previously determined regional boundaries by the National Urban League. The regions are divided into four categories: East, South, Central and West. A map of the 105 affiliates is provided in the appendix.

There are twenty seven agencies in the Eastern region and thirty agencies in the Southern region. There are thirty-eight agencies in the Central region and ten agencies in the Western region. Of the fifty-eight respondents to the survey, only six agencies (10%) responded to the survey from the East. The Southern region had the largest respondents with twenty-two participants (38%) followed by the central region with twenty-one participants (36%). The Western region had nine out of the ten agencies in their region to complete the survey, and they represented sixteen percent of the respondents in the survey.

Evaluation Practices Revealed

<u>Extent of Nonprofit Evaluation</u>

This study seeks to: determine the extent to which nonprofit organizations are attempting to evaluate their performance; identify which nonprofit organizations are doing evaluations; why they are performing evaluations; and which nonprofit agencies evaluate the most.

To answer these questions, the participants in this study were asked if they performed evaluations in the past twelve months. The National Urban League affiliate study finds there were more affiliates conducting evaluations than originally predicted. The survey reveals that fifty out of the fifty-eight agencies (86%) have evaluated something in the past twelve months. Almost half of the

115

agencies surveyed (46%) conducted between four and six evaluations within the past twelve months. Another thirty-eight percent of the affiliates said they evaluated between one and three evaluations. The percentages are much lower (16%) for agencies who conducted seven or more evaluations. The agency's characteristics were examined to determine differences among the agency in frequency of evaluations. The study shows the agencies in the Southern and Western regions conducted more evaluations and found agencies with less revenue performed evaluations. Agencies with twenty-five or more full time staff equivalents conducted the most evaluations even though the smaller revenue generating agencies produced more evaluations. Table 3 shows the number of evaluations the agencies conducted.

Table 3 – Percentage of Evaluations Conducted in Past 12 Months

	1 – 3 Evaluations	4 -6 Evaluations	7 or More Evaluations
Overall (N = 50)	38%	46%	16%
Region (N = 50):			
East (n = 5)	60	20	20
South (n = 19)	37	53	10
Central (n = 17)	35	41	24
West (n = 9)	33	56	11
Revenue			
< $1 million (n = 13)	38	54	8
$ 1 million - $3 million (n = 24)	37	46	17
> $3 million (n = 12)	42	42	16
Full time Equivalents			
< 10 FTEs (n = 14)	50	36	14
11 – 25 FTEs (n = 19)	26	48	26
> 25 FTEs (n = 12)	42	58	0

The study examined which agencies conducted evaluations based on the characteristics of geographic region, agency revenues, and the number of full time equivalents.

Evaluation Practices Conducted Regionally by Affiliates

Geographically, the study examines the percentage of agencies who conducted evaluations within the past twelve months. The data reveals that the Southern region and the Western region evaluated at a higher percentage than the Eastern or Central regions. Ninety percent of the agencies in the Southern region report conducting between one and six evaluations. This is followed by the Western

117

region, who report eighty-nine percent of the agencies conducting between one and six evaluations. The agencies in the Central region had the lowest evaluation participation. Only seventy-six percent of the agencies reported conducting between one and six evaluations, and forty-one percent of those conducted between four and six evaluations.

Impact of Agency Revenue on Evaluation Practices

Agency revenue was selected as an important agency characteristic in determining whether or not financial resources made an impact on the number of evaluations conducted by agencies. The study finds that the agencies conducted between four to six evaluations at a higher percentage regardless of agency revenue. The data also reveals that agencies with less revenue performed more evaluations. Ninety-two percent of the agencies with less than $1 million in revenue report conducting between one and six evaluations, and fifty-four percent of this group of agencies conducted between four to six evaluations. However, fewer agencies (8%) who had revenues of less than $1 million conducted seven or more evaluations.

Eighty-three percent of the agencies with revenues between $1 million and $3 million conducted between one to six evaluations, with forty-six percent conducting between four to six evaluations. Eighty-four percent of the agencies with revenues of $3 million or more conducted one to six evaluations within the

118

past twelve months, and forty-two percent of this group of agencies performed between four to six evaluations.

Impact of Staff Capacity on Evaluation Practices

Several authors have noted the amount of staff capacity of an agency is a determinant of whether an agency will perform evaluations. Looking at the number of full-time equivalents (FTEs), the study finds that agencies with more staff performed more evaluations. Eighty-six percent of the agencies with ten or less FTEs respond that they evaluated between one to six evaluations in the past twelve months. Agencies with twenty-five or more FTEs reveal that forty-two percent of this group conducted between one and three evaluations and fifty-eight percent conducted between four and six evaluations within the past twelve months.

One would suspect that the more staff an agency has, the more evaluations the agency would perform. This did not proved to be the case in this study. There are no agencies, with twenty-five or more FTEs, who report conducting seven or more evaluations in the past twelve months. Agencies who conducted seven or more evaluations (26%) had between eleven and twenty-five FTEs or ten or less FTEs (14%).

119

Evaluations Required by Type of Organization

When nonprofit organizations evaluated their programs, the question was asked, why did they evaluate and what did they evaluate? If the evaluations were required, who required the evaluations? Was it the federal government, state government, local government or foundations? The agencies were asked to check every type of organization that required an evaluation by the agency. The results of this study find that thirty-nine out of the fifty-eight respondents (67%) evaluated something because it was a requirement of a funder. Most of the evaluations were conducted to comply with governmental grants and 16% were conducted for a United Way funding source. This is surprising because United Way of America has advocated that they are one of the main reasons why nonprofit organizations have moved toward program evaluation. Table 4 summarizes the type of organization that required the nonprofit agencies to conduct evaluations.

Table 4 – Evaluations Required by Type of Organization

Type of Organization	Evaluations Required By Percent (%)	Number of Evaluations Required (N)
Overall	67	39
Local	52	30
Federal	40	23
State	36	21
Foundation	36	21
United Way	16	9
National Urban League	7	4

The results of this study suggest that governmental and foundation funding plays a much more significant role than the United Way in motivating nonprofit organizations to conduct evaluations. Of the thirty-nine agencies required to evaluate their programs, sixty-four percent were required by their funder to conduct between one and three evaluations and twenty-six percent were required to conduct between four and six evaluations.

Evaluations Required by Funders by Agency Characteristics

The data in this study also suggest that nonprofit organizations are conducting evaluations. However, sixty-seven percent of the evaluations conducted by

agencies in this study are mandated by their funders. Funders that require evaluations in the study were examined by the agency characteristics of geographic region, agency revenue, and the number of agency full-time equivalents. The agencies in the Southern and Western regions are required by funders to conduct more evaluations. Middle-sized agencies, with revenues between one and three million, are required by funders to conduct more evaluations. Likewise, middle-size agencies, that have between eleven and twenty-five full time equivalents were required by funders to conduct more evaluations. Table 5 illustrates the number and percent of evaluations required by funders and the characteristics of the agency.

Table 5 – Percent of Evaluations Required by Funders

	1 – 3 Evaluations	4 -6 Evaluations	7 or More Evaluations
Overall (N = 39)	64%	26%	10%
Region :			
East (n = 3)	0	2	1
South (n = 17)	76	18	6
Central (n = 12)	58	25	17
West (n = 7)	71	29	0
Revenue			
< $1 million (n = 10)	60	30	10
$ 1 million - $3 million (n = 18)	67	28	5
> $3 million (n = 10)	70	20	10
Full time Equivalents			
< 10 FTEs (n = 10)	60	30	10
11 – 25 FTEs (n = 16)	56	38	6
> 25 FTEs (n = 11)	82	9	9

A regional glance at the agencies that were required to conduct evaluations shows that more evaluations were required by funders in the Southern (94%) and Western (100%) regions. Seventy-six percent of the agencies in the South were required by their funders to conduct between one and three evaluations, and seventy-one percent of the agencies in the West were required to conduct between one and three evaluations.

The Central region had fewer funders require evaluations, but still had fifty-eight percent of their funders require between one and three evaluations. The Eastern region only had five agencies report that they evaluated their programs. Of the five agencies, only three reported that they were required by their funders to conduct evaluations and the funders required between four or more evaluations.

Agency revenue did not make a difference in the amount of evaluations conducted by the agencies. Both groups, agencies with a $1 million or less and agencies with $3 million or more, reported (90%) performing between one and six evaluations. Agencies with a $1million or less reported that sixty percent of the agencies conducted between one and three evaluations.

Agencies with revenues of $3 million or more (90%) report conducting between one and six evaluations in the past year. The agencies with revenues between $1 million and $3 million were required by more funders to conduct evaluations (95%). Like the agencies with revenues of $3 million or more, approximately seventy percent of the agencies with revenues between $1 million and $3 million were required to conduct between one to three evaluations.

Agencies that employed between eleven and twenty-five full-time equivalents (FTEs) were required to conduct evaluations more than the other group of agencies. Ninety-four percent of this group was required by their funders to conduct between one and six evaluations, and fifty-six percent were required by their funders to conduct between one to three evaluations. However, ninety percent of the agencies with less than ten FTEs and ninety percent of the agencies with twenty-five or more FTEs were also required by their funders to conduct between one and six evaluations. Most of the agencies with twenty-five or more FTEs (82%) were required to conduct between one and three evaluations.

Types of Nonprofit Evaluations

To answer the question about the extent of nonprofit evaluations and how nonprofit organizations conduct their evaluations, a closer look is needed into their evaluation activities and the type of evaluation information the agency

124

collected. Specifically, I want to address the question concerning what types of evaluations are done the most. Are nonprofit organizations conducting evaluations based on output, demographic, outcomes or financial data? Are nonprofit organizations evaluating ongoing programs, or only specific projects or services? Most agencies in the study evaluated ongoing programs and the type of evaluation information they collect the most is on their programs and projects. Table 6 shows the agencies' evaluation activity and table 7 illustrates the type of evaluation data collected by the agencies.

Table 6 – What Agencies Evaluated In Past 12 Months

Evaluation Activity	Percent	Number
On-going Programs	81%	42
Overall Impact	57	33
Agency Board	55	32
Fundraising Activity	55	32
Specific Projects	53	31
Products Developed	31	18
Volunteers	24	14

Table 7 – Type of Evaluation Information Collected

Data Collected	Percent	Number
Project/Program	86%	50
Outcomes	78	45
Client	66	38
Costs	60	35

More than eighty percent of the respondents in the survey note that they evaluated their ongoing programs. Only fifty-seven percent respond that they evaluated their agency's overall impact. Agency volunteers (24%) were evaluated the least.

With the push from funders for agencies to conduct outcomes data, the survey results are lower than expected. Only seventy-eight percent of the agencies reported they collected outcomes data. While this amount is significant, it is less than the percentage of agencies (86%) who evaluated their projects or programs. It was interesting to find that only sixty percent of the agencies say that they evaluated their costs.

Agency Motivation to Evaluate

What is an agency's motivation to conduct an evaluation, and when do they perform the evaluations? The study reveals that most nonprofit agencies in the study routinely conduct evaluations but conduct evaluations more often to satisfy the funders' requirement. Table 8 outlines the frequency in which nonprofit organizations performed evaluations.

Table 8 – Frequency in Conducting Nonprofit Evaluations

Nonprofit Agencies Generally Conducts Evaluations:	Percent	Number (N = 53)
Routinely	75%	40
As Required By Funder	15	8
For Specific Issues	4	2
Other Reasons	6	3

With responses from fifty-three of the fifty-eight agencies, the table shows that more than seventy percent of the agencies say they conduct evaluation as a routine function and evaluation is a part of the organization's regular activities.

127

However, fifteen percent of the agencies indicate that the evaluations were conducted only because it was required and not because it was routine function of their operations.

Although thirty-nine agencies state they were required to conduct evaluations, twenty-six agencies report that the requirement was their motivation for conducting evaluations. The remaining six of out thirty-nine agencies required by the funders to conduct evaluations were motivated by their staff, board or accreditation reasons.

Table 9 – Motivation to Conduct Evaluations

Evaluations conducted by Nonprofits because:	Percent	Number (N = 53)
Required by the funder	49%	26
Staff Reasons	34	18
Board Reasons	13	7
Accreditation	4	2

The assumption prior to the study was that nonprofit organizations are conducting evaluations because they are required or mandated by the funder to conduct the evaluation. The difference in the two responses may be attributed to the motivation that comes from the funder to require evaluations from nonprofit organizations. Grant makers find themselves in a grey area of ethical uncertainty when they begin making grants for evaluation and organizational effectiveness. Whether the funder underwrites the cost of a consultant or directly advises the grantee, the situation has the potential for conflict in terms of power, boundaries, and unintended consequences (GEN and GEO 2000).

The board of directors of the agencies and staff (47%) also motivated the agencies to conduct evaluations. Only four percent were motivated to conduct evaluations for accreditation purposes.

Funders Expectations and Support

One of the questions this study seeks to address is if there is an increasing demand on nonprofit organizations to evaluate based on the funders' expectations. In other words, are funders asking for more evaluation information from the nonprofit organizations than the dollars they provide to implement the programs? If the funders are asking for more evaluation information, is the increase in funders' expectation a problem for nonprofit organizations? If funders are asking

nonprofit organizations for frequent evaluations and more evaluation information, are funders providing the necessary support for nonprofit organizations to meet the expectations?

The frequent mandating of evaluations by government institutions and other funders does not indicate a widespread valuing of evaluation that will continue to appear even in the absence of legislative or board mandates (Sanders 2002). The practical application of systematic evaluations for nonprofit organizations suggests most policy makers, managers, and funding agencies need answers to relatively straightforward questions that do not require great methodological sophistication. Agencies and their funders often face the pressures of working under budget and time constraints, which often result in a lack of attention to sound research design with limited attention being given to identifying and addressing factors affecting the validity of the findings (Bamberger, Rugh, Church and Fort 2004).

Table 10 provides a snapshot of funders' expectations of nonprofit agencies in conducting evaluations. The study found nonprofit agencies collect evaluation data they would not have collected if it was not required by the funder. The agencies also found the amount of evaluation information has increased from the funder but the agencies find the information beneficial. The agencies also find there is a greater demand on the agencies to conduct outcomes evaluations.

130

Table 10 – Funders' Evaluation Requirements for Nonprofit Agencies

Degree to which Nonprofit agencies: N = 27	Not at all	Not very Much	Somewhat	A Great Deal
Required to collect data not otherwise collected	15%	22%	41%	22%
Used Funders' requirements for other purposes	0%	18%	52%	30%

Agencies were asked to what degree the funder requires their organization to collect information that they would not have otherwise collected. Twenty-seven out of the thirty-nine agencies responded to the question concerning the funders' requirement to collect additional data. Sixty-three percent of the agencies said they had to collect information that they would not have collected if they were not required to collect it for the funder. Only fifteen percent said they were not required to collect information that they would not have otherwise collected.

The data shows that the agencies did collect more information than what they would have collected if they were not required to do so. Did this cause a problem for the agencies? Sixty percent of the respondents say that it did not create a problem, and forty percent state that it created a small or moderate problem.

Did the collection of the required data by funders benefit the agencies? Did they use the required data for purposes other than for reporting the information to

131

the funder? One hundred percent of the agencies respond that they use the information for more purposes than to report to the funder. In fact, thirty percent of the agencies state they use the required data a great deal, and only eighteen percent say they do not use the data very much.

Since the agencies report they benefited from the funders' requirements, were the expectations from the funders reasonable to the agencies? Did the collection of additional information cause a hardship for the agencies? The nonprofit agencies in this study, found the expectations in reporting evaluation data to funders beneficial and reasonable. An increase in funders' expectations can be attributed to the increase in the funders' request for outcomes data by the agencies. Ninety-three percent of the respondents felt the funders' expectations were reasonable, and of the ninety-three percent, approximately twenty percent, indicated the funders' expectations were very reasonable. These results indicate a willingness by the agencies to collect the information requested conduct evaluations.

Table 11 – Funders' Expectations on Providing Evaluation Data

N = 57	Increased	Decreased	Remained the Same	Don't Know
Funders' Expectations on Required Information	67%	0%	19%	14%
Funders' Expectations on Outcomes Data	63%	0%	26%	11%

Along the same lines, the study wanted to find out if funders' expectations about the amount of evaluation information the nonprofit organizations provide has changed over the past three years. Table 11 shows that sixty-seven percent of the agencies responded that the amount of information requested by the funder has increased and approximately twenty percent said the amount of information provided did not change.

The nonprofit sector has been heavily persuaded to perform evaluations, and they have been heavily persuaded to perform outcome-based evaluations. Many state and local government agencies, foundations, managed care systems and accrediting bodies have added outcome measurement to the list of performance and accountability measures they require of nonprofit organizations within their sphere (United Way of America 2000). If funders are requiring agencies to report more information, have the funders' expectations changed regarding whether or

not the agency should provide information to them about the outcomes or impacts of their programs or services? Did the funders want to know if the agencies' programs produce a change in their clients or users?

Table 11 shows that sixty-three percent of the agencies said the funders were requiring more outcomes information. This validates the United Way of America (2000) statement that more funders are requesting outcomes measurement. Only twenty-six percent of the agencies say that they provided the same outcomes data. The agencies also reported that they found the outcomes information either somewhat useful (56%) or very useful (44%). No agency reported finding the outcomes data not useful for their organization.

Since the demands on the agencies are increasing from the funders, one would hope that the funders' support would also increase. Table 12 shows the support from funders to help nonprofit conduct evaluations is limited. Most of the support comes in the form of resource materials and training.

Table 12 – Funder Support in Conducting Nonprofit Evaluations

Types of Funder Support	Yes	No
Tools and Resources	59%	41%
Training	41%	59%
Advice	26%	74%
Other Support	0%	100%

Table 12 shows that fifty-nine percent of the agencies received tools or resource materials from funders to conduct evaluations. Another forty-one percent received training from the funders. None of the agencies responded that additional revenue was given to support their evaluation efforts.

Since the demand was greater on nonprofit agencies to perform outcome evaluations, the study sought to find if the agencies had any difficulties in identifying, collecting, analyzing or interpreting the outcomes data. The findings (Table 13) reveal that the agencies had very little difficulty with outcome evaluation. The agencies report the most difficulty they experience is in collecting outcomes data. The least difficulty found with outcomes data in the study in interpreting outcomes data.

135

Table 13 – Perceived Difficulties in Conducting Evaluations

Evaluation Problems N = 50	Very Difficult	Somewhat Difficult	Not Difficult
Identifying Outcomes	4%	28%	68%
Collecting Outcomes Data	4	34	62
Analyzing Outcomes Data	0	33	67
Interpreting Outcomes Data	2	25	73

Only thirty-eight percent state they find it very difficult (4%) or somewhat difficult (34%) in collecting outcomes data. The other area of slight difficulty was in identifying and analyzing of outcomes data. Thirty-two percent say that they find identifying outcomes data very difficult (4%) while others find it somewhat difficult (28%). Only thirty-three percent of agencies say that analyzing outcomes data is somewhat difficult.

Fifty out of fifty-eight agencies in the survey state that they evaluated something in the past twelve months (see Table 2). Only thirty-nine out of the fifty agencies say they were required to produce their evaluation results for a funder (see Table 5). This study wanted to find out what the agencies perceived as their strengths and weaknesses in conducting evaluations. What do nonprofit agencies think about their level of evaluation skills and knowledge? What did the

affiliates find the most difficult in the evaluation process? Do nonprofit agencies think they have the ability to perform evaluations without outside assistance like evaluation consultants? What were the top three barriers perceived by nonprofit organizations in conducting evaluation?

Evaluation Barriers for Nonprofit Organizations

Perceived Agency Strengths and Weaknesses

To get a better understanding on some possible reasons why nonprofit organizations were not conducting more evaluations, the researcher examined the strengths and weaknesses to determine possible barriers in conducting evaluations by nonprofit organizations. The data (table 14) reveals that the greatest perceived strength the nonprofit organizations feel they have, as it relates to evaluation, is in understanding their programs and projects (95%). This is followed by ninety-one percent of the respondents that feel their strengths lie within having a positive attitude towards evaluation.

Table 14 - Strength with Respect to Evaluation

Type of Evaluation Strength: (N = 57)	Strength	Small Strength	No Strength
Understanding of Programs/projects	95%	3%	2%
Positive Attitude	91	7	2
Understanding of Community Needs	87	9	4

Additional strengths identified in the study are the perception the agencies felt about the accuracy of their evaluation results and their confidence in their organization's ability to perform evaluations effectively. Ninety-six percent of the agencies report that they are somewhat accurate (51%) or very accurate (45%) in their evaluation results in terms of reflecting what was going on in their projects or programs.

Ninety-three percent of the agencies feel that they are somewhat confident (53%) or very confident (40%) in their organization's ability to perform evaluation effectively. Therefore, with the proper training and funds to systematically conduct evaluations, nonprofit agencies are confident that they can perform evaluations effectively.

The largest weakness for nonprofit organizations with respect to evaluation is funding. This weakness is followed by limited skills and knowledge on

evaluations. Table 15 shows the largest weakness identified by the nonprofit organizations is funding for evaluation (56%).

Table 15 – Nonprofit Weaknesses in Evaluations

Types of Barriers (N = 57)	Weakness	No Weakness
Funding for Evaluation	56%	44%
Evaluation Skills and Knowledge	20	80
Effective Communication of Evaluation Results	17	83

Fifty-six percent of the agencies report that they could not secure funding for evaluations at all (23%), or they have little success (33%). Eighty percent of the agencies believed they did possess a moderate to strong level of evaluation skills or knowledge. Of the eighty percent, twenty percent of the agencies believe they have strong evaluation skills and knowledge. Another interesting note is that once the nonprofit agencies have the information, they feel comfortable in effectively communicating the evaluation results. Eighty-three percent of the agencies report that they effectively communicate their evaluation results with approximately forty percent stating that they did a great job in communicating their evaluation results. It appears that more education and training in this area could improve how nonprofit organizations communicate their evaluation data.

139

If nonprofit agencies believe they have good evaluation skills and knowledge, and they are able to communicate their evaluation results effectively, is funding for systematically evaluations the only problem for nonprofits organization?

Most nonprofits have far too little time, money, or in-house skills for developing evaluation systems and maintaining them over time (Murray 2001). Additionally, the more extensive and comprehensive the evaluation, the more it will divert the time, energy, and resources of the nonprofit organization away from the primary purpose of the project (Carson 2000). Nonprofit agencies acknowledge there are barriers to conducting evaluations that are driven by money. They also acknowledge there are psychological barriers to conducting evaluations as well. The psychological barriers that tend to motivate nonprofit organizations to conduct evaluations are less of a problem than the problems that are driven by finances. The top three barriers perceived by nonprofit organizations in conducting evaluations are lack of money, lack of staff time, and lack of tools such as resource manuals and guides.

Table 16 illustrates nonprofit organizations' perceived evaluation problems in which nonprofit organizations can place some monetary value.

Table 16 - Perceived Evaluation Barriers (Monetary Driven)

Perceived Problem (N = 55)	Problem	Small Problem	No Problem
Lack of Money	75%	16%	9%
Lack of Staff Time	62	20	18
Lack of Tools/Resources	57	23	20
Access to College Students	30	24	46
Technology	22	16	62

Seventy-five percent of the agencies perceive funding to be the number one barrier in conducting evaluations. Thirty-seven percent of the agencies state that this is a moderate problem while thirty-eight percent state that this is big problem. Sixty-two percent of the agencies perceive lack of staff time to be a major hindrance to evaluations. More than twenty-five percent of the agencies state that lack of staff time is a moderate problem, and thirty-five percent say lack of staff time is a big problem. If there is difficulty finding money for evaluation, the logic follows that it would be difficult to secure tools and resources for evaluations. This is true in this study. Fifty-seven percent of the agencies state that the lack of tools is a problem. Of the fifty-seven percent, forty-seven percent stated that this is a moderate problem.

141

Another interesting finding from the study is that technology is not a problem for the agencies. There has been an assumption that nonprofit organizations are limited in conducting evaluations because of the lack of technology to collect, store, process and maintain evaluation information. Almost eighty percent of the agencies in the study report that technology is neither a problem (62%) or only a small problem (16%).

When asked if the agencies had access to resources like evaluation training, professional associations or informal networks that could assist them with their evaluation efforts, more than seventy percent (71%) answered affirmative. Approximately seventy percent of the agencies note that they have between one and five staff people trained in evaluation, and thirty-one percent report they have more than five staff members trained in evaluation. This is also reflected in Table 17, where the agencies report that approximately seventy percent perceive to have no problems (27%), or only a small problem (40%) with evaluation skills and knowledge in conducting evaluations.

Along with skills and knowledge, the nonprofits agencies were asked if there are other psychological barriers that hamper the motivation or prevent nonprofit agencies from conducting evaluations. What nonprofit agencies perceive as not being problems in conducting evaluations is an interesting facet in this study. The agencies report that working with evaluation consultants, having evaluation

objectivity and having an understanding of the value of evaluation are not problems for the organizations.

Table 17 - Perceived Evaluation Barriers (Psychological Driven)

Perceived Problem (N = 55)	Problem	Small Problem	No Problem
Limited Skills/Knowledge	33%	40%	27%
Unclear Expectations from Funders	32	38	30
Confidence to conduct Evaluations	26	27	47
Understanding the Value of Evaluation	16	31	53
Evaluation Objectivity	16	33	51
Working with Consultants	7	36	57

Ninety-three percent of the agencies state they have no problem (57%) or a small problem (36%) when working with evaluation consultants, while only seven percent stating they had a moderate problem working with evaluation consultants.

Understanding the value of evaluation was not a problem for the group of agencies in this study. Eighty-four percent of the agencies report that they had no problem (53%) or a small problem (31%) with understanding the value of evaluation. This finding is important because it shows that nonprofit organizations

143

understand the importance of evaluation, but several factors prevent systematic evaluations.

Several of the agencies interviewed for this study, stated that evaluations were very important to their agency's operations. James Shelby, President and CEO of the Greater Sacramento Urban League (Shelby 2005), states that "Without evaluations, you don't know if you are achieving the objectives. Evaluation is a way to look for best practices. It is also a tool for the funders."

Shelby (2005) stated that his agency has been conducting systematic evaluations for the past three years. He stated that the evaluations help his agency become more aware of the effectiveness of his programs. When asked how the Greater Sacramento Urban League conducts their evaluations Shelby (2005) stated, "Weekly we enter the evaluation information into a database. We give a copy to the staff and to the board of directors to improve our programs."

The agencies in this study were asked if they lack evaluation objectivity because they want to present their activities or programs in a good light. Eighty-four percent of the nonprofit agencies report that they have no problem (51%) or a small problem (33%) with evaluation objectivity. Only two percent of the agencies report that they have big problems with evaluation objectivity.

Factor Analysis: Perceived Evaluation Barriers

Reviewing the responses from the agencies on the perceived barriers to evaluations, there appears to be certain variables that clustered together. A factor analysis was conducted to see if certain variables would cluster together. The analysis indicated that people are responding in clusters based on certain perceived barrier categories. There are variables that cluster together based upon the agency's perception of their lack of skills and resources. These variables included lack of resources and tools; lack of confidence in evaluation ability; lack of access to college students; lack of money for evaluation; lack of evaluation skills and knowledge; lack of internal capacity and lack of technology. These variables were clustered together into a new variable name called lack of skills and resources.

There are variables that cluster together based on the agency's perception of their lack of proper focus. These variables include lack of objectivity of evaluation; lack of difficulty working with consultants; lack of understanding evaluation value; and the lack of clear expectations from funders. These variables were clustered together into a new variable name called lack of proper focus.

I ran two-factor regression analysis to see if there was a causal relationship between any of the variables based on the percentage of evaluations that were

145

required by funders. The percentage of evaluations by funders is calculated by taking the number of evaluations required by the funder and dividing it by the total evaluations conducted. In the regression analysis, the dependent variable is the percentage of evaluations that were required by the funders. The independent variables are the lack of proper focus and lack of skills and knowledge. The R squared in the analysis was .222. This indicates that approximately 22% of the variability of each of the variable is shared with the other variable. The new cluster variable, lack of proper focus variable, is significant at .010.

Scatterplot graphs were run on the lack of skills variable and the lack of focus variable. The dependent variables in both graphs are the percent of evaluations that are required by the funders. The variables in both graphs did not appear to cluster. The graphs are displayed in Appendix C.

The lack of objectivity of evaluation (.833) and lack of difficulty working with consultants (.796) appear to be variables that have an impact on the number of evaluations taking place by nonprofits. It appears that the more nonprofit agencies learn about conducting evaluations, the more they become aware of the need to be objective in their evaluations. Also, the increase in difficulty the nonprofit organizations may be experiencing in working with consultants may be the result of mixed messages between nonprofit agencies and the consultants based on the level of evaluation knowledge the nonprofit agency may possess.

146

Evaluation training needs to be very specific based on the evaluation skills of each agency.

I ran a three-factor regression analysis to see if there was a causal relationship between any of the variables based and the percentage of evaluations that were required by funders. In the regression analysis, the dependent variable is the percentage of evaluations that were required by the funders. The independent variables are the lack of skills and knowledge; lack of resources and lack of understanding. The R squared in the analysis was .232. This indicates that approximately 23% of the variability of each of the dependent variables is shared with the other variable.

The lack of skills and knowledge cluster included the lack of access to college students; lack of access to technology; lack of unclear expectations from funders; lack of confidence in evaluation ability and lack of skills and knowledge. Lack of understanding variable included lack of objectivity of evaluation; lack of difficulty working with consultants and lack of understanding evaluation value. The lack of resources variable included the lack of money for evaluation, lack of internal capacity and lack of tools and resources. The two variables that emerged again were the lack of objectivity of evaluation (.823) and the difficulty in working with consultants (.817).

147

There appears to be a need from nonprofit agencies for more evaluation training. It appears that nonprofit agencies are experiencing some level of difficulty working with consultants which could be the result of not having the prerequisites to completely understand the consultant from the initial training. Nonprofit agencies may also experience a sense of being overwhelmed initially by consultants with the large commitment of time and resources needed to conduct evaluations and evaluation training. It appears that the more the agencies know about evaluation, the more they are aware about the need to be more objective about their evaluations and the more they are aware of their evaluation skill limitations.

The argument can be made that some nonprofit agencies think they are conducting evaluation, but once they are trained in evaluation they realize the need to improve on the quality of the evaluation. The greater the exposure to evaluation training, the greater the agency's focus is on systematic evaluations. The more they are trained in evaluation, the more confidence the agencies have in conducting quality evaluations.

Evaluation Needs for Nonprofit Organizations

The needs identified in this study for nonprofit organizations are more money, staff and tools (guides and manuals). Seventy-five percent of the NUL agencies perceive funding to be the number one barrier in conducting evaluations. If funding is a problem, where do the majority of nonprofit organizations get their funding for evaluations, and is the amount of money adequate for systematic evaluations? Approximately sixty percent of the agencies say they covered the cost of the evaluations through their general operating fund, and only sixteen percent state they received funding for evaluation from an external funder.

Forty-eight percent of the agencies say that the funds they had to conduct the evaluations were not adequate. However, forty percent of the agencies say the funds they have for evaluation were adequate and twelve percent say the fund were very adequate. Only twelve percent of the agencies have someone dedicated onsite to conduct evaluations.

Clearly, resources are needed to assist with more and better systematic evaluations, and more funding should be sought from program sources. More funding is needed to hire evaluators or to hire staff devoted to evaluation. In addition to money and staff, more evaluation training materials are needed. Most tools are not readily accessible and many fail to take into account the diversity of

149

nonprofit organizations in terms of the skill sets of their staff and volunteers, their financial resources and the types of programs they deliver (Bozzo and Hall 1999).

Therefore, the study finds that nonprofit organizations are conducting evaluations, but many of the agencies are conducting evaluations because they are required by the funders. The nonprofit organizations value evaluations, but have limitations due to the lack of time, money and staff. Since nonprofit organizations value evaluations and funders require the agencies to conduct them, how do the nonprofit organizations use the data they collect?

Evaluation Use for Nonprofit Organizations

Evaluation literature has given a great deal of attention to use (Patton 1997; Hatry 1999; Weiss 1998). This study also seeks to find out how nonprofits use their evaluation results. Do nonprofit organizations use their evaluation data to improve their services? Do they use the evaluation results more to improve their services? Do the agencies use the data to market their programs or for fund development? Do the agencies use evaluation data for strategic planning or to share with other agencies?

Overall, the agencies report using the evaluation information to report to their funders. This study also finds that nonprofit agencies use evaluation data to improve their programs. The evaluation data is used less by nonprofit

150

organizations for strategic planning and for fund development. Table 18 illustrates evaluation utilization for nonprofit organizations.

Table 18 – Reasons Nonprofit Agencies Use Evaluation Data

Type of Use (N = 58)	Not at All	Not Very Much	Used Evaluation Data
Report to Funder	84%	7%	9%
Improve Programs & Services	83	12	5
Increase Agency Awareness	72	16	12
Fund Development	70	10	20
Strategic Planning	69	19	12
Share with Other Agencies	40	26	34

Nonprofit organizations collect the data to track the outcomes in order to respond to funder requirements (Hatry and Lampkin 2001). This study suggests, to this point, that the agencies are motivated by their funders to conduct evaluations. But, did the agencies only use this data to report to the funders, or did they use their evaluation information for other purposes?

151

Approximately eighty percent of the agencies in the study state that they use their evaluation data to report to the funder. Of the eighty percent, seventy-one percent state they used the evaluation information a great deal to report to the funder. This suggests that the agencies are strongly motivated by their funders to conduct evaluation, and their funders influence how their information is used.

The agencies in this study were asked to what degree they used the evaluation information to improve their programs and services. This is important because the evaluation literature suggests that the key reasons for conducting evaluations is to see if the program or services are performing as expected; and to find ways to improve the program; or become more efficient. Patton states that evaluations are to be conducted based upon the need of the user. Therefore, if the evaluation is driven with the funder in mind to be the end user, the agency's focus may not be to use the information to improve programs and services

Bernett Williams from the Akron Ohio Urban League (2005) suggested that finding the dollars to run the agency is more important for some agencies than evaluating the programs. She states the following:

> "Affiliates have not made it (evaluation) a priority at this time. We eliminate programs because we do not have the funds. We assume that the program must not be working because we have no more funds to support the program, even though the program may be producing the desired results. We chase the dollars."

152

Warren Logan (2005) also suggested that evaluation is valued by the agencies but not a high priority in their daily activities. He states, "Agencies are more interested in funding programs. They are more interested in replicating programs that make revenue for an agency than finding out if the program is a good program that produces the intended outcomes. Performing evaluations is not a high priority for discussion."

Evaluation results are rarely intended by evaluators or evaluation sponsors to be "knowledge for knowledge's sake." Rather, they are intended to be useful, and to be used, by those with responsibility for making decisions about the program, whether at the day-to-day management level or at broader funding or policy levels (Rossi, Freeman and Lipsey 1997). Eighty-three percent of the agencies in the study report that they use the evaluation information to improve their programs and services. Of the eighty-three percent, sixty-two percent report using the information a great deal to improve their programs and services. This would suggest that there are some residual benefits in collecting evaluation information for nonprofit organizations that are required by their funders. The data suggests that the agencies find the information useful for improving the programs and services.

Is there any evaluation use for general administration or back office functions for the nonprofit agencies? Can evaluation be useful in marketing the agencies'

program and services? Can evaluation aid in program development or strategic planning? Along with assessing the value and worth of a program, evaluations should aid decision-making, support change management, and improve processes (Hubelbank 2001).

Approximately seventy percent of the agencies used the evaluation information to increase their agency's awareness, or to market their agency's programs or services. Some of the possible marketing uses are to promote how effective the program is in changing the lives of their clients. Another use is in demonstrating the effectiveness of a program in reaching a difficult target market. Of the seventy-two percent, approximately fifty percent of the respondents state they use the information a great deal to increase awareness at their agency.

Strategic planning is another agency activity that funders are requesting nonprofit organizations to conduct regularly. Funders sometimes request that an agency submit a three to five year strategic plan demonstrating to the funding agency the sustainability of a program or service once the funding has ended. Evaluation information collected systematically by nonprofit agencies can help with this process. In this study, evaluation data is used approximately seventy percent in strategic planning for the agencies, with forty-five percent of the agencies saying they used their evaluation data a great deal for strategic planning. With the push for nonprofit organizations to conduct strategic planning,

incorporating evaluation information in the strategic planning process is an opportunity for nonprofit organizations. Dr. Walter Brame, President and CEO of the Grand Rapids Michigan Urban League, (Brame 2005) states, "there needs to be recognition by affiliates that the evaluation results are needed in the planning process."

Another possible use for evaluation information for nonprofit organization is in fund development. Demonstrating successes through evaluation data to past, current and future funders will increase funds for the agency. Several of the Presidents and CEOs stated that the funders were very receptive to their evaluation information. Warren Logan, President and CEO of the Chattanooga Tennessee Urban League (Logan 2005) stated, "More funders are surprised when they have the outcomes data? Evaluation makes my affiliate more competitive."

Having a track record by measuring what has worked and what needs improvement helps funders determine to continue funding for a program or to increase funding for a program. It also helps in matching potential funders with an agency that can deliver the need to a targeted client base. Theory-based evaluations test the links between what programs assume their activities are accomplishing and what actually happens at each small step along the way. Although the agencies state they use their evaluation data approximately seventy

percent for fund development, only forty-five percent state they use their information a great deal for fund development.

The evaluation data is used less by the agencies for fund development and to share with other agencies. The lowest utilization of evaluation information by the agencies indicated in the study is sharing the evaluation data with others (40%). There are several reasons why nonprofit organizations may not share their evaluation information with other agencies. First, there may be turf issues among agencies, locally and regionally. The funding among nonprofit organization is extremely competitive. Sharing what works and what needs improvement by an agency that is currently delivering the program or services could open the agency up for a hostile takeover. Although the clients or customers who receive the programs or services would benefit more from agency collaborations, the reality is that agencies are very reluctant to share their weaknesses. Dr. Walter Brame (2005) says, "Agencies operate in a competitive mode. Agencies are not anxious to share best practices. Sometimes we do not share because of the expectations. Sharing information maybe perceived as a liability."

The second reason why agencies may not share the information with other agencies is due to the lack of staff time. Most staff time is built into program delivery and reporting. Agencies who find the time to conduct and report their evaluation findings usually report their finding to the funder or the board of

directors. The third reason for not sharing the information with others may be the lack of an established evaluation network among agencies. Without this network, agencies may not know who to seek out to share the information with or who to look for with similar programs.

In this study, sixty percent or the agencies state that they do not share their information at all (34%), or not very much (24%). This is a missed opportunity for nonprofit organizations as they look for additional ways to leverage evaluation resources.

Agency Characteristics and Evaluation Utilization

It is important to look at the agency's characteristics as they relate to evaluation utilization. Specifically, the study analyzes if geographic location, agency revenue, program budget or staff size have an impact on how the agencies' used their evaluation information.

The agencies in the Western region used their evaluation data more (89%) to improve their programs than any other region. The Western region also used their evaluation data for reporting to funders (100%) and for fund development (100%). Agencies with more revenue use the evaluation data more than the agencies with smaller revenues. Also, agencies with more employees or full-time

157

equivalents, use their evaluation data more than agencies with less full-time equivalents. However, the larger number of full-time equivalent staff did not make a difference in strategic planning and sharing evaluation data with others. Table 19 illustrates the nonprofit evaluation utilization by agency characteristics.

Table 19 – Evaluation Utilization by Agency Characteristics by Percent

Evaluation Use (N = 58)	Improve Programs (Yes)	Improve Programs (No)	Agency Awareness (Yes)	Agency Awareness (No)	Funders Report (Yes)	Funders Report (No)	Strategic Planning (Yes)	Strategic Planning (No)	Share with Others (Yes)	Share with Others (No)	Fund Development (Yes)	Fund Development (No)
West Region (N = 9)	89	11	87	13	100	0	55	45	50	50	100	0
South Region (N = 22)	86	14	82	18	82	18	73	27	41	59	67	33
East Region (N = 6)	83	17	67	33	83	17	83	17	33	67	60	40
Central Region (N = 21)	76	24	57	43	78	22	67	33	37	63	62	38
< $1 million	76	24	80	20	79	21	74	26	28	72	76	24
$1 - $3 million	78	22	65	35	80	20	60	40	48	52	70	30
> $3 million	94	6	80	20	93	7	80	20	40	60	54	46
< 10 FTEs	81	19	75	25	80	20	69	31	33	67	78	22
11 – 25 FTEs	82	18	67	33	90	10	68	32	50	50	79	21
> 25 FTEs	86	14	80	20	86	14	66	34	40	60	50	50

159

Regional Difference in Evaluation Utilization

As previously stated, the National Urban League affiliates or agencies were chosen for this survey because they are diverse in size but also because the agencies are dispersed in more than one hundred cities around the United States. The affiliates of the National Urban League provide services based on the geographic need of its target population. The agencies are affiliated based on their agency brand, but their services and the size of the affiliates vary depending on local community needs.

The geographic location of the agency is an important characteristic of this study for several reasons. First, the National Urban League affiliates are geographically dispersed in more than one hundred cities and urban communities around the nation. This geographic dispersion provides an indication about what nonprofit organizations are doing concerning systematic evaluation around the country. It also helps to identify where certain regions of the country are conducting evaluations and where there may be opportunities for sharing resources in other regions to increase systematic evaluations. Geographic location was grouped into previously determined regional boundaries by the National Urban League. The regions are divided into four categories: Eastern, Southern, Central and Western. There are twenty seven agencies in the Eastern region and thirty agencies in the Southern region. There are thirty-eight agencies in the

160

Central region and ten agencies in the Western region. Of the fifty-eight respondents to the survey, only six agencies (10%) responded to the survey from the East. The Southern region had the largest respondents with twenty-two participants (38%), followed by the central region with twenty-one participants (36%). The Western region had nine out of the ten agencies in their region to complete the survey and they represented sixteen percent of the respondents in the survey.

Evaluation utilization was analyzed by geographic locations to see if the location of an agency made an impact on how they used their services. The study looks to see if the location of an agency makes a difference in who use the evaluation information to improve their programs and services.

The study also examined whether the location has an impact, if the agency uses the information to increase agency awareness, or if they use the evaluation data just to report to the funder. Geographic location is also analyzed to see if location makes a difference in whether the agencies use evaluation information for their strategic plans or fund development. The study also seeks to find if geographic proximity of an agency has an impact on who share their evaluation information the most.

The study found that there are slight differences among the agencies based on geographic location and whether or not the agencies use their evaluation information to improve their programs and services. More agencies use their evaluation information in the Western and Southern regions than in the Eastern and the Central regions. Eighty-nine percent of the agencies in the Western region report using their evaluation information either somewhat or a great deal to improve their programs and services. This contrasts with the Central region with only seventy-six percent of the agencies reporting using the information somewhat (14%) or a great deal (62%) to improve their programs and services. Using evaluation information to improve the programs and services is an opportunity for the Central region.

The study looked to see if there are regional differences among the agencies that use the evaluation information to increase their agency's awareness. Did they use the information to promote their programs, services or agency? Overall, seventy-one percent of the agencies reported they use the evaluation information to increase awareness about their agency. The findings also reveal that there are regional differences among the agencies that used the data to increase agency awareness.

The Western region is the clear leader with approximately ninety percent of the agencies saying they use the evaluation information to increase their agency's

awareness. Of the eighty-seven percent of the agencies in the west who use the information for agency awareness, seventy-five percent of the agencies said they used the information a great deal.

The Southern region followed the Western region and reported eighty-two percent of the agencies in the south used their evaluation information to increase agency awareness. Of the eighty-two percent, fifty-five percent state that they use the information a great deal to improve the agency's programs and services. The Central region, again, shows another opportunity to utilize evaluation information. Approximately forty percent of the agencies in the Central region indicate that they do not use the evaluation information to market their programs and services. Of the forty-three percent who indicated they did not use the evaluation data, twenty-four percent state they do not use the information very much, and approximately twenty percent state they do not use the information at all to increase the agency's awareness. Although the Eastern region reports that the agencies use evaluation to increase agency awareness (67%), thirty-three percent of the agencies in this region state they do not use the information very much to increase agency awareness.

The study looked to see if any of the regions use the data more than the other regions to comply with funder requirements. Regional differences emerged from the findings that indicate some regions use the evaluation information more than

others to report to their funders. Approximately eighty percent of the agencies report that they use their evaluation information either somewhat (12%), or a great deal (67%) to report to their funders. However, one hundred percent of the Western region said they use their evaluation information to report to their funders. The Western region stated that eighty-nine percent of their agencies use the information a great deal and eleven percent of the agencies said they use it somewhat to report to their funders.

The Southern and Eastern regions are similar in their utilization of evaluation information to report to the funders. The Eastern region reports that eighty-three percent of the agencies use their evaluation information somewhat (17%), or a great deal (18%) to report to their funders. The Southern region almost mirrors the Eastern region and reports eighty-two percent of agencies' evaluation utilization is to report to funders. Of the eighty-two percent, eighteen percent state they somewhat use the information while the other sixty-four percent find that they use the information a great deal to report to funders.

Strategic planning has been discussed as an essential component in the operations of nonprofit organizations. Funders are increasingly requesting that agencies have a strategic plan that outlines the agency's program expectations to predict program sustainability. The study finds that overall only sixty-eight

percent of the agencies are using their evaluation information to conduct strategic planning, but there are clear differences between the regions.

The Eastern region has more success using their evaluation information for strategic planning. Eighty-three percent of the agencies in the east state they use the information somewhat (33%,) or a great deal (50%) for strategic planning. However, only six of the fifty-eight agencies in this study represent agencies from the eastern region. This compares to the Western region who report only fifty-five percent of the agencies use the evaluation information either somewhat (22%), or a great deal (33%) in their strategic planning process. The Southern region has the largest number of agencies reporting that they did not use their evaluation information for strategic planning. Eighteen percent of the agencies said they did not use the evaluation information at all and nine percent state they did not use their evaluation information very much in their strategic planning.

It was noted previously that there are opportunities for nonprofit organizations to use their evaluation information for sharing with other agencies. Overall, the agencies reported that approximately sixty percent did not use their evaluation information to share with other agencies. The study hoped that there were some regions that collaborated more with their evaluation information than other regions.

165

The finding reveals that the region most collaborative in their evaluation efforts is the Western region, with fifty percent of the agencies stating that they use their information to share with their colleagues. The Eastern region (67%) and the Southern region (59%) are the two regions less likely to share their evaluation information. Using evaluation information collaboratively between agencies is an area for further study and research. Pooling valuable resources to address evaluation needs could lead to more sharing of evaluation information.

Using evaluation data to assist nonprofit organizations in their fund development has potential opportunities. Overall, approximately forty percent of the agencies do not use their evaluation information to help them with their fund development. Looking at the data regionally, one clearly sees some differences among the regions. Again, the Western region has seized the opportunity to increase their funding sources by utilizing the evaluation information in their fund development. One hundred percent of the agencies in the Western region state that they use the information somewhat (43%) or great deal (57%) to conduct their fund development activities. By contrast, forty percent of the agencies in the Eastern region and thirty-eight percent of the agencies in the Central region do not report using the evaluation information for fund development purposes.

166

Evaluation Utilization: Differences by Agency Revenue

The size of an agency based on annual revenues was an important characteristic in this study. The study looked to see if there are differences in how the agencies use their evaluation data based on the size of the agency. One of the ways to analyze the size of an agency is by its annual revenues. Annual revenue is one of the determinant factors funders look at in determining agency capacity to administer certain programs and services. Table 19 illustrates the agency's evaluation utilization based on the size of the agency's revenue. The agencies in this study were grouped into three categories according to the size of their annual revenues. Fifteen agencies (26%) of the respondents have annual revenues of $1 million or less, and twenty-seven agencies (48%) have revenues between $1 million and $3 million. The remaining agencies (26%) have annual revenues of more than $3 million.

The data revealed the more agency revenue, the more the agency used the information to improve their programs and services. Ninety-four percent of the agencies with $3 million or more state they used the evaluation information somewhat (7%) or a great deal (87%) to improve their programs and services. Agencies with revenues between $1 million and $3 million have fewer agencies (78%) than the agencies with $3 million or more to use their information to help their programs and services. This group of agencies reports that they use their

information somewhat (33%), or a great deal (45%) to improve their programs and services.

However, the fewer the resources, the less likely the agencies use their evaluation data. Agencies with revenues of $1 million or less have fewer agencies report that they used their evaluation information to help improve their programs. Of the seventy-six percent of the agencies with revenues of $1 million or less who use the information to improve their programs and services, seventy-three percent state they used the information a great deal.

Agency revenue did make a difference in whether or not an agency used the evaluation data to help improve their programs and services. There is approximately a fifteen percent or more difference between the evaluation utilization of agencies with $3million or greater and the other agencies with less agency revenues. Could agency revenue be a predictor of how the agencies used the evaluation information to increase agency awareness?

There are some interesting findings in the data. Logically, one would assume that the more revenue the agency had, the more evaluation information they would use to increase awareness of the programs or agency. The data reveals that the agencies with revenues of $1 million or less and the agencies with revenues of $3 million or more use their evaluation information at the same level (80%) to increase their agency's awareness.

Agencies with revenues between $1 million and $3 million did not use their evaluation information as much as the other agencies in the study. For this group of agencies, only sixty-five percent of the agencies used their information to increase agency awareness. This percentage is lower than the seventy-one percent of the overall agencies who said they used the evaluation information to increase agency awareness.

The percentage of agencies that used the evaluation information to report to their funders looked similar to the percentage of agencies that used the data to improve their programs and services. Ninety-three percent of the agencies with revenue of $3 million or more stated that they used their evaluation information to report to the funder. Of the ninety-three percent, eighty percent stated they used the information a great deal to report to the funder.

Agencies with annual revenues between $1 million and $3 million reported that eighty percent of the agencies used the information to report to the funder, with sixty percent of the agencies reporting that they used the information a great deal to report to the funder. Similarly, agencies with $1 million or less reported that seventy-nine percent of the agencies not only used the information, but used the information a great deal (79%) to report to the funder. One could infer that the less revenue the agency had, the more the funding source required the agency to conduct the evaluation.

Overall, sixty-eight percent of the agencies stated they used their evaluation information for strategic planning. However, the data looks differently when it is analyzed by agency revenue. Agencies with revenues between $1 million and $3 million reported using the evaluation data the least for strategic planning. Only sixty percent of the agencies in this group reported they used the information for strategic planning with thirty percent of this group stating they used it a great deal for strategic planning.

Approximately seventy-five percent of the agencies who had revenues of $1 million or less reported that they used the evaluation data for strategic planning. This finding was surprising because even with fewer dollars to conduct evaluations, the agencies found value in the evaluation information and used the data for planning and future capacity building. As suspected, the agencies with $3 million or more reported that eighty percent of the agencies used the information for strategic planning.

Sharing information among the agencies for evaluation has proven to have very little success in this study. This theme remains consistent when looking at how the agencies share information based on agency revenues. The less revenue the agencies reported, the less likely they were to share their evaluation information. Agencies with revenues of $1 million or less were the least likely (72%) to share their evaluation information. Of the seventy-two percent, forty-

three percent stated they did not share their information at all, while twenty-nine percent stated they did not share very much their evaluation information.

Thirty-six percent of the agencies who had revenues between $1 million and $3 million reported they did not share at all their evaluation information. Although, thirty-six percent of the agencies with revenues between $1 million and $3 million reported they did not share their evaluation information; this group had the highest percentage (48%) who shared their information with other agencies.

Surprisingly, agencies with $3 million or more did not share their evaluation information. Sixty percent of this group reported that they did not share at all (27%), or not very much (33%), their evaluation data. Overall, the agencies reported that approximately sixty percent of the agencies did not share their evaluation data. The size of the agency based on agency revenues does not make a difference in whether an agency will share their evaluation information with other agencies. Large, medium and small nonprofit organizations should look for opportunities to share evaluation information. Clearly there is a need for more agencies to share evaluation information based on the results of this study.

This study examined if the size of the agency's revenue made a difference in whether or not the agency used the evaluation information for fund development. More than fifty percent of the agencies overall said they used evaluation information for fund development. Based on this study, agencies with more

revenue used evaluation information less for fund development than did agencies with smaller annual revenues. Fifty-four percent of the agencies with annual revenue of $3 million or more reported they used the information somewhat (18%), or a great deal (36%) to assist in their fund development efforts. This contrasts with the agencies with revenues of $1 million or less. Seventy-six percent of this group reported they used the information somewhat (26%) or a great deal (50%) to conduct fund development. Agencies with revenue between $1 million and $3 million reported seventy percent of their agencies used the information somewhat (26%) or a great deal (44%) to improve to assist with fund development.

Evaluation Utilization: Differences by Staff Size

The size of an agency's staff is also a good indicator of agency capacity. Most of the research points to the lack of staff capacity as another barrier to conducting evaluations. Staff size in this study was measured in full-time equivalents (FTEs). The FTEs were grouped into three categories: agencies with less than ten FTEs; agencies who had between eleven and twenty-five FTEs; and agencies with more than twenty-five FTEs. A cross-tabulation of the uses of evaluations with staff size within the agencies provided interesting findings in table 19 (See page 159).

The assumption is that the fewer the number of FTEs an agency has, the fewer opportunities the agency has to use their evaluation results.

More than eighty percent of the agencies with less than ten or less full-time equivalents (FTEs) stated they used the evaluation information to improve their programs and services. Of the eighty percent, seventy-five percent said they used the information a great deal to improve their programs and services. Agencies with between eleven and twenty-five employees reported similar findings (82%) but stated that the agencies used the information a great deal by fifty percent. However, agencies with the most employees, more than twenty-five employees, reported almost ninety percent of their agencies used the information to improve their programs services.

Interestingly, agencies with more than twenty-five employees stated that eighty percent of them used the information either somewhat (33%, or a great deal (47%) to increase the agency's awareness. The agencies with staff between eleven and twenty-five FTEs used the evaluation information less (67%) than the other group of agencies to increase agency awareness. However, approximately fifty percent of the agencies who had between eleven and twenty-five FTEs stated they used the information a great deal to increase agency awareness.

Unexpectedly, seventy-five percent of the agencies with ten or less FTEs reported that they used the information to increase agency awareness. Of the

seventy-five percent, fifty percent indicated they used the information a great deal. However, agencies with FTEs between eleven and twenty-five used the evaluation data only sixty-seven percent to increase agency awareness.

Most of the findings to this point have demonstrated a strong utilization of the data to report to the funder. Analyzing the data based on staff size revealed similar results. Agencies with FTEs between eleven and twenty-five reported ninety percent of the agencies used their information to report to the funder with more than seventy percent indicating they used the data a great deal. Agencies with more than twenty five percent noted that eighty-six percent of the agencies used the data to report to funders, with over seventy percent of the agencies using the information a great deal. Lastly, eighty percent of the funders who had ten or less FTEs reported using the evaluation information a great deal to report to funders.

Does staff size make a difference in the utilization of evaluation information for strategic planning or fund development? Does staff size make a difference in whether or not an agency uses the data to share with other agencies? There was very little difference in evaluation utilization based on staff size for strategic planning. There were some differences between agencies when staff size was considered in the use of evaluation for fund development. There were small

differences between staff sizes when agencies used evaluation information to share with other agencies.

Sixty-nine percent of the agencies with ten or less FTEs stated they used the information for strategic planning, with fifty-six percent of those agencies stating they used the information a great deal for strategic planning. Agencies with the number of FTEs between eleven and twenty-five reported sixty-eight percent of the agencies using the evaluation data for strategic planning, but only thirty-two percent of the agencies reported using the information a great deal. Slightly lower percentages were found for agencies that had more than twenty-five FTEs. This group of agencies reported sixty-six percent of the agencies used the evaluation information for strategic planning, with fifty-three percent of them using the information a great deal.

There were interesting differences between staff size and the use of evaluation information for fund development. Agencies with more than twenty-five FTEs or more reported the lowest percentage of usage for fund development. This group of agencies reported that only fifty percent of the agencies used their evaluation information and only twenty-five percent used the data a great deal for fund development. Seventy-nine percent of the agencies with FTEs between eleven and twenty five stated they used the information somewhat (32%), or a great deal (47%). Even agencies with ten or less employees noted they used the evaluation

information for fund development (78%) with sixty-four percent of the agencies using the data a great deal for fund development.

The data to this point has not shown that agencies used the evaluation information to share with other agencies. An analysis by staff size did not show the agencies used the evaluation information to share with other agencies. Fifty percent of the agencies with between eleven and twenty-five FTEs share the evaluation information with other agencies. Even fewer agencies (33%) that had ten or less FTEs, used the information to share with other agencies. Likewise, agencies with more than twenty-five employees reported forty percent of the agencies used the evaluation information to share with other agencies, and of the forty percent, only thirteen percent used the data a great deal for this purpose.

Several agency presidents and chief executives officers were interviewed and asked what where the reasons why some agencies did not evaluate. N. Charles Anderson (2005), President and CEO for the Detroit Urban League said, "Agencies are too busy doing their work to evaluate. They do not see having the time or bearing the expense of evaluations." He stated that nonprofit agencies should do evaluations internally or have someone from the outside conduct the evaluation.

Howard Henderson (2005), President and CEO of the Baltimore Urban League suggested other reasons why some of the affiliates may not conduct evaluations.

He says, "There are some agencies that have been getting traditional funding that does not require evaluation. However, new funding will require agencies to do evaluations, particularly outcome evaluations." Henderson (2005) also suggests that if agencies do not currently conduct evaluations, that they should start. He offers collaborating with universities to conduct evaluations.

Warren Logan (2005) suggested the three reasons why agencies may not conduct evaluations he states:

> "First, it takes more effort to do evaluations, so some agencies will not do them. Second, some agencies, do not have a clear understanding or knowledgeable about evaluations. Lastly, the funding source the agency may report to may not stress the importance of conducting evaluations."

I asked several Presidents and CEOs for advice they would give to an agency that was not conducting systematic evaluations. James Shelby (2005) says, "Agencies are creating a death trap if you don't evaluate. Evaluation allows you to focus on areas where you need the greatest attention."

This chapter presented three characteristics of the population surveyed and findings on evaluation practices for nonprofit organizations. This chapter also explored the extent of nonprofit evaluation. It revealed nonprofit funders' expectations and how funders support nonprofit organizations. The chapter provided an analysis of nonprofit organization's perceived strengths and weaknesses along with uncovering the needs that nonprofits may have to conduct

evaluations. Lastly, the chapter discussed evaluation utilization for nonprofit organizations based on agency size and geographic location.

Urban League of Greater Oklahoma City

The previous sections in the data analysis chapter discussed the findings of the fifty-eight agencies who participated in the study. One of the affiliates that participated in the study was the Urban League of Greater Oklahoma City. As a participant-observer in the study, I was able to participate and observe the extent of nonprofit evaluations and understand the barriers to conducting systematic evaluations. The participant observer's role entails being present in everyday settings that enhance the awareness and curiosity about the interactions taking place around the researcher (Glesne and Peshkin 1992).

The Urban League of Greater Oklahoma City was a good agency to review in this study. Approximately three years ago, the agency started the first steps in learning how to evaluate its programs. The agency, based on the characteristics of this study, could be considered a medium size agency with between eleven and twenty-five full-time equivalents. Its annual revenues are between $1 million and $3 million. The agency's program budgets range from $5000 to $200,000 dollars. The agency has a diversity of funding sources: Thirty percent of the agency's funding comes from governmental grants (federal, state and local).

178

Approximately twenty percent of funds come from foundations with the remaining funds coming from private sources.

As previously stated, the Urban League of Greater Oklahoma City began evaluating its programs approximately three years ago with the encouragement from the one of its funders, the United Way of Central Oklahoma. The United Way of Central Oklahoma asked ten agencies to participate in a pilot program to teach nonprofit organizations how to conduct outcome measurement.

For many of the seventy-plus agencies in the United Way of Central Oklahoma network, many agencies were reluctant and even terrified about learning how to evaluate their agencies. Many agencies felt it was an unfunded mandate that they could not afford. Many of the agencies were reluctant because they felt they did not have the staff, time or capacity to evaluate their programs. The Urban League of Greater Oklahoma City volunteered to become on of the pilot agencies.

The agencies that participated in the pilot program were different sizes (staff size and agency revenues) and different in their level of knowledge and skills. Some of the agencies in the pilot program where affiliated with a national organization like the Urban League of Greater Oklahoma City, and some of the agencies were private nonprofit organizations. Lastly, some of the agencies in the pilot program had some type of evaluation process established, and some, like the Urban League, had no systematic evaluation process in place at that time.

The first step in the pilot process was to educate the nonprofit organizations on outcome evaluations. Two training meetings consisting of three to four hours were conducted to start the educational process. The agencies were also given a training manual as a resource to take back to their agencies.

The Urban League of Greater Oklahoma City focuses the majority of its evaluation training on its senior program staff. This included the vice president of operations and the department heads for all of the five departmental programs. The five departments at the Urban League of Greater Oklahoma City include Education and Workforce Development; Housing and Community Development; Community Health; Youth and Family Assistance; and Economic Development. Within each of the departments, the number of programs range from four to ten programs.

Challenges and Opportunities

In the initial phase of the evaluation process, the Urban League of Greater Oklahoma City was eager to undertake evaluating all of its programs and services simultaneously. We later learned that we did not have the time or resources for the massive undertaking. We re-shifted our focus and began evaluating three of the ten programs within the Youth and Family Assistance department. We found it was better to evaluate something on a small scale than not to evaluate at all. It

was important to get started, even on a small scale, by evaluating a single program or a discrete set of activities with the plan to expand the evaluation efforts at a later date (Rosenbaum 2003).

The second challenge was maintaining the level of engagement by staff in the process. The evaluation process is very time-consuming, and after one or two evaluation meetings, staff wanted to conclude the process. Program staff is very focused on serving the clients and have been trained to count the number of people served as a measure of success. Participating in a process that takes time to reach consensus among the agency staff was initially viewed by staff as counterproductive. At the end of the evaluation process, the staff found value in conducting the evaluations and are using the data in several ways.

The Urban League of Greater Oklahoma City took a theory-based approach to evaluation. The agency conducted logic models for three programs below:

Urban League of Greater Oklahoma City, Inc

LOGIC MODEL – "Give Back" Community Service Program

Long-term OUTCOME

Youth have improved social and school behaviors

Intermediate OUTCOMES

| Youth improve academic performance | Youth improve attendance at school | Youth improve attitudes about themselves, life & school | Parents increase involvement with their children |

Initial OUTCOMES

| Youth complete homework in timely manner | Youth complete community service | Youth learn new positive life skills | Parents learn parenting skills |

OUTPUTS

| # of youth involved monthly / yearly, by: school, grade, subject, age, gender, ethnicity | # hrs of community services by # of kids | # hrs of youth classes | # of tutors, mentors, staff and time spent | # of parents attending | # hrs of parent classes |

ACTIVITIES: Three project streams:
1."Diversion" life skills training for at-risk youth/1st time offenders (and their parents) referred to come to UL by court, parents, schools, churches, and/or other community groups
2. Juvenile Offenders Community Services Mentoring Program (JOCSMP) for youth sent to UL by City courts for tutoring and community service
3. Community Service Opportunities for youth mandated to do CS and who choose to do so with UL

INPUTS: UL staff, volunteers, contractors, reading materials, supplies, computers, advertisement, transportation and food. Partners including: courts, probation officers, case workers and judges. The program is supported through (donors). At-risk and adjudicated youth.

Urban League of Greater Oklahoma City, Inc
LOGIC MODEL – NULITES Leadership Development Program

Long-term OUTCOME

> Growing numbers of African American and minority youth serve as leaders and role models

Intermediate OUTCOMES

> - Youth plan to apply to and attend college
> - Youth improve attitudes about themselves, life & school
> - Youth volunteer in the community

Initial OUTCOMES

> - Youth learn positive life skills
> - Youth learn leadership skills
> - Youth learn about civic & community responsibility

OUTPUTS

> - # of youth involved monthly / yearly, by: school, grade, subject, age, gender, ethnicity
> - # of hours and/or sessions conducted
> - # of youth serving as tutors mentors
> - # of youth service projects / hours

ACTIVITIES:
Monthly meetings providing training (in positive life skills, leadership, financial fitness, ACT prep, interviewing, public speaking) and volunteer service opportunities (tutoring, mentoring, other voluntary projects)

INPUTS: Agency has 1 UL staff, volunteers, reading materials, supplies, computers, advertisement, transportation, and food. The program is supported through partnering agencies. African-American and other minority youth from various socio-economic backgrounds, ages 13 – 18 years.

183

Urban League of Greater Oklahoma City, Inc
LOGIC MODEL – UL Stars Program

Long-term OUTCOME

Youth are promoted to next grade in school

Intermediate OUTCOMES

Youth improve academic performance	Youth improve attitudes about themselves, life & school	Youth improve attendance at school	Parents increase involvem ent with their children

Initial OUTCOMES

Youth complete homework in timely manner	Youth learn new life skills	Parents learn parenting skills

OUTPUTS

# of youth involved monthly / yearly, by: school, grade, subject, age, gender, ethnicity	# of hours and/or sessions conducted	# of tutors, mentors, staff and time spent	# of parents attending	# of parent training sessions

ACTIVITIES:
Supplemental education in reading, math, science, English and chemistry to youth grades K-12
Mentoring and life skill modeling, including male to male, female to female
Parent-training and support for parents of youth in the program tutoring and mentoring

INPUTS: Agency has 1 UL staff, volunteers, reading materials, supplies, computers, advertisement, transportation, and food. The program is supported through partnering agencies and Public School districts. African-American and other minority youth from various socio-economic backgrounds, needing to improve grades and self-esteem

184

This process was a difficult for the staff because their focus had been on implementation or making sure the client received a service. With the theory-based evaluation process, the focus was on analyzing the program or services to see if the program produced the desired outcome of the program. The paradigm shift in thinking for staff was a slow and difficult process, but a process that was valued in the end.

The Urban League of Greater Oklahoma City did distribute surveys to their clients to access the level of satisfaction with the agency and its programs and services prior to the evaluation training. In the pilot program, new surveys were developed to determine if the program was operating as designed by the staff (see Appendix for Urban League of Greater Oklahoma City data collection tools). Once the tools were designed, the difficulty that followed for the agency was in the implementation process.

Although the department heads saw the importance of collecting the evaluation data, the program staff or, street level bureaucrats, found the data collection process to be time-consuming or a nuisance. Like so many other organizations, it was met by resistance from staff and the problem of managing institutional change (Giudice and Bolduc 2004). Continuous training of program staff and maintaining consistent messages on the importance of data collection by

those in the field is very important in changing the culture of program staff. It is very important to incorporate the evaluation process in the daily activities of staff.

With the successful implementation of the evaluation process for a few of the agency's programs, the Urban League of Greater Oklahoma City is looking to expand the process by continuing to evaluate the remaining programs in the Youth and Family Assistance Department with an expansion to other departments in the agency. Staff is aware that this process will not occur overnight, and resources will need to be implemented in the budgeting process to cover the cost of evaluation. It is possible that it will take several years to fully implement an agency-wide evaluation process. Putman (2004) reported that one of the evaluation experts estimated that a foundation should have three to five years to get everything in place.

Evaluation Utilization: Urban League of Greater Oklahoma City

The Urban League of Greater Oklahoma City is in its first year in the data collection and analysis process. The results have been very promising and the utilization of the evaluation data has been high. Staff used the data a great deal to report to funders. They also use the information a great deal in fund development in their proposal writings. The staff found the data very useful in marketing their

programs to the community. The agency is now incorporating evaluation in the strategic planning process.

Lessons Learned: Urban League of Greater Oklahoma City

Evaluation training is needed for nonprofit organizations. The lessons from participating in the United Way's outcome measurement pilot program for the Urban League of Greater Oklahoma City were many. First, there needs to be agency-wide training on what evaluation is and why it is important for all staff. There needs to be a commitment and understanding from the top down. This buy-in within the agency needs to start at the board of directors level and flow to the program staff and, even possible, to the client level.

Secondly, the training needs to be on-going and flexible to meet the needs of all of the agency staff. The third lesson is the board, staff and volunteers need to understand that evaluation is not an instant process. There may be several years before an agency-wide evaluation can occur. The key is to start evaluating something, then expand the circle to include other programs and departments.

The fourth lesson learned is that evaluation is costly. The United Way of Central Oklahoma provided the initial training for the outcome measurement process. However, funds for data collection and analysis parts to evaluations were not available. The Urban League of Greater Oklahoma City was able to utilize its own resources to continue the process. The United Way of Central Oklahoma is

187

providing the training for the logic model process but there is a concern by the researcher that many agencies will not have the funds to implement the data collection and maintain an evaluation system once they have been trained. It is very important for nonprofit organizations to incorporate the cost of evaluation into the budgeting process.

The Urban League of Greater Oklahoma City learned that systematic evaluation is a valuable tool that can be used to improve the agency for fund development and to increase agency awareness. Agencies that choose not to conduct evaluation will soon be agencies of the past.

CHAPTER 7

FINDINGS IN CONTEXT

The Canadian Centre for Philanthropy (CCP) study (2003) and this study on the National Urban League (NUL) affiliates seeks to uncover how to increase the number of quality evaluations among nonprofit organizations. The CCP study (2003) cites many challenges with the nonprofit evaluation sector. First, they note the task of evaluation has become more complex. The researchers suggest that as part of a general trend on the part of governments and other funders toward results-based management, there has been a shift away from input and output measurement to outcome measurement. Contributing to the complexity is the need to measure organization effectiveness as well as program effectiveness; the desire to involve stakeholders in the evaluation process; and the difficulties of measurement, such as trying to measure long-term effects in the short timeframes available for evaluation.

Second, the researchers state that evaluation requires an investment of time, money and staff expertise that many organizations are unable to make. Third, the evaluation techniques that are available are largely designed for business and government. Nonprofit organizations have different and more complex concerns, particularly when it comes to assessing those aspects of performance that are tied

189

to their social missions. Nonprofit organizations often have to demonstrate their effectiveness to a variety of stakeholders including beneficiaries, funders, partners, volunteers and employees, all of who may have somewhat different criteria for evaluating performance. Fourth, the current policy and political environment, which is highly risk-adverse and does not appreciate mistakes or failures, creates considerable pressure to "look good to avoid blame." In this context, the researchers suggest that it is difficult to report bad results, so nonprofit organizations have to "square sound evaluation with political imperatives." Given their missions aimed at serving community needs, nonprofit organizations may face a natural tendency to equate program needs with program effectiveness, thus limiting how well they make use of evaluation. The researchers also suggest that nonprofit organizations are often expected to respond to sometimes unrealistic evaluation demands by funders.

The CCP study (2003) stated that although nonprofit organizations appear to be under increasing pressure to provide information about their performance, there has been little understanding on how organizations approach evaluations. The CCP's in-depth research of nonprofit evaluation practices along with funders' expectations produced some very enlightening findings

Findings: Canadian Centre for Philanthropy and the National Urban League

<u>Extent of Nonprofit Evaluation</u>

A comparison of the two studies provided some interesting insights into the practices and needs in nonprofit evaluations. A detailed comparison of the findings between the Canadian Centre for Philanthropy (CCP) study (2003) and the National Urban League (NUL) study is provided below.

191

Table 20

Canadian Centre for Philanthropy (CCP) Data and
National Urban League (NUL) Affiliate Data
Summary

CCP Study	NUL Study
Key Characteristics of Survey Population	Key Characteristics of Survey Population

1. Organization's purpose/activity
2. Size defined by annual revenue:
  • 28% have revenues less than $125,000
  • 25% have revenues $125,000-$499,000
  • 27% have revenues $500,000 - $1.5 million
  • 20% have revenues over $1.5 million
3. Geographic Location –divided into six regions
	1. Organization's purpose/activity
	2. Size defined by annual revenue:
	  • 26% have revenues less than $1 million
	  • 48% have revenues $1 million - $3 million
	  • 26% have revenues over $3 million
	3. Geographic Location –divided into four regions

Funders' Expectations from nonprofits' view: • 44% expected more information • 49% expected more information on outcomes	Funders' Expectations from nonprofits' view: • 67% expected more information • 63% expected more information on outcomes
Reasonableness of Funders' expectations for nonprofits' view: • 91% said expectations were reasonable • 7% said expectations were unreasonable	Reasonableness of Funders' expectations for nonprofits' view: • 93% said expectations were reasonable • 7% said expectations were unreasonable
What was evaluated in past 12 months? • 75% conducted some type of evaluation in past 12 months • 66% evaluated on-going	What was evaluated in past 12 months? • 86% conducted some type of evaluation in past 12 months • 81% evaluated on-going

192

programs • 50% evaluated overall effectiveness • 50% evaluated projects • 45% evaluated fundraising activity • 40% evaluated volunteers • 33% evaluated board activity • 23% evaluated products • 23% did no evaluation	programs • 57% evaluated overall • effectiveness • 53% evaluated projects • 55% evaluated fundraising activity • 24% evaluated volunteers • 55% evaluated board activity • 31% evaluated products • 14% did no evaluation
What drives (motivates) evaluations? • 73% performed evaluations routinely • 21% performed evaluation to address specific issues • 4% performed evaluations only when required by funders	What drives (motivates) evaluations? • 75% performed evaluations routinely • 4% performed evaluation to address specific issues • 15% performed evaluations only when required by funders • 6% listed other reasons
Evaluations conducted: • 49% conducted between 1 and 5 • 51% conducted 6 or more	Evaluations conducted: • 38% conducted between 1 and 3 • 46% conducted between 4 and 6 • 16% conducted 7 or more
Main reasons for evaluations: • 73% board decision • 16% funder requirement • 4% for accreditation • 4% for new funding requirement	Main reasons for evaluations: • 13% board decision • 49% funder requirement • 4% for accreditation • 4% for staff reasons
Used data that would not have collected: • 48% used evaluation data • 76% used funder mandated data • 91% used data to improve programs	Used data that would not have collected: • 48% used evaluation data • 63% used funder mandated data • 83% used data to improve programs
Evaluation data collected: • 76% -projects/programs/activities	Evaluation data collected: • 86% --projects/programs/activities

• 66% --outcomes • 65% --client information • 54% --financial cost	• 78% --outcomes • 66% --client information • 60% --financial cost
Outcomes difficulty: • 70% --easy to identify outcomes data • 26% --difficulty to identify outcomes data • 73% --easy to collect outcomes data • 26% --difficult to collect outcomes data • 77% --easy to analyze outcomes data • 19% --difficult to analyze outcomes data • 80% easy to interpret outcomes data • 19% difficult to interpret outcomes data	Outcomes difficulty: • 68% --easy to identify outcomes data • 32% --difficulty to identify outcomes data • 62% --easy to collect outcomes data • 38% --difficult to collect outcomes data • 67% --easy to analyze outcomes data • 33% --difficult to analyze outcomes data • 73% easy to interpret outcomes data • 27% difficult to interpret outcomes data
Resources provided by funders: • 50% tools/manuals • 38% advice • 16% training	Resources provided by funders: • 59% tools/manuals • 74% advice • 41% training
Funding to conduct evaluations: • 57% internal sources • 19% external sources • 18% both internal/external sources	Funding to conduct evaluations: • 61% internal sources • 16% external sources • 23% both internal/external sources
Was funding adequate for evaluation purposes? • 50% adequate funding • 11% not adequate funding • 31% had no opinion	Was funding adequate for evaluation purposes? • 52% adequate funding • 48% not adequate funding

Lack of internal capacity by revenue size: Evaluation Strengths: • 94% --Understanding of projects/programs • 95% --Understanding of community needs • 91% --Positive attitude towards evaluation • 89% --Effective communication results • 82% --Evaluation skills/knowledge	Lack of internal capacity by revenue size: Evaluation Strengths: • 95% --Understanding of projects/programs • 87% --Understanding of community needs • 91% --Positive attitude towards evaluation • 83% --Effective communication results • 80% --Evaluation skills/knowledge
Evaluation Challenges (Barriers) • 67% --Internal capacity • 66% --Lack of money • 31% --Skills/knowledge • 28% --Activities in a good light • 25% --Value of evaluation • 20% --Confidence in ability • 12% --Difficulty in working with consultants	Evaluation Challenges (Barriers) • 62% --Internal capacity • 74% --Lack of money • 33% --Skills/knowledge • 49% --Activities in a good light • 16% --Value of evaluation • 53% --Confidence in ability • 43% --Difficulty in working with consultants
Unclear expectations from funders by revenue size: • 11% -- < $125,000 • 11% -- $125,000 - $499,000 • 11% -- $500,000 - $1.5 million • 9% -- >$1.5 million	Unclear expectations from funders by revenue size: • 26% -- < $1 million • 47% -- $1 million - $3 million • 27% -- >$3 million
Lack of money by revenue size: • 32% -- <$125,000 • 32% --$125,000 - $499,000 • 38% --$500,000 - $1.5 million • 37% --$>1.5 million	Lack of money by revenue size: • 26% -- < $1 million • 47% -- $1 million - $3 million • 27% -- >$3 million
• 35% -- <$125,000 • 36% --$125,000 - $499,000 • 36% --$500,000 - $1.5 million	• 27% -- < $1 million • 47% -- $1 million - $3 million • 26% -- >$3 million

• 39% --$>1.5 million	

One purpose of this study was to determine to what extent nonprofit organizations are conducting evaluations. The NUL survey revealed that fifty out of the fifty-eight agencies (86%) responded that they evaluated something within a timeframe of twelve months. Overall, forty-six percent of the agencies reported that they conducted between four and six evaluations, and thirty-eight percent indicated that they conducted between one and three evaluations within twelve months. The percentages are much lower (16%) for agencies who conducted seven or more evaluations.

More than three quarters of the nonprofit organizations in the Canadian Centre for Philanthropy (2003) study report that they have done some type of evaluation. Close to three quarters of the CCP agencies said they have evaluated as a matter of routine and not merely in response to funders demands. Most of the nonprofit organizations in the CCP study report using a variety of means to evaluate their work, ranging from staff and volunteer meetings to focus groups, interviews and

surveys. The study also finds that there was a substantial level of board involvement in the performance assessment process in most organizations.

The data revealed that agencies with less revenue performed more evaluations. Ninety-two percent of the agencies with revenues of less than $1 million reported conducting between one and six evaluations, and fifty-four percent of this group of agencies conducted between four to six evaluations. However, fewer agencies (8%) who had revenues of $1 million or less conducted seven or more evaluations.

Eighty-three percent of the agencies, with revenues between $1 million and $3 million, conducted between one to six evaluations, with forty-six percent conducting between four to six evaluations. Eighty-four percent of the agencies, with revenues of $3 million or more, conducted one to six evaluations within the twelve months and forty-two percent of this group of agencies performed between four to six evaluations.

The CCP study (2003) found that forty-nine percent of the agencies reported conducting between one and five evaluations and fifty-one percent of the agencies reporting conducting six or more evaluations. The CCP study (2003) also found a liner relationship between annual revenue and evaluation performance. Sixty-eight percent of the agencies with revenues of less than $125,000 performed evaluations. Seventy-six percent of the agencies with revenues between $125,000

and $499,999 conducted evaluations. Eighty-one percent of the agencies, with revenues of $500,000 or more, conducted evaluations.

The study NUL study found that agencies with more staff performed more evaluations. Eighty-six percent of the agencies with ten or less FTEs responded that they evaluated between one to six evaluations in the past twelve months. Agencies with twenty-five or more FTEs revealed that forty-two percent of this group conducted between one and three evaluations, and fifty-eight percent of the agencies conducted between four and six evaluations within a twelve month timeframe.

More than eighty percent of the respondents in the NUL survey noted that evaluated their ongoing programs. The eighty percent is higher than the sixty-six percent of the agencies in the CCP study (2003) who agencies evaluated their on-going programs. Only fifty-seven percent of the organizations in the NUL study responded they evaluated their agency's overall impact. Agency volunteers (24%) were evaluated the least in the NUL study but evaluated at a higher percentage in the CCP study (40%). With the push from funders for agencies to conduct outcomes data, the survey results were lower than expected; only seventy-eight percent of the NUL agencies reported they collected outcomes data and only sixty-six percent of the agencies in the CCP study.

Funders' Expectations and Support

The CCP study (2003) confirms a rise in expectations with respect to nonprofit evaluations and showed that although nonprofit organizations understand the need for increased evaluation practices, they were not well-equipped to meet these demands. Almost half of the nonprofit organizations surveyed report that funder expectations had increased over the previous three years. The survey results from the funders reveal that nearly half of the funders required the organizations they fund to do evaluations. The funders report that they expect more evaluation information than they have in the past, and they are increasingly looking for evaluations that report on the outcomes of the programs and projects they fund rather than those that report on outputs. Although funders are increasingly expecting outcome evaluations, it appears that nonprofit organizations may not fully understand what this entails.

Doug Easterling (2000) suggests that foundations are changing their focus from process evaluations to outcome evaluations. He shares that although foundations are becoming clearer about their expectations for outcome evaluation, the task facing grantees is often clouded and confused. He states that funders are no longer interested in activity based progress reports that count how many meals were served, how many beds were filled, how many addicts were treated, or how many offenders showed up for group counseling. He suggests that nonprofit

organizations must gather data that accommodates bottom-line thinking, showing concrete, tangible changes that resulted from the funder's support.

Bernett Williams, President and CEO for the Akron Ohio Urban League (Williams 2005) suggest that there is a strong possibility that many nonprofit agencies may not really understand the evaluation process. Williams says, "For some people, we think we are doing program evaluation but we are really not performing evaluation."

Even more revealing in CCP study (2003) than the increase in funders' expectations is the unfunded mandates the funders openly admit to as being a standard operating procedure in their business. The researchers report that the funders' increased expectations do not appear to be accompanied with increased financial support. Only about one in five funders reported increasing their funding support for evaluation. Less than half of the funders reported providing funding for evaluation or allowing project funds to be used for evaluation purposes. However, approximately six in ten funders reported offering advice on evaluation, and about half provide nonprofit organizations with evaluation tools and resources. Less than one fifth offer any training. These findings validates Nancy Rosenbaum's (2003) claim that funders want nonprofit organizations to be able to demonstrate the effectiveness of the programs they fund, but few are willing to subsidize the work.

The CCP study (2003) finds that the vast majority of nonprofit organizations is satisfied with the quality of their evaluations and believes that they use their results effectively. Like the Fine study (1998), the CCP study (2003) reveals that satisfaction and perceived accuracy of evaluation results appear to be higher when nonprofit organizations used external evaluators to conduct their evaluations and when evaluations are done in response to funder requirements.

The researchers also report that both nonprofit organizations and funders agree that nonprofit organizations have many strengths that they bring to the evaluation process. The ability of nonprofit organizations to understand program and project activities and their ability to understand community needs is a principal strength. However, the report states that nonprofit organizations tend to have a more positive view about the strengths they possess than do their funders. The nonprofit organizations in the CCP study (2003) also reports having problems in a number of areas. They point to the lack internal capacity, such as staff or time, and a lack of money as being particularly big problems. Other problems include unclear directions from funders about what is expected in an evaluation and lack of skills and knowledge in conducting evaluations. From these challenges, the nonprofit organizations identify a need for more financial resources, better access to technology, and a greater consistency on the part of funders in their use of terminology. The funders also identify several needs to

improve the evaluation process for funders. They suggest that funders ask for similar evaluation information in cases where there are multiple funders, the need for more financial resources and the need for greater consistency on the part of funders in their use of evaluation terminology.

The NUL affiliate study also finds that nonprofit organizations are conducting evaluations. However, sixty-seven percent of the evaluations conducted are mandated by their funders, and there is an increasing demand on nonprofit organizations based on the funders' expectations. The results of the NUL study revealed that governmental and foundation funding plays a much more significant role than the United Way in motivating nonprofit organizations to conduct evaluations. Most of the evaluations conducted by the agencies were to comply with governmental grants (74%), and less than twenty percent (16%) were conducted for a United Way funding source.

An assessment by regions in the NUL study, showed more evaluations were required by funders in the Southern (94%) and Western (100%) regions. Seventy-six percent of the agencies in the Southern region were required by their funders to conduct between one and three evaluations, and seventy-one percent of the agencies in the Western region were required to conduct between one and three evaluations.

Overall, sixty-three percent of the agencies said they had to collect information that they would not have collected if they were not required to collect it for the funder. Only fifteen percent said they were not required to collect information that they would not have otherwise collected.

Sixty-seven percent of the agencies responded that the amount of information requested by the funder has increased, and approximately twenty percent said the amount of information provided did not change. However, ninety-three percent of the respondents felt the funders' expectations were reasonable, and of the ninety-three percent, approximately twenty percent indicated the funders' expectations were very reasonable. Similar results were found in the CCP (2003) study on the reasonableness of funders' expectations. Ninety-one percent of the agencies in the CCP (2003) study said the funders' expectations were reasonable. These results indicate a willingness by the agencies to collect the information requested to conduct evaluations.

In terms of funders' support, seventy-four percent of the NUL agencies said they were given advice from the funder, and approximately sixty percent of the agencies said they were given resource guides or manuals on evaluations. An interesting statistic was that only forty-one percent of the agencies were provided training from the funders.

The NUL study percentages were higher in terms of funders' support than the CCP (2003) study. Agencies in the CCP (2003) study state that fifty percent received tools from their funders. While thirty-eight percent of the agencies received advice and another sixteen percent received training from funders. Training is the lowest area of support provided by funders. This is a possible opportunity for collaborative efforts with funders and evaluation experts who help nonprofit organizations.

Agency Strengths and Weaknesses

The greatest perceived strength nonprofit organizations felt they had, as it relates to evaluation, was in understanding their programs and projects (95%). This was followed by ninety-one percent of the respondents who believe they have a positive attitude toward evaluation and eighty-seven percent felt their strengths lay within understanding community needs. Only a small number (12%) of agencies believed their strength was funding for evaluation. Fifty-six percent of the agencies reported they could not secure funding for evaluations either none at all (23%), or they had little success (33%). Eighty percent of the agencies believed they did possess a moderate to strong level of evaluation skills or knowledge.

The findings from the NUL study revealed that the agencies had very little difficulty with outcome evaluation. Only thirty-eight percent stated they found it very difficult (4%) or somewhat difficult (34%) in collecting outcomes data. The other area of slight difficulty was in identifying and analyzing of outcomes data. Thirty-two percent said that they found identifying outcomes data very difficult (4%), while others found it somewhat difficult (28%). Only thirty-three percent of agencies said that analyzing outcomes data was somewhat difficult. The level of difficulty with outcomes data was similar in the CCP (2003) study. This may indicate that once the agencies are properly trained in outcomes evaluation, they find outcome evaluation less difficult.

Another interesting note is that, once the nonprofit agencies had the information, they felt comfortable in effectively communicating the evaluation results. Eighty-three percent of the agencies reported they effectively communicated their evaluation results, with approximately forty percent stating they did a great job in communicating their evaluation results.

Like other authors, Emmett Carson (2000) states that evaluation places considerable burdens on the nonprofit organization whose project is being assessed. He notes that the more extensive and comprehensive the evaluation, the more it will divert the time, energy and resources of the nonprofit organization away from the primary purpose of the project. He stresses that it is important for

205

both the foundation and the grantee to have a clear understanding of what the evaluation is going to measure, what constitutes strong or weak performance, how the findings will be used, and with whom the results will be shared.

Limited funding and limited staff for evaluations were the greatest barriers to systematic evaluations for nonprofit organizations in both studies. Seventy-four percent of the agencies in the NUL study perceive funding to be the number one barrier in conducting evaluations. Thirty-six percent of the agencies state that this was a moderate problem while thirty-eight percent stated that this was a big problem. Sixty-two percent of the agencies perceive lack of staff time to be a major hindrance to evaluations. More than twenty-seven percent of the agencies stated that lack of staff time was a moderate problem, and thirty-five percent said lack of staff time was a big problem. Fifty-six percent of the agencies stated the lack of tools was a problem. Of the fifty-six percent, forty-four percent stated this was a moderate problem. The CCP (2003) study also noted the lack of internal capacity or lack of staff time and limited funding as the biggest barriers to evaluations.

However, understanding the value of evaluation was not a problem for the group of agencies in this study. Eighty-four percent of the agencies reported they had no problem (53%), or a small problem (31%) with understanding the value of evaluation. This is an important finding because it shows that nonprofit

organizations understand the importance of evaluation but several factors prevent systematic evaluations

The CCP (2003) study and the NUL study had similar results pertaining to funding their evaluation efforts. The CCP (2003) study found that fifty-seven percent of the agencies they surveyed secure funding for evaluation internally. Nineteen percent use external funding. Eleven percent of the agencies note that the funding they have available for evaluation is not adequate.

Approximately sixty percent of the agencies in the NUL study said they cover the cost of the evaluations through their general operating fund, and only sixteen percent stated they receive funding for evaluation from an external funder. Forty-eight percent of the agencies said the funds they have to conduct the evaluations are not adequate. Only twelve percent of the agencies have someone dedicated onsite to conduct evaluations.

Clearly, resources are needed to assist with more and better systematic evaluations and more funding should be sought from program sources. Agencies need to look for opportunities to incorporate evaluation into their program and agency's budgets. More funding is needed to hire evaluators or to hire staff devoted to evaluation. In addition to money and staff, more evaluation training materials are needed.

Evaluation Utilization

Utilization of evaluation information for nonprofit organizations is rarely discussed. Sawhill and Williamson (2001) provide two good uses for nonprofit evaluations: marketing and management. The authors believe that measures of success can serve as powerful marketing tools for nonprofit organizations. More importantly, nonprofits that have adopted and implemented effective systems of performance measures will be well positioned to take advantage of this trend. They also found that performance measures help establish a culture of accountability within nonprofits.

Both the NUL study and the CCP (2003) study found that when the nonprofit organization conducted the evaluations, they did use the data to improve their programs, increase agency awareness, and for strategic planning. Although eighty-three percent of the agencies in this study stated they use their evaluation data to report to the funder, they also use the information for other purposes.

Sixty-two percent of the agencies reported using the information to improve their programs and services. In the CCP study, ninety-one percent of the agencies said they used the data to improve programs. This would suggest that there are some residual benefits in collecting evaluation information for nonprofit organizations that are required by their funders.

Approximately seventy percent of the agencies in the NUL study use the evaluation information to increase their agency's awareness or to market their agency's, programs or services. The agencies in the CCP study (2003) used the evaluation data most for strategic planning (87%) and to increase awareness (74%). The agencies used the information less to report to funders (56%), fundraising (52%) and sharing information with other agencies (47%).

Evaluation data is used by the NUL agencies approximately seventy percent in strategic planning for the agencies, with forty-four percent of the agencies saying they use their evaluation data a great deal for strategic planning. Approximately seventy-five percent of the agencies who have revenues of less than $1 million reported they used the evaluation data for strategic planning. The agencies find value in the evaluation information and use the data for planning and future capacity building. As suspected, the agencies with $3 million or more report eighty percent of the agencies use the information for strategic planning. With the push for nonprofit organizations to conduct strategic planning, incorporating evaluation information in the strategic planning process is an opportunity for nonprofit organizations.

The evaluation data is used less by the agencies for fund development and to share with other agencies. Although the agencies state they use their evaluation data approximately seventy percent for fund development, only forty-five percent

stated they used their information a great deal for fund development. The lowest utilization for evaluation information by the agencies indicated in the study was sharing the evaluation data with others.

Sixty percent of the agencies state they do not share their information at all (35%), or not very much (25%). This is a missed opportunity for nonprofit organizations as they look for additional ways to leverage evaluation resources. There are several reasons why nonprofit organizations may not share their evaluation information with other agencies.

First, there maybe turf issues among agencies locally and regionally. The funding among nonprofit organization is extremely competitive. Sharing what works and what needs improvement by an agency that is currently delivering the program or services could open the agency up for a hostile takeover. Although the clients or customers who receive the programs or services would benefit more from agency collaborations, the reality is that agencies are very reluctant to share their weaknesses.

The second reason why agencies may not share the information with other agencies is due to the lack of staff time. Most staff time is built into program delivery and reporting. Agencies who find the time to conduct and report their evaluation findings usually report their findings to the funder or the board of directors. The third reason for not sharing the information with others may be the

lack of an established evaluation network among agencies. Without this network, agencies may not know who to seek out to share the information or who to look for with similar programs.

The NUL study found that there were slight differences among the agencies based on geographic location and whether or not the agencies used their evaluation information to improve their programs and services. More agencies used their evaluation information in the Western and Southern regions than in the Eastern and Central regions. The findings also revealed that there were regional differences among the agencies that used the data to increase agency awareness. The Western region was the clear leader with approximately ninety percent of the agencies saying they used the evaluation information to increase their agency's awareness. Of the eighty-seven percent of the agencies in the Western region who use the information for agency awareness, seventy-five percent of the agencies said they use the information a great deal.

Regional differences emerged from the findings indicated that some regions use the evaluation information more than others to report to their funders. Approximately eighty percent of the agencies report they use their evaluation information either somewhat (12%) or a great deal (67%) to report to their funders. However, one hundred percent of the Western region said they use their evaluation information to report to their funders

211

The study found that there were differences in how the agencies use their evaluation data based on the size of the agency. The data revealed the more agency revenue, the more the agency uses the information to improve their programs and services. Ninety-four percent of the agencies with $3 million or more state they use the evaluation information somewhat (7%) or a great deal (87%) to improve their programs and services. Agencies with revenues between $1 million and $3 million had fewer agencies (78%) than the agencies with $3 million or more to use their information to help their programs and services. This group of agencies reported that they use their information somewhat (33%) or a great deal (45%) to improve their programs and services.

Overall, the agencies reported that approximately sixty percent did not use their evaluation information to share with other agencies. The region that is most collaborative in their evaluation efforts was the Western region with fifty percent of the agencies stating that they used their information to share with their colleagues. The Eastern region (67%) and the Southern region (59%) were the two regions less likely to share their evaluation information. What is proven is that the size of the agency based on agency revenues does not make a difference in whether an agency will share their evaluation information with other agencies. Large, medium and small nonprofit organizations should look for opportunities to

share evaluation information. Based on the results of the study, it shows there is a need for more agencies to share evaluation information.

Using evaluation information collaboratively between agencies is an area for further study and research. Pooling valuable resources to address evaluation needs could lead to more sharing of evaluation information.

The results from both studies indicate that evaluation is taking place among nonprofit agencies but the quality and quantity of the evaluation may be subject to the amount of funding, internal capacity and training available to the agencies.

CHAPTER 8

CONCLUSIONS

Summary

The literature on evaluations in the nonprofit sector is limited. In general, when the literature has been studied, the evaluation efforts have been focused on output measures. Knowing how many people were served has been the focal point in evaluation measures on nonprofits. Nonprofit organizations that lack evaluations are limited in their ability to demonstrate positive change or results and do not have information necessary to make informed decisions about day-to-day management on future programming and resource allocations. The lack of evaluations performed by nonprofit organizations not only affects the nonprofit organization, but also the clients or customers they serve.

Nonprofit organizations need to conduct systematic evaluations that assess the process, outcomes and the cost of the programs they administer. These organizations need information to replicate good programs, determine outcomes, measure the cost of a success, and address areas that can be improved. Evaluations for nonprofit organizations need to be as comprehensive and rigorous as possible.

214

The purpose of this dissertation was to assess the current status of evaluation activities and the utilization of their results for the nonprofit sector in the United States. The dissertation examined how much evaluation is done by nonprofits, what types of evaluations are being conducted, and how evaluations are being used. The dissertation looked to see if evaluations were being used. The dissertation uncovered and answered the question of why nonprofit organizations do not use evaluations.

Replicating the voluntary sector portion of the Canadian Centre for Philanthropy's study (2003), this dissertation answered questions concerning evaluation practices for nonprofit organizations in the United States. Specifically, the study examined the evaluation practices of over 100 nonprofit organizations in the United States that are affiliates of the National Urban League.

The problem in the nonprofit sector is making evaluation an essential function in the daily operations for the nonprofit organization. Similar to management and fiscal skills, nonprofit organizations need to be trained in evaluation. Training for nonprofit organizations has focused on leadership (board and management) training and accounting or fiscal training. The next development step for nonprofit organizations is in evaluation. A paradigm shift is needed in the nonprofit sector in order for nonprofit organizations to take the next development step. Nonprofit organizations need to recognize they will not be able to continue to operate

215

without a systematic evaluation system. The greater accountability by government and private sector funders for results will necessitate the nonprofit sector to become evaluation savvy. This will require more funds for training and implementation of evaluation systems.

Nonprofit agencies find evaluation information useful but have limited training, tools and money to conduct evaluations. Nonprofit organizations need to find methods to conduct evaluations given their limitations. Pooling valuable resources to address evaluation needs could lead to more sharing of evaluation information. Using evaluation information collaboratively between agencies is an area for further study and research.

General Recommendations for Increasing Systematic Evaluations

The findings showed the nonprofit organizations in this study did evaluate something within a twelve month timeframe, but they conducted evaluations mostly because they were required to comply with the funders' requirements. However, a residual benefit from the required evaluation was that the nonprofit organizations used the data to improve their programs, increase agency awareness and for strategic planning. One of the disappointments in the study was the lack of evaluation collaboration among nonprofit organizations.

Several recommendations arise from the dissertation findings. First, nonprofit organizations should negotiate with funders to add an evaluation component in the program delivery costs. In other words, when writing program grants, one should try to incorporate the cost of evaluation into the program.

Second, nonprofit organizations should think of evaluation as an essential component of their operations and budget the cost of conducting evaluations within the program and agency budget. Nonprofit organizations should allocate a percentage of their general operating budget for research and evaluation that fits within the agency's strategic plan. They should build the evaluation into the program's general and operating proposal budgets as a fixed-cost line item (Rosenbaum 2003).

I also recommend that agencies look for opportunities to collaborate on the cost of evaluation and the use of evaluation data. The agencies reported that approximately sixty percent did not use their evaluation information to share with other agencies. Large, medium and small nonprofit organizations should look for opportunities to share evaluation information. Based on the results of this study, there is a need for more agencies to share evaluation information.

Second, I recommend that foundations and government funders should fund evaluations as part of their grant allocation to nonprofit agencies. More funds are needed to fund and sustain systematic evaluations. Funders should look for

217

opportunities to be flexible in the use of program funds for evaluation efforts, and leverage other resources to assist in funding evaluations.

My final recommendation is for the nonprofit sector to have access to more training materials or resources for evaluation purposes. Although there are numerous tools designed for use by nonprofit organizations, many of these tools exist in unpublished forms and are not widely distributed. Many tools also fail take into account the diversity of nonprofit organizations in terms of their skills sets of their staff and volunteers, their financial resources, and the types of programs they deliver (Bozzo 2000).

What is needed are tools with practical applications for a variety of nonprofit organizations with ease of distribution. More online manuals and simple "how to evaluate for nonprofit organizations" publications are needed.

Recommendations for Increasing Systematic Evaluations for National Urban League Affiliates

There are several recommendations I make to assist the National Urban League affiliates in increasing the number of quality systematic evaluations. There should be:

218

1. ***A commitment from National Headquarters to assist in evaluation training & resources.*** There should be a commitment from the National Urban League headquarters to the affiliates to provide training opportunities, resource materials and possible funding opportunities for affiliates. There should be a commitment from the National Urban League headquarters that every affiliate should be conducting some system of evaluation within five years. Systematic evaluations should be on of the criteria incorporated within the National Urban League Performance Assessment tool.

2. ***A training curriculum needs to be implemented into phases.*** There should be training on:

   - The basic overview of evaluation that provides education on what is evaluation; the importance of evaluation; benefits to evaluations; funding for evaluation, evaluation utilization, etc.

   - How to design an evaluation plan

   - The design and implementation of a process evaluation

   - The design and implementation of outcome evaluation

   - Data collection methods that include some level of design training and implementation

   - Analyzing and interpreting evaluation information

219

The Affiliate Development Department of the National Urban League should utilize evaluation consultants that specialize in nonprofit evaluations to train the NUL CEO's, Affiliate Board Chairs and auxiliary groups.

3. *Ongoing training for all National Urban League Affiliate board & staff and auxiliaries*

There should be training opportunities for affiliate board members; Chief Executive Officers (CEOs) affiliate staff and the volunteer auxiliaries on evaluation. The training should be incorporated into the current training forums each year. Presidents and Chief Executive Officers should be trained on evaluation in "New CEO training" and at every Mid-winter conference. Board members should be trained at regional conferences and national conferences. The training format for board members should be different than the training for affiliate staff. Board members should be trained on the basics of evaluation; what type of evaluation to expect from their affiliate given the affiliate's resources and capacity; and how to use evaluation for policy and strategic planning. Staff should receive evaluation training at the Whitney M. Young Institute. Staff training should focus more data collection and implementation. However, staff needs to be trained on the basics of evaluation before they can master evaluation methods.

4. ***Tools and templates for NUL affiliates.*** The National Urban League should recommend or develop evaluation tools and/or templates for affiliates to implement once they have received the basic training and have a general understanding on how to implement an evaluation plan.

What is critical to the success of the recommendations listed above is the collective buy-in by the National Urban League network that evaluations will increase affiliate capacity and improve programs.

Recommendations for Future Research

There is a need in the evaluation community to conduct more research on systematic nonprofit evaluations. Future research is needed on how to train nonprofit organizations effectively in conducting systematic evaluations and how to make training materials more accessible for nonprofit organizations. Research is needed on the type of evaluation tools that are most effective for nonprofit organizations, and which tools can be applied in the nonprofit sector.

Using evaluation information collaboratively between agencies is an area for further study and research. Pooling valuable resources to address evaluation needs could lead to more sharing of evaluation information. Research is needed to determine why nonprofit organizations have not used evaluation information

more collaboratively and how to identify possible opportunities for sharing evaluation resources.

Bibliography

Abramson, Alan. "Improving Nonprofit Performance: Key Issues and Promising Approaches." *Snapshots: Research Highlights from the Nonprofit Sector Research Fund.* 31 (Sept 2003): 1-6.

Anderson, N. Charles. President and Chief Executive Officer of the Detroit Urban League. Interview by author, 9 April 2005, Legislative Policy Conference, Washington, D.C.

Bamberger, Michael, Jim Rugh, Mary Church, and Lucia Fort. "Shoestring Evaluation: Designing Impact Evaluations under Budget, Time and Data Constraints." *American Journal of Evaluation,* 25 (1), (2004): 5-37.

Bozzo, Sandra L. "Evaluation Resources for Nonprofit Organizations: Usefulness and Applicability." *Nonprofit Management and Leadership,* 10 (4) (Summer 2000): 463-470.

Bozzo, Sandra L., and Michael H. Hall. *A Review of Evaluation Resources for Nonprofit Organizations.* Canadian Centre for Philanthropy. (March 1999): 1 -18.

Brame, Walter. President and Chief Executive Officer of the Grand Rapids, Michigan Urban League. Interview by author, 7 April 2005, Legislative Policy Conference, Washington, D.C.

Brett, Belle, Lynnae Hill-Mead, and Stephanie Wu. "Perspectives on Evaluation Use and Demand by Users: The Case of City Year". *New Directions of Evaluation,* Jossey-Bass. 88 (Summer 2000).

Brooks, Arthur C. "Can Nonprofit Management Help Answer Public Management's Big Questions"? *Public Administrative Review;* 62,3 (May 2002): 259-266.

Canadian Centre for Philanthropy. Assessing Performance: Evaluation Practices and Perspectives in Canada's Voluntary Sector (2003): 1-37.

Carman, Joanne G. "Evaluation theory vs. practice: What are nonprofits really doing?" Paper submitted to the Association for Research on Nonprofit Organizations and Voluntary Action, 32nd Annual Conference. (22 November 2003): 1-13.

Carson, Emmett D. On Foundations and Outcome Evaluation. *Nonprofit and Voluntary Sector Quarterly*, 29 (3), (September 2000): 479-481.

Center of Effective Philanthropy. Indicators of Effectiveness: Understanding and Improving Foundation Performance. Boston, Mass. (2002).

Chen, Huey-tsyh. Current Trends and Future Directions in Program Evaluation. Evaluation Practice, 15 (3) (1994): 229-238.

Chen, Huey-tsyh and Peter Rossi. Integrating Theory into Evaluation Practice in Using Theory to Improve Program and Policy Evaluations. Greenwood Press, Westpost, CT. (1992).

Cronbach, Lee J. and Associates. Toward Reform of Program Evaluation. San Francisco: Jossey-Bass.(1980).

DeVita, Carol J., Cory Fleming, and Eric C. Twombly, "Building Nonprofit Capacity: A Framework for Addressing the Program." *Center for Nonprofit and Philanthropy*, The Urban Institute. (June 2001): 5-91.

Dibella, Anthony. "The Research Manager's Role in Encouraging Evaluation Use," *Evaluation Practice*, 11(2) (1990): 119.

Dinesh, D., Palmer, E. "Management By Objectives and the Balanced Scorecard: Will Rome Fall Again?" *Management Decision*. 36 (1998) ; 363-369.

Easterling, Doug. Using Outcome Evaluation to guide grant making: Theory, reality and possibilities. *Nonprofit and Voluntary Sector Quarterly.* 29 (3) (Sept. 2000) 482-487.

Fine, Allison H. Colette E. Thayer and Anne Coghian. Program Evaluation Practice in the Nonprofit Sector. Innovation Network Inc. (1998).

Fitzpatrick, Jody L., James Sanders and Blaine R. Worthen. Program Evaluation: Alternative Approaches and Practical Guidelines. 3rd edition. Pearson Education (2004).

Forbes, Daniel. Measuring the Unmeasurable: Empirical Studies of Nonprofit Organization Effectiveness from 1977 to 1997. *Nonprofit and Voluntary Sector Quarterly*, Sage Publications, Inc., 27 (2) (June 1998): 183-202.

Formative Evaluation Research Associates. 2005. http:/www.feraonline.com.

Franklin, Aimee. Managing for Results in Arizona: A Fifth-Year Report Card. Public
    Productivity & Management Review, 23 (2) (1999): 194-209.

Giudice, Phil and Kevin Bolduc. Assessing Performance at the Robert Wood Johnson
    Foundation: A Case Study. *The Center for Effective Philanthropy.* (2004): 1 –
    15.

Glasrud, Bruce. The muddle of outcome measurement. *Nonprofit World.* 19 (6): 3-5.

Glesne, Corrine and Alan Peshkin. Becoming Qualitative Researchers: An Introduction.
    Longman. White Plains, New York. (1992).

Grantmakers Evaluation Network (GEN) and Grantmakers for Effective Organizations
    (GEO) conference. 2000. High-Performance Organizations: Linking Evaluation
    & Effectiveness. Report on the Grantmakers Evaluation Network (GEN) and
    Grantmakers for Effective Organizations (GEO). Fairmont Hotel. Kansas City
    Missouri. (2002).

Greene, Jennifer C. Evaluation Extrapolations. *American Journal of Evaluation,* 22 (3)
    (2001): 397-402.

Hall, Peter Dobkin. "Historical Perspectives on Nonprofit Organizations", in the
    Handbook of Nonprofit Leadership and Management. San Francisco, California:
    Jossey-Bass Inc. Publishers (1994).

Hall, Peter Dobkin. (2003). "A solution is a product in search of a problem": A History
    of Foundations and Evaluation Research. Retrieved Mar. 2003 from
    http://ksghome.harvard.edu/~.phall.hauser.ksg/EVALUATION%20ESSAY.pdf.

Hatry, H.P., Van Houten, T., Plantz, M.C., Greenway, M.T. *Measuring Program
    Outcomes: A Practical Approach.* Alexandria, VA.: United Way of America.
    (1996).

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

Hatry, Harry and Linda Lampkin. An Agenda for Action: Outcome management in
   Nonprofit Organizations. *The Urban Institute*. (Dec. 2001): 1-21

Henderson, Dale A. Bruce W Chase; Benjamin M Woodson. **Performance measures
   for NPOs.** *Journal of Accountancy;* 193, 1: (Jan. 2002): 63.

Henderson, Howard. President and Chief Executive Officer of the Baltimore, Maryland
   Urban League. Interview by author, 8 April 2005, Legislative Policy Conference,
   Washington, D.C.

Herman, Robert D., and David O. Renz. (2002). Nonprofit Organizational
   Effectiveness: Practical Implications of Research on an Elusive Concept: 1-11.
   Retrieved from www.bloch.umkc.edu/cookingham.

Hubelbank, Jeanne H. Evaluation as a Catalyst for Change. Paper presented at the
   Annual Assessment Conference of the American Association of Higher Education
   (June 23, 2001): 3-5.

Johnson, R.B. Toward a theoretical model of evaluation utilization. Evaluation and
   Program Planning. 21 (1) (1998): 93 – 110.

Kellogg, W.K. Foundation. 1998. Evaluation Handbook. Battle Creek, MI.

Lewis, Annelle. Senior Executive Vice President for Affiliate Development for the
   National Urban League. Interviewed by author in March 2002 and February
   2004.

Light, Paul C. *Making nonprofits work: A report on the tides of nonprofit management
   reform.* Washington, D.C. The Aspen Institute Brooking Institution Press.
   (2000).

Lipsky, Michael and Steven Rathgeb Smith. *Nonprofits for Hire.* Cambridge,
   Massachusetts: Harvard University Press. (1993).

Logan, Warren. President and Chief Executive Officer of the Chattanooga, Tenn. Urban
   League. Interview by author, 7 April 2005. Legislative Policy Conference,
   Washington, D.C.

Macpherson, Malcolm. Performance Measurement in not-for-profit and public-sector organizations. *Measuring Business Excellence*; 5 (2) ABI/INFORM Global (2001): 13-17.

Mattessich, Paul W. The Manager's Guide to Program Evaluation: Planning, Contracting, and Managing for Useful Results. Saint Paul, MN. Wilder Publishing Center. Amherst H. Wilder Foundation. 2003.

McKinsey & Company. . "Effective Capacity Building in Nonprofit Organizations." *Venture Philanthropy Partners*.(2001): 2-113.

Mott, Andrew. Evaluation: Good News for Funders. *Neighborhood Funders Group*. Washington, D.C. (2003): 1.

Murray, Vic. "The State of Evaluation Tools and Systems for Nonprofit Organizations." *New Directions for Philanthropic Fundraising*, 31 (Spring 2001) 39-49.

Murray, Vic. "Types of Evaluation: Choosing The Right One For You." School of Public Administration, University of Victoria (2001).

National Urban League. "Policies, Standards and Procedures Manual for Affiliates of the National Urban League, Inc." (2002).

New Pittsburgh Courier/Pittsburg Post-Gazette, Community Service Supplement. July 23, 2003, pages UL4 – UL6.

O'Sullivan, Elizabethann., and Steven Lesky. Community, Nonprofit and Program Evaluation. A paper presented to 2000 Annual Conference of the American Evaluation Association, Washington, D.C., (1November 2000).

O'Sullivan, Elizabethann., and Edwin D. Miller. *One Program, Many Funders –How Many Evaluations?* A paper presented to 2002 Annual Conference of the American Evaluation Association, Washington, D.C., (6 November 2002).

Patton, Michael Quinn. Practical Evaluation, Newbury Park, CA: Sage Publications. (1982).

Patton , Michael Quinn. . Utilization-Focused Evaluation: The New Century Text. 3rd Edition. Thousand Oaks, Calif. Sage Publications. (1997).

Patton, Michael Quinn. On Evaluation Use: Evaluative Thinking and Process Use. "Where We've Been and Where We're Going: Experts Reflect and Look Ahead." *The Evaluation Exchange: Harvard Family Research Project.* 9 (4) (Winter 2003): 4-5.

Plantz, Margaret C., Martha Taylor Greenway, and Michael Hendricks. "Outcome Measurement: Showing Results in the Nonprofit Sector". *New Directions For Evaluation,* Jossey-Bass Publishers. 75 (1997).

Poole, Dennis L. Jill K. Davis, Jane Reisman and Joan E. Nelson. Improving the Quality of Outcome Evaluation Plans. *Nonprofit Management and Leadership,* 11 (4) (Summer 2001: 405-421.

Posavac, Emil J. and Raymond G. Carney. Program Evaluation: Methods and Case Studies. Fourth Edition. Prentice Hall Englewood Cliffs, NJ. (2002).

Putnam, Kirsten. Measuring Foundation Performance: Examples from the Field. California Healthcare Foundation. (2004).

Rist, Ray. On Evaluation Utilization: From Studies to Stream. "Where We've Been and Where We're Going: Experts Reflect and Look Ahead." *The Evaluation Exchange: Harvard Family Research Project.* 9 (4) (Winter 2004): 5-6.

Rogers, Patricia J. Timothy A. Hacsi, Anthony Petrosino and Tracy Huebner. Program Theory in Evaluation: Challenges and Opportunities. *New Directions for Evaluations.* 87 (Fall 2000): 5-12.

Rosenbaum, Nancy. "Evaluation Myth I. Evaluation is too Expensive". Support Center for Nonprofit Management. 1-3. Retrieved Dec. 29, 2003 from http:/www.suportctr.org.

Rossi, Peter H., and Howard E. Freeman. Evaluation: A Systematic Approach, Newbury Park, CA: Sage Publications. (1999).

Rossi, Peter H., Mark Lipsey, and Howard E. Freeman. Evaluation: A Systematic Approach, 7th Edition. Thousand Oaks. CA: Sage Publications. (2004).
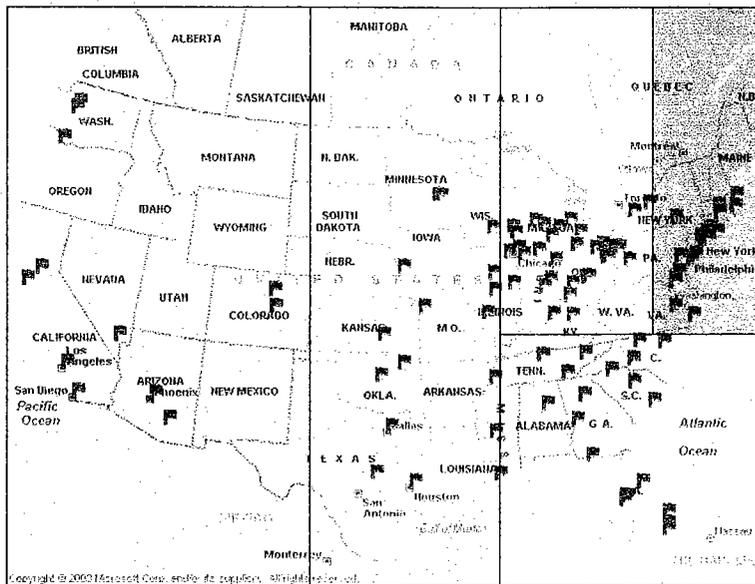
Sabo, Kim and Dana Fusco. "Participatory Evaluation for Continuous Improvement."
    The Evaluation Exchange: Harvard Family Research Project. 8 (2) (Fall 2002):
    12.

Sanow, Susan. "Searching and Discovering the Best in Nonprofit Management." *The
    Public Manager.* (Fall 2003). 23-26.

Sanders, James R. "A Vision for Evaluation." *American Journal of Evaluation.* 22 (3)
    (2002). 363-366.

Sawhill, John C. and David Williamson. "Mission Impossible? Measuring Success in
    Nonprofit Organizations." *Nonprofit Management & Leadership,* 11 (3) (Spring
    2001): 371-386.

Scriven, Michael. Evaluation Thesaurus. 4th Edition. Thousand Oaks, CA Sage
    Publications. (1991.

Shelby, James. President and Chief Executive Officer of the Sacramento, California
    Urban League. Interview by author, 9 April 2005, Legislative Policy Conference,
    Washington, D.C.

Smith, Steven R. "Nonprofit Organizations in Urban Politics and Policy in Nonprofit in
    Urban America by Richard Hula and Cynthia Jackson-Elmoore. (2000).

Spriggs, William E. and Suzanne Bergeron. The National Urban League and Social
    Welfare Policy: A Historical Perspective. The State of Black America of 2002.
    National Urban League. (2002).

The State of Black America 2004. National Urban League. (2004).

United Way of America. (2000). Agency Experiences with Outcome Measurement:
    Survey Results. Alexandria, VA. 16. Retrieved March 2005 from
    http:/www.national.unitedway.org.

Weiss, Carol H. On Theory-based Evaluation: Winning Friends and Influencing People.
    "Where We've Been and Where We're Going: Experts Reflect and Look
    Ahead." *The Evaluation Exchange: Harvard Family Research Project.* 9 (4)
    (Winter 2004) 2-3.

229

Weiss, Carol H. "Which Links in Which Theories Shall We Evaluate?" *New Directions for Evaluation.* 87 (Fall 2000): 35-45.

Weiss, Carol H. "Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families.": In James Connell, et al. (ed.), New Approaches to Evaluating Community Initiatives: Concepts, Methods and Contexts, Washington D.C.: The Aspen Institute. (1995).

Weiss, Carol H. Evaluation: Methods for Studying Programs and Policies. 2nd Edition. Prentice-Hall, Englewood Cliffs, New Jersey. (1998).

Wholey, Joseph S., Henry Hatry, and Kathryn E. Newcomer. Handbook of Practical Program Evaluation. San Francisco, CA. Jossey-Bass. (1994).

Wildavsky, Aaron. "The Self-Evaluating Organizations." *Public Administration Review.* (1972): 382-393.

Young, Whitney M. "The Urban League and Its Strategy." *Annals of the American Academy of Political and Social Science, The Negro Protest,* 357 (Jan. 1965): 102-107.

Zimmerman, Reid. (2003) Measuring Success and Making Change with Evaluation. 1 -4. Retrieved Dec. 19, 2003 from http://www.mncn.org/info/basic_eval.htm.

APPENDIX A

REGIONAL MAP OF
NATIONAL URBAN LEAGUE AFFILIATES

National
Urban League

Empowering Communities.
Changing Lives.

Regional Representation

4

APPENDIX B

LETTER FROM NATIONAL URBAN LEAGUE

NATIONAL URBAN LEAGUE

July 21, 2004

Dear Affiliate President & CEO:

It is with great pleasure that I write to support a research study developed by Valerie Thompson, President and CEO of Urban League of Greater Oklahoma City, Inc. entitled "Systematic Evaluation for Nonprofit Organizations: Problems, Prospects and Recommendations".

I encourage you to participate in the study during our national conference. Our goal is to further the understanding of evaluation utilization of nonprofit organizations.

The study is being provided to us at no cost and will greatly benefit the National Urban League and its affiliates as as we look forward to the applications and recommendations that will come from the research.

Thank you for your participation.

Very truly,

Marc H. Morial
President & CEO
National Urban League

MHM/geb

Barrier: Lack of Skills

factor 1 scale Lack of Skills

Barrier: Lack of Proper Focus

Lack of proper focus
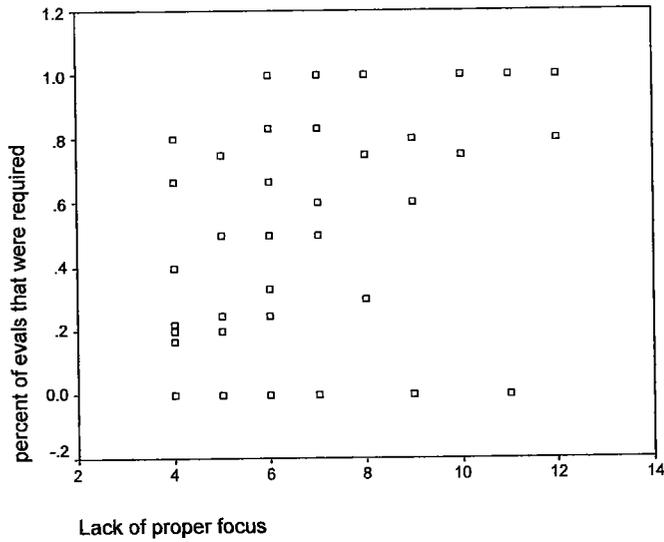
APPENDIX D

NONPROFIT EVALUATION SURVEY
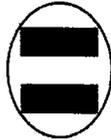
Telephone: (405) 424-5243
Fax: (405) 424-3382

Urban League of Greater Oklahoma City, Inc.
Urban League Community Center
3017 N. Martin Luther King Ave.
Oklahoma City, OK 73111

July 2004   Nonprofit Evaluation Survey

Dear Colleagues:

I am a graduate student under the direction of Professor Tom James in the Political Science Department at the University of Oklahoma-Norman Campus. I invite you to participate in a research study being conducted under the auspices of the University of Oklahoma-Norman Campus, entitled *Systematic Evaluation for Nonprofit Organizations: Problems, Prospects and Recommendations* and IRB #3191

The purpose of the study is to assess the current status of evaluation activities and the utilization of their results in the nonprofit sector. The study will seek to discover how much evaluation is done by nonprofits, what types of evaluations are being conducted, and how evaluations are being used. If evaluations are not being used, this study will try to answer the question why nonprofit organizations do not use evaluations. Additionally, the study will recommend how to promote the systematic evaluation of nonprofit organizations in light of the challenges and conditions in which they operate.

Your participation will involve the completion of a survey and should only take about 10 to 15 minutes to complete. Your involvement in the study is voluntary, and you may chose not to participate or to stop at any time. The results of the research study may be published, but your name or the name of your organization will not be linked to responses in publications that are released from the project. In fact, the published results will be presented in summary form only. All information you provide will remain strictly confidential.

The findings from this project will provide information on the theoretic understanding of evaluation utilization of nonprofit organizations by examining the use or non-use of systematic evaluations by examining the National Urban League affiliates. It also will make a contribution to the practical application of evaluation by providing a set of recommendations for nonprofit organizations, that don't use evaluations, to begin to systematically evaluate their programs. If you have questions about this research project, please feel free to call me at (405) 424-5243 extension 112 or email me at vrthompson@urbanleagueok.org. You can also contact Dr. Tom James at (405) 325-6622 or email Dr. James at tjames@ou.edu. Questions about your rights as a research participant or concerns about the project should be directed to the Institutional Review Board at the University of Oklahoma-Norman Campus at (405) 325-8110 or irb@ou.edu.

234

By returning this questionnaire in the envelope provided, you will be agreeing to participate in the above described project. **Please return the completed questionnaire to the CEO conference registration sites at the hotel or delegate assembly.**

Thanks for your consideration!

Sincerely

Valerie Thompson, President/CEO
Urban League of Greater Oklahoma City, Inc.
3017 N. Martin Luther King Avenue
Oklahoma City, OK 73111
(405) 424-5243 extension 112
Vrthompson@urbanleagueok.org

Please complete information and sign below.

Name_____ Affiliate Name_____

Address_____City_____State_____Zip_____Phone_____

Signature_____Date _____

For the purposes of this survey, the term evaluation is used quite broadly. It includes a variety of activities, such as: program evaluation, client or member satisfaction studies, impact analysis, outcome measurement, cost analysis and any other activities that help funders and nonprofit organizations assess *program* performance. The evaluation could have been conducted by your staff or could have been done through a contract with an individual or organization.

Section I: General Questions

1. In the past year, has your organization evaluated any of the following? (Please check all that apply)
   □ Products (anything developed by the agency for the public)
   □ Projects (one-time funded agency service)
   □ Ongoing programs or services (long-term programs existing at the agency (5 years or more))
   □ Fundraising activities
   □ Volunteer experiences
   □ Overall performance of the Board
   □ Your organization's overall effectiveness or impact
   □ Other, please specify _____
   □ None of the above (Skip to Question 25)

2. Approximately how many evaluations have been conducted by your organization over the 12 months? _____

3. Approximately how many of the evaluations that you conducted in the past 12 months were required by a funder? _____

4. What type of funder(s) required the evaluation?
   □ Federal Grant
   □ State Grant
   □ Local Grant
   □ Foundation Grant
   □ Other, please specify _____

5. Which of the following best describes how your organization generally conducts evaluations:
   □ Our evaluations are performed routinely and are a regular part of our organization's activities.
   □ We perform them only if we need to address specific issues (e.g. problems with a project or program)
   □ We perform them only when required to do so by funders.
   □ Other, please specify

   _____

   _____

6. If you had all of the resources needed to conduct evaluations, would evaluations be an on-going function at your organization? If not, why not?

_____

_____

Section II. Specific Project Questions

For the following series of questions, I would like you to refer to the last evaluation conducted by your organization, *excluding* the National Urban League's Performance Assessment.

7. What types of evaluation information were collected? Please check all that apply. Information about:
   - ◻ Project or program activities or outputs. For example, information about the number of clients served, number of hours served, or products produced.
   - ◻ Client or user information (demographic information)
   - ◻ Outcomes or impacts. For example, information about the changes the program or service produced in its clients or users.
   - ◻ Financial costs.
   - ◻ Other kinds of information? (Please specify)_____

_____

8. To what degree did your organization use the evaluation information that was obtained?

	Not at All	Not Very Much	Somewhat	A Great Deal
a. To improve program and services	◻	◻	◻	◻
b. To increase awareness of your organization or its cause	◻	◻	◻	◻
c. To report to the funder	◻	◻	◻	◻
d. For strategic planning purposes (e.g. program development, planning the next cycle of activity)	◻	◻	◻	◻
e. To share information with other similar organizations	◻	◻	◻	◻
f. For funding purposes	◻	◻	◻	◻
g. Other (please specify)				

_____

237

9. Which of the following is the main reason that your organization conducted this evaluation?
   a. It was required by funders.   □
   b. It was encouraged by funders.   □
   c. For accreditation purposes.   □
   d. It was a decision by our board for internal reasons.   □
   e. It was a decision by our staff for internal reasons.   □

If you selected answers c, d, or e, skip to question 15.

10.   How reasonable were the expectations that the funder had for the evaluation?
        Not at all   Not very   Reasonable   Very reasonable
           □   □   □   □
11.   To what degree did the funder require you to collect evaluation information that you otherwise
        would not have collected?

        Not at all   Not very much   Some what   A great deal
           □   □   □   □
12.   Was this a problem for your organization's evaluation efforts?
        Not a problem   Small problem   Moderate Problem   Big problem
           □   □   □   □

13.   To what extent did you use the evaluation information that the funder required for purposes other than
        to report to the funder?

        Not at all   Not very much   Some what   A great deal
           □   □   □   □
14.   Please check all of the following provided by the funder:

        □   Evaluation training support for your staff
        □   Evaluation tools and resources such as manuals or guidelines
        □   Advice about how to do evaluation
        □   Other types of support, please specify _____

238

15. Who was responsible for managing the evaluation?
- ☐ President/CEO
- ☐ Senior Administrative staff
- ☐ Program staff
- ☐ External evaluator
- ☐ Volunteers
- ☐ Student intern

16. Who was responsible for conducting the evaluation?
- ☐ President/CEO
- ☐ Senior Administrative staff
- ☐ Program staff
- ☐ External evaluator
- ☐ Volunteers
- ☐ Student intern

17. How satisfied were you with the quality of evaluation that was performed?

Very dissatisfied	Somewhat Dissatisfied	Somewhat Satisfied	Very Satisfied
☐	☐	☐	☐

18. How effectively do you think your organization used the evaluation information?

Not at all effectively	Not very effectively	Somewhat effectively	Very effectively
☐	☐	☐	☐

19. What was the approximate overall budget for the project, program, or service that was evaluated?

20. How adequate was the amount of money you had for evaluation purposes?

Not at all adequate	Not very adequate	Somewhat adequate	Very adequate
☐	☐	☐	☐

21. Where did the funding come from to enable your organization to undertake your last evaluation?

a. Internal funds from general revenue sources ☐
b. Internal evaluation fund ☐
c. From an external funder (e.g., government or foundation) ☐
d. From both internal and external sources ☐

239

22. An example of outcome or impact information is the change that programs or services produce in their clients or users. If your last evaluation collected such information, to what degree was it useful to your organization?

Not at all useful	Not very useful	Somewhat useful	Very useful
□	□	□	□

23. When conducting an outcome evaluation, how difficult was the following:

	Very Difficult	Somewhat Difficult	Not Very Difficult	Not at all Difficult
a. Identifying outcome measures	□	□	□	□
b. Collecting the outcome information	□	□	□	□
c. Analyzing the outcome information	□	□	□	□
d. Interpreting the outcome information	□	□	□	□

24. In your opinion, how accurate were the evaluation results in terms of reflecting what was going on in the project or program?

Not at all accurate	Not very accurate	Somewhat accurate	Very accurate
□	□	□	□

Section III: Evaluation Strengths and Barriers

The following series of questions address your organization's strengths with respect to evaluation and potential barriers limiting performance and evaluation effectiveness.

25. To what degree does your organization possess each of the following with respect to evaluation:

	No extent	Small Extent	Moderate Extent	Great Extent
a. Positive attitude towards evaluation	□	□	□	□
b. Evaluation skills and knowledge	□	□	□	□
c. Understanding of the activities of the projects and program	□	□	□	□
d. The ability to effectively communicate results	□	□	□	□
e. Understanding the needs of the community	□	□	□	□
f. Secure funding for evaluation	□	□	□	□

240

26. In general, thinking about all of your evaluation activities, how confident are you in your organization's ability to perform evaluation effectively?

Not at all confident	Not very confident	Somewhat Confident	Very confident
☐	☐	☐	☐

27. In your opinion, how much of a problem are each of the following for your organization?

	No problem	Small problem	Moderate problem	Big problems
a. Lack of money for evaluation	☐	☐	☐	☐
b. Lack of evaluation resources or tools (i.e., books or guides)	☐	☐	☐	☐
c. Unclear expectations from funders about what is expected in an evaluation.	☐	☐	☐	☐
d. Lack of confidence in the ability to perform evaluations	☐	☐	☐	☐
e. Lack of skills and knowledge in conducting evaluations.	☐	☐	☐	☐
f. Lack of internal capacity such as limited staff or lack of time	☐	☐	☐	☐
g. Lack of understanding the value of evaluation	☐	☐	☐	☐
h. Have difficulty working with evaluation consultants	☐	☐	☐	☐
i. Lack objectivity because organizations want to present their activities in a good light	☐	☐	☐	☐
j. Lack of access to university/college students who could help them with evaluation.	☐	☐	☐	☐
k. Lack of access to technology like computer software	☐	☐	☐	☐

l. Other, please specify _____

_____

Section IV. General Organization Questions

28. How many individuals are classified as full time equivalents? _____

29. In your opinion, how has funders' expectations about the amount of evaluation information that organizations should provide changed over the past three years?

Increased	Decreased	Remain the same	Don't know
☐	☐	☐	☐

30. How has funders' expectations changed about whether or not organizations should provide information to them about the outcomes or impacts of their programs or services. For example, the changes these programs produce in their clients or users?

Increased	Decreased	Remain the same	Don't know
☐	☐	☐	☐

Section V: General Questions about Nonprofit Organizations

The final set of questions address the background questions about your organization that will help us in the overall analysis of the results of this survey.

31. It is important to have an appropriate number of different size organizations responding to this survey. To help ensure this, please check which of the following categories of total approximate revenue from all sources during the last fiscal year describes your organization:

☐ <$125,000
☐ $125,000 to $249,000
☐ $250,000 to $349,000
☐ $350,000 to $499,000
☐ $500,000 to $749,000
☐ $750,000 to $899,000
☐ $900,000 to $999,999
☐ $1,000,000 to 1,249,000
☐ $1,250,000 to $1,499,000
☐ $1,500,000 to $1,749,999
☐ $1,750,000 to 1,999,999
☐ $2,000,000 to $3,000,000
☐ >$3,000,000

242

Section IV. General Organization Questions

28. How many individuals are classified as full time equivalents? _____

29. In your opinion, how has funders' expectations about the amount of evaluation information that organizations should provide changed over the past three years?

Increased	Decreased	Remain the same	Don't know
☐	☐	☐	☐

30. How has funders' expectations changed about whether or not organizations should provide information to them about the outcomes or impacts of their programs or services. For example, the changes these programs produce in their clients or users?

Increased	Decreased	Remain the same	Don't know
☐	☐	☐	☐

Section V: General Questions about Nonprofit Organizations

The final set of questions address the background questions about your organization that will help us in the overall analysis of the results of this survey.

31. It is important to have an appropriate number of different size organizations responding to this survey. To help ensure this, please check which of the following categories of total approximate revenue from all sources during the last fiscal year describes your organization:
- ☐ <$125,000
- ☐ $125,000 to $249,000
- ☐ $250,000 to $349,000
- ☐ $350,000 to $499,000
- ☐ $500,000 to $749,000
- ☐ $750,000 to $899,000
- ☐ $900,000 to $999,999
- ☐ $1,000,000 to 1,249,000
- ☐ $1,250,000 to $1,499,000
- ☐ $1,500,000 to $1,749,999
- ☐ $1,750,000 to 1,999,999
- ☐ $2,000,000 to $3,000,000
- ☐ >$3,000,000

243

32. Do you have access to any evaluation support such as resources (e.g. professional associations, informal networks, etc.), training, or advice besides anything available from your funders? Yes ▢ or No ▢.

33. How many staff have received evaluation training? _____

34. For those who have been trained, please describe the nature of the training (e.g. brief training (1 – 5 days), seminar/workshop, college course, etc. _____

35. Does your organization have an evaluator on site? Yes ▢ or No ▢.

36. Rank, 1 to 3, with number 1 being the most important based on your organization's mission, the following descriptions which best reflects the work that is done by your organization?
    a. Housing (multifamily development, single family development, counseling, etc.)   _____
    b. Education (pre-school, elementary, primary and secondary education, after-school programs, Summer school programs, etc.)   _____
    c. Health Services (public health and wellness education, screenings, health fairs, crisis intervention, treatment, etc.)   _____
    d. Employment Services (job training, job placement, career fairs, job referrals, etc.)   _____
    e. Economic Development (business plan development, feasibility studies, procurement services, etc.)   _____
    f. Civil Rights (advocacy, diversity training, etc.)   _____
    g. Other, please specify _____   _____

Thank you again for your generous time

This information will be provided with no cost to you other than the time it takes to complete the survey.
Please return the completed questionnaire to the CEO conference registration sites at the hotel or delegate assembly.

Sincerely,

Valerie Thompson, President/CEO
Urban League of Greater Oklahoma City, Inc.
3017 N. Martin Luther King Avenue
Oklahoma City, OK 73111
(405) 424-5243 extension 112
Vrthompson@urbanleagueok.org